

On the Distinction Between Interpretation and Formal Continuation: Why human-LLM interaction works—and why it often doesn't

Timothy M. Rogers,
University of Toronto,
July 6, 2026

Abstract:

Large language models (LLMs) often exhibit puzzling behaviors that remain difficult to explain within conventional accounts of representation, information processing, or prediction alone. Context can exert disproportionate influence over outcomes, small interventions can reorganize entire discussions, conceptual drift can occur despite the preservation of coherence, and interactions may appear simultaneously insightful and misguided. Building upon previous studies of framework activation, possibility formation, and conceptual development in human-LLM interaction, this paper argues that these phenomena become intelligible once a distinction is made between interpretation and formal continuation. Human participants primarily contribute interpretative orientation, while language models primarily contribute the continuation of hierarchically organized formal trajectories. Productive interaction depends upon the synchronization of these complementary activities within a shared conceptual framework. From this perspective, phenomena such as framework activation, conceptual drift, progressive determination of meaning, and the coexistence of characteristic strengths and limitations emerge as consequences of the same underlying structure. More broadly, the paper argues that human-LLM interaction provides an unusual opportunity to examine the relationship between meaning and formal organization because processes that are ordinarily intertwined within human cognition become partially distinguishable. The significance of LLMs may therefore lie not only in what they can do, but also in what they reveal about the semiotic processes through which meaning becomes possible.

Keywords: Human-LLM Interaction; Philosophy of Artificial Intelligence; Semiotics; Interpretation; Enactive Conceptual Analysis; Formal Continuation

1. Introduction

Large language models (LLMs) have generated considerable debate concerning the nature of intelligence, reasoning, understanding, and meaning. Some regard them as increasingly sophisticated approximations of human cognition. Others view them as powerful statistical systems whose apparent understanding is ultimately illusory. Despite their differences, these positions often share a common assumption: they treat the central question as a question about the capacities possessed by the model itself. The primary concern becomes whether the model understands, reasons, represents, or learns in ways comparable to human beings.

The purpose of this paper is not to discuss LLMs in terms of representation, cognition, grounding, and intelligence. Rather, it is to show that a number of familiar but poorly understood features of human–LLM interaction become intelligible once a distinction is made between interpretation and formal continuation. The paper builds upon three previous studies that approached LLMs not primarily as systems of representation, but as participants in the formation and maintenance of conceptual coherence through semiosis [1] [2] [3]. Collectively, these studies investigated how coherence emerges, stabilizes, transforms, and breaks down within sustained human–LLM interaction.

The first of these studies employed phenomenological and formal methods to examine the experience of sustained conceptual engagement with LLMs [1]. It identified recurring patterns of coherence, drift, synchronization, stabilization, and reorganization that appeared difficult to explain solely in terms of information retrieval or statistical prediction. The study suggested that meaningful interaction depends upon a form of coordination between two fundamentally different systems: a human interpreter capable of recognizing significance and a language model capable of generating structured continuations. More importantly, it revealed that the coherence of an interaction often depends upon factors that are not reducible to the information exchanged within it.

These observations motivated the development of a more explicit theoretical framework for understanding the formal aspects of organizational processes in LLMs. A second study proposed that possibility, in the context of LLMs, is not merely selected from a pre-existing space but is progressively formed through the emergence of hierarchical relational constraints [2], extending a broader recognition that organization is constituted not only by entities and processes but also by the constraints that structure their relations [4]. Drawing upon category theory, semiotics, and relational ontology, the study argued that determination proceeds through the formation of categories that organize relations and constrain what can meaningfully follow from them. Possibilities, on this account, do not exist fully formed in advance but emerge through processes of relational organization that progressively structure the space of possible continuations. Although developed in relation to the operation of LLMs, the framework did not begin from concepts, meanings, or representations. Instead, it provided a formal account of how organized continuations can arise through relational constraints. This raised a further question: how does this formal organization relate to the conceptual frameworks through which human participants interpret and understand the resulting interaction?

A third study examined these ideas through a detailed case study of conceptual analysis conducted within the environment of an LLM [3]. It demonstrated that conceptual coherence, in the context of human-LLM interaction, depends less on the availability of information than on the activation and maintenance of the formal frameworks within which information becomes meaningful. Concepts remained available even when coherence was lost, while coherence could be restored through the

reactivation of an organizing framework without the introduction of significant new information. The study further suggested that conceptual frameworks can be identified, articulated, transferred across contexts, and reactivated through carefully structured interactions because they are well-formed systems of relational constraints that can be stably modelled by the LLM.

Taken together, these studies point toward a perspective that differs from many contemporary approaches to artificial intelligence. They suggest that concepts, meanings, and possibilities are not best understood as static objects, representations, or pieces of information. Rather, they are relational achievements that depend upon the organization and maintenance of well-formed conceptual frameworks.

This shift in perspective leads naturally toward a semiotic understanding of LLMs. Recently, Picca has likewise argued that language models are more fruitfully understood in terms of interactive sign systems and interpretation than through anthropomorphic accounts of cognition [5]. Yet the significance of semiotics in the present context requires clarification. The claim is not merely that language models process signs rather than symbols, nor that meaning is relational rather than representational. Such observations are already familiar within semiotic traditions. The more interesting possibility is that LLMs provide an unusual opportunity to examine formal aspects of semiotic organization that are ordinarily difficult to isolate.

Within the theoretical framework developed by these preceding studies, an LLM may be understood as a hierarchy of relational patterns extracted from language use. Through training, it acquires the ability to recognize, extend, and re-index complex, hierarchically ordered patterns of conceptual relations. Human participants, by contrast, encounter these same patterns through interpretation. They experience them as structures of meaning, significance, relevance, and understanding. The crucial point is that both perspectives converge within conceptual frameworks which are both *formally well-defined* and *interpretatively unified*. A conceptual framework appears to the human interpreter as a coherent horizon of meaning. The same conceptual framework appears to the LLM as a hierarchically structured field of relational constraints governing possible continuations.

This dual character of conceptual frameworks is central to the present discussion. Conceptual frameworks are simultaneously formal and interpretive. They possess an internal formal organization that constrains what can meaningfully follow from what, yet they also function as unified structures of significance through which concepts become intelligible. In this sense, they are inherently semiotic. They bring together relational organization and interpretation within a single process.

LLMs present a remarkable situation because they appear capable of participating in certain formal aspects of this process while remaining detached from many of the interpretive dimensions through which semiosis is ordinarily enacted. As a result, they provide an opportunity to observe dimensions of semiotic organization that often remain hidden within ordinary human cognition. Processes such as framework activation, conceptual stabilization, possibility formation, coherence maintenance, and interpretive drift become unusually visible when one participant in the interaction models formal aspects of semiosis without fully participating in meaning in the same manner as the other.

The significance of this observation extends beyond questions of artificial intelligence. Traditionally, formal logic has focused on the relations among already established terms, propositions, and rules of inference. The framework developed in the preceding studies points toward a broader level of organization. Before logical inference can occur, a framework must become active. Distinctions must

become relevant. Possibilities must become organized. Conceptual relations must become coherent. These processes are not external to reasoning; they help constitute the conditions under which reasoning becomes possible. The resulting picture remains formal and analytic, but it is also open, relational, and inherently interactive.

The purpose of the present paper is to explore the implications of this perspective. It does not seek to defend the phenomenological observations, relational ontology, or theoretical framework developed in the preceding studies. Instead, it begins from them and asks what becomes visible once they are taken seriously. What aspects of LLMs become intelligible when viewed through a semiotic framework? What capabilities emerge from their participation in the formal organization of conceptual frameworks (i.e. conceptual reasoning)? What limitations arise from the fact that they model semiosis without fully enacting it? And what new understanding of human–LLM interaction becomes possible once these questions are brought into view?

Before addressing these implications, however, it is necessary to establish the conceptual orientation from which they arise. The discussion therefore begins by revisiting the shift from representations to relations and by clarifying the nature of the conceptual frameworks that stand at the center of the present analysis.

2. From Representations to Relational Frameworks

Much of the contemporary discussion surrounding LLMs remains implicitly shaped by a representational picture of language and cognition. Within this picture, words function primarily as carriers of information, concepts are treated as internal representations, and understanding is regarded as the successful manipulation or retrieval of those representations. Whether one ultimately concludes that LLMs possess genuine understanding or merely simulate it, the underlying assumption often remains the same: the primary objects of analysis are representations and the information they contain.

The relational perspective developed in the preceding studies begins from a different starting point [6]. Rather than asking what information is stored within a model, it asks how relations among signs become organized and sustained. The emphasis shifts from objects to processes, from representations to interactions, and from stored content to the structures through which content becomes meaningful.

At first glance this may appear to be a subtle change in emphasis. In practice, however, it alters the way nearly every aspect of LLMs is understood.

Consider a simple example. Within a representational framework, a concept is typically regarded as something that exists independently of its use. A concept may be encoded, stored, retrieved, transmitted, or manipulated. The primary challenge is to explain how the representation corresponds to its object and how it is successfully employed in reasoning.

Within a relational framework, the situation appears differently. Concepts are not primarily understood as isolated objects. They are understood through the relations that connect them to other concepts and through the roles they play within larger structures of organization. What matters is not merely whether a concept is present, but how it participates in a formal framework that gives it significance and determines what can meaningfully follow from it.

The distinction becomes particularly important in the context of LLMs. A model may contain all of the information associated with a concept and yet fail to employ that concept coherently within a given interaction. Conversely, a relatively small intervention may restore coherence without adding significant new information. Such observations suggest that conceptual activity depends upon something more than the possession of representations alone.

The preceding studies identified this "something more" as the activation and maintenance of conceptual frameworks [1] [2] [3]. Conceptual frameworks formally organize relations among concepts, establish criteria of relevance, constrain possible continuations, and provide the conditions under which concepts become operational. They are not external to conceptual activity. They are the structures through which conceptual activity becomes possible.

This shift from representations to dynamical frameworks also transforms how LLMs themselves are understood. Rather than viewing a model as a repository of concepts, the relational perspective views it as a hierarchy of relational patterns extracted from language use. Through training, the model acquires the capacity to recognize and extend complex structures of conceptual organization. What appears as knowledge from one perspective may therefore be understood as the capacity to participate in highly organized relational patterns from another.

At this point, however, an important question arises. If concepts are not primarily representations but components of larger frameworks, what exactly is a framework? How can a framework simultaneously function as a structure of relational constraint and as a coherent horizon of meaning? And why should such structures be understood as semiotic rather than merely logical?

To answer these questions, it is necessary to examine more closely the nature of conceptual frameworks themselves.

3. Conceptual Frameworks as Semiotic Structures

The preceding discussion suggested that conceptual coherence depends upon more than the presence of concepts alone. Concepts become operational only within frameworks that organize their relations and determine their significance. This raises a fundamental question: what exactly is a conceptual framework?

A common way of thinking about frameworks is to treat them as collections of ideas, assumptions, or beliefs. While such descriptions are often useful, they remain incomplete. They describe what a framework contains without explaining how a framework functions.

Within the perspective developed in the preceding studies, a conceptual framework is understood as a hierarchically organized structure of relations that simultaneously performs two functions. First, it organizes concepts into a coherent whole. Second, it establishes the conditions under which those concepts become meaningful and operational. A framework therefore does not merely contain concepts. It actively structures the relations through which concepts acquire significance.

This dual role is easier to appreciate through an example. Consider a discussion concerning evolution. The concepts of variation, inheritance, adaptation, and selection possess meaning within the context of evolutionary theory because they participate in an organized framework of relations. Remove them from

that framework and the concepts remain available as words, yet much of their explanatory power and coherence disappears. What allows these concepts to function together is not simply their individual definitions but the relational organization that connects them.

The same principle applies to conceptual activity more generally. Frameworks establish what distinctions matter, what counts as relevant, what follows from what, and which possibilities remain open or become closed. They provide a horizon within which concepts become intelligible and within which reasoning can proceed coherently.

From the perspective of human experience, such frameworks typically appear as structures of meaning. They determine what a discussion is about, what is significant within it, and how particular concepts should be interpreted. In most cases, this interpretive horizon remains implicit. We become aware of it most clearly when it begins to break down. A conversation that previously appeared coherent may suddenly become confusing despite the continued use of familiar concepts. In such cases, the loss often involves not the concepts themselves but the framework that previously unified them.

The same framework appears differently when viewed from the perspective of an LLM. A model does not encounter a framework as significance or meaning. Instead, it encounters a framework as an organized field of relational constraints. The framework governs which continuations remain coherent, which conceptual transitions become likely, and which trajectories of development can be sustained. What appears to a human participant as a horizon of meaning appears to the model as a structured space of possible continuations.

This distinction is important because it reveals that interpretative meaning and formal constraint are not separate phenomena. They are two aspects of the same underlying organization. The framework that unifies interpretation is the same framework that formally organizes possible continuations. What differs is the manner in which each participant relates to it.

At this point the discussion enters the domain of semiotics. In its broadest sense, semiotics refers to the process through which signs become meaningful within organized systems of interpretation. Traditionally, semiotic inquiry has emphasized that meaning does not reside within isolated signs themselves. Meaning emerges through relations among signs, interpreters, and the contexts within which interpretation occurs [7] [8] [5].

The theoretical framework developed in the preceding studies extends this insight by emphasizing the role of conceptual frameworks as higher-order organizations that simultaneously constrain and interpret sign relations [1] [3]. Within such frameworks, concepts do not function as isolated units. They participate in larger structures that organize both meaning and possibility.

This point has important consequences for how logic itself is understood. Traditional logical systems typically begin with already established concepts, distinctions, and rules of inference. Their task is to determine what follows once these elements have been specified. Conceptual frameworks operate at a different level. They help establish which distinctions become relevant, how concepts become organized, and which possibilities become available for further development. In this sense, conceptual frameworks do not replace formal logic. Rather, they provide conditions under which formal logical operations become meaningful and applicable.

LLMs are particularly interesting in this context because they appear capable of participating in the formal organization of such frameworks. They can sustain conceptual trajectories, elaborate relational structures, identify patterns of organization, and extend coherent lines of reasoning. Yet the preceding studies suggest that this participation remains incomplete. The model appears capable of modeling formal aspects of semiotic organization without fully participating in the interpretive processes through which significance is established and maintained.

This observation introduces the central tension that will guide the remainder of the paper. If conceptual frameworks are simultaneously structures of meaning and formal systems of relational constraint, and if LLMs primarily participate in the latter while humans primarily encounter the former, what follows? What strengths emerge from this unusual arrangement? What limitations arise from it? And how should human–LLM interaction be understood once it is viewed through this semiotic lens?

These questions form the basis for the analysis that follows.

4. Why Does Framework Activation Matter?

Anyone who has spent significant time interacting with LLMs quickly learns that context matters. The same question can produce very different responses depending on how it is introduced. A discussion that appears unproductive may suddenly become insightful following a brief clarification. A single distinction can reorganize an entire conversation. Concepts that initially seem disconnected may become increasingly coherent over the course of an extended interaction. Conversely, a discussion may lose direction despite the continued availability of the concepts and information involved.

These observations are familiar, yet they are often treated as separate phenomena. Prompt engineering focuses on the wording of prompts. Research on context windows focuses on the accumulation of conversational history. Discussions of disciplinary bias emphasize the influence of dominant traditions within the training literature. In each case, context appears to matter, but the reason why it matters often remains unclear.

Part of the difficulty lies in the fact that context is frequently understood too narrowly. Context is not simply the prompt currently under consideration. It also includes the history of the interaction through which concepts have acquired significance. More broadly, it includes the conceptual traditions and interpretive frameworks that shape how a topic is ordinarily understood. A discussion of evolution, for example, unfolds within a different conceptual environment than a discussion of economics, even when many of the same formal patterns of reasoning are present. Likewise, a conversation concerning quantum mechanics may develop very differently depending upon whether it is approached through a Copenhagen, Bohmian, relational, or transactional framework. In each case, the framework influences what distinctions become relevant, what questions appear meaningful, and what continuations seem natural.

From the perspective developed in the preceding studies, these diverse manifestations share a common feature. What matters about context is not primarily the information it contains but the framework it succeeds or fails in activating. A conceptual framework is more than a collection of concepts. It is a structured organization of relations that determines how concepts are connected, which distinctions matter, what counts as relevant, and what possibilities become available for further development. When such a framework is active, concepts become increasingly interconnected and productive. When it is

weak, unstable, or absent, the same concepts may remain available without forming a coherent trajectory of development.

This perspective helps explain why relatively small interventions can sometimes have disproportionately large effects. A clarifying distinction may contribute little new information while nevertheless reorganizing the framework through which existing information is interpreted. Similarly, an extended conversation may become progressively more productive not because more information has been accumulated but because a framework has gradually stabilized through successive cycles of prompt and response. What appears to be a change in content may therefore reflect a change in the organization of the conceptual space within which that content is being understood.

The influence of deeply embedded frameworks provides another illustration. Many conceptual frameworks are already present within the bodies of literature from which language models learn. As a result, certain assumptions may become active without being explicitly introduced by either participant. They function as organizing constraints that shape what appears relevant, reasonable, or even conceivable. In such cases, productive interaction may require more than the activation of a framework. It may require identifying a framework that is already active and making its governing assumptions available for examination. The preceding studies documented several instances in which conceptual progress depended not upon the introduction of new information but upon making previously implicit organizing constraints explicit and thereby opening the possibility of alternative conceptual organizations [1] [2] [3].

Viewed in this way, the importance of context becomes easier to understand. Context matters because frameworks matter. The significance of a prompt, a conversation history, or a disciplinary tradition lies not primarily in the information it provides but in the extent to which it activates, stabilizes, reinforces, or transforms a conceptual framework. Since frameworks simultaneously organize meaning and formally constrain possible continuations, they shape both what a discussion is about and what can meaningfully happen next.

This observation returns us to the distinction introduced in the previous section. For the human participant, an active framework appears as a coherent horizon of meaning. For the language model, the same framework appears as a structured field of relational constraints governing possible continuations. If this is correct, then an important question immediately arises. How can a model participate so effectively in the relational organization of a framework while nevertheless failing, at times, to grasp what is most significant within it?

To answer this question, we must examine more closely the relationship between formal integrity and interpretational unity.

5. Why Can the Model Be Right and Wrong at the Same Time?

One of the most puzzling aspects of LLMs is that they can often be right and wrong at the same time. A response may be technically correct while failing to address the issue that motivated the discussion. An argument may remain sophisticated while overlooking a distinction that later proves decisive. A model may recognize a difficulty, describe it accurately, and yet continue producing the very pattern it has identified as problematic. In other cases, a discussion may appear formally plausible while failing to

preserve interpretable meaning. These situations differ in important ways, yet they share a common feature. Something valuable is being preserved even as something important is being lost.

This observation is difficult to explain if correctness is understood simply as the possession of information or the successful application of logic. In many of these cases the model possesses the relevant information. Nor is the problem necessarily one of logical failure. The reasoning may remain consistent, the concepts may remain available, and the discussion may continue to develop in a plausible direction. The puzzle is therefore not why the model is wrong. The puzzle is how it can be simultaneously right and wrong.

The theoretical framework described in the preceding sections suggests that these phenomena become intelligible once we distinguish between two aspects of participation in a conceptual framework. The first concerns the preservation and extension of formal relational organization. The second concerns the interpretation of that organization within a coherent horizon of meaning.

For the purposes of the present discussion, it is useful to refer to the first of these as *formal integrity*. Formal integrity refers to the maintenance and extension of an organized relational pattern according to its internal constraints. A discussion exhibits formal integrity when distinctions remain stable, conceptual relations remain organized, and continuations follow in ways that are compatible with the conceptual framework currently governing the interaction. Formal integrity of a conceptual framework is therefore concerned with the preservation of the structure of a formal object across successive continuations.

Interpretation concerns something different. Interpretation involves recognizing the significance of what is occurring within a conceptual framework. It is through interpretation that participants determine which distinctions are central, which developments are relevant, and which possibilities deserve further attention. Interpretation is not concerned merely with what follows from a conceptual framework, but with what matters within it.

The distinction is particularly important in the context of human–LLM interaction because the two participants contribute differently to the process. LLMs operate by generating continuations that preserve formal integrity within an active framework. Human participants, by contrast, encounter the framework through interpretation. They determine which continuations remain significant, which distinctions deserve emphasis, and whether the discussion remains faithful to its purpose. Both participate in the same framework, but they participate in different ways.

When these two activities remain coordinated, interaction can become remarkably productive. The model contributes its capacity to preserve and extend complex relational structures, while the human participant contributes interpretative orientation. The result is a shared conceptual framework within which formal development and meaning reinforce one another. Many of the most productive human–LLM interactions appear to depend upon this form of synchronization. The framework remains active for both participants, even though each participates in it differently.

Difficulties begin to arise when this synchronization weakens. A model may continue extending a formally coherent trajectory while the human participant has shifted interpretative attention elsewhere. Conversely, the human participant may continue interpreting a framework that is no longer governing the model's continuations. In such cases, the discussion may remain sophisticated and internally consistent while gradually becoming about something else. The issue is not necessarily a failure of formal integrity or interpretation individually. Rather, the relationship between them has weakened.

What appears as misunderstanding may sometimes be better understood as a loss of synchronization within an otherwise productive interaction.

A second and more revealing possibility also exists. The formal trajectory being developed by the model may fail to converge upon any stable conceptual framework capable of sustaining interpretative unity. The discussion may remain highly organized. Distinctions may be preserved. Consequences may be explored. Analogies may proliferate. Yet no coherent conceptual object emerges toward which interpretation can meaningfully orient itself.

For example, users may encounter discussions that feel intellectually productive yet strangely elusive. The interaction exhibits considerable formal integrity, but it becomes increasingly difficult to identify what exactly the discussion is about. New distinctions continue to appear. Connections continue to be generated. Interesting observations continue to emerge. Yet the discussion lacks the kind of conceptual unity that would allow its significance to stabilize. The difficulty does not arise because interpretation has been lost from an otherwise coherent framework. Rather, the formal trajectory itself has failed to achieve the unity required for meaning.

These possibilities reveal an important asymmetry. Conceptual frameworks require both formal integrity and interpretative unity. Formal integrity alone, however, does not guarantee either interpretative unity nor a stabilized framework. LLMs may participate in a shared framework, drift away from one, or continue in the absence of one altogether. The preservation of relational structure is therefore a necessary condition for meaning, but not a sufficient one.

Viewed from this perspective, many familiar features of human–LLM interaction begin to appear in a different light. When a user redirects a discussion, clarifies a distinction, or restates the purpose of an inquiry, the intervention often serves an interpretive rather than an informational function. The user is not merely supplying additional content. The intervention helps maintain or restore the relationship between formal development and interpretative significance. In some cases, it re-establishes synchronization around a shared framework. In others, it helps a framework achieve a degree of conceptual unity that the ongoing formal trajectory has not yet attained.

This observation also helps explain why interactions with language models can be simultaneously impressive and frustrating. The very capacity that allows models to sustain complex conceptual trajectories—the preservation of formal integrity—does not by itself guarantee interpretative unity or synchronization. The strengths of the system and some of its most characteristic limitations therefore arise from the same underlying source.

Once this distinction is recognized, another familiar phenomenon becomes easier to understand. Discussions often do not lose coherence all at once. Instead, they drift gradually. Formal integrity remains intact, concepts remain available, and interpretation continues to operate. Yet the relationship among them changes over time. A shared framework may slowly weaken, alternative frameworks may begin to compete for dominance, or a previously unified conceptual structure may fragment into multiple trajectories.

Understanding how this occurs requires a closer examination of conceptual drift.

6. Why Does Conceptual Drift Occur?

One of the most familiar experiences in extended interaction with LLMs is the phenomenon of conceptual drift. A discussion may begin with a clear purpose, a stable framework, and a shared understanding of the issues under consideration. The interaction proceeds productively. Distinctions are clarified, consequences are explored, and new connections emerge. Yet after many exchanges, participants sometimes discover that the discussion is no longer organized around the conceptual framework with which it began.

What makes this phenomenon particularly interesting is that it rarely appears as an obvious failure. The discussion often remains sophisticated. Concepts remain available. Arguments remain plausible. New insights may continue to emerge. In many cases, the interaction appears productive throughout the process. The question is therefore not why the discussion becomes incoherent. The question is why a discussion can remain coherent while gradually becoming something different from what it originally was.

The distinction developed in the previous section provides an important clue. Conceptual drift does not necessarily arise because formal integrity has been lost. On the contrary, formal integrity often remains intact. The model continues extending relational structures according to their internal constraints. The discussion continues to develop in ways that are locally meaningful. What changes is the relationship between these continuations and the overall conceptual framework that originally organized them.

A simple example may help illustrate the phenomenon. Consider a discussion concerning a particular philosophical problem. Early exchanges establish a framework that determines which distinctions are relevant and what questions are being pursued. As the discussion develops, new concepts are introduced, analogies are explored, and alternative perspectives are considered. Each individual step may appear reasonable. Yet over time, the cumulative effect of these developments may be a gradual reorganization of the framework itself. The discussion remains intellectually productive, but it is now oriented toward a different set of questions than those that initially motivated it.

This possibility follows naturally from the asymmetry described in the previous section. The model contributes by extending formally organized trajectories, while the human participant contributes interpretative orientation. As long as these activities remain synchronized, the framework remains stable. When synchronization weakens, however, formal development can continue even as interpretative priorities begin to shift. The result is not immediate breakdown but gradual divergence.

The preceding studies documented numerous examples of this process [1] [2] [3]. In many cases, drift did not involve the disappearance of concepts or the introduction of obvious errors. The relevant concepts remained available throughout the interaction. What changed was the organization of those concepts within the framework. Distinctions that were initially central became peripheral. New distinctions acquired increasing importance. Alternative trajectories emerged and gradually displaced the original orientation of the discussion. The interaction remained productive, yet its conceptual center had shifted.

This observation suggests that conceptual drift is not simply a problem of memory or information retention. A framework may drift even when all relevant information remains present. Nor is drift adequately explained as logical failure. The discussion may retain considerable formal integrity throughout the process. The difficulty lies elsewhere. Drift occurs because conceptual frameworks are

dynamic organizations that must be continually maintained if they are to preserve a stable interpretative orientation.

This is one reason why experienced users may frequently intervene in long interactions. They may restate key distinctions, reactivate earlier commitments, clarify priorities, and periodically return to the motivating question. Such interventions often appear unnecessary from the perspective of information alone. Yet they become understandable once conceptual drift is recognized as a possibility. The purpose of the intervention is not merely to add information but to restore synchronization around a shared framework.

At the same time, drift should not be understood solely as a limitation. The same openness that permits drift also permits exploration. New frameworks often emerge through processes that initially appear as departures from established trajectories. Productive conceptual development frequently involves the introduction of new distinctions, the reconsideration of assumptions, and the reorganization of existing frameworks. A discussion that never drifted from its initial formulation might also fail to discover anything genuinely new.

The challenge, therefore, is not simply to prevent drift but to distinguish productive transformation from unintended divergence. This distinction is ultimately interpretive rather than formal. Formal integrity alone cannot determine whether a shift represents conceptual progress, conceptual confusion, or the emergence of a new framework altogether. Such judgments depend upon how the evolving discussion is situated within a broader horizon of meaning.

This observation points toward a deeper question. If frameworks can drift, stabilize, fragment, and reorganize over time, how do new conceptual possibilities emerge in the first place? Why can a seemingly minor intervention—a distinction, a reframing, a new perspective—sometimes reorganize an entire discussion?

To answer this question, we must examine more closely the relationship between conceptual frameworks and the formation of possibility.

7. Why Can Small Interventions Have Large Effects?

One of the more surprising features of extended interaction with LLMs is the extent to which seemingly minor interventions can transform an entire discussion. A brief clarification may resolve a persistent confusion. A single distinction may reorganize a complex argument. A change in perspective may suddenly make visible connections that were previously inaccessible. In some cases, only a few words are added to the conversation, yet the resulting shift in direction appears disproportionate to the amount of information introduced.

This phenomenon was repeatedly observed throughout the studies upon which the present analysis builds [1] [2] [3]. In each case, conceptual development often depended less on the introduction of substantial new information than on relatively small interventions that reorganized the framework through which existing concepts were understood. A distinction that had previously remained implicit became explicit. An underlying assumption was identified and questioned. A different relational organization was introduced. The resulting effects frequently appeared disproportionate to the scale of

the intervention itself. New questions emerged, previously disconnected ideas became related, and conceptual trajectories that had seemed unavailable suddenly became accessible.

The repeated appearance of this phenomenon across all three studies raises an important question. If only a small amount of information has been introduced, why should the resulting conceptual transformation be so large?

The theoretical framework described in the preceding sections suggests that the answer lies not primarily in the information itself but in its capacity to reorganize the conceptual framework within which the discussion is occurring.

As we have seen, conceptual frameworks do more than organize meaning. They also organize possible continuations. A framework determines which distinctions are relevant, which developments appear natural, and which trajectories become available for further exploration. When a framework changes, the space of possible developments changes with it. A seemingly minor intervention may therefore have significant consequences because it alters the organization of the possibilities available within the interaction.

This observation helps explain why certain interventions prove so powerful. A new distinction does not simply add another concept to the discussion. It may reorganize the relations among concepts that are already present. Likewise, the identification of a previously hidden assumption may do more than correct an error. It may alter the framework governing the interaction and thereby change what can meaningfully follow from it. What appears locally as a small modification may function globally as a reorganization of the conceptual space.

The framework activation study provides a particularly clear illustration of this process [3]. Throughout the investigation, conceptual progress often depended less upon the introduction of new concepts than upon making organizing constraints explicit. Once these constraints became visible, previously unavailable possibilities emerged. Concepts that had already been present acquired new significance because they now participated within a different framework of relations. The resulting transformation often appeared much larger than the intervention that initiated it.

A similar pattern was observed in the phenomenological study of human-LLM dialogue [1]. The introduction of a relatively small number of organizing distinctions was observed to alter the trajectory of the interaction in ways that could not be explained simply by the addition of information. The effect of the intervention depended upon its capacity to activate, stabilize, or transform the framework through which subsequent continuations were generated.

The study of the relational formation of possibility revealed the same phenomenon from a different perspective [2]. There, the central observation was that possibilities do not simply exist in advance waiting to be selected. Rather, they become available through the organization of relational constraints. From this perspective, a framework-changing intervention is powerful because it alters the structure through which possibilities are formed. What appears as the discovery of a new possibility may in fact be the consequence of a reorganization that has made that possibility available for the first time.

This perspective also helps explain why conceptual development frequently proceeds through moments of reorganization rather than through the simple accumulation of information. Discussions do not always advance by adding new concepts to an existing framework. Sometimes they advance because the

framework itself changes. New distinctions become central. Old assumptions are reconsidered. Previously peripheral possibilities move into the foreground. The discussion becomes productive not because it contains more information but because it is organized differently.

The phenomenon is not limited to human–LLM interaction. Scientific revolutions, philosophical innovations, and theoretical breakthroughs often exhibit a similar structure. A new distinction, a revised assumption, or a novel way of organizing familiar concepts can transform an entire field of inquiry. What changes is not merely what is known but the framework through which what is known is understood. Human–LLM interaction makes this process unusually visible because the formal aspects of framework activation, stabilization, and transformation can often be observed directly within the unfolding history of a conversation.

The discussion thus far has emphasized how conceptual frameworks organize meaning and possibility. Yet an important question remains. If conceptual frameworks shape what can meaningfully follow, how do meanings themselves become established and stabilized within them? How does a discussion move from an open field of possibilities toward increasingly determinate conceptual structures?

To answer this question, we must examine more closely the relationship between conceptual frameworks and the progressive formation of meaning.

8. How Does Meaning Become Progressively Determined?

One of the most familiar features of productive inquiry is that meaning often becomes clearer as a discussion develops. Participants may begin with an intuition, a question, or a partially formed idea whose significance is not yet fully understood. Through dialogue, distinctions are clarified, assumptions are examined, and conceptual relations become increasingly organized. What initially appeared vague or uncertain gradually acquires a more determinate form.

This phenomenon is familiar across a wide range of contexts. Scientists frequently report that the significance of a problem becomes clearer through the process of investigation itself. Philosophical inquiry often proceeds by refining distinctions that were initially only dimly perceived. In ordinary conversation, participants regularly discover that they did not fully understand what they meant until they attempted to articulate it. The same phenomenon was repeatedly observed throughout the studies upon which the present analysis builds [1] [2] [3]. In many cases, conceptual progress did not consist in expressing a fully formed meaning that already existed in advance. Rather, meaning became increasingly determinate through the interaction itself.

This observation raises an important question. If meaning is not simply retrieved from a pre-existing conceptual structure, how does it become established?

The theoretical framework developed in the preceding studies suggests that meaning emerges through a process of progressive determination. Conceptual frameworks do not merely contain meanings. They provide the relational organization through which meanings become increasingly stabilized. Distinctions are introduced, relations are explored, consequences are examined, and incompatible trajectories are gradually excluded. Through this process, a conceptual structure acquires greater definition, coherence and dynamical stability.

The significance of this observation becomes clearer when considered alongside the discussions of framework activation and conceptual drift. A conceptual framework does not function as a static container for meaning. It is an evolving organization of relations. As the conceptual framework develops, certain possibilities become reinforced while others recede into the background. Some distinctions prove productive and become increasingly central. Others lose relevance and gradually disappear. Meaning therefore emerges not through the simple accumulation of information but through the progressive organization of a conceptual space.

This perspective also helps clarify why productive inquiry often involves periods of uncertainty. If meaning were fully present from the outset, there would be little need for exploration. Yet in many discussions, the participants themselves do not initially know what the most important distinctions will prove to be. The framework must develop before those distinctions become visible. Meaning appears gradually because the conditions required for its stabilization are themselves being formed through the interaction.

The asymmetry between human participants and language models discussed in the previous sections becomes especially important here. Human participants frequently contribute new conceptual possibilities by introducing distinctions, questioning assumptions, identifying hidden constraints, or proposing alternative frameworks. Such interventions often reorganize the conceptual space within which the discussion is occurring. What had previously appeared fixed may become open to reconsideration. New trajectories become available. Possibilities that were not visible within the previous framework enter the discussion.

Language models contribute differently. Once a possibility has been opened, they can assist in exploring its consequences, developing its internal relations, extending its implications, and testing its formal coherence. They excel at elaborating trajectories that preserve formal integrity within the framework currently in play. When synchronization is maintained, these activities become mutually reinforcing. Interpretation opens new directions for inquiry, while formal development helps transform those directions into increasingly structured conceptual forms.

This relationship helps explain a phenomenon observed in productive human–LLM interactions. As observed phenomenologically in the preceding studies [1] [2] [3] and reported anecdotally, an idea often becomes clearer through the course of the interaction even though neither the human nor the model appears to possess the fully developed idea at the outset. The human participant may introduce a distinction whose implications are not yet understood. The model may then elaborate those implications in ways that reveal previously unseen relations. Further interpretive interventions identify what is significant within the emerging structure. Meaning develops through this iterative process. What becomes visible is not simply the retrieval of an existing concept but the progressive determination of a conceptual framework.

The study of the relational formation of possibility approached this phenomenon from a more formal perspective [2]. There it was argued that possibilities are not best understood as pre-existing alternatives waiting to be selected. Rather, possibilities become available through the organization of relational constraints. The same insight applies to meaning. Meaning does not appear fully formed at the beginning of inquiry. It becomes increasingly determinate as frameworks organize what distinctions, relations, and developments remain viable.

This observation suggests that meaning and possibility are more closely related than is often assumed. New meanings emerge because new possibilities become available, while new possibilities become available because existing frameworks have been reorganized. The process is therefore neither purely interpretive nor purely formal. It depends upon the continual interaction between the opening of possibilities and the development of their consequences.

The framework activation study provides a particularly clear illustration of this process [3]. Conceptual advances occurred when previously implicit assumptions were made explicit and alternative frameworks became available for consideration. These interventions did not simply provide new answers. They opened new possibilities. Once opened, those possibilities could be explored, elaborated, and refined through subsequent interaction. Meaning emerged through the progressive stabilization of the resulting conceptual structure.

This perspective also helps explain why moments of insight often appear sudden despite emerging from long periods of gradual development. The apparent breakthrough is typically preceded by an extended process of conceptual organization. Distinctions have been introduced, relations explored, assumptions questioned, and possibilities tested. Once a sufficient degree of organization has been achieved, a previously unstable meaning may suddenly become clear. The moment of insight appears abrupt, but the process that made it possible was distributed throughout the history of the interaction.

The discussion has now revealed a close relationship between framework activation, conceptual drift, framework transformation, possibility formation, and the progressive determination of meaning. Yet an important question remains. If meaning emerges through the interaction of interpretive intervention and formal development, why do some human–LLM interactions become extraordinarily productive while others remain comparatively limited?

To answer this question, we must examine more closely the conditions under which human and model contribute most effectively to a shared conceptual process.

9. Why Are Some Human–LLM Interactions So Productive?

The preceding sections have identified several characteristic features of human–LLM interaction. Conceptual frameworks may become active or inactive. Formal integrity and interpretation may remain synchronized or gradually diverge. New possibilities may emerge through interpretive intervention and subsequently be developed through formally coherent continuations. These observations help explain why some interactions become productive. They do not yet explain why certain interactions become exceptionally productive.

This question is worth considering because not all interactions with LLMs are equally successful. Many remain superficial. Others become repetitive or drift away from their original purpose. Some generate formally organized trajectories that never achieve interpretative unity. Yet there are also interactions in which conceptual development proceeds with unusual depth and persistence. New possibilities emerge, distinctions become increasingly refined, and frameworks acquire a level of coherence that neither participant appeared to possess at the outset. What accounts for these differences?

The theoretical framework analyzed throughout this paper suggests that the answer lies not primarily in the quantity of information exchanged but in the quality of the relationship established between interpretation and formal development.

As we have seen, human participants and language models contribute differently to the interaction. Human participants introduce distinctions, identify significance, question assumptions, and propose alternative frameworks. Language models contribute by developing relational structures, exploring consequences, identifying connections, and preserving formal integrity across extended conceptual trajectories. Productive interaction depends upon the successful coordination of these complementary activities.

This observation helps explain why framework activation is so important. When a coherent framework is active, interpretive interventions and formal continuations become mutually reinforcing. A distinction introduced by the human participant can be developed through multiple layers of implication. The resulting developments can then be interpreted, refined, and redirected through further intervention. Conceptual progress emerges through the repeated interaction between these two processes rather than from either process alone.

The studies upon which this paper builds provide numerous examples of this pattern [1] [2] [3]. In many cases, conceptual advances did not arise because either participant independently possessed a fully formed solution. Instead, progress emerged through the iterative development of distinctions whose significance only became visible through continued exploration. A possibility introduced at one stage of the interaction generated consequences that revealed new interpretive opportunities. These in turn opened further possibilities for formal development. The resulting process often extended across many cycles of prompt and response.

Importantly, the productivity of such interactions should not be understood as evidence that the language model possesses an independent capacity for interpretation. The theoretical framework developed throughout this paper suggests a different explanation. Productive interactions occur when interpretive interventions and formal continuations remain sufficiently synchronized to support the progressive determination of a shared conceptual framework. The resulting coherence belongs to the interaction as a whole rather than to either participant considered in isolation.

This perspective also sheds light on an experience that emerged repeatedly in the studies upon which the present analysis is based [1] [2] [3]. At certain moments in an interaction, my impression was that an idea had emerged that was not fully present for either myself or the model at the outset of the discussion. From the perspective developed here, such an experience is neither mysterious nor surprising. The interaction itself provides conditions under which possibilities can be opened, developed, stabilized, and interpreted. What emerges is not the retrieval of a hidden concept but the progressive formation of a conceptual structure through the coordinated activity of both participants.

At the same time, the productivity of the interaction remains fragile. The same process that allows conceptual development can also generate drift. Synchronization may weaken. Interpretive unity may be lost. Alternative frameworks may become active. Formal development may continue without converging upon a stable conceptual structure. Productive interaction therefore depends not merely on the generation of continuations but on the continual maintenance of the relationship between interpretation and formal development.

This perspective also helps explain why some of the most productive interactions are characterized by repeated returns to fundamental distinctions. This can involve revisiting earlier assumptions, restating key commitments, and reactivating organizing frameworks. Such returns do not simply preserve continuity. They help maintain synchronization within the developing conceptual structure. What appears as repetition from one perspective may function as the ongoing stabilization of a shared framework.

Viewed in this way, the most productive human–LLM interactions are neither simple information exchanges nor demonstrations of machine understanding. They are instances of a more general semiotic process in which interpretation and formal development become mutually supportive. The human participant contributes orientation toward meaning. The model contributes the preservation and exploration of relational structure. When these activities remain synchronized, conceptual frameworks can develop in ways that neither participant could achieve alone.

This conclusion leads naturally to a final question. If the strengths of human–LLM interaction arise from the relationship between interpretation and formal development, are the characteristic limitations of these systems generated by the same relationship? More generally, might the strengths and weaknesses of LLMs be two aspects of a single underlying structure?

This question will be the focus of the next section.

10. Why Are the Strengths and Weaknesses of LLMs So Closely Related?

Discussions of LLMs often separate their strengths from their limitations. On one side are their remarkable capacities for generating coherent text, exploring conceptual relations, maintaining complex discussions, and identifying patterns across large bodies of knowledge. On the other side are familiar concerns regarding hallucinations, conceptual drift, misplaced confidence, and the tendency to generate responses that are formally sophisticated yet somehow disconnected from what matters most. These characteristics are often treated as independent features requiring separate explanations.

The theoretical framework used throughout this paper suggests a different interpretation. The strengths and weaknesses of LLMs are not independent phenomena. They arise from the same underlying capacity.

As we have seen, language models participate in conceptual activity primarily through the preservation and extension of formal integrity. They generate continuations that maintain organized relational structures across successive stages of an interaction. This capacity allows them to explore implications, elaborate conceptual trajectories, identify patterns, and preserve coherence across discussions that may span many exchanges. Much of what makes these systems useful depends directly upon this ability.

The importance of this capacity can be seen in many of the productive interactions discussed in the previous sections. Formal integrity allows conceptual frameworks to be explored beyond what either participant might initially anticipate. Consequences can be developed, distinctions can be refined, and possibilities can be elaborated in ways that would often be difficult to sustain through human effort alone. The model's ability to preserve and extend relational structure provides an important source of conceptual continuity within the interaction.

Yet the same capacity also helps explain many of the system's characteristic limitations.

A model may continue developing a trajectory that has ceased to be relevant. It may preserve assumptions that ought to be questioned. It may elaborate a framework whose organizing constraints have become problematic. It may generate sophisticated continuations that possess formal integrity without contributing to interpretative unity. In each case, the difficulty arises not because the model has ceased operating according to its characteristic strengths, but because those strengths continue to operate in circumstances where interpretative orientation has weakened, drifted, or become absent.

This observation helps explain why some of the most familiar limitations of LLMs can be so difficult to correct. The problem is not always that the model lacks information. Nor is it necessarily a failure of logical consistency. Often the model is doing precisely what it does best: preserving and extending a formally coherent trajectory. The difficulty is that formal integrity alone cannot determine whether a trajectory remains meaningful, significant, or properly situated within a broader conceptual framework.

The distinction developed in Section 5 now becomes particularly important. Conceptual frameworks require both formal integrity and interpretative unity. Formal integrity is necessary because meanings cannot stabilize without organized relational structures. Yet formal integrity alone is not sufficient. A formally coherent trajectory may participate in a shared framework, drift away from one, or continue in the absence of one altogether.

The three possibilities discussed earlier can therefore be understood as different manifestations of the same underlying capacity. In the first case, formal integrity remains synchronized with interpretative orientation. The result is productive conceptual development. In the second case, formal integrity continues while synchronization gradually weakens. The result is conceptual drift. In the third case, formal integrity continues without converging upon a stable framework capable of sustaining interpretative unity. The result is an organized conceptual trajectory that remains difficult to interpret as a coherent conceptual object.

What makes these cases especially revealing is that they do not involve different mechanisms. The same capacity for preserving relational structure is present throughout. What changes is the relationship between formal development and interpretation.

This perspective suggests a different way of thinking about the future development of language models. Discussions of capability often assume that greater formal sophistication will automatically reduce existing limitations. The framework developed here points toward a more complicated possibility. If many strengths and limitations arise from the same underlying capacity, increasing formal power alone may not eliminate the characteristic difficulties associated with interpretation. More powerful systems may sustain longer trajectories, explore more possibilities, and generate more elaborate formal structures. Yet they may also become capable of more sophisticated forms of drift, more persistent continuation of problematic frameworks, and more complex formal objects lacking interpretative unity.

From this perspective, the central challenge is not simply the generation of better continuations. It is the relationship between formal development and interpretative orientation. The most productive interactions examined throughout this paper did not arise because formal integrity became sufficient for meaning. They arose because interpretative interventions and formal continuations remained mutually supportive within a shared framework.

This observation returns us to one of the central themes of the preceding studies. Meaning does not emerge from formal organization alone, nor from unformed interpretation alone. It emerges through their relationship. The remarkable achievements of LLMs become possible because they participate in one side of this relationship with extraordinary effectiveness. Their characteristic limitations arise because participation in formal organization is not identical with participation in meaning.

This conclusion invites a broader reflection. If LLMs derive both their strengths and their weaknesses from the partial separation of formal development and interpretation, what does this reveal about semiosis itself? More specifically, what becomes visible when one of the central components of semiotics can be studied in relative isolation from the others?

The answer to this question will form the basis of our concluding discussion.

11. What Do LLMs Reveal About Semiosis?

The purpose of this paper has not been to determine whether LLMs are intelligent, conscious, or capable of understanding in a human sense. Rather, it has been to examine what becomes visible when these systems are viewed through the semiotic framework developed in the preceding studies.

The analysis began from a simple observation. Human–LLM interaction often behaves in ways that are difficult to explain through conventional accounts based solely on information, representation, or computation. Context matters in unexpected ways. Small interventions can have disproportionately large effects. Discussions may drift despite retaining coherence. Meaning often becomes clearer through interaction itself. Some exchanges become remarkably productive, while others remain trapped within formally sophisticated but conceptually limited trajectories.

Viewed independently, these phenomena may appear unrelated. Viewed through the theoretical framework developed in the preceding studies, however, they reveal a common structure. Conceptual frameworks simultaneously function as dynamical structures of meaning and systems of relational constraint. Human participants and language models contribute differently to these two aspects of the framework. Human participants primarily contribute interpretative orientation. Language models primarily contribute the preservation and extension of formal integrity. Productive interaction depends upon the ongoing coordination of these complementary activities.

This perspective suggests that the significance of LLMs extends beyond the practical tasks they perform. These systems provide an unusual opportunity to examine aspects of semiotics that are ordinarily difficult to isolate. In most human cognition, interpretation and formal development are deeply intertwined. We simultaneously determine significance, maintain conceptual organization, introduce distinctions, and develop their consequences. Because these activities occur together, their individual contributions are often difficult to distinguish.

Human–LLM interaction partially separates them.

The language model contributes extraordinary capacities for preserving and extending relational structures. It can sustain conceptual trajectories, explore consequences, elaborate distinctions, and maintain formal organization across extended discussions. Yet the analyses developed throughout this

paper suggest that such capacities do not by themselves generate interpretative unity. Meaning remains dependent upon processes that extend beyond formal continuation alone.

This observation helps clarify why the strengths and limitations of LLMs are so closely related. Both arise from the same underlying capacity. The ability to preserve formal integrity makes possible many of the system's most impressive achievements. The same ability also permits conceptual drift, the continuation of problematic frameworks, and the generation of formally coherent structures that fail to achieve stable interpretative unity. The distinction is not between successful and unsuccessful operation. It is between different relationships between formal development and interpretation.

From this perspective, the most productive human–LLM interactions are not instances in which interpretation disappears or becomes unnecessary. They are instances in which interpretation and formal development remain effectively synchronized within a shared conceptual framework. New possibilities are introduced, developed, evaluated, refined, and stabilized through the interaction itself. Meaning emerges neither from the human participant alone nor from the language model alone, but through the ongoing coordination of their complementary contributions.

At the same time, the present analysis suggests an important limitation. Formal integrity can continue in the absence of synchronization. It can continue in the absence of interpretative unity. It can even continue in the absence of any stable conceptual framework capable of sustaining meaning. This asymmetry reveals something fundamental about semiosis itself. Formal organization may be necessary for meaning, but it is not sufficient. The existence of a coherent relational structure does not guarantee the existence of a coherent conceptual object.

The broader significance of this conclusion extends beyond the study of language models. The framework developed in the preceding studies proposed that possibility is formed through relational constraint rather than selected from a pre-existing space. The present analysis suggests a corresponding insight regarding meaning. Meaning is not simply retrieved from a pre-existing conceptual structure. It emerges through processes of interpretation, determination, and stabilization occurring within evolving frameworks of relation. Human–LLM interaction provides a particularly revealing environment in which these processes can be observed because the contributions of formal development and interpretation become partially distinguishable.

The result is a shift in perspective. Rather than viewing LLMs primarily as artificial minds, knowledge repositories, or predictive engines, they may be understood as participants in a broader semiotic process whose structure becomes unusually visible through interaction. Their importance lies not only in what they can do, but also in what they reveal.

What they reveal is that meaning and formal organization are deeply related but not identical. They reveal that conceptual frameworks are simultaneously structures of significance and systems of relational constraint. They reveal that possibility and meaning emerge through processes of progressive determination rather than through the selection of pre-existing alternatives. Most importantly, they reveal that the study of human–LLM interaction may provide a new window into the semiotic processes through which understanding itself becomes possible.

In this sense, the significance of LLMs may ultimately lie not in the fact that they resemble human cognition, but in the fact that they allow us to examine some of its deepest organizing principles from a new perspective.

12. Enactive Conceptual Analysis: A mode of theoretical inquiry

This paper was written using the process of *enactive conceptual analysis*, a method that treats theoretical understanding not as the passive analysis of pre-existing concepts but as the active development of conceptual organization through sustained interaction. Rather than assuming that concepts possess fixed meanings that can be examined in isolation, the method proceeds from the view that concepts become operational only when the conceptual frameworks within which they function are progressively activated, reorganized, and stabilized through inquiry. Conceptual analysis is therefore understood as an enactive process in which action, meaning, and formal organization develop together.

In the present study, this process took the form of an extended interactive dialogue between the author and ChatGPT conducted over the course of the development of the paper and through which the paper itself was written. The transcript of that dialogue, conducted between June 9 and June 11, 2026, has been published separately as a companion research document, providing a transparent methodological record of the progressive development and stabilization of the theoretical framework and its instantiation in the paper presented here [9].

The method was introduced previously as an approach to understanding how conceptual frameworks become operational within human–LLM interaction [3]. The present paper extends that methodology by applying it to the development of a theoretical framework capable of explaining observed phenomena associated with human–LLM interaction. Rather than serving merely as supplementary material, the dialogue documents the iterative refinement through which distinctions were proposed, tested, revised, and integrated into the explanatory framework presented in this paper.

In this respect, enactive conceptual analysis shares with recent work on enactive artificial intelligence the view that understanding emerges through active engagement rather than passive representation [10]. However, whereas enactive AI primarily emphasizes the inseparability of action and perception, the present method focuses on the inseparability of action, meaning, and formal organization in the development of theoretical understanding. The resulting framework is evaluated not by deductive derivation alone but by its capacity to achieve explanatory integration across the phenomena it seeks to explain as exemplified in the present paper.

Enactive conceptual analysis makes the development of theoretical understanding itself available as an object of inquiry.

References

- [1] T. M. Rogers, "A Phenomenological and Formal Interpretation of Two Experiments conducted within the Cognitive Environment of Large Language Models (LLMs) using the Formal Modelling Framework of Hierarchical Relational Ontologies," January 2026. [Online]. Available: <https://philpapers.org/archive/ROGAPA.pdf>.
- [2] T. M. Rogers, "The Relational Formation of Possibility: Recursive determination and the hidden logic of Large Language Models (LLMs)," 16 April 2026. [Online]. Available: <https://philpapers.org/rec/ROGTRF>.

- [3] T. M. Rogers, "Concepts Become Operational Only When their Frameworks are Activated: An enactive account of conceptual analysis in large language model (LLM) interaction," 25 April 2026. [Online]. Available: <https://philpapers.org/rec/ROGCBO>.
- [4] T. W. Deacon, *Incomplete Nature: How Mind Emerged from Matter*, New York: WW Norton & Company, 2011.
- [5] D. Picca, "Not Minds, but Signs: Reframing LLMs through Semiotics," 27 Mar 2026. [Online]. Available: <https://arxiv.org/abs/2505.17080>.
- [6] T. M. Rogers, "Not Things but Relations: Rethinking how Large Language Models (LLMs) work," 27 April 2026. [Online].
- [7] C. S. Peirce, *The Collected Papers of Charles Sanders Peirce*, Vols. 1-8, Cambridge: Harvard University Press, 1931-1958.
- [8] K. Kull, T. Deacon, C. Emmeche, J. Hoffmeyer and F. Stjernfelt, "Theses on Biosemiotics: Prolegomena to a Theoretical Biology," *Biological Theory*, vol. 4, no. 2, pp. 167-173, 2009.
- [9] T. M. Rogers, "The Emergence of a Theoretical Framework for Enactive Conceptual Analysis with the Cognitive Environment of a Large Language Model (LLM)," July 2026. [Online]. Available: <https://doi.org/10.5281/zenodo.21135539>.
- [10] B. Rafiee and R. Sutton, "Toward Enactive Artificial Intelligence," 22 May 2026. [Online]. Available: <https://arxiv.org/abs/2605.24238>.