FORMALIZING THE CANONICAL NARRATIVE SCHEMA: A SEMIOTIC NEURO-SYMBOLIC APPROACH TO NARRATIVE ANALYSIS

Nicolas Chiappucci
University of Bologna, Italy
nicolas.chiappucci2@unibo.it

Abstract

Since the advent of large language models (LLMs), computational narratology has begun to use them to analyse narrative discourse, with the aim of capturing deeper dimensions such as modality and the passions. In this context, the renewed momentum of neuro-symbolic approaches opens up new possibilities for integrating semiotics into computational methodologies. By combining logical schemas with LLMs, a neuro-symbolic framework enables a more articulated understanding of narrative patterns, bringing the interpretive richness of discourse into dialogue with structural stability and coherence. This paper formalizes Greimas's canonical narrative

Digital Age in Semiotics & Communication, Vol. IX, 2026, 123–142
Under CC BY 4.0
https://doi.org/10.33919/dasc.26.9.6

schema as an OWL ontology – the CNS ontology – and integrates it with an LLM to analyse Edgar Allan Poe's short story "The Pit and the Pendulum". The ontology-informed analysis is compared with an LLM-only baseline, each condition being run five times, according to four criteria: theoretical adherence, analytical granularity, explanatory value and replicability. The ontology-informed outputs prove more consistent with the Greimasian model, more finely segmented and more stable across runs, while also showing a tendency to extend actantial categories beyond the elements they can justifiably describe. The paper thus shows how semiotics can support finer-grained narrative analysis through a neuro-symbolic approach and discusses both the potential and the limitations of a semiotic agent for the analysis of narrative and creative discourse.

Keywords: semiotics, neuro-symbolic AI, ontology, narrative analysis, large language models

1. For a computational semiotics

Attempts to establish a computational semiotics in dialogue with new technologies have long been both a challenge and a persistently explored field of study. This trajectory can be traced back to Dmitry Pospelov (1976), who pioneered applied semiotics and semiotic modelling. With the renewed interest in fuzzy systems and neural networks from the 1990s onwards, computational semiotics re-emerged as a potential theoretical framework for knowledge representation. A distinct line of research then developed within computer semiotics, particularly through the work of Peter Bøgh Andersen at Aarhus University, which shifted the focus to the communicative and design-oriented dimensions of human–computer interaction (Andersen 1990). The advent of deep learning and chatbots seemed to put an end to a potential resurgence of computational semiotics, giving way to a semiotic inquiry increasingly concerned, on the one hand, with the socio-cultural transformations brought about by AI and, on the other, with the semiotic status of these systems, their agency and their modes of enunciation.

The most recent developments in computer science have broadened the methodological possibilities of research, allowing qualitative and interpretive approaches to be combined with advanced computational tools. Approaches from the social sciences and the humanities, such as the ethnographic method (Seniutis et al. 2025) or the anthropological perspective

(De Ninno, Inverardi, and Belotti 2025), can thus be rethought as operational resources for the design of intelligent systems: from the construction of prompts sensitive to their context of use, through the formalization of ontologies, to the integration of explicit knowledge into neural architectures. Some experiments have indeed attempted to bring the semiotic perspective into computational systems. Dong et al. (2025), for example, propose GLASS (Greimas Literary Analysis via Semiotic Square), a framework that integrates the formal model of the semiotic square, in the form of a structured prompt, into the interaction with a Large Language Model (LLM) in order to obtain robust analyses and interpretations of literary texts. Generative theory, in particular, presents itself as a stratification of levels – configured "like an onion", as Ferraro (2012: 128) notes – and thus condenses theoretical models on different planes. It works through the interaction of levels that make it possible to detect first configurations of values and then their successive externalizations; in a preliminary integration into machine systems and programming languages, these planes can be conceived, experimentally, as standardized functions. This article sets out to integrate the semiotic perspective into a neuro-symbolic approach and to evaluate its possibilities for the analysis of narrative texts. More specifically, it is guided by the following research question: can a semiotic ontology improve the stability, explicitness and theoretical controllability of LLM-based narrative analysis? To address this question, after defining a common vocabulary, we present an experimental analysis of Edgar Allan Poe's "The Pit and the Pendulum" from two different perspectives. This comparison allows us to reflect on the possibility of an automatic semiotic agent for the evaluation of narrative and creative discourse.

1.1. Research trajectory

The approach presented here recalls the formalization projects developed within semiotics, specifically the Greimasian project. In the *Dictionnaire* (Greimas & Courtés 1979), Greimas defines the metalinguistic character of his theory and proposes, through the model of the generative trajectory, a hierarchical system organized along syntagmatic and paradigmatic dimensions, in which a strictly syntactic component appears alongside the semantic one at each level. This line of inquiry introduces scientific rigour through the construction of an effective metalanguage. Although the project has raised several reservations (cf. Bremond 1982), particularly about the potential logical weakness of its methodology, the search for a structure within the object of analysis, or within the conditions under which meaning is produced or perceived, remains one of the general characteristics of semiotics

(Batu 2012). The computational formalization of a semiotic model can therefore be understood as a continuation of this modelling vocation: semiotic theoretical frameworks such as the canonical narrative schema lend themselves to translation into formal structures consisting of classes, properties and logical relations.

2. Operational lexicon for a semiotic neuro-symbolic approach to narrative analysis

A preliminary integration of semiotic theories into programming languages forces us to make a cut, which future, more in-depth research may well repair. A further consideration, which limits our scope, is that we treat the levels of the generative trajectory of meaning as a series of instances of logical rules. In this way, we obtain a semiotic grid that can be used as a magnifying glass on a given domain.

2.1. Ontologies

Our first step is to transform semiotic theories into computational ontologies. We define an ontology as "an explicit specification of a conceptualization" (Gruber 1993: 199), which provides a common vocabulary for a domain. An ontology makes explicit the assumptions about a domain and separates operational knowledge from domain knowledge, which is then investigated by means of the operational knowledge. Imagine, then, a schema made up of concepts relevant to a domain and used as a grid to obtain specific information about that same domain: an ontology built on a common vocabulary for the domain of wines, for example, can be used to obtain precise information on the types of wine, their fermentation or their vintage (Noy & McGuinness 2001). In constructing an ontology, domain knowledge can be represented with formal structures based on description logics, a family of knowledge representation formalisms closely related to decidable fragments of first-order logic. Languages based on description logics structure domain knowledge through unary and binary predicates. The Web Ontology Language (OWL), standardized by the World Wide Web Consortium (W3C) in 2004 and extended as OWL 2 in 2009, provided a standard language for the formal representation of ontologies. It allows the concepts, properties and relations of a specific domain of knowledge to be represented formally, making these structures interpretable and processable by computational systems. Although it does not follow a rigid method, ontology engineering begins by defining the domain to be modelled and the purpose the ontology is to serve. This can be done, for example, by formulating competency questions (Grüninger & Fox 1995), which help to determine the level

of representation of the domain and its conceptualization. In short, building an ontology makes it possible to formalize a research domain as a system of concepts, relations and logical axioms grounded in a theory. The potential of generative theory lies precisely in its ability to provide the architecture of such a vocabulary, which can ultimately be applied to almost any domain. The configuration of Greimas's narrative programme, for example, offers a series of abstract terminological labels that bring out the conditions under which action unfolds in a text.

2.2. Large language models

The other component of our experiment is an LLM: a deep learning model, typically based on a transformer architecture, that can process and generate natural language. An LLM undergoes a pre-training phase on large collections of texts that have been retrieved and passed through various quality filters. During this phase, the model learns statistical regularities and linguistic patterns, developing capacities that allow it to perform a variety of linguistic and inferential tasks and to display forms of socio-cultural knowledge derived from its training data. Text generation, in turn, is a classic task of computational linguistics: each word is generated on the basis of the probabilities that the model assigns to the possible continuations.

2.3. The neuro-symbolic approach

The integration of symbolic approaches, grounded in explicit knowledge representation, with neural architectures based on statistical learning is the foundation of what is generally called neuro-symbolic artificial intelligence. Although attempts to combine the two paradigms predate LLMs, the spread of generative systems has renewed interest in introducing symbolic structures capable of supporting their reasoning. Despite the remarkable performance of neural models, several limitations persist with regard to structured reasoning, the integration of expert knowledge and the reliability of inferences (Belle & Marcus 2026). These difficulties cannot necessarily be overcome by scaling models further; they may instead require a reconsideration of the relationship between statistical learning and explicitly represented knowledge. The relevance of this integration becomes particularly evident when we consider the nature of the knowledge represented in LLMs: several studies have highlighted difficulties with systematic compositional generalization (Zhao et al. 2024) and with the consistency of the ontological commitments expressed in model-generated responses (Köhler & Neuhaus 2025). Unlike a formal ontology, whose categories and relations are specified by logical axioms, the knowledge mobilized by an LLM is represented

implicitly in a distributed statistical architecture, without necessarily being organized into a coherent and accessible ontological framework.

This distinction is particularly relevant to the relationship between implicit knowledge and interpretive processes. As Gangemi and Nuzzolese (2025) observe, training large-scale neural systems allows them to acquire patterns and relations associated with forms of tacit knowledge (cf. Polanyi 1966), including cultural conventions, metaphors and context-dependent inferences. Such knowledge, however, is not explicitly codified in a formal conceptual structure but represented through distributed patterns learned during training. As Confalonieri and Guizzardi (2025) argue, ontological structures can contribute to AI systems in reference modelling, the integration of common-sense knowledge and the management of different levels of conceptual abstraction. In narrative analysis, these possibilities become particularly significant when the ontological representations are grounded in a semiotic metalanguage whose categories can be selectively formalized and then used to orient the interpretive processes of neural models. Semiotics can thus once again become a reference framework for knowledge representation, defining the conceptual conditions under which specific dimensions of meaning can be identified, related and analysed. The resulting semiotic neuro-symbolic approach can be understood as an attempt to bring the generative and inferential capabilities of neural models into dialogue with the structural coherence of explicitly formalized semiotic knowledge.

3. Building the CNS ontology

Having outlined the method, we can now make semiotics operational, bearing in mind the cuts and potential imperfections of an approach still in its early stages. Greimas's narrative programme offers a framework for examining the production of meaning through narrative action, understood in terms of a subject's orientation towards an object of value. It breaks complex situations down into phases – manipulation, competence, performance, sanction – in order to understand how desire and transformations produce meaning within a system. The configuration of the canonical narrative schema (Greimas 1970) introduces abstract analytical categories for describing the circumstances under which action unfolds in a text. Elements such as competence, manipulation, performance and sanction are articulated through modal configurations that establish the possibilities and constraints of the subject's action: a subject may be able to perform an action, know how to perform it (for instance, know the conditions under which a given move can be made) or simply want to perform it. On this basis, we

modelled the ontology of the canonical narrative schema (CNS) as actant-centred, comprising a series of classes and of properties that link them.[1]

3.1. Classes

The classes of the CNS ontology (Figure 1) define the fundamental conceptual categories through which the elements of the canonical narrative schema are represented.

gre:canonicalNarrativeSchema: root class representing a canonical narrative schema. It is the main domain of the relations that connect a narrative configuration to its subjects, objects, states, phases and actantial positions.

gre:actantPosition: general class representing the actantial positions provided by the narrative model. Its subclasses are:

gre:subject ⊑ gre:actantPosition: position of the narrative subject, i.e. the actant oriented towards the attainment of the object.

gre:object ⊑ gre:actantPosition: position of the narrative object, i.e. that towards which the subject's programme tends.

gre:sender ⊑ gre:actantPosition: position of the sender, i.e. the instance that founds, motivates or legitimizes the narrative programme.

gre:receiver ⊑ gre:actantPosition: position of the receiver, i.e. the instance that benefits from, receives or is transformed by the attainment of the object.

gre:helper ⊑ gre:actantPosition: position of the helper, which facilitates the subject's path towards the object.

gre:opponent ⊑ gre:actantPosition: position of the opponent, which hinders the narrative programme.

gre:anti-subject ⊑ gre:actantPosition: position of the anti-subject, i.e. an antagonistic subject pursuing a competing or contrary programme.

gre:phase: general class representing the phases of the canonical narrative schema. Its subclasses are:

gre:manipulation ⊑ gre:phase: phase in which the narrative programme is instituted, through causing-to-do (*faire faire*), persuasion, mandate or the motivation of the subject.

gre:competence ⊑ gre:phase: phase in which the subject acquires the modal conditions necessary for action.

[1] For a previous application of the proposed neuro-symbolic approach to screenplay interpretation using an ontology of passions (GreP), see Chiappucci & Lechelt (2026).

gre:performance ⊑ gre:phase: phase of the actual realization of the narrative programme.

gre:sanction ⊑ gre:phase: evaluative phase, in which the action performed is recognized, judged or sanctioned.

gre:modality: general class for the modalities that qualify the subject's competence. Its subclasses are:

gre:wantingToDo ⊑ gre:modality: wanting-to-do modality.

gre:havingToDo ⊑ gre:modality: having-to-do modality.

gre:knowingHowToDo ⊑ gre:modality: knowing-how-to-do modality.

gre:beingAbleToDo ⊑ gre:modality: being-able-to-do modality.

gre:state: class representing a narrative state. A state may be initial or final with respect to a canonical narrative schema and can be qualified by a junction type.

gre:junctionType: class representing the type of junctive relation between subject and object. It is defined as an enumeration of two individuals[2]:

gre:junctionType ≡ { **gre:conjunction, gre:disjunction** }

gre:conjunction: individual of **gre:junctionType**; indicates a state of conjunction between subject and object.

gre:disjunction: individual of **gre:junctionType**; indicates a state of disjunction between subject and object.

gre:value: class representing the value invested in the object. The narrative object is not merely a thing or an entity, but an object insofar as it carries value.

3.2. Object properties

The object properties of the CNS ontology connect the canonical narrative schema to its constituent elements and make it possible to model the relations between phases, actants, states, objects and values.

gre:hasSubject: connects a canonical narrative schema to the subject of the narrative programme.

domain: **gre:canonicalNarrativeSchema**

range: **gre:subject**

[2] Whereas classes define general conceptual categories, individuals denote concrete or particular entities belonging to those categories; an individual is thus a specific instance of a class.

gre:hasObject: connects a canonical narrative schema to the object of value towards which the subject's programme tends.

domain: **gre:canonicalNarrativeSchema**

range: **gre:object**

gre:assignsPosition: associates a canonical narrative schema with an actantial position. It is a general property that connects the schema to the different positions provided by the actantial model.

domain: **gre:canonicalNarrativeSchema**

range: **gre:actantPosition**

gre:hasPhase: connects the canonical narrative schema to its constituent phases.

domain: **gre:canonicalNarrativeSchema**

range: **gre:phase**

gre:hasInitialState: connects a canonical narrative schema to the initial state of the narrative programme.

domain: **gre:canonicalNarrativeSchema**

range: **gre:state**

gre:hasFinalState: connects a canonical narrative schema to the final state of the narrative programme.

domain: **gre:canonicalNarrativeSchema**

range: **gre:state**

gre:hasJunctionType: qualifies a narrative state by the type of junction that characterizes it.

domain: **gre:state**

range: **gre:junctionType**

gre:hasValue: connects a narrative object to the value with which it is invested.

domain: **gre:object**

range: **gre:value**

3.3. Enumeration of junction types

The class **gre:junctionType** is defined as the closed set of the two individuals **gre:conjunction** and **gre:disjunction**:

gre:junctionType ≡ { **gre:conjunction, gre:disjunction** }

where:

gre:conjunction rdf:type gre:junctionType

gre:disjunction rdf:type gre:junctionType

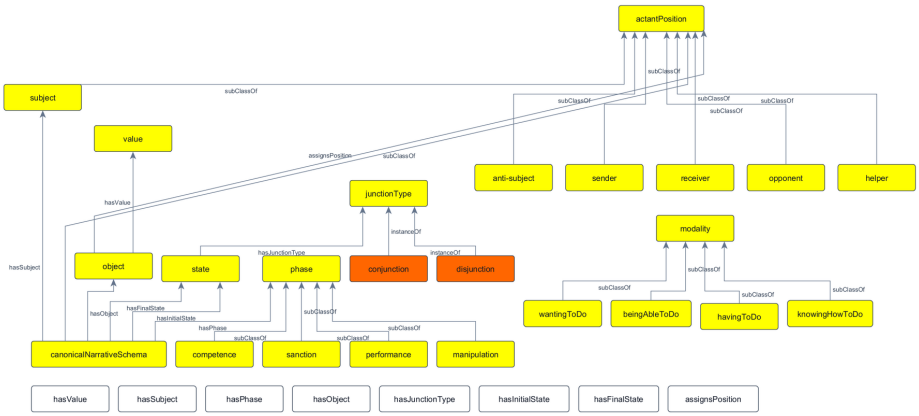

Figure 1: The CNS ontology visualized through Graffoo

4. Evaluation

4.1. Experimental set-up

To assess how an LLM analyses narrative discourse when guided by a semiotic ontology, we conducted an experiment on Edgar Allan Poe's short story "The Pit and the Pendulum". The experiment had two conditions:

> Condition 1 (ontology-informed): using a prompting framework, the model was asked to analyse the text by applying the attached ontology.
>
> Condition 2 (LLM-only baseline): the model was asked to analyse the narrative structure of the story without the semiotic ontology.

The experiments were run in Google Colab in a Python 3 environment, with the LLM accessed remotely via its API. We had previously checked the consistency of the CNS ontology in Protégé 5.6.4 with the HermiT 1.4.3.456 reasoner, which completed the check in 122 ms. The model used was Claude Sonnet 4.6 (Anthropic 2026), with max_tokens set to 12,000 and the effort parameter set to "high". The runs were carried out in May 2026. Each condition was run five times. To avoid post hoc selection of the

most favourable outputs, the primary qualitative analysis was restricted to the first two runs of each condition, generated under identical settings.[3]

The story recounts the experience of a prisoner of the Spanish Inquisition who has been sentenced to death and locked in a dark cell. It follows the protagonist's growing awareness of the space around him and of the forms of torture to which he is subjected: a hidden pit at the centre of the cell and a blade that swings like a pendulum as it slowly descends towards his body. The narration is built almost entirely on the protagonist's subjective perception, on the relationship between body, fear, space and time, and on the alternation between loss and recovery of consciousness. Poe's text was chosen for this strong tensive and perceptual organization around the subject, which lends itself to a semiotic analysis of the narrative structure and of the transformations of the subject's experience, and for a narrative structure that is not immediately transparent. The semiotic neuro-symbolic approach could therefore help the LLM to make the narrative and modal articulations of the text explicit.

Table 1: Excerpts from the outputs of the two conditions, grouped into four parts of the story: the narrative overall, the opening act, the changes following the discovery of the pendulum, and the final sequence (Anthropic 2026)

	Semiotic-neurosymbolic LLM approach	**LLM-only baseline**
Narrative overall	The narrative opens in a condition of radical dispossession. The protagonist finds himself severed from every value that might sustain a self: freedom, knowledge of his surroundings, physical integrity, even the certainty of being alive. This initial state is one of absolute disjunction — the subject is cut off from everything that could constitute a livable existence. The sentence pronounced by the judges is not merely a legal verdict; it is the formal act through which an external authority, cold and implacable, sets the entire narrative machinery in motion.	The narrative begins at the outermost edge of consciousness, where the central figure exists in a condition of near-total dissolution. He has been condemned, and the first thing the text establishes is not a place or an action but a state: the senses are failing, sound is collapsing into undifferentiated hum, and vision, though still briefly operative, has become hallucinatory and distorted. The judges' lips appear white, moving without sound, fashioning his name into a sentence he can see but no longer hear. This opening moment is crucial because it establishes the fundamental dynamic that will

[3] The CNS ontology, outputs and input prompts for both conditions are available at https://github.com/ChiappucciN/Formalizing-the-Canonical-Narrative-Schema.

(continued)

	Semiotic-neurosymbolic LLM approach	**LLM-only baseline**
		organize the entire narrative: the figure is a subject who perceives but cannot act, who receives information about his own destruction without possessing any means to alter it.
Opening act	The subject's initial want is not heroic acquisition but bare survival, and even this minimal desire is not freely formed — it emerges under extreme duress, shaped by the very torment that threatens to extinguish it. The darkness of the dungeon, the disorientation of the body, the progressive erosion of rational faculties: all of these function as obstacles that do not merely impede action but attack the very capacity to act.	The prisoner reasons his way through this possibility with deliberate effort, and his conclusion — that he is not entombed — is reached by logic, by working through what he knows of how condemned prisoners are treated. This early exercise of reason under extreme duress is the first of many such moments, and it establishes a rhythm that will govern the entire narrative: panic giving way, through effort, to some measure of rational control, followed by a new discovery that undoes that control and demands it again.
Pendulum changes	The pendulum introduces a different temporal logic. Where the pit threatened sudden, accidental death, the pendulum operates through inexorable gradualness — a slow, measured descent that transforms time itself into an instrument of torture. The subject is now bound, immobilized, stripped of the ability to move that had constituted his only resource. His competence, so painfully assembled, is negated. What remains is perception and thought, and even these are under assault: the thirst deliberately provoked, the food laced with drugs, the psychological architecture of despair carefully maintained by unseen observers.	The pendulum introduces a new temporal structure into the narrative. Where the pit was a danger of sudden, unforeseeable arrival, the pendulum is a danger of slow, measurable approach. The figure can see it, can count its vibrations, can calculate approximately when it will reach him. This visibility does not reduce the horror; it transforms it. He is now forced to watch his own death approaching, to live inside the interval between the present moment and the moment of destruction, with no possibility of action.
Final sequence	At the moment of absolute extremity, when the subject teeters on the brink	The rescue that arrives at this exact moment is structurally abrupt,

(continued)

	Semiotic-neurosymbolic LLM approach	**LLM-only baseline**
	of the pit with no remaining foothold and no conceivable resource, the resolution arrives from entirely outside the narrative's internal logic. The French army, General Lasalle, the collapse of the Inquisition's power —these are not outcomes that the protagonist has worked toward or could have influenced. They arrive as pure external intervention, a sender of an entirely different order who dissolves the entire system of constraint in an instant. The conjunction that closes the narrative — the subject reunited with life, with freedom, with the possibility of continued existence — is not earned through the protagonist's competence or will alone. It is granted, at the last possible moment, by a force whose existence the narrative has given no prior indication of. This closing movement reveals something essential about the deep structure of the story: the subject's heroism, such as it is, consists entirely in enduring long enough for rescue to become possible. The value recovered is not triumph but mere survival, and even that is not fully the subject's own achievement. The narrative ends in conjunction, but the conjunction is fragile, contingent, and shadowed by the recognition of how completely the subject was, throughout, at the mercy of forces — destructive and saving alike — that operated entirely beyond his control.	almost violent in its suddenness. Trumpets, voices, the walls rushing back, an arm catching him as he falls. An external force, entirely outside the logic of the dungeon, breaks into the closed system of torment and dissolves it. The narrative has been organized throughout as a sealed world in which the only variables are the figure's endurance, his intelligence, and the ingenuity of his tormentors. The rescue does not emerge from any of these internal dynamics; it arrives from outside, from a historical event that the dungeon's architects could not have anticipated or controlled. This ending does not resolve the tension so much as it terminates it by removing the conditions that produced it. The figure is saved not because he found a way out but because the world outside changed in a way that made the dungeon's power suddenly obsolete. What the narrative has traced, across all its stages, is the progressive reduction of a conscious being to the very limit of existence, and the equally progressive discovery that at each apparent limit, something remains. Consciousness survives the swoon. Reason survives the darkness. Intelligence survives the pendulum. Endurance survives the rats. And at the final moment, when all of these have been exhausted and the figure is genuinely about to fall, something outside the self arrives to catch him. The text does not present this as triumph or as the reward of virtue; it presents it as the last possible contingency, the one accident that could not have been planned for, arriving at the one

(continued)

	Semiotic-neurosymbolic LLM approach	LLM-only baseline
		moment when nothing else remained.

4.2. Evaluation of the outputs

Comparing the outputs (Table 1) allows us to assess whether the LLM, guided by the CNS ontology, applied the Greimasian model correctly. The outputs are compared according to four criteria: theoretical adherence, analytical granularity, explanatory value and replicability. The comparative evaluation builds on the criteria adopted in Chiappucci & Lechelt (2026), extending them by considering explanatory value as a distinct analytical dimension.

Theoretical adherence. The ontology-informed output adheres more closely and more consistently to the Greimasian model. The analysis opens with a radical disjunction: the protagonist is reconstructed as a narrative subject separated from freedom, understood as a fundamental object of value. The Inquisition is defined not by its thematic background but as an actantial instance that has imposed a programme of annihilation on the subject. The LLM-only baseline, by contrast, privileges the protagonist's perceptual and cognitive dimension. From the outset, its analysis focuses on the progressive collapse of the protagonist's perception and on the terror and despair that follow. Its categories are mostly phenomenological and experiential – consciousness, sensation, terror, despair, endurance, perception – and it does not attempt to apply a formal narrative grammar.

Analytical granularity. The ontology-informed output appears better able to distinguish the different levels of the narrative programme. It separates the initial state of disjunction, the acquisition of competence, the roles of helpers and opponents, the partial performance represented by the escape from the pendulum, and the final sanction produced by an external intervention. The protagonist's actions are therefore treated not as isolated episodes but as transformations within a modal and actantial structure. The attempts to explore the cell are read as attempts to acquire competence; the pit is interpreted as a mechanism that turns movement into self-destruction, and the pendulum as a device that negates the subject's bodily competence while preserving perception and thought; the rats are reconstructed as an anomalous helper, since their abject nature is turned into a resource for liberation. The

LLM-only baseline offers a more continuous and less segmented interpretation. It follows the story through the progressive modulation of consciousness: the swoon, the return of perception, the reconstruction of space, the discovery of the pit, the terror of the pendulum and the final collapse of the cell. The text, however, is not interpreted through a standardized narrative articulation derived from a specific theoretical model. The pendulum, for example, is defined primarily by the effects it produces on the protagonist, namely anticipation, terror, temporal suspension and bodily resistance. The analysis is therefore precise at the level of experience but less granular at the level of narrative structure.

Explanatory value. The ontology-informed output brings out structures that are not immediately evident in a purely descriptive reading. Its main contribution is to show that the story is not only a sequence of tortures, nor only the narration of a psychological ordeal, but a structured conflict between actantial positions. The Inquisition operates as an anti-subject or opposing instance; the dungeon functions as an articulated apparatus of manipulation; the protagonist's survival depends on unstable and partial acquisitions of competence; the rats become helpers only through a reversal of their expected value; and the final rescue appears as an external actantial intervention that dissolves the system from outside. In this sense, the ontology guides the LLM towards a more formalized reading of the deep structure of the story.

At the same time, this formalization reveals possible limitations. The model tends to extend the actantial reading too far, turning into actants, for instance, elements that could be interpreted at the enunciative level. Not every element of the dungeon should correspond to an actantial role; some may operate more properly at the discursive level. This tendency recalls a problem already debated in generative semiotics, namely the risk that narrative models extend their descriptive scope beyond the phenomena to which they can justifiably be applied. Rastier (1971), for instance, observes that excessively broad definitions of the elementary narrative structure may also account for phenomena that are not conventionally regarded as narratives, and stresses the need to distinguish between narrative and non-narrative manifestations. This tendency could be mitigated by defining the limits and conditions of applicability of the formal categories more explicitly – for instance, by requiring the model to justify each actantial attribution with reference to a specific narrative programme and allowing elements to remain unclassified when such a relation is not sufficiently supported. This would be particularly useful if the system were given a complementary ontology formalizing the enunciative plane through a distinct set of categories, so

that the model could distinguish between narratively functional dimensions and those more properly related to enunciative organization. The LLM-only baseline has a different explanatory value. It renders well the subjective density of Poe's narration, especially the alternation between loss and recovery of consciousness and perception, and the progressive transformation of space into a threatening environment. Precisely because it remains closer to the figurative and experiential surface of the text, however, it does not fully formalize the narrative structure. Combining the two approaches, starting from better-designed prompts and ontologies, could address these shortcomings.

Replicability. To assess inter-run consistency, each condition was executed five times with the same input and prompt configuration. Across the ontology-informed runs, the core actantial relations remained stable: the prisoner was consistently positioned as the central subject, the Inquisition as the opposing force, the rats as instrumental helpers and the final military intervention as an external rescuing force. Some secondary assignments, particularly the sender–object configuration, varied between runs. The LLM-only condition likewise preserved the main narrative relations and the sequence of transformations, although its interpretive emphasis varied from run to run. The baseline depends more heavily on the model's general interpretive capacities and on the rhetorical construction of the prompt. Although it may produce a refined analysis, its categories therefore remain relatively unstable: a different prompt could easily steer the model towards a different interpretation, depending on what the analyst needs. The ontology-informed output, by contrast, is more easily reproducible because it is constrained by a predefined vocabulary. The ontology does not eliminate interpretation, but it stabilizes the analytical field by making explicit the categories through which the model is expected to read the text.

5. Discussion

Building a semiotic ontology allows an LLM to sharpen its analytical gaze. This formal layer is placed alongside an interpretive dimension constituted by the specific functions and qualities of the neural network in question. LLMs can indeed be described as "agents of symbolic recombination" (Picca 2025: 5), insofar as their outputs result from a statistical yet guided recombination based on probabilistic proximities of semantic adjacency and relatedness, which actualizes a field of interpretive possibilities. This does not mean siding with a semiotics concerned exclusively with a metalanguage of simulation, but rather favouring forms of interpretation in which the canonical

narrative schema plays an integrative role, constructing a local commensurability through the formalization of theory. From a semiotic perspective, the neuro-symbolic approach thus appears as an attempt to combine an interpretive approach – suited to investigating LLMs as instruments capable of performing acts of cultural translation and of embedding new codes – with a formalizing dimension that exploits the rigour of generative semiotics.

The interpretive process of an LLM, particularly in high-context interpretive tasks such as narrative analysis, depends not only on linguistic comprehension but also on forms of cognitive alignment, such as epistemic awareness, rhetorical cohesion and cross-cultural adaptability (Yu, Ruiz-Dolz, and Yi 2025). In our view, semiotic theory can address these shortcomings by offering a cohesive and stable framework that allows the model to simulate the adoption of a given epistemology and that, being semiotic, tends to reduce the friction arising from the intercultural translation of texts.

Such an approach has several consequences. First, by specifying the theory under examination and building the corresponding ontology, we could arrive at a genuine semiotic agent capable of carrying out in-depth analyses of the themes of interest. The possibility of intervening on parameters and structures opens up ample perspectives for future research. A hypothetical semiotic agent also invites us to reconsider potential relations of co-agency with the user. In such a paradigm, the reiteration of the analytical process is a fundamental component, as is the factor of time: slightly modifying the parameters, the formulation of the request or the semiotic ontology can lead to an output with a different informational value. Such agency, however, should not be understood in terms of pragmatic interaction. The agent contributes to reorganizing the interpretive field, working on textual saliences and reading hypotheses, starting precisely from the indexicality proper to language, which refers back to the extralinguistic dimension (Violi 2001). The risk remains that ontological formalization may rigidify theoretical categories originally designed to operate dynamically and contextually. The value of a semiotic agent would therefore depend not only on the solidity of its ontologies but also on how they are adopted in context by the user.

Second, there is the social impact of such an approach. Fabbri's semiotic gaze, which allows us to search for untrodden paths in the human sciences (Cervelli 2017), requires rigour and study. The effort needed to acquire a critical awareness of stereotypes is an example of this (cf. Cosenza 2023): certain interesting textual configurations become fully visible only through a progressive refinement of the gaze. In this sense, a semiotic agent based

on formal ontologies could support the analyst's workflow by helping to reiterate analyses, compare different outputs, formulate the theoretical categories employed explicitly and stabilize interpretive procedures that would otherwise be difficult to replicate. Moreover, building standardized models of semiotic theories would make it possible to form a repository of reusable ontologies applicable to different texts, serving as a support infrastructure for semiotic research.

6. Conclusion

The aim of this article has been to use semiotics functionally, proposing a framework for integrating sign theory into computational systems. The experiment explored the possibility of an expert semiotic agent that could support a critic or an analyst. It also revealed certain limitations: for example, constraints would need to be placed on the way an LLM uses a given theory, in order to limit the interpretive hallucinations that might otherwise emerge. Nevertheless, rethinking the potential of intelligent agents in this way allows us to reconsider the potential of semiotics as an active semiotics: one capable of moving beyond a superficial analysis of artificial intelligence and of positioning itself as a functional theory in the construction of the systems themselves. Future work will focus on a systematic experimental evaluation of the semiotic neuro-symbolic approach, assessing both its interpretive capabilities and its potential integration into professional workflows.

Declaration on the use of generative AI

Apart from the experiment described in Section 4, OpenAI's ChatGPT (GPT-6) was used for academic English language editing and corrections.

References

Andersen, P. B. 1990. *A Theory of Computer Semiotics: Semiotic Approaches to Construction and Assessment of Computer Systems*. Cambridge: Cambridge University Press.

Anthropic. 2026. *Claude* (Claude Sonnet 4.6) [Large language model], accessed via the Anthropic API. Available at: https://www.anthropic.com/claude (accessed May 2026).

Batu, B. 2012. "An Overview of the Field of Semiotics". *Procedia – Social and Behavioral Sciences*, Vol. 51, 464–469. Available at: https://doi.org/10.1016/j.sbspro.2012.08.190.

Belle, V., G. Marcus. 2026. The Future Is Neuro-Symbolic: Where Has It Been, and Where Is It Going? In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 40, No. 48. Association for the Advancement of Artificial Intelligence, 40954–40961. Available at: https://doi.org/10.1609/aaai.v40i48.42130.

Bremond, C. 1982. "Sémiotique d'un conte mauricien: sur deux études de Clément Légaré". *Recherches sémiotiques / Semiotic Inquiry*, Vol. 2, No. 4, 405–423.

Cervelli, P. 2017. "Elogio dell'interazione: lo sguardo semiotico sulle micropolitiche del quotidiano". *Actes Sémiotiques*, No. 120, 1–8.

Chiappucci, N., S. Lechelt. 2026. "A Neurosymbolic Architecture for Semiotic Screenplay Interpretation". Proceedings of the 2026 ACM Creativity and Cognition (C&C '26). New York, NY: Association for Computing Machinery. Available at: https://doi.org/10.1145/3803784.3816849.

Confalonieri, R., G. Guizzardi. 2025. "On the Multiple Roles of Ontologies in Explanations for Neuro-Symbolic AI". *Neurosymbolic Artificial Intelligence*, Vol. 1, 1–14. Available at: https://doi.org/10.3233/NAI-240754.

Cosenza, G. 2023. "Stereotypes and Prejudices: From Social Sciences to Semiotics". *E|C*, No. 37, 243–257. Available at: https://mimesisjournals.com/ojs/index.php/ec/article/view/2767.

De Ninno, G., P. Inverardi and F. Belotti. 2025. "Value Elicitation: Towards Moral Profiles in Early Requirements Engineering via Role-Playing Games and Anthropologist LLMs". arXiv preprint 2510.01189. Available at: https://doi.org/10.48550/arXiv.2510.01189.

Dong, F., Y. Zeng, Y. Sang and H. Shen. 2025. Structuralist Approach to AI Literary Criticism: Leveraging Greimas Semiotic Square for Large Language Models. In Barner, D., N. R. Bramley, A. Ruggeri and C. M. Walker (eds.). *Proceedings of the 47th Annual Conference of the Cognitive Science Society*. Cognitive Science Society, 3148–3155.

Ferraro, G. 2012. *Fondamenti di teoria sociosemiotica*. Rome: Aracne.

Gangemi, A., A. G. Nuzzolese. 2025. "Logic Augmented Generation". *Journal of Web Semantics*, Vol. 85, Article 100859. Available at: https://doi.org/10.1016/j.websem.2024.100859.

Greimas, A. J. 1970. *Du sens: essais sémiotiques*. Paris: Seuil.

Greimas, A. J., J. Courtés. 1979. *Sémiotique: dictionnaire raisonné de la théorie du langage*. Paris: Hachette.

Gruber, T. R. 1993. "A Translation Approach to Portable Ontology Specifications". *Knowledge Acquisition*, Vol. 5, No. 2, 199–220. Available at: https://doi.org/10.1006/knac.1993.1008.

Grüninger, M., M. S. Fox. 1995. Methodology for the Design and Evaluation of Ontologies. In *Proceedings of the IJCAI-95 Workshop on Basic Ontological Issues in Knowledge Sharing*. Montreal.

Köhler, N., F. Neuhaus. 2025. "The Mercurial Top-Level Ontology of Large Language Models". *Applied Ontology*, Vol. 20, No. 1, 5–15. Available at: https://doi.org/10.1177/15705838251336685.

Noy, N. F., D. L. McGuinness. 2001. *Ontology Development 101: A Guide to Creating Your First Ontology*. Technical Report KSL-01-05. Stanford: Stanford Knowledge Systems Laboratory.

Picca, D. 2025. "Not Minds, but Signs: Reframing LLMs through Semiotics". arXiv preprint 2505.17080. Available at: https://doi.org/10.48550/arXiv.2505.17080.

Polanyi, M. 1966. *The Tacit Dimension*. Garden City: Doubleday.

Pospelov, D. A. 1976. "Semiotic Models: Successes and Perspectives". *Cybernetics and Systems Analysis*, Vol. 12, No. 6, 929–937. Available at: https://doi.org/10.1007/BF01070425.

Rastier, F. 1971. "Situation du récit dans une typologie des discours". *L'Homme*, Vol. 11, No. 1, 68–82. Available at: https://doi.org/10.3406/hom.1971.367154.

Seniutis, M., V. Gružauskas, A. Sas, V. Navickas and M. Švažas. 2025. "Designing a Framework for Ethnography-Driven Prompt Engineering in Social Work". *Human Technology*, Vol. 21, No. 1, 106–127. Available at: https://doi.org/10.14254/1795-6889.2025.21-1.5.

Violi, P. 2001. *Meaning and Experience*. Bloomington: Indiana University Press.

Yu, H., R. Ruiz-Dolz and Q. Yi. 2025. A Structured Framework for Evaluating and Enhancing Interpretive Capabilities of Multimodal LLMs in Culturally Situated Tasks. In *Findings of the Association for Computational Linguistics: EMNLP 2025*. Suzhou: Association for Computational Linguistics, 1945–1971. Available at: https://doi.org/10.18653/v1/2025.findings-emnlp.103.

Zhao, J., J. Tong, Y. Mou, M. Zhang, Q. Zhang and X. Huang. 2024. Exploring the Compositional Deficiency of Large Language Models in Mathematical Reasoning through Trap Problems. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*. Miami: Association for Computational Linguistics, 16361–16376. Available at: https://doi.org/10.18653/v1/2024.emnlp-main.915.