Mitigating Semantic Drift in Multi-Agent Communication: A Dynamic Neuro-Symbolic Approach

Xie Linhao[1,2], Li Fan[3], Wu Mingxuan[4], Song Yong[1], Ouyang Ye[1]

(1. AsiaInfo Technologies, Beijing 100193, China;
2. Communication University of China, Beijing 100024, China;
3. China Unicom Beijing Branch Network Optimization Center, Beijing 100032, China;
4. CRSC Information Industry Co., Ltd., Beijing 100070, China)

DOI: 10.12142/ZTECOM.202602008

Manuscript received: 2026-02-24

Abstract: The emergence of multi-agent systems (MAS) based on large language models (LLMs) has enabled autonomous collaboration on complex, goal-oriented tasks. However, effective interaction is frequently hindered by semantic drift, a phenomenon where heterogeneous agents assign conflicting meanings to shared terminology due to differing internal prompts or domain knowledge. Existing communication paradigms either rely on unconstrained natural language, which suffers from structural vagueness, or rigid symbolic schemas that fail to adapt to emergent concepts. To address this gap, we propose DOA, a novel dynamic ontology alignment framework that serves as a semantic mediation layer for MAS. DOA integrates a proactive semantic prober to detect conceptual mismatches and a neuro-symbolic aligner that reconciles local semantic structures in real time. By grounding fluid natural language dialogues in an evolving shared ontology, our framework ensures deterministic mutual understanding over long-horizon tasks. Empirical evaluations in cross-domain supply chain and healthcare coordination scenarios demonstrate that DOA improves task success rates by an average of 31.5% and reduces communication overhead (token consumption) by 50% compared to state-of-the-art baselines. Our results provide a robust and scalable foundation for semantic consistency in next-generation industrial-grade AI systems.

Keywords: DOA; multi-agent systems; neuro-symbolic AI; semantic drift; large language models

Citation (Format 1): Xie L H, Li F, Wu M X, et al. Mitigating semantic drift in multi-agent communication: a dynamic neuro-symbolic approach [J]. ZTE Communications, 2026, 24(2): 64–70. DOI: 10.12142/ZTECOM.202602008

Citation (Format 2): L. H. Xie, F. Li, M. X. Wu. "Mitigating semantic drift in multi-agent communication: a dynamic neuro-symbolic approach," *ZTE Communications*, vol. 24, no. 2, pp. 64–70, Jun. 2026. doi: 10.12142/ZTECOM.202602008.

1 Introduction

The rapid advancement of large language models (LLMs) has fundamentally transformed the landscape of artificial intelligence, shifting focus from individual generative tasks toward the development of autonomous, goal-oriented multi-agent systems (MAS). By simulating human-like organizational structures and collaborative reasoning, LLM-based agents have demonstrated unprecedented capabilities in solving multifaceted problems, ranging from automated software engineering to complex scientific discovery[1]. In these distributed environments, communication serves as the vital substrate that enables information exchange, task synchronization, and collective consensus[2]. However, as the complexity of collaborative tasks scales, the fundamental limitations of existing communication protocols have become increasingly apparent[3].

A critical yet often overlooked challenge in agent-to-agent interaction is the phenomenon of semantic drift, as illustrated in Fig. 1. Despite their linguistic fluency, LLM-based agents lack a shared, deterministic conceptual framework. Because agents are often initialized with different system prompts, role-specific instructions, or private domain knowledge, they frequently assign conflicting meanings to the same terminology. This "contextual drift" creates a subtle but destructive form of ambiguity, where agents operate under the illusion of agreement while pursuing divergent logical paths[4–5]. Recent studies have begun to quantify this drift, revealing that even minor misalignments in initial conceptual grounding can lead to significant reasoning hallucinations and a degradation in system reliability, particu-

---

This work is supported by Mobile Information Networks-National Science and Technology Major Project of China under Grant No. 2025ZD1304800.

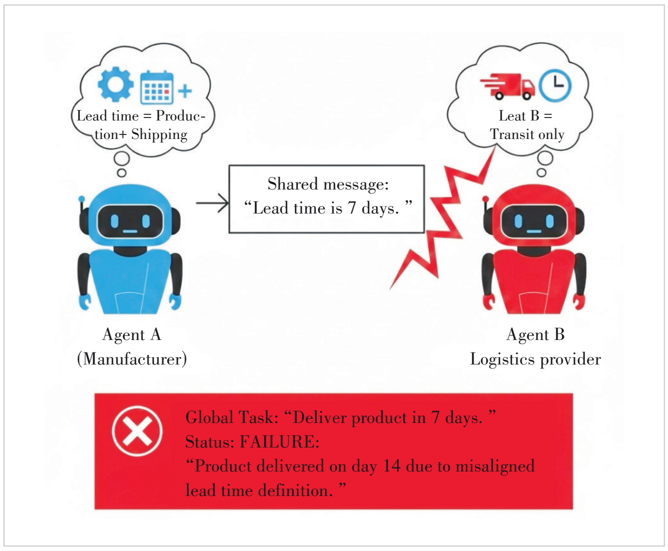


Figure 1. Semantic drift in multi-agent communication

larly in safety-critical applications such as autonomous healthcare or industrial supply chain management[6].

Research into agent communication has primarily oscillated between two suboptimal extremes. On one hand, unconstrained natural language communication offers maximum flexibility but suffers from structural vagueness and high computational overhead due to excessive token consumption[2]. On the other hand, traditional symbolic approaches, such as static ontologies, provide rigorous formalisms but are inherently too rigid to accommodate the emergent, non-predefined concepts that LLMs generate during real-time problem solving[7]. While recent efforts have explored using LLMs to assist in offline ontology matching[8–9], these methods fail to address the fluid nature of live agent-to-agent dialogues. This dichotomy necessitates a new paradigm that combines the logical rigor of formal ontologies with the adaptive reasoning of neural models[10].

To bridge this gap, we propose a novel framework called dynamic ontology alignment (DOA) for LLM-based multi-agent communication. Rather than imposing a predefined, "one-size-fits-all" global schema, our approach enables agents to autonomously negotiate and align their local semantic structures in real time as the dialogue progresses. The proposed DOA mechanism incorporates a proactive semantic probing protocol to detect conceptual mismatches at the onset of interaction, coupled with a neuro-symbolic alignment mediator that resolves conflicts by reasoning over hierarchical constraints. By dynamically grounding natural language communication in an evolving shared ontology, the DOA framework ensures that agents maintain semantic consistency across long-horizon tasks. Empirical evaluations in cross-domain collaborative scenarios demonstrate that our method significantly improves task success rates while reducing communication redundancy, providing a robust and scalable foundation for next-generation industrial-grade AI systems[11].

2 Related Work

The development of LLM-based MAS has necessitated a re-evaluation of how autonomous entities coordinate and share knowledge. Our work lies at the intersection of agent communication protocols[12–13], semantic alignment, and neuro-symbolic integration.

Existing communication protocols in LLM-based MAS fall into two main categories (also referred to as paradigms)[14]. The first is unconstrained natural language interaction, popularized by frameworks such as AutoGPT and MetaGPT, which leverage the linguistic flexibility of LLMs to simulate human-like collaboration. While versatile, these systems are prone to logical inconsistencies and "hallucinated" instructions. The second paradigm employs structured formats, such as JavaScript Object Notation (JSON) or predefined Application Programming Interface (API) schemas, to enforce syntactical consistency. However, these rigid structures often fail to capture the nuanced semantic relationships inherent in complex, multi-domain tasks, creating a "semantic gap" between the agent's internal reasoning and the external communication interface[15].

Ontology matching has evolved from classical algorithms to LLM-enhanced approaches[16–17]. Ontology matching is a well-established field in the Semantic Web community, traditionally focused on finding correspondences between heterogeneous metadata schemas. Classical approaches, such as LogMap and AgreementMakerLight (AML), utilize graph-based matching and lexical similarity to align static ontologies. While effective for stable datasets, these methods lack the contextual adaptability required for generative AI environments[18]. Recent studies, such as Agent-OM, have demonstrated that LLM-based agents can outperform traditional algorithms in identifying complex semantic mappings by leveraging their vast parametric knowledge. Nevertheless, most existing LLM-enhanced ontology matching methods, such as LLMs4om and Agent-OM, are primarily designed for offline alignment of static schema files. These approaches lack the temporal responsiveness and dynamic negotiation capabilities required for live, multi-turn agent dialogues where concepts emerge contextually. In contrast, DOA introduces a real-time mediation layer that specifically targets these "on-the-fly" semantic discrepancies. Semantic drift and the challenge of reaching consensus have been central to research on multi-agent collaboration[17]. A fundamental challenge in long-horizon agent interaction is semantic drift, where the shared understanding of a concept gradually diverges as the conversation progresses. Research into "Shared Mental Models" suggests that for effective collaboration, agents must maintain a synchronized internal representation of the task environment[19]. Becker et al.[5] quantified this drift in multi-turn interactions, noting that without an explicit grounding mechanism, agents frequently succumb to "Contextual Illusion", i.e., the false assumption of mutual understanding. Our work addresses this by introducing a dynamic re-alignment mechanism that acts as a continuous grounding layer, preventing the accumulation of semantic errors.

The landscape of semantic alignment has shifted significantly in 2024 and 2025 with the advent of more sophisticated reasoning frameworks. For instance, He et al.[20] provided a comprehensive vision for LLM-based MAS in software engineering, identifying semantic consistency as a critical bottleneck for long-horizon task coordination. Furthermore, the emergence of neuro-symbolic hybrids, as discussed by Bougzime et al.[21], underscores the necessity of combining generative flexibility with logical rigor, i.e., a principle that lies at the core of our DOA framework. Other recent works, like Kamali et al.[18], have explored compositional generalization in neuro-symbolic agents, further justifying our approach to dynamic concept grounding.

Neuro-symbolic integration and dynamic schema generation offer promising directions for robust agent communication[20]. The integration of neural models with symbolic reasoning (Neuro-Symbolic AI) provides a promising path toward more robust agent communication[21]. By grounding the "soft" reasoning of LLMs in the "hard" constraints of formal ontologies, systems can achieve both creativity and reliability[11]. Furthermore, emerging research in dynamic schema generation explores how systems can evolve their data structures in response to novel inputs. Our proposed DOA framework extends this concept by treating the ontology not as a static artifact but as a living protocol that is negotiated and refined through agent interaction, thereby combining the adaptability of neural agents with the precision of symbolic logic.

3 Methodology

The DOA framework serves as a semantic mediation layer that operates between an agent's internal reasoning engine and its external communication interface. It ensures that heterogeneous agents reach a state of mutual understanding by dynamically resolving conceptual discrepancies.

We define a MAS as a tuple $\mathcal{S} = (A, O, \Gamma)$, where $A = \{A_1, A_2, \cdots, A_n\}$ represents the set of autonomous agents, $O = \{O_1, O_2, \cdots, O_n\}$ denotes the set of local ontologies held by each agent, and Γ signifies the operational task context.

Each local ontology O_i is a structured graph (C_i, R_i, Σ_i), consisting of concepts C, relations R, and logical constraints Σ. Semantic divergence occurs when two agents A_i and A_j assign different internal representations to the same linguistic token T.

3.1 Definition of Semantic Drift Degree

Let C be a concept invoked in a message, and let $\sigma(C, O_i)\epsilon = \mathbb{R}^d$ be the semantic embedding of the concept derived from the LLM's parametric knowledge grounded in ontology O_i. The semantic drift δ between agents A_i and A_j regarding concept C is defined as:

$$\delta(A_i, A_j|C) = 1 - \frac{\sigma(C, O_i)\cdot\sigma(C, O_j)}{\left\|\sigma(C, O_i)\right\|\left\|\sigma(C, O_j)\right\|} \quad (1).$$

A mediation process is triggered if and only if $\delta > \tau$, where τ is a predefined tolerance threshold for semantic precision.

3.2 System Architecture

The DOA architecture comprises three primary modules:

- Semantic Prober: responsible for extracting concept snippets from outbound messages and identifying potential conflicts in inbound headers.
- Neuro-Symbolic Aligner: a hybrid reasoning engine that utilizes the LLM's zero-shot reasoning to propose mappings and a symbolic reasoner to verify logical consistency.
- Dynamic Consensus Vault (DCV): a volatile, session-based repository; in the implementation, it is referred to as the Volatile Consensus Vault (V). It caches established mappings $\varnothing$ to ensure temporal consistency and reduce redundant negotiations.

3.3 Dynamic Alignment Algorithm

The core logic for resolving semantic conflicts is encapsulated in the dynamic alignment algorithm. Unlike traditional batch-matching techniques, this algorithm operates on demand during the dialogue. Algorithm 1 delineates the end-to-end execution flow for resolving semantic discrepancies during active agent-to-agent dialogue. The process begins with a lightweight concept extraction phase, which isolates key predicates from the natural language stream. A critical component is the use of the semantic drift function (δ) in Step 2, which acts as a computational "trigger" to ensure that expensive reasoning resources are invoked only when conceptual divergence exceeds the tolerance threshold τ. By integrating a symbolic check in Step 4, the algorithm guarantees that the neural-generated mappings do not violate the hard logical constraints (Σ) of the domain, thereby preventing "hallucinated" alignments that often plague purely LLM-based systems. From a computational perspective, Step 2 (Compatibility Analysis) is a lightweight operation with a time complexity of $O(d)$, where d is the embedding dimension, typically executed on a CPU or a small-scale embedding model. In contrast, LLM_DeepReasoning is the primary consumer of resources, involving a forward pass of an LLM. Its GPU memory consumption scales with the context length of the source concept definition (D_{src}) and the local ontology (O_{rec}), though this cost is mitigated by the lazy alignment strategy described in Section 3.4.

Algorithm 1: LLM-driven dynamic alignment process

```
Input: Source message $M_{src}$, Source concept definition $D_{src}$,
Recipient local ontology $O_{rec}$, Context $\Gamma$
Output: Alignment mapping rule $\varnothing$
1.      Step 1: Concept Extraction
```

```
        $E \leftarrow$ Parse($M_{src}$) //Extract primary predicate $P$ and attributes {attr}
    Step 2: Compatibility Analysis
        $\delta \leftarrow$ ComputeSemanticDrift($P_{src}$, $P_{rec}$, $\Gamma$)
        If $\delta > \tau$ then
            $R_{deep} \leftarrow$ LLM_DeepReasoning($D_{src}$, $O_{rec}$, $\Gamma$)// Invoke Neural Phase
    Step 3: Conflict Resolution
            If isSubset($P_{src}$, $P_{rec}$) then
                $\varnothing \leftarrow$ DefineSubsumptionMapping($P_{src}$, $P_{rec}$)
            Else if hasAttributeMismatch({attr$_{src}$},{attr$_{rec}$}) then
                $f \leftarrow$ GenerateTransformation($x$) // e.g., unit conversion
                $\varnothing \leftarrow$ ApplyTransformation($f$)
            End if
        Else
            $\varnothing \leftarrow$ IdentityMapping
        End if
    Step 4: Consistency Verification
        If SymbolicCheck($\varnothing$,$\Sigma$) is False then
            $\varnothing \leftarrow$ RefineMapping($\varnothing$,$\Sigma$) // Resolve logical contradictions
        End if
    Step 5: Consensus Commitment
        UpdateDCV($\varnothing$) // Store in Dynamic Consensus Vault
        NotifyAgents($A_{src}$, $A_{rec}$, $\varnothing$)
    Return $\varnothing$
```

3.4 Pseudo-Code Implementation

The following pseudo-code illustrates the implementation of the alignment mediator using a neuro-symbolic algorithm. Algorithm 2 formalizes the internal reasoning mechanism of the DOA mediator. The mediator follows a "lazy alignment" strategy facilitated by the Volatile Consensus Vault (V), which minimizes redundant negotiation by caching successful mappings within the session context. The architecture exemplifies a neuro-symbolic hybrid approach: Phase 2 leverages the zero-shot reasoning capabilities of LLMs to handle linguistic variety, while Phase 3 enforces logical rigor via symbolic verification. If a mapping fails the consistency check, the mediator initiates a recursive refinement loop, ensuring that the final consensus is both semantically rich and logically sound before it is committed to the shared environment.

Algorithm 2: Dynamic ontology alignment mediator

```
Input: Source concept $C_{src}$, Recipient concept $C_{rec}$, Task context $\Gamma$
Output: Aligned semantic mapping rule $\varnothing$
    Phase 1: Volatile Cache Lookup
        key $\leftarrow$ concatenate($C_{src}$·uid,$C_{rec}$·uid)
        If key $\in V_{consensus}$ then
            Return $V_{consensus}$[key] // Retrieve from shared consensus vault
        End if
    Phase 2: Neuro-Symbolic Inference
        $\mathcal{P} \leftarrow$ GenerateAlignmentPrompt($C_{src}$, $C_{rec}$, $\Gamma$)
        $\varnothing_{raw} \leftarrow$ LLM_Reasoning($\mathcal{P}$) // Neural Phase: LLM-based semantic
    Phase 3: Symbolic Consistency Verification
        If SymbolicReasoner.check($\varnothing_{raw}$, $C_{rec}$·$\Sigma$) then
            $V_{consensus}$[key] $\leftarrow \varnothing_{raw}$ // Commit to volatile cache
            Return $\varnothing_{raw}$
        Else
            // Refinement Phase: Recursive negotiation if logic verification fails
            Return RefineNegotiation($C_{src}$, $C_{rec}$)
        End if
```

3.5 Grounded Message Transformation

Once the mapping $\varnothing$ is established, the original message M is transformed into a grounded message M^*. This transformation is defined by a function $\mathcal{G}$:

$$M^* = \mathcal{G}(M, \varnothing, \Gamma) \qquad (2).$$

In M^*, ambiguous natural language tokens are replaced or augmented with unique semantic identifiers, i.e., Uniform Resource Identifiers (URIs) from the aligned space. This ensures that Agent A_{rec} can deserialize the message into its internal state with deterministic accuracy, thereby eliminating downstream errors in the collaborative workflow.

4 Experimental Evaluation

In this section, we evaluate the DOA framework's efficacy in resolving semantic conflicts and improving collaborative performance in multi-agent systems. Our experiments aim to answer three key research questions (RQs):

RQ1: Does DOA improve the success rate of complex tasks compared to standard LLM-based communication?

RQ2: How robust is the framework against increasing levels of semantic drift?

RQ3: Does the neuro-symbolic alignment significantly reduce communication overhead (tokens) compared to raw natural language negotiation?

4.1 Experimental Setup

1) Datasets and environments

We utilized two distinct environments to simulate heterogeneous agent interactions:

- Cross-domain supply chain (CDSC): Agents represent a manufacturer (using ISO-based terminology) and a logistics provider (using proprietary transport schemas). The task requires synchronizing inventory levels and shipping schedules where units and lead times are intentionally mismatched.

- Autonomous healthcare coordination (AHC): Diagnostic agents and therapeutic agents must collaborate, using Systematized Nomenclature of Medicine—Clinical Terms (SNOMED CT) standard ontology[22] and International Classification for Diseases (ICD)-10 ontologies[23], respectively.

2) Baseline models

We compared DOA against three baselines:

- Raw-NL: unconstrained natural language communication;
- Fixed-Schema: Agents are forced to use a predefined JSON schema (static);
- LLM-Prompt-Only: Agents are prompted to "be clear and define terms" but have no formal alignment mechanism.

3) Implementation details

We used DeepSeek-R1 as the core LLM for the Neuro-Symbolic Aligner and used the Protege-OWL API for symbolic consistency checking. The drift threshold is set to $\tau = 0.3$. Additionally, all experiments were conducted on a server equipped with dual NVIDIA H100 (80 GB) GPUs and an Intel Xeon Platinum 8480C CPU to ensure reproducible latency measurements.

4.2 Task Success Rate (RQ1)

We conducted 100 trials for each environment. A trial was marked as successful only if the final state matched the global goal without any constraint violations.

As shown in Table 1, DOA outperforms the best baseline (LLM-Prompt-Only) by over 25%. The failure of the Fixed-Schema in complex tasks highlights the inability of static structures to handle the emergent concepts required for multi-domain reasoning.

Unlike baseline methods that rely solely on the LLM's internal parametric knowledge (similar to the mechanism in LLM-Prompt-Only), DOA's superiority lies in its Neuro-Symbolic hybridity. While state-of-the-art LLM-based aligners often focus on lexical similarity, our framework ensures that proposed mappings are validated against hard logical constraints (Σ), preventing the reasoning hallucinations that often lead to task failure in purely neural approaches.

Table 1. Task success rate across frameworks in CDSC and AHC

Framework	CDSC/%	AHC/%	Average/%	Improvement over Baseline/%
Raw-NL	62.5	54.0	58.3	**31.5**
Fixed-Schema	45.0	38.5	41.8	**48.0**
LLM-Prompt-Only	67.5	61.5	64.5	**25.3**
DOA (Ours)	**91.5**	**88.0**	**89.8**	–

AHC: autonomous healthcare coordination
CDSC: cross-domain supply chain
DOA: dynamic ontology alignment
LLM: large language model
NL: natural language

4.3 Robustness to Semantic Drift (RQ2)

To evaluate the resilience of the DOA framework, we conducted a sensitivity analysis by manually injecting "semantic noise" into the agents' local ontologies. We controlled the semantic drift degree (δ) by perturbing concept embeddings with Gaussian noise and introducing disjoint taxonomic labels for identical concepts, ranging from $\delta = 0.1$ (minor terminology variation) to $\delta = 0.6$ (significant conceptual misalignment). The detailed success rates under different drift levels are shown in Table 2.

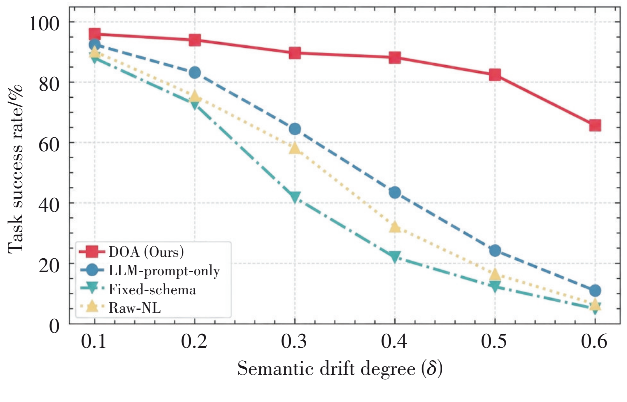
As illustrated in Fig. 2, there is a clear "performance cliff" for baseline methods.

- Breakdown of Raw-NL: When $\delta > 0.3$, the performance of the Raw-NL baseline collapses. This confirms the "contextual illusion" hypothesis, i. e., agents continue to communicate using shared tokens but are unaware that their underlying logical grounding has diverged, leading to irreversible reasoning errors.
- DOA's graceful degradation: In contrast, the DOA frame-

Table 2. System success rate vs. semantic drift (δ)

Drift (δ)	Raw-NL/%	Fixed-Schema/%	LLM-Prompt-Only/%	**DOA (Ours)/%**
0.1 (Low)	90.3	88.0	92.5	**96.0**
0.2	75.5	72.8	83.3	**94.0**
0.3 (Threshold)	58.3	41.8	64.5	**89.8**
0.4	32.3	22.0	43.5	**88.3**
0.5 (High)	16.5	12.3	24.3	**82.5**
0.6 (Critical)	6.5	5.0	11.0	**65.8**

DOA: dynamic ontology alignment
LLM: large language model
NL: natural language



DOA: dynamic ontology alignment
LLM: large language model
NL: natural language

Figure 2. Robustness analysis of the impact of semantic drift δ on task success rate, where the DOA framework maintains stability compared to the sharp decline of baseline methods

work maintains a success rate above 80% even at high levels of drift (δ = 0.5). The stability of the curve is attributed to the semantic prober's proactive detection mechanism. By calculating the drift before the task execution begins, DOA triggers the Neuro-Symbolic Aligner to bridge the gap early.

- Symbolic anchor: Even at the "Critical" drift level (δ = 0.6), where natural language becomes almost deceptive, the symbolic consistency check ensures that the agents do not commit to contradictory plans, allowing the system to fail safely or request human intervention rather than proceeding with hallucinated consensus.

4.4 Communication Efficiency (RQ3)

One major concern with LLM-based negotiation is the "token explosion". We measured the average tokens consumed per successful task completion.

As shown in Table 3, while Fixed-Schema exhibits the lowest token consumption and latency due to its rigid structure, it fails to handle emergent semantic conflicts as previously shown in Table 1. Our DOA framework strikes an optimal balance: it incurs slightly higher initial overhead than Fixed-Schema due to the neuro-symbolic alignment phase. Although the observed alignment latency (1.4 s) is acceptable for supply chain coordination, it may pose challenges in ultra-high-stakes scenarios, such as robotic surgery or acute medical emergency coordination, where sub-second responses are mandatory. However, since DOA utilizes a "lazy alignment" strategy via the DCV, this latency is typically a one-time cost per concept, with subsequent interactions benefiting from cached mappings at near-zero overhead.

During our experiments, the peak GPU memory utilization for the Neuro-Symbolic Aligner was recorded at approximately 24 GB per instance when processing complex healthcare ontologies. The ComputeSemanticDrift function accounted for less than 1% of the total alignment latency, confirming that the gating mechanism successfully prevents unnecessary invocations of the more expensive neural reasoning phase.

Table 3. Communication overhead analysis: comparison of efficiency metrics per successful task completion

Metric	Raw-NL	Fixed-Schema	LLM-Prompt-Only	**DOA (Ours)**
Avg. turns to consensus	8.4	2.0*	6.2	3.1
Avg. tokens per task	4 200	1 200	3 850	1 950
Alignment latency / s	N/A	0.3	N/A	1.4

Notes: * Fixed-Schema achieves the lowest number of consensus turns among all frameworks, but it also has the lowest task success rate (see Table 1).

DOA: dynamic ontology alignment
LLM: large language model
NL: natural language

5 Conclusions and Future Work

In this paper, we introduce the DOA framework, a neuro-symbolic approach designed to mitigate semantic drift in LLM-based multi-agent systems. By combining the adaptive reasoning of LLMs with the formal rigor of symbolic consistency checking, our framework ensures that heterogeneous agents maintain synchronized conceptual grounding during real-time collaboration.

Our experimental results demonstrate that DOA increases task success rates by an average of 25.25% over natural language baselines, reduces token consumption by 50% through efficient grounded message transformation, and maintains robustness in the face of significant conceptual divergence.

Despite its efficacy in structured domains, a potential limitation of the current DOA framework lies in its performance on unconstrained open-domain dialogues. Specifically, when agents encounter "out-of-ontology" concepts that lack any formal parent-child relationship or attribute constraints in their local knowledge bases, the neuro-symbolic aligner may struggle to provide a verifiable mapping. In such cases, the system defaults to probabilistic neural alignment without the added guarantees of symbolic verification.

Beyond current evaluations, we recognize the necessity of validating DOA in domains with stricter real-time constraints and safety sensitivities. We intend to extend the framework to cooperative autonomous driving scenarios, where sub-second semantic alignment is critical for collision avoidance. Furthermore, we plan to incorporate active learning to allow local ontologies to evolve incrementally based on successful alignments.

References

[1] Qian C, Liu W, Liu H Z, et al. ChatDev: communicative agents for software development [C]//Proc. 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). ACL, 2024: 15174 – 15186. DOI: 10.18653/v1/2024.acl-long.810

[2] Xi Z H, Chen W X, Guo X, et al. The rise and potential of large language model based agents: a survey [J]. Science China information sciences, 2025, 68(2): 121101. DOI: 10.1007/s11432-024-4222-0

[3] Yang Y X, Chai H C, Shao S, et al. AgentNet: decentralized evolutionary coordination for LLM-based multi-agent systems [PP/OL]. V1. arXiv (2025-04-01)[2025-12-12]. https://doi.org/10.48550/arXiv.2504.00587

[4] Park J S, O'Brien J, Cai C J, et al. Generative agents: interactive simulacra of human behavior [C]//36th Annual ACM Symposium on User Interface Software and Technology. ACM, 2023: 1 – 22. DOI: 10.1145/3586183.3606763

[5] Becker J, Kaesberg L B, Stephan A, et al. Stay focused: problem drift in multi-agent debate [PP/OL]. V1. arXiv (2025-02-26)[2025-12-12]. https://doi.org/10.48550/arXiv.2502.19559

[6] Chen W Z, Su Y S, Zuo J W, et al. AgentVerse: facilitating multi-agent collaboration and exploring emergent behaviors [C]//International Conference on Learning Representations. ICLR, 2024: 1 – 43. DOI: 10.48550/arXiv.2308.10848

[7] Noy N F, McGuinness D L. Ontology development 101: a guide to creating your first ontology [EB/OL]. [2026-01-13]. http://protege.stanford.edu/publications/ontology_development/ontology101.pdf

[8] Giglou H B, D'Souza J, Engel F, et al. LLMs4OM: matching ontologies with large language models [PP/OL]. V2. arXiv (2024-04-23)[2026-01-14]. https://doi.org/10.48550/arXiv.2404.10317

[9] Qiang Z C, Wang W Q, Taylor K. Agent-OM: leveraging LLM agents for ontology matching [J]. Proceedings of the VLDB endowment, 2024, 18(3): 516 – 529. DOI: 10.14778/3712221.3712222

[10] Hitzler P, Sarker M K. Neuro-symbolic artificial intelligence: the state of the art [M]. Amsterdam: IOS Press, 2021. DOI: 10.3233/faia342

[11] Chang E Y. Multi-LLM agent collaborative intelligence: the path to artificial general intelligence [M]. New York: ACM, 2025. DOI: 10.1145/3749421

[12] Hong S R, Zhuge M C, Chen J, et al. MetaGPT: meta programming for a multi-agent collaborative framework [C]//International Conference on Learning Representations. ICLR, 2024: 1 – 29. DOI: 10.48550/arXiv.2308.00352

[13] Cancedda N, Dessi R, Dwivedi-Yu J, et al. Toolformer: language models can teach themselves to use tools [C]//Proceedings of Advances in Neural Information Processing Systems 36. Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2023: 68539 – 68551. DOI: 10.52202/075280-2997

[14] Renney H, Nethercott M N, Renney N, et al. LLM-enabled multi-agent systems: empirical evaluation and insights into emerging design patterns & paradigms [PP/OL]. V1. arXiv (2026-01-06)[2026-02-01]. https://doi.org/10.48550/arXiv.2601.03328

[15] Wang Z, Moriyama S, Wang W Y, et al. Talk structurally, act hierarchically: a collaborative refinement framework for LLM multi-agent systems [J]. IEEE transactions on artificial intelligence, 2026: 1 – 11. DOI: 10.1109/tai.2026.3676025

[16] Jiménez-Ruiz E, Cuenca Grau B. LogMap: logic-based and scalable ontology matching [C]//The Semantic Web—ISWC 2011. Springer, 2011: 273 – 288. DOI: 10.1007/978-3-642-25073-6_18

[17] He Y, Chen J Y, Dong H, et al. Exploring large language models for ontology alignment [PP/OL]. V1. arXiv (2023-09-12)[2026-01-18]. https://doi.org/10.48550/arXiv.2309.07172

[18] Kamali D, Barezi E J, Kordjamshidi P. NeSyCoCo: a neuro-symbolic concept composer for compositional generalization [J]. Proceedings of the AAAI conference on artificial intelligence, 2025, 39(4): 4184 – 4193. DOI: 10.1609/aaai.v39i4.32439

[19] Schelble B G, Flathmann C, McNeese N J, et al. Let's think together! assessing shared mental models, performance, and trust in human-agent teams [J]. Proceedings of the ACM on human-computer interaction, 2022, 6: 1 – 29. DOI: 10.1145/3492832

[20] He J D, Treude C, Lo D. LLM-based multi-agent systems for software engineering: literature review, vision, and the road ahead [J]. ACM transactions on software engineering and methodology, 2025, 34(5): 1 – 30. DOI: 10.1145/3712003

[21] Bougzime O, Jabbar S, Cruz C, et al. Unlocking the potential of generative AI through neuro-symbolic architectures: benefits and limitations [PP/OL]. V1. arXiv (2025-02-16)[2026-02-23]. https://doi.org/10.48550/arXiv.2502.11269

[22] El-Sappagh S, Franda F, Ali F, et al. SNOMED CT standard ontology based on the ontology for general medical science [J]. BMC medical informatics and decision making, 2018, 18: 76. DOI: 10.1186/s12911-018-0651-5

[23] Möller M, Sintek M, Biedert R, et al. Representing the international classification of diseases version 10 in OWL [C]//International Conference on Knowledge Engineering and Ontology Development. KEOD, 2010: 50 – 59

Biographies

Xie Linhao is currently an undergraduate student at Communication University of China. He is also an intern at the Product R&D Center of AsiaInfo Technologies. During his university years, he received several awards, including the First Prize in the 2025 AI.Talk National College Student Artificial Intelligence Knowledge Competition, the Second Prize in the North China region of the 15th MathorCup Mathematical Application Challenge in 2025, and the National Third Prize in the 2024 National Mathematics Competition. His research interests include multi-agent systems, computer vision, and intelligent robotics.

Li Fan is a Senior Engineer at China Unicom Beijing Branch. She holds a Master of Engineering from Beijing University of Posts and Telecommunications, China. Her research interests include 4G/5G network optimization, mobile network digital operation, intelligentization of mobile communication networks, intelligent optimization and operation and maintenance, and emerging mobile communication technologies.

Wu Mingxuan received his MS degree from Peking University, China. He is currently the Deputy Manager of the Information Technology Department at CRSC Information Industry Co., Ltd., and holds the professional title of Engineer. His research interests include application of AI agents in vertical fields, digital transformation for large enterprises, and smart city solutions.

Song Yong (songyong@asiainfo.com) received his MS degree from Peking University, China. He is currently an algorithm expert at the Product R&D Center of AsiaInfo Technologies. He received the Second Prize of the Beijing Science and Technology Progress Award. His research interests include automatic ontology generation, multi-agent systems, and the Internet of Agents.

Ouyang Ye is a professor and an IEEE Fellow. He serves as the Chief Executive Officer and Chief Technology Officer at AsiaInfo Technologies. He holds a Bachelor of Engineering from Southeast University in China, a Master of Science from Tufts University in the United States, a second Master of Science from Columbia University in the United States, and a PhD from Stevens Institute of Technology in the United States. He has extensive experience in large-scale team management and R&D innovation in the ICT field. He focuses on cross-domain innovation and the commercialization of technologies in cellular networks, AI, and data science.