Online Boundary-Aware Memory for Case-Based Reasoning Agents

Zheng Dong [1] **Luming Shang** [1]

Abstract

Large language model (LLM) agents increasingly operate in streaming case-based reasoning (CBR) settings, where continuous improvement from past experience is crucial. Existing methods achieve this by storing past cases and retrieving similar ones as few-shot examples. This strategy fails near decision boundaries, where highly similar cases have conflicting outcomes and the discriminative factors are not explicit. We introduce Online Boundary-Aware Memory (OBAM), an agent memory architecture that discovers and refines decision-boundary knowledge progressively from the case stream. Unlike offline boundary analysis over a static dataset, OBAM detects boundaries online as the agent encounters contrasting cases and stores structured boundary memory entries encoding shared patterns and discriminative rules. These entries are refined as evidence accumulates, transforming ambiguous case experience into reusable decision knowledge. Across legal, medical, and fraud reasoning tasks, OBAM outperforms in-context learning and state-of-the-art agent memory baselines, demonstrating the value of boundary-aware online memory for continual improvement in streaming CBR agents.

1. Introduction

Large language models (LLMs) have been widely adopted as case-based reasoning (CBR) agents that make structured decisions repeatedly over sequential streams of cases, including medical differential diagnosis (Tang et al., 2024), legal holding identification (Cui et al., 2023), and fraud detection (Singh et al., 2025). In the human counterpart of these tasks, professionals improve through repeated exposure to cases. They remember prior decisions, recognize recurring patterns, and gradually learn which subtle distinctions determine different outcomes. Existing LLM agents typically operationalize this process through standard CBR cycle, where they store past cases in memory, and the most similar cases are retrieved as few-shot examples to guide future decisions via in-context learning (ICL).

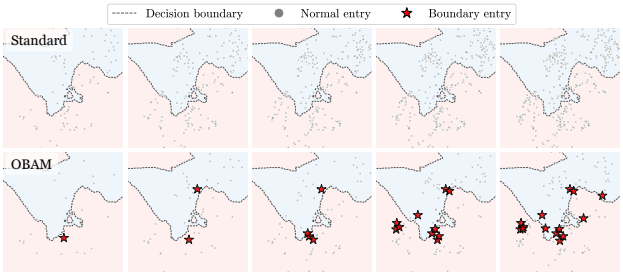

Figure 1. Comparison of memory evolution between standard memory in CBR (top) and OBAM (bottom) on a fraud reasoning task. Standard CBR memory accumulates entries indiscriminately and stores cases in isolation. OBAM stores boundary entries (red stars) near the decision boundary (dashed line) while keeping interior regions sparse, achieving comparable region coverage with fewer entries concentrated where disambiguation is needed. Background shading indicates KNN-inferred class regions.

Despite its simplicity, this strategy has a fundamental limitation in streaming CBR that it treats memory as a repository of isolated cases rather than as an evolving representation of decision knowledge. This limitation is pronounced near *decision boundaries*, where highly similar cases can require different correct decisions. At such boundaries, the agent naturally surfaces cases from both sides of the boundary and is exposed to contradictory precedents. However, they provide no explicit representation of the relationship between conflicting cases. The agent is therefore forced to infer the minimal distinguishing feature from scattered examples while simultaneously solving the current decision problem. Meanwhile, unlike supervised training where the full dataset is available for offline boundary analysis, streaming CBR requires the agent to discover and represent boundaries incrementally from individual cases as they arrive. The same boundary can be re-encountered repeatedly, yet the agent never accumulates the disambiguation because the simple case-based memory lacks the capacity for such experience storage. This becomes an unreliable basis for continual improvement of agent performance.

We propose Online Boundary-Aware Memory (OBAM), a memory architecture for streaming CBR that enables continual improvement by discovering and refining decision boundary knowledge progressively. Inspired by the estab-

[1]Amazon, Seattle, USA. Correspondence to: Zheng Dong <zhengo@amazon.com>.

Published at the Second Workshop on Agents in the Wild: Safety, Security, and Beyond (AIWILD) at ICML 2026.

lished principle that boundary cases carry disproportionate decision-relevant information (Vapnik, 1995; Smyth & McKenna, 1999) and such information can be learned incrementally (Cauwenberghs & Poggio, 2000), OBAM detects boundaries online via empirical neighbor agreement and stores *boundary entries*, a structured memory unit that encodes cross-case contrastive knowledge rather than a single case alone. This allows agents to retrieve the discrimination rule directly based on similar shared patterns and eliminate the need for inference-time contrastive reasoning. Critically, these entries are refined as being discovered continuously, enabling the memory to develop precise disambiguation through continued exposure. As shown in Figure 1, OBAM concentrates learned experience at decision boundaries while keeping interior regions sparse, achieving effective coverage with fewer entries focused where they are most informative. Our contributions are:

1. We propose OBAM, an online boundary-aware memory architecture for streaming CBR agents that enables boundary-centric experience accumulation. It concentrates memory resources at decision boundaries where the agent's errors occur, while maintaining sparse coverage of well-understood regions.
2. We realize this architecture through a structured memory design comprising boundary memory entries, which are refined progressively as the agent encounters additional boundary evidence from the case stream.
3. We evaluate on three domain tasks (legal, medical, fraud) across four datasets with different LLM backbones, demonstrating that OBAM achieves superior continual improvement over both non-streaming and streaming CBR baselines as well as state-of-the-art agent memory systems designed for other task types.

2. Method

We consider an LLM agent $\mathcal{A}$ that processes a sequential stream of cases $c_1, c_2, \ldots$, where each case $c_t = (x_t, y_t)$ pairs a natural-language case description x_t with a ground-truth outcome $y_t \in \mathcal{D}$ in a finite decision space. At each time t, the agent observes x_t, produces a decision d_t informed by its current memory state, and subsequently receives the true label y_t as feedback. The agent maintains an evolving external memory $\mathcal{M}_t$ that accumulates structured experience from the case stream. The objective is to maximize cumulative decision success rate over the stream.

2.1. OBAM Memory Schema

Central to OBAM is a dual-entry memory schema that distinguishes between individual case experience and cross-case boundary knowledge.

Region entries. A region entry records the experience from a single case: $m_i = (s_i, y_i, l_i, \mathbf{e}_i)$, where s_i is a natural-language summary of the case context, y_i is the ground-truth decision, l_i is a lesson distilled by the LLM through self-reflection, and $\mathbf{e}_i \in \mathbb{R}^d$ is the embedding of s_i.

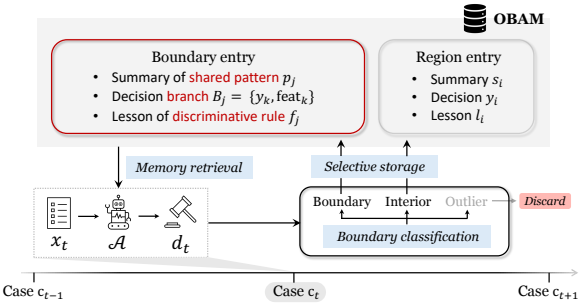

Figure 2. The OBAM lifecycle.

Boundary entries. A boundary entry extends from region entries to encode a decision boundary instead of single case lesson: $m_j = (p_j, B_j, f_j, \mathbf{e}_j)$, where p_j is the shared pattern summarized from the context of cases on both sides of the boundary, $B_j = \{(y_k, \text{feat}_k)\}_{k=1}^{|B_j|}$ is a set of decision branches each pairing an outcome type with its distinguishing features, f_j is the discriminative rule capturing the minimal factor that determines which branch applies, and $\mathbf{e}_j \in \mathbb{R}^d$ is the embedding of p_j.

Note that boundary entries reflect the structured relationship between similar cases and how their decisions diverge, while a region entry describes a single case and its outcome. This representation enables the memory to serve disambiguation directly at retrieval time, rather than leaving the agent to infer it from conflicting individual entries.

2.2. The OBAM Lifecycle

During the CBR streaming, the agent memory $\mathcal{M}_t$ interacts with the incoming case c_t through the OBAM lifecycle, as illustrated in Figure 2. The agent first summarizes the case context into s_t and performs memory retrievals based on embedding $\mathbf{e}_t$. We retrieve top-K_1 region entries and the top-K_2 boundary entries from $\mathcal{M}_t$, both filtered by a minimum cosine similarity threshold θ_{sim}. Let R_t and R_t^b denote the retrieved region and boundary sets, respectively. Both are presented in the agent context window together with case input x_t for the agent to produce decision d_t.

The next phase of boundary classification plays the key role in OBAM. Upon receiving ground-truth y_t and the distilled case lesson l_t, the system classifies the case based on the internal agreement rate among its retrieved neighbors:

$$\text{agree}(R_t) = \frac{|\{(i,j) : y_i = y_j,\ m_i, m_j \in R_t\}|}{\binom{|R_t|}{2}}.$$

This yields a three-way classification of (1) *Boundary* if $\text{agree}(R_t) < \theta_{\text{agree}}$, indicating a divided neighborhood. (2) *Interior* with $\text{agree}(R_t) \geq \theta_{\text{agree}}$ and y_t consistent with

the neighbor majority, suggesting the case conforms to the neighborhood. (3) *Outlier* if agree$(R_t) \geq \theta_{\text{agree}}$ and y_t inconsistent, suggesting the case disagrees with the neighborhood. Particularly, we determine the case to fall in a *novel region* if $|R_t| < K_1$ and $|R_t^b| = 0$, indicating insufficient neighbors exist for reliable classification. We note that this online classification operates efficiently on empirical label distributions and requires no LLM judgment.

OBAM applies class-specific memory storage strategies. For a Boundary case, it either creates a new boundary entry or refines an existing one, depending on whether boundary entries were retrieved. If $|R_t^b| = 0$, OBAM identifies the most contrasting region entry $m^* \in R_t$ whose ground-truth decision differs from y_t and whose embedding is most similar to $\mathbf{e}_t$. The LLM then contrasts c_t with m^* to extract the shared pattern p, decision branches B covering each outcome type and distinguishing features, and discriminative rule f that separates the outcomes, forming a new boundary entry in $\mathcal{M}_t$. If $|R_t^b| \neq 0$, the most relevant boundary entry b^* is refined accordingly: (1) If y_t matches an existing branch in B, the LLM strengthens the discriminative rule f using additional evidence; (2) If y_t represents a unseen outcome, a new branch $(y_t, \text{feat}_{\text{new}})$ is added to B and the discriminative rule f is updated to accommodate the additional decision path. This refinement enables boundary entries to improve progressively as the agent encounters more cases near the boundaries, analogous to how domain experts sharpen their intuitions about edge cases over time. We always store cases in novel region and introduce a sampling strategy to store Interior cases with probability $\alpha < 1$, as they carry redundant messages that are already well-understood in the past experience, maintaining high concentration on the boundary memories. We discard Outliers as storing them would introduce noise (Wilson, 1972).

2.3. Warm-up

The OBAM lifecycle requires populated memory to compute neighbor agreement. Thus, OBAM begins with a warm-up phase that stores the first W cases as region entries (defaults to 1% of the stream length in experiments). During warm-up, the agent operates in zero-shot mode: stored entries are excluded from retrieval, and boundary classification is disabled. After warm-up, these entries become retrievable, providing sufficient neighborhood density for reliable agreement computation and genuine boundary distinction.

3. Experiments

We adopt the streaming evaluation setup (Wu et al., 2024) that feeds testing set into the model sequentially and evaluate model performance using aggregate success metric at the final time step T. For example, the aggregate accuracy is calculated as $\frac{\sum_1^T \mathbb{I}(y_t = d_t)}{T}$. This marks how success the agent meet as many user requirements as possible over time.

Table 1. Model testing performance. Best result in **bold**. OBAM outperforms all baselines across datasets and LLM backbones.

Backbone	Method	CaseHOLD	DDXPlus	Phish	Review
Claude Sonnet 4	Zero-shot	0.788	0.753	0.805	0.593
	Few-shot	0.796	0.775	0.847	0.625
	AWM	0.805	0.807	0.882	0.635
	Mem0	0.812	0.851	0.891	0.648
	Self-StreamICL	0.807	0.862	0.894	0.650
	OBAM	**0.825**	**0.919**	**0.926**	**0.663**
Mistral Large 3	Zero-shot	0.722	0.650	0.738	0.573
	Few-shot	0.736	0.772	0.822	0.606
	AWM	0.732	0.780	0.840	0.603
	Mem0	0.740	0.805	0.882	0.627
	Self-StreamICL	0.742	0.813	0.875	0.618
	OBAM	**0.749**	**0.862**	**0.917**	**0.653**

Datasets. We evaluate on four datasets across three CBR task types (details in Appendix B): legal holding identification (CaseHOLD), medical differential diagnosis (DDXPlus), and fraud detection (DiFraud-Phish and DiFraud-Product Reviews). These datasets cover major real-world CBR domains, and accuracy is used as the success metric throughout. Since the original datasets have no temporal dependency, we randomly shuffle and serialize them for the streaming setup. Appendix C shows that performance is stable across random permutations.

Baselines. We compare OBAM against both non-streaming and streaming state-of-the-arts memory baselines. Non-streaming methods include (1) **Zero-shot** with the instruction-only LLM inference; (2) **Few-shot** with fixed selected exemplars. Streaming methods include (3) **AWM** (Wang et al., 2025) that stores procedural workflows for long-horizon tasks; (4) **Mem0** (Chhikara et al., 2025) that stores atomic conversational facts for general use; (5) **Self-StreamICL** (Wu et al., 2024) that accumulates and retrieves examples from a case store for streaming tasks, which is the closest to a standard CBR memory. We evaluate all methods on Claude Sonnet 4 (Anthropic, 2024) and Mistral Large 3 (Mistral AI, 2025). See Appendix B for the hyper-parameter setup and Appendix C for sensitivity analysis that confirms our model robustness.

3.1. Baseline Comparisons

Table 1 reports aggregated test accuracy. OBAM achieves the best performance across datasets and both LLM backbones, outperforming memory baselines designed for other task types and highlighting the need for CBR-specific memory. Its consistent gains over Self-StreamICL further show the value of boundary memories for complex decision boundaries, such as DDXPlus with 49 classes. All streaming methods outperform non-streaming baselines, confirming that accumulated knowledge benefits online decision-making.

Figure 3 traces the continual improvement of model performance across checkpoints. OBAM improves faster than other baselines in the early stages, as structured bound-

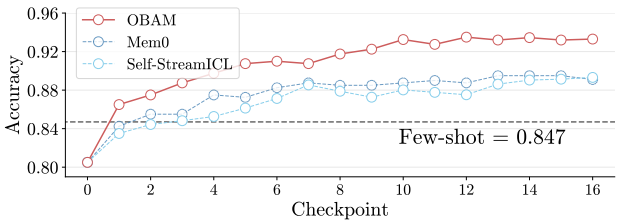

Figure 3. Test accuracy at different checkpoints during the case stream. OBAM improves faster and sustains a widening gap over baselines as boundary entries accumulate and refine.

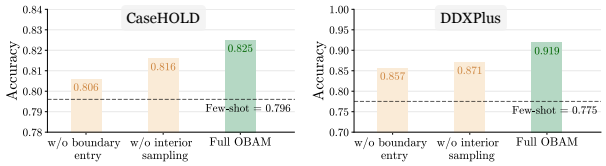

Figure 4. Ablation study on CaseHOLD and DDXPlus.

ary entries provide effective disambiguation from the first boundary detection onward. Notably, the performance gap between OBAM and baselines widens over time while the baseline performance becomes stagnant. This widening reflects the advantages of our progressive boundary refinement as additional cases arrive near previously identified boundaries, yielding continued accuracy gains without proportional memory growth. Overall, the results demonstrate that our boundary-aware memory enables more effective continual improvement than existing agent memory designs.

3.2. Ablation Study

To validate the contributions of boundary entry storage and selective accumulation via interior sampling, we conduct an ablation study on CaseHOLD and DDXPlus, with results shown in Figure 4. Removing boundary entries reduces OBAM to a standard CBR memory and causes substantial performance drops on both datasets, confirming that structured boundary representation is the main source of improvement. Without such entries, the agent remains vulnerable to conflicting precedents near decision boundaries. In contrast, storing all interior cases increases memory size linearly while slightly degrading accuracy. This indicates that indiscriminate accumulation introduces retrieval noise and dilutes boundary knowledge in agent memory. Thus, selective storage both controls memory growth and preserves the signal-to-noise ratio of retrieved context.

3.3. Case Study: Boundary Entries in Action

Figure 5 shows a DDXPlus example where a boundary entry corrects an otherwise wrong prediction. In DDXPlus, the agent receives a patient profile describing symptoms, medical history, and antecedents, and must select the correct pathology from 49 possible diagnoses. Respiratory symptoms such as shortness of breath and productive cough, together with COPD history, lead the baseline agent to predict acute COPD exacerbation. However, the patient also has

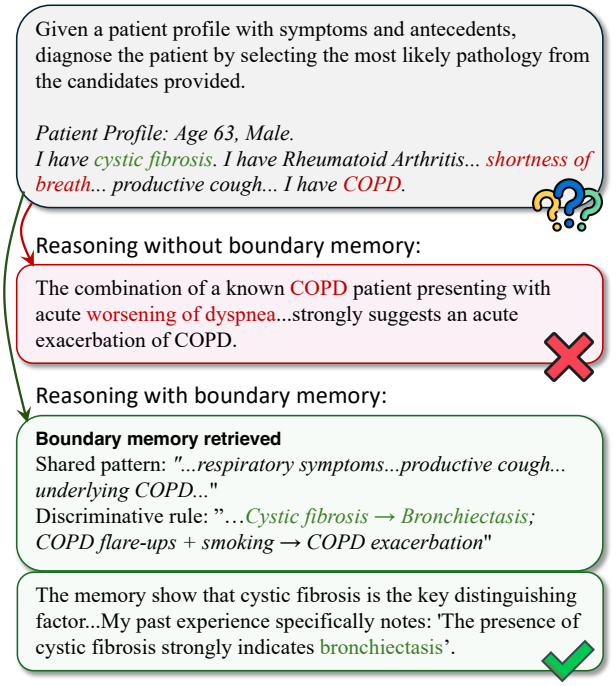

Figure 5. Example of a boundary entry correcting agent decision. The patient's COPD history misleads the baseline but the boundary entry identifies the discriminative factor of cystic fibrosis.

cystic fibrosis, the decisive factor distinguishing bronchiectasis from COPD exacerbation in overlapping respiratory presentations. The boundary entry successfully captures this rule from prior contrasting cases and guides the agent to the correct diagnosis of bronchiectasis. Additional examples from other datasets are provided in Appendix D.

4. Conclusion

We presented OBAM, an online boundary-aware memory architecture for streaming CBR agents that discovers and refines decision boundary knowledge progressively. OBAM provides the agent with cross-case disambiguation that isolated case retrieval cannot supply and addresses the limitation of existing CBR agents at decision boundaries. Experiments across legal, medical, and fraud reasoning tasks demonstrate that OBAM achieves superior continual improvement over both non-streaming baselines and state-of-the-art agent memory systems. These results establish that organizing agent memory around decision boundaries, rather than storing experience uniformly, is a principled and effective strategy for continual improvement of LLM agents in streaming case-based decision-making. This conclusion, however, demands caution in high-stakes domains such as medical (Tang et al., 2024), legal reasoning (Cui et al., 2023), and public health (Dong et al., 2023b; Liao et al., 2026), where a boundary rule induced from noisy or biased feedback may entrench a spurious distinction. The boundary entries can be better advocated as auditable decision support subject to expert review rather than as autonomous authority.

Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none which we feel must be specifically highlighted here.

References

Aamodt, A. and Plaza, E. *Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches*, volume 7. IOS Press, 1994.

Aljundi, R., Belilovsky, E., Tuytelaars, T., Charlin, L., Caccia, M., Lin, M., and Page-Caccia, L. Online continual learning with maximal interfered retrieval. *Advances in Neural Information Processing Systems*, 32, 2019.

Anthropic. The Claude model family. https://www.anthropic.com, 2024.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. Language models are few-shot learners. In *Advances in Neural Information Processing Systems*, volume 33, pp. 1877–1901, 2020.

Buzzega, P., Boschini, M., Porrello, A., Abati, D., and Calderara, S. Dark experience for general continual learning: A strong, simple baseline. In *Advances in Neural Information Processing Systems*, volume 33, pp. 15920–15930, 2020.

Cauwenberghs, G. and Poggio, T. Incremental and decremental support vector machine learning. In *Advances in Neural Information Processing Systems*, volume 13, 2000.

Chhikara, P., Khullar, D., and Baranwal, T. S. Mem0: Building production-ready AI agents with scalable long-term memory. *arXiv preprint arXiv:2504.19413*, 2025.

Cui, J., Li, M., Gao, J., Deng, S., and Shang, R. ChatLaw: Open-source legal large language model with integrated external knowledge bases. *arXiv preprint arXiv:2306.16092*, 2023.

Dong, Z., Cheng, X., and Xie, Y. Spatio-temporal point processes with deep non-stationary kernels. In *The Eleventh International Conference on Learning Representations*, 2023a.

Dong, Z., Zhu, S., Xie, Y., Mateu, J., and Rodríguez-Cortés, F. J. Non-stationary spatio-temporal point process modeling for high-resolution covid-19 data. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 72 (2):368–386, 2023b.

Dong, Z., Fan, Z., and Zhu, S. Conditional generative modeling for high-dimensional marked temporal point processes. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 1*, pp. 248–259, 2025a.

Dong, Z., Shang, L., and Olinto, G. GreenTEA: Gradient descent with topic-modeling and evolutionary auto-prompting. In *First International KDD Workshop on Prompt Optimization, 2025*, 2025b.

Fansi Tchango, A., Goel, R., Wen, Z., Martel, J., and Ghosn, J. DDXPlus: A new dataset for automatic medical diagnosis. *Advances in Neural Information Processing Systems*, 35, 2022.

Guo, Q., Wang, R., Guo, J., Li, B., Song, K., Tan, X., Liu, G., Bian, J., and Yang, Y. Connecting large language models with evolutionary algorithms yields powerful prompt optimizers. In *The Twelfth International Conference on Learning Representations*, 2024.

Hart, P. E. The condensed nearest neighbor rule. *IEEE Transactions on Information Theory*, 14(3):515–516, 1968.

Lewis, D. D. and Gale, W. A. A sequential algorithm for training text classifiers. In *Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 3–12. Springer-Verlag, 1994.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., et al. Retrieval-augmented generation for knowledge-intensive NLP tasks. In *Advances in Neural Information Processing Systems*, volume 33, pp. 9459–9474, 2020.

Li, J., Liu, X., and Montreuil, B. A multi-agent system for operating hyperconnected freight transportation in the physical internet. *IFAC-PapersOnLine*, 56(2):7585–7590, 2023.

Li, J., Liu, X., Dahan, M., and Montreuil, B. Stochastic service network design with different operational patterns for hyperconnected relay transportation. *Proceedings of 9th International Physical Internet Conference (IPIC)*, 2024.

Liao, C.-Y., Dong, Z., Garcia, G.-G. P., Paynabar, K., Xie, Y., and Jalali, M. S. Tides need stemmed: A locally operating spatiotemporal mutually exciting point process with dynamic network for improving opioid overdose death prediction. *Manufacturing & Service Operations Management*, 28(2):577–593, 2026.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. Focal loss for dense object detection. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2980–2988, 2017.

Liu, X., Li, J., and Montreuil, B. Logistics hub capacity

deployment in hyperconnected transportation network under uncertainty. In *IISE Annual Conference Proceedings*. Institute of Industrial and Systems Engineers (IISE), 2023.

Liu, X., Li, J., Dahan, M., and Montreuil, B. Multi-period stochastic logistic hub capacity planning for relay transportation. In *Institute of Industrial and Systems Engineers (IISE)*, 2024.

Liu, X., Li, J., Dahan, M., and Montreuil, B. Dynamic hub capacity planning in hyperconnected relay transportation networks under uncertainty. *Transportation Research Part E: Logistics and Transportation Review*, 194:103940, 2025a.

Liu, X., Muthukrishnan, P., and Montreuil, B. Network design and capacity management in hyperconnected urban logistic networks. *Proceedings of 11th International Physical Internet Conference (IPIC)*, 2025b.

Liu, X., Klibi, W., and Montreuil, B. Modular and mobile capacity planning for hyperconnected supply chain networks. *arXiv preprint arXiv:2601.11107*, 2026.

Mistral AI. Introducing Mistral 3, 2025. URL https://mistral.ai/news/mistral-3. Blog post, December 2, 2025.

Ouyang, S., Yan, J., Hsu, I.-H., Chen, Y., Jiang, K., Wang, Z., Han, R., Le, L., Daruki, S., Tang, X., Tirumalashetty, V., Lee, G., Rofouei, M., Lin, H., Han, J., Lee, C.-Y., and Pfister, T. Reasoningbank: Scaling agent self-evolving with reasoning memory. In *The Fourteenth International Conference on Learning Representations*, 2026.

Packer, C., Wooders, S., Lin, K., Fang, V., Patil, S. G., Stoica, I., and Gonzalez, J. E. MemGPT: Towards LLMs as operating systems. *arXiv preprint arXiv:2310.08560*, 2023.

Rolnick, D., Ahuja, A., Schwarz, J., Lillicrap, T., and Wayne, G. Experience replay for continual learning. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

Shinn, N., Cassano, F., Gopinath, A., Narasimhan, K., and Yao, S. Reflexion: Language agents with verbal reinforcement learning. *Advances in Neural Information Processing Systems*, 36, 2023.

Shrivastava, A., Gupta, A., and Girshick, R. Training region-based object detectors with online hard example mining. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 761–769, 2016.

Singh, G., Singh, P., and Singh, M. Advanced real-time fraud detection using rag-based llms. *arXiv preprint arXiv:2501.15290*, 2025.

Smyth, B. and McKenna, E. Remembering to forget: A competence-preserving case deletion policy for case-based reasoning systems. In *Proceedings of the 16th International Joint Conference on Artificial Intelligence*, pp. 377–382, 1999.

Song, K., Tan, X., Qin, T., Lu, J., and Liu, T.-Y. Mpnet: Masked and permuted pre-training for language understanding. *Advances in neural information processing systems*, 33:16857–16867, 2020.

Tang, X., Zou, A., Zhang, Z., Zhao, Z., Zhang, Y., Cohan, A., and Gerstein, M. MedAgents: Large language models as collaborators for zero-shot medical reasoning. In *Findings of the Association for Computational Linguistics: ACL 2024*, pp. 599–621, 2024.

Vapnik, V. N. *The Nature of Statistical Learning Theory*. Springer-Verlag, New York, 1995.

Wang, Z. Z., Mao, J., Fried, D., and Neubig, G. Agent workflow memory. In *Forty-second International Conference on Machine Learning*, 2025.

Wilson, D. L. Asymptotic properties of nearest neighbor rules using edited data. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-2(3):408–421, 1972.

Wu, C.-K., Tam, Z. R., Lin, C.-Y., Lin, Y.-N., and Tsai, H.-y. StreamBench: Towards benchmarking continuous improvement of language agents. In *Advances in Neural Information Processing Systems*, 2024.

Xu, Z. et al. Domain-agnostic adapter architecture for deception detection: Extensive evaluations with the DIFrauD benchmark. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pp. 5260–5274, 2024.

Yuksekgonul, M., Bianchi, F., Boen, J., Liu, S., Lu, P., Huang, Z., Guestrin, C., and Zou, J. Optimizing generative ai by backpropagating language model feedback. *Nature*, 639(8055):609–616, 2025.

Zhao, A., Huang, D., Xu, Q., Lin, M., Liu, Y.-J., and Huang, G. ExpeL: LLM agents are experiential learners. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 19632–19642, 2024.

Zheng, L., Guha, N., Anderson, B. R., Henderson, P., and Ho, D. E. When does pretraining help? Assessing self-supervised learning for law and the CaseHOLD dataset of 53,000+ legal holdings. In *Proceedings of the 18th International Conference on Artificial Intelligence and Law*, pp. 159–168, 2021.

Zhu, S., Wang, H., Dong, Z., Cheng, X., and Xie, Y. Neural spectral marked point processes. In *International Conference on Learning Representations*, 2022.

A. Related Work

Case-based reasoning and LLM agents for CBR. Case-based reasoning (CBR) solves new problems by retrieving and adapting solutions from similar past cases, following a retrieve-reuse-revise-retain cycle (Aamodt & Plaza, 1994). Classical CBR maintenance research established that not all cases contribute equally to system competence: boundary cases are disproportionately important, while interior cases are largely redundant (Smyth & McKenna, 1999). LLM agents have recently been applied to CBR tasks such as medical diagnosis and legal reasoning, where the agent processes a sequential case stream and improves via accumulated experience (Tang et al., 2024; Cui et al., 2023). These agents inherit the CBR structure but replace symbolic similarity with embedding-based retrieval and replace rule adaptation with in-context prompting. However, none incorporate the classical insight that memory should be organized around decision boundaries; they store all cases uniformly regardless of their competence contribution.

Retrieval-augmented in-context learning for streaming tasks. LLMs learn from examples provided in context (Brown et al., 2020) or via feedback-based prompt optimization (Guo et al., 2024; Dong et al., 2025b; Yuksekgonul et al., 2025), and retrieval-augmented generation (RAG) extends this by fetching relevant documents at inference time (Lewis et al., 2020). In streaming CBR settings, this translates to retrieving similar past cases as few-shot demonstrations: Self-StreamICL (Wu et al., 2024) retrieves raw input-output pairs from a growing case store; Reflexion (Shinn et al., 2023) augments retrieval with verbal self-reflections on past failures. The shared assumption is that similarity-based retrieval provides useful guidance, which holds when retrieved neighbors agree on the outcome. At decision boundaries, however, retrieval surfaces contradictory examples from both sides without explaining why similar cases diverge.

Memory architectures for LLM agents. Several systems structure agent memory beyond flat case retrieval. MemGPT (Packer et al., 2023) introduces an OS-inspired hierarchy between working and archival memory, designed for long-context conversation management. Mem0 (Chhikara et al., 2025) extracts and stores atomic facts from conversations in a key-value format optimized for personal assistants. ExpeL (Zhao et al., 2024) distills cross-task insights from success and failure trajectories for procedural agent tasks. AgentWorkflowMemory (Wang et al., 2025) and ReasoningBANK (Ouyang et al., 2026) stores reusable procedural workflows discovered across task episodes. These systems encode how to act in an environment, such as action trajectories, executable skills, or verbal plans, rather than how to decide given a pattern of evidence, making them a poor fit for CBR. `OBAM` introduces a qualitatively different entry type, boundary entries, designed specifically for the CBR setting where the critical challenge is disambiguating similar cases with conflicting outcomes.

Boundary-aware learning and continual improvement. The principle that decision boundaries carry disproportionate information is well-established in traditional ML. SVM theory defines classifiers entirely by support vectors at the boundary (Vapnik, 1995); active learning queries points near decision surfaces for maximum informativeness (Lewis & Gale, 1994); the Condensed Nearest Neighbor rule (Hart, 1968) reduces training sets to boundary-defining subsets. In parametric settings, several works (Shrivastava et al., 2016; Lin et al., 2017) invest in concentrating gradient updates on difficult (boundary-adjacent) examples, while incremental SVMs (Cauwenberghs & Poggio, 2000) maintain the support vector set as data arrives in a stream. Online continual learning extends these ideas to non-stationary streams (Rolnick et al., 2019; Aljundi et al., 2019; Buzzega et al., 2020), a challenge also studied in statistical models of non-stationary streaming data (Zhu et al., 2022; Dong et al., 2023a; 2025a).. All of these methods, however, operate in parameter space or assume access to a trainable model. StreamBench (Wu et al., 2024) recently formalized the streaming evaluation setting for LLM agents and measure continual improvement as the agent processes sequential feedback, without considering boundary learning. `OBAM` applies the boundary-concentration principle to *memory space* for a frozen LLM. An analogous principle appears in operations research: stochastic capacity planning concentrates resources at high-uncertainty nodes rather than allocating uniformly across a network (Liu et al., 2023; 2025a; 2024), with extensions to modular and containerized resource allocation (Liu et al., 2026) and urban network design (Liu et al., 2025b). Similarly, adaptive multi-agent coordination in networked systems must handle conflicting operational signals under evolving conditions (Li et al., 2024; 2023).

B. Experiment details

Datasets. We evaluate on four datasets spanning three CBR task types. Table 2 summarizes the dataset statistics.

- **CaseHOLD** (Zheng et al., 2021) is a legal holding identification benchmark derived from U.S. case law. Each instance presents a citing context from a judicial opinion with the holding statement masked, and the agent must select the correct holding from five candidates. The task requires distinguishing between legally similar but substantively different holdings, making it a natural testbed for decision boundaries in legal reasoning. We random sample 1,000 cases for streaming test.

- **DDXPlus** (Fansi Tchango et al., 2022) is a medical differential diagnosis dataset where each case presents a patient profile (symptoms, medical history, antecedents) and the agent must identify the correct pathology from 49 possible diagnoses. The high class count and overlapping symptom presentations create numerous decision boundaries where patients with similar profiles require different diagnoses. We use the test split (1,764 cases) as the streaming set.
- **DiFraud** (Xu et al., 2024) provides two binary classification tasks for deception detection. *DiFraud-Phishing* requires classifying emails as genuine or deceptive (phishing attempts), where sophisticated phishing emails closely mimic legitimate corporate communications. *DiFraud-Product Reviews* requires classifying product reviews as genuine or deceptive (fake), where deceptive reviews may include surface authenticity markers that mimic real customer feedback. For each configuration, we randomly sample 1,600 cases as the streaming set.

Table 2. Dataset statistics. CaseHOLD uses semantic label matching ($\theta_{\text{label}} = 0.4$); DDXPlus and DiFraud use exact-match labels.

Dataset	Domain	Label type	Classes	Test size
CaseHOLD	Legal holding identification	5-way MC	5	1,000
DDXPlus	Medical differential diagnosis	Multi-class	49	1,764
DiFraud-Phishing	Email fraud detection	Binary	2	1,600
DiFraud-Reviews	Fake review detection	Binary	2	1,600

Label consistency modes. The agreement function in boundary classification phase supports two modes depending on the nature of ground-truth labels. In *exact-match mode* (used for DDXPlus and DiFraud, where labels are meaningful category names), two entries agree if their ground-truth decisions are identical. In *semantic mode* (used for CaseHOLD, where labels are opaque holding texts), two entries agree if the cosine similarity between their label embeddings exceeds a threshold θ_{label}. This distinction ensures that boundary classification operates correctly regardless of whether the decision space consists of categorical labels or free-form text.

Lesson extraction. After the agent produces decision d_t and receives ground-truth y_t, the lesson l_i is generated via LLM self-reflection with outcome-conditioned prompting. If $d_t = y_t$ (correct), the LLM is prompted: "You decided [d_t] and it was correct. What key evidence supports this decision?" If $d_t \neq y_t$ (incorrect), the LLM is prompted: "You decided [d_t] but the correct answer was [y_t]. What did you miss? What evidence should have led to [y_t]?" This asymmetric prompting produces lessons that capture confirmatory evidence for correct decisions and diagnostic error signals for incorrect ones.

Memory embedding. Memory entries are embedded using all-mpnet-base-v2 (Song et al., 2020) (a 768-dimensional sentence embedding model) with FAISS inner-product search for retrieval. When a boundary entry is refined (reinforce or extend), the updated shared pattern p_j is re-embedded and the retrieval index is updated to reflect the new vector. This ensures that refined boundary entries remain retrievable for future cases in the same region, even as the shared pattern description evolves.

Hyperparameters. Table 3 shows our setup for hyperparameters.

Table 3. Hyperparameter settings for all experiments.

Symbol	Value	Description
W	1% of train set	Warm-up period (cases stored unconditionally)
θ_{sim}	0.5	Minimum cosine similarity for retrieval
K_1	5	Region entries retrieved
K_2	2	Boundary entries retrieved
θ_{agree}	0.5	Agreement threshold for boundary detection
α	0.3	Interior case sampling rate

Presentation templates. Retrieved entries are formatted with tone-appropriate framing before injection into the agent's prompt:

- **Boundary entries:**

  `Cases with this pattern can lead to different outcomes. When [features`$_1$`] → outcome D`$_1$`. When [features`$_2$`] → outcome D`$_2$`. The key distinguishing factor is: [discriminative rule].`

- **Interior region entries:**

  `A similar case had outcome D because [lesson].`

- **novel region entries:**

  `A prior case in this area had outcome D. Evidence observed: [lesson].`

The neutral tone for entries in novel regions avoids biasing the agent in under-explored regions where boundary status is unknown.

C. Additional Results

C.1. Standard deviation of model performance

To show the standard deviation of each model's performance and confirm that the performance gains of OBAM in Table 1 are statistically significant, we run streaming experiments over 5 independent resamples of CaseHOLD and DDXPlus test set, with Claude Sonnet 4 as the backbone. We report the 95% confidence intervals for OBAM and the baseline of Mem0 and Self-StreamICL in Table 4.

Table 4. 95% confidence intervals of model testing performance (5 independent runs) using Claude Sonnet 4.

Method	**CaseHOLD**	**DDXPlus**
Mem0	[0.809, 0.815]	[0.849, 0.853]
Self-StreamICL	[0.805, 0.811]	[0.859, 0.865]
OBAM	[0.821, 0.829]	[0.916, 0.922]

As shown in Table 4, the confidence intervals of OBAM do not overlap with those of either baseline on both datasets, confirming that the observed performance gains are statistically significant and not attributable to random variation in the test set composition. The narrow intervals (spanning 0.6–0.8 percentage points) further indicate that OBAM's performance is stable across different random permutations of the case stream, validating that the streaming serialization does not introduce order-dependent artifacts.

C.2. Sensitivity analysis

Table 5 reports the sensitivity of OBAM to each hyperparameter on DDXPlus and DiFraud-Phishing, varying one parameter at a time while holding others at their default values.

Performance is robust across reasonable ranges for all parameters. The warm-up size W has minimal impact: even at 0.1% of the stream, the system achieves within 1–2 points of the default, indicating that a small seed suffices for boundary detection to activate reliably. The number of retrieved region entries K_1 shows the largest sensitivity: reducing to $K_1 = 1$ degrades accuracy substantially (by 7–8 points), as a single neighbor provides insufficient signal for agreement computation. Performance saturates at $K_1 = 5$, with no further gain at $K_1 = 7$. The number of boundary entries K_2 peaks at the default of 2; retrieving more boundary entries ($K_2 = 4$) slightly degrades performance, likely due to less relevant boundary entries introducing noise. The interior sampling rate α shows that very sparse interior storage ($\alpha = 0.1$) hurts performance by 3–7 points, as some minimum interior coverage is needed for reliable agreement computation. The default $\alpha = 0.3$ balances coverage with boundary concentration.

D. Qualitative Examples

We present additional case studies demonstrating how boundary entries correct agent decisions across different domains. Each example shows a case where the baseline agent (without boundary memory) makes an incorrect decision, and the

Table 5. Hyper-parameter sensitivity analysis on DDXPlus and DiFraud-Phishing (accuracy %). Default values in **bold**. Performance is robust within reasonable ranges. Backbone LLM: Claude Sonnet 4.

Param	Dataset	Value			
W		0.1%	0.5%	**1%**	2%
	DDXPlus	0.906	0.912	0.919	0.916
	Phish	0.908	0.917	0.926	0.928
K_1		1	3	**5**	7
	DDXPlus	0.840	0.901	0.919	0.919
	Phish	0.876	0.921	0.926	0.923
K_2		1	**2**	3	4
	DDXPlus	0.903	0.919	0.917	0.910
	Phish	0.912	0.926	0.925	0.917
α		0.1	0.2	**0.3**	0.5
	DDXPlus	0.849	0.903	0.919	0.914
	Phish	0.879	0.911	0.926	0.918

boundary entry provides the discriminative knowledge needed to arrive at the correct answer.

CaseHOLD (Legal Holding Identification). The agent must select the correct legal holding from five candidates given a citing context from a judicial opinion.

Input (excerpt): "...of notice in a 'major local newspaper'. At the same time, it should not be defined so broadly as to include those who would lack any real interest in commenting on it. To require Aviall to solicit comments from such persons would be to impose the same type of rigid formality that the NCP seeks to avoid..."

Without boundary memory: Holding 0 ("holding that public meetings held after implementation of final remedial action were not meaningful and that the only pu...") ✗

Boundary entry retrieved:
Shared pattern: Both cases involve legal notice requirements and whether alternative forms of notice satisfy due process and statutory requirements.
Branches:

- Notice to attorney suffices → established legal relationship creates a known channel
- Publication notice adequate → party is unknown/unreachable, constructive notice needed

Discriminative rule: Adequacy depends on whether a known communication channel exists between the parties.

With boundary memory: Holding 4 ("holding that public meetings were not meaningful where first was held without adequate notice and second was held after...") ✓
The boundary entry's notice-adequacy framework guides the agent to the holding that matches the procedural context of inadequate notice.

DiFraud-Phishing (Email Fraud Detection). The agent must classify an email as genuine or deceptive (phishing).

Input (excerpt): "Held Spam Report for January 29, 2014. Daily Held Spam Report mirandal@dnc.org. Spam: 4. View — Release frankr@emcgti.com — Try new working in the new direction! $15000 10:03p from Argentina. [...] Copyright ©2003-2016 AppRiver. Unsubscribe."

Without boundary memory: "deceptive" ✗
(Money amounts and foreign origins trigger phishing signals.)

Boundary entry retrieved:
Discriminative rule: Emails with crude formatting and obvious fraud indicators are genuine (legitimate services reporting spam), whereas emails with plausible backstories and coherent narratives are deceptive.

Region entries also retrieved (3× genuine): "Authentic email structure, proper forwarded message headers"; "Sent to specific subscriber, included clear unsubscribe."

With boundary memory: "genuine" ✓
The email is a legitimate spam filter report from AppRiver, not spam itself. The boundary entry and supporting region entries together identify the authentic service format.

DiFraud-Product Reviews (Fake Review Detection). The agent must classify a product review as genuine or deceptive.

Input: "I bought as gift for my son. He says he loves his new headphones. The quality is very good. His only complaint is that they are a little bit of an awkward fit. I offered to send them back but he said absolutely not. Thinking of getting myself a pair."

Without boundary memory: "genuine" ✗
(Mentioning a complaint and offering a return appears authentic.)

Boundary entry retrieved:
Shared pattern: Product reviews for headphones discussing sound quality, comfort, and overall assessment.
Branches:

- Genuine → specific product model name (e.g., "JVC HARX700"), admits drawbacks with technical detail, mentions specific usage contexts, comparative references to other products
- Deceptive → vague and generic language, no model names or specifications, overly enthusiastic without substantive detail

Discriminative rule: Genuine reviews demonstrate authentic ownership through specific product identification and practical usage details; deceptive reviews rely on generic language regardless of surface authenticity markers.

With boundary memory: "deceptive" ✓
The review contains no product model name, no technical specifications, and only generic praise. Despite surface authenticity signals (complaint, return offer), the absence of specific product details matches the deceptive pattern identified by the boundary entry.