DESCENDANT CLOSURE KNOCKOUT FOR CONTEXT MANAGEMENT IN LONG-HORIZON SEARCH AGENTS

Anonymous authors
Paper under double-blind review

ABSTRACT

Long-horizon search agents accumulate tool calls and web observations over extended trajectories, causing append-only histories to mix useful evidence with redundant content and eventually exhaust the context window. In existing long-horizon search agents, retention decisions are typically made by the summarizer or the search policy itself, without a separate observation-level predictor trained from gold-answer-conditioned influence. A natural way to construct such a signal is single-observation knockout, but it can underestimate delayed influence because later queries may reuse entities introduced by early evidence, leaving downstream observations that continue to carry information from the root in the counterfactual context. We introduce **D**escendant **C**losure **K**nockout (DCK), a lightweight retention layer that ranks observations for protected compaction while keeping the search agent, compaction trigger, and summarizer frozen. Offline, DCK constructs a temporal entity-lineage proxy from entity reuse in later queries and jointly masks a root observation with its currently addressable lineage descendants. It measures the change in the frozen agent's teacher-forced gold-answer log-likelihood and distills this horizon-conditioned closure-knockout influence into an agent-specific lightweight head. At deployment, when the host triggers compaction, DCK preserves bounded verbatim blocks from the highest-scoring observations within the fixed reset-context budget, while the host summarizes the remaining history. We evaluate DCK across frozen search agents of different scales, multiple summary tools, with two context-compaction instantiations, using matched context and tool-call budgets within each paired comparison.

1 INTRODUCTION

Long-horizon web search requires language-model agents to resolve complex questions through reasoning, querying, browsing, and evidence synthesis. ReAct Yao et al. (2023) formalized the interleaved thought, action, and observation loop underlying such agents Li et al. (2025b); Wu et al. (2025a); Liu et al. (2025), while later systems expanded search capability through synthesized trajectories and post-training Li et al. (2025b;a). Benchmarks like BrowseComp test multi-step navigation and hard information retrieval Wei et al. (2025). Every round appends reasoning traces, tool calls, and webpage returns Wu et al. (2025c); Lu et al. (2026), so the history accumulates redundant material Li et al. (2026a); Lu et al. (2026) and approaches the context limit, making context management essential to sustained search Wu et al. (2025c); Li et al. (2026a); Lu et al. (2026).

Existing context-management work spans fixed heuristics, summary-based memory, and active editing. Sliding-window and discard-at-threshold schemes bound growth by retaining only recent steps or flushing accumulated history once a limit is reached Bai et al. (2026); DeepSeek-AI et al. (2025). Periodic and learned approaches condense trajectories into summaries or fixed-size states, spanning recurrent memory agents Zhou et al. (2025b); Yu et al. (2025); Yuan et al. (2026), external summary tools Wu et al. (2025c), and trained context folding Shao et al. (2025); Sun et al. (2025). Adaptive systems place timing and selection inside the agent: SelfCompact uses a rubric-guided inline summarization tool Li et al. (2026a), AgentFold and ARC make proactive compression decisions Ye et al. (2025); Yao et al. (2026), and LongSeeker expands the policy's operations to compression, deletion, rollback, and exact snippet retention Lu et al. (2026). Across these designs, retention value is assigned by a fixed rule Bai et al. (2026); DeepSeek-AI et al. (2025), a summarizer Wu et al. (2025c); Li et al. (2026a), or the search policy itself Lu et al. (2026); Zhou et al. (2025b); Ye et al.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

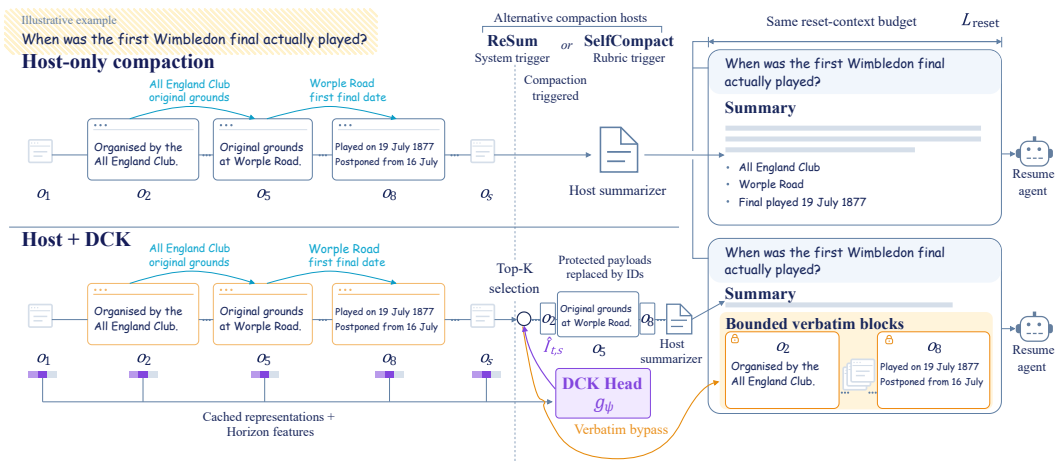


Figure 1: An illustrative search example compares host-only compaction with DCK-augmented compaction. After triggers compaction, DCK ranks currently addressable observations using cached representations and horizon features, preserving selected text as bounded verbatim blocks while the host summarizes the remaining history.

(2025). This leaves frozen search agents without a separate observation-level predictor trained from gold-answer-conditioned influence.

This remaining gap becomes concrete when lossy compaction must preserve precise web evidence. A summary can omit a date, entity relation, source, quotation, or qualifying condition Lu et al. (2026), and poorly timed compression can erase partial results that later reasoning needs Li et al. (2026a); Zhang et al. (2026). Consider an early observation introducing person A; a later query reuses A to discover organization B, and a subsequent query uses B to find the final date. Masking only the early observation leaves the later queries and observations visible, so organization B and the date can still carry information initiated by the root. The resulting single-observation knockout therefore underestimates delayed influence. A useful retention signal must account for the downstream evidence lineage that remains in the counterfactual context, while remaining computable from recorded trajectories and usable at deployment, where the gold answer is unavailable.

We introduce **D**escendant **C**losure **K**nockout (DCK), an agent-specific retention layer for protected compaction. Offline, DCK first constructs a temporal entity-lineage proxy by tracing entities introduced in an observation and recursively reused in later queries. At a fixed horizon, it jointly masks a root observation and the lineage descendants that remain independently addressable, then measures the change in the frozen agent's teacher-forced gold-answer log-likelihood on the recorded trajectory prefix at that horizon. DCK distills this horizon-conditioned closure-knockout influence into a lightweight head using cached observation representations and horizon features. At deployment, the head runs only after the host triggers compaction; it ranks the currently addressable candidate observations, and DCK selects the highest-scoring roots under a fixed protection budget. Bounded verbatim blocks from selected observations enter the reset context, while the host preserves its trigger and summarizes the remaining history. We instantiate this interface with ReSum as the main training host and SelfCompact as a fixed-head transfer setting with no relabeling or head retraining, keeping the search agents and summary tools frozen and matching context and tool-call budgets within each paired comparison. Our contributions are threefold:

- We propose Descendant-Closure Knockout (DCK), a rollout-free method that uses temporal entity-lineage closure knockout to estimate gold-answer-conditioned observation influence for protected compaction.
- We distill the horizon-conditioned influence into a lightweight agent-specific head and integrate it through a budget-preserving interface that keeps the search agent, compaction trigger, and summarizer frozen.
- We evaluate DCK across multiple agent scales, summary tools, and instantiate it with two representative context-compaction frameworks, showing the effectiveness of DCK.

2 Related Works

2.1 Search Agent

Iterative web-search agents must choose queries, interpret observations, and continue until evidence suffices. ReAct established an interleaved reasoning, action, and observation loop as an interaction pattern Yao et al. (2023), while WebGPT trained a text-browser policy from demonstrations and human preferences to gather cited evidence Nakano et al. (2021). WebSailor couples synthesized high-uncertainty tasks with post-training Li et al. (2025b), and BrowseComp supplies a deliberately hard benchmark for multi-step browsing Wei et al. (2025). Search-o1 lets a reasoning model trigger retrieval at uncertain points and refines retrieved documents before reinsertion Li et al. (2025c), while WebWalker pairs a web-traversal benchmark with explorer and critic agents Wu et al. (2025b). Recent systems optimize whole trajectories: Search-R1 trains multi-turn search with outcome-based reinforcement learning Jin et al. (2025), R1-Searcher uses two-stage outcome-based reinforcement learning to elicit retrieval calls Song et al. (2025), and WebDancer combines trajectory data, supervised fine-tuning, and reinforcement learning Wu et al. (2025a). WebExplorer extends the operating regime to 128K context and up to 100 tool turns Liu et al. (2025), whereas Beyond Ten Turns uses large-scale asynchronous reinforcement learning for longer interactions Gao et al. (2025). Data and training pipelines include SearchArt's synthesis and verification of tasks and trajectories Mei et al. (2026), OpenResearcher's large collection of synthesized trajectories Li et al. (2026b), WebShaper's formalization-guided task synthesis Tao et al. (2025), DeepDive's knowledge-graph question generation with multi-turn reinforcement learning Lu et al. (2025b), and Tongyi DeepResearch's agentic mid-training and post-training Li et al. (2025a). Evaluation also spans general tool-using assistance in GAIA Mialon et al. (2024) and challenging multi-hop Chinese web browsing in BrowseComp-ZH Zhou et al. (2025a). Together, these works broaden the mechanisms, training pipelines, and evaluation settings for iterative search; explicitly long-horizon systems Liu et al. (2025); Gao et al. (2025); Li et al. (2025a) specifically motivate managing growing histories.

2.2 Context Management for Agents

Agent context managers differ in controller and retained representation. ReSum uses host-triggered periodic replacement by an external summary Wu et al. (2025c), whereas SelfCompact lets the model invoke rubric-gated inline summarization without additional training Li et al. (2026a). In policy-controlled approaches, FoldAct and context folding train the search agent to decide how completed history is condensed Shao et al. (2025); Sun et al. (2025), and AgentFold adds proactive multi-scale folding Ye et al. (2025). ARC decouples action generation from a reflection-driven context manager that revises memory Yao et al. (2026), while LongSeeker broadens the action space to Skip, Compress, Rollback, Snippet, and Delete Lu et al. (2026). Other learned managers couple context control to optimization: SUPO co-trains periodic summarization with tool use Lu et al. (2025a), ContextBudget learns compression timing and intensity under an explicit budget Wu et al. (2026), AdaCoM trains an external LLM manager while freezing the acting agent Yi et al. (2026), and MemSearcher jointly trains multi-turn search with bounded rolling-memory updates Yuan et al. (2026). Fixed recency and threshold-reset policies bound context growth Bai et al. (2026); DeepSeek-AI et al. (2025). Evidence removal is nevertheless regime dependent: stale-observation masking can help or hurt when useful evidence leaves the active context Zhang et al. (2026).

3 Method

3.1 Problem Setup and Compaction-Host Interface

Let x denote a search question and π_θ be a frozen search agent. Following ReAct Yao et al. (2023), each tool-interaction round t contains a reasoning trace τ_t, a parsable action a_t, and a message-level observation $o_t = \mathcal{R}(a_t)$. Terminal-answer rounds receive no observation index. Let $Q_t = \mathrm{QueryArgs}(a_t)$ collect any natural-language search queries in the action, including batched queries. We write H_s for the *actual working context* after appending o_s and before compaction at horizon s. Importantly, H_s need not be the raw trajectory prefix: after earlier compaction events, it may contain prior summaries, protected evidence blocks, and interactions generated since the latest reset.

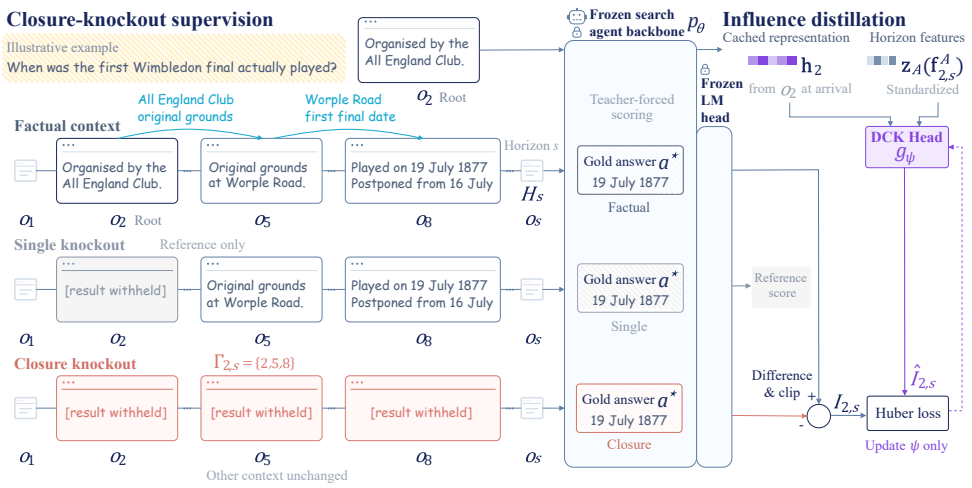


Figure 2: Overview of DCK. Left: offline supervision on compaction snapshots. Each root is jointly masked with its entity-lineage descendants, and the resulting shift in teacher-forced gold-answer log-likelihood yields a closure-knockout influence label, distilled into an agent-specific head. Right: online protected compaction. Once the host triggers, the head ranks addressable observations and the top-K roots enter the reset context as bounded verbatim blocks under the shared budget, while the host summarizes the remaining history. Agent, trigger, and summarizer stay frozen throughout.

An observation is *addressable* at horizon s when it remains as a uniquely identifiable text block in H_s. We denote the corresponding index set by

$$\mathcal{A}_s = \{t \le s : o_t \text{ can be independently located and replaced in } H_s\}. \tag{1}$$

This set includes raw tool responses still present in the working context and evidence blocks preserved by an earlier compaction. Information already absorbed into a summary is not addressable and cannot be restored from an external archive. Such an archive may retain raw payloads and stable identifiers for offline labeling and auditing, but it is never exposed to the agent as a memory store.

We represent a lossy compaction host by

$$\mathcal{H} = (\chi_{\mathcal{H}}, \mathcal{S}_{\mathcal{H}}, \mathrm{Compose}_{\mathcal{H}}), \tag{2}$$

where $\chi_{\mathcal{H}}(H_s) \in \{0,1\}$ is the host trigger, $\mathcal{S}_{\mathcal{H}}(H_s; L)$ is the host's summary-generation procedure under output budget L, and $\mathrm{Compose}_{\mathcal{H}}(x, S, P)$ constructs the reset context from the original question, a summary, and a protected evidence sequence P. DCK neither modifies $\chi_{\mathcal{H}}$ nor trains or replaces $\mathcal{S}_{\mathcal{H}}$. At a triggered horizon, it outputs a protected subset $B_s \subseteq \mathcal{A}_s$ under the same reset-context budget as the host-only system.

3.2 Temporal Entity-Lineage Closure

Let $\eta(\cdot)$ be a frozen entity extraction and normalization map. For agent A and language ℓ, let $\mathcal{G}_A^\ell$ be a stop-entity set constructed only from the corresponding head-training split. The entities first introduced by observation o_t are

$$E_t = \eta(o_t) \setminus \left(\eta(x) \cup \bigcup_{j<t} \eta(o_j) \cup \mathcal{G}_A^\ell \right). \tag{3}$$

Thus, an entity belongs to E_t only when it is absent from the question, earlier observations, and the frozen stop set.

For a possibly batched Search action, let $\eta(Q_u) = \bigcup_{q \in Q_u} \eta(q)$. At horizon s, define the forward entity-reuse graph $G_s = (V_s, \mathcal{E}_s)$ with $V_s = \{1, \dots, s\}$ and

$$(v,u) \in \mathcal{E}_s \iff v < u \ \wedge \ \eta(Q_u) \cap E_v \neq \emptyset. \tag{4}$$

The horizon-conditioned descendant closure of root t is the set of nodes reachable from t,

$$D_{\le s}(t) = \{u \in \{t+1, \dots, s\} : t \rightsquigarrow u \text{ in } G_s\}. \tag{5}$$

Equivalently, a later step joins the closure when its query reuses an entity introduced either by the root or by an already admitted descendant. Because only independently replaceable text can be intervened on, the executable knockout set is

$$\Gamma_{t,s} = \{t\} \cup (D_{\le s}(t) \cap \mathcal{A}_s), \qquad t \in \mathcal{A}_s. \tag{6}$$

The lineage is an observable proxy for query-mediated dependence; it is not an exact causal graph.

Lemma 1 (Forward closure). The recursive temporal lineage is the unique reachability set of G_s and is computable by a chronological scan. Every edge satisfies $v < u$, so G_s is acyclic. Induction over $u = t+1, \dots, s$ shows that the recursive admission rule includes u exactly when a path $t \rightsquigarrow u$ exists. A complete proof is given in the appendix.

3.3 Horizon-Conditioned Closure-Knockout Influence

For a gold answer $a^\star$, let $\mathcal{Y}_A(a^\star)$ denote its content-token positions under the tokenizer and serialization of agent A. We define the teacher-forced gold-answer log-likelihood

$$\mathcal{J}_A(H_s; a^\star) = \sum_{j \in \mathcal{Y}_A(a^\star)} \log p_{\theta_A}\left(a^\star_j \mid H_s, a^\star_{<j}\right). \tag{7}$$

Let $\mathsf{K}_\Gamma(H_s)$ replace the semantic payload of every observation indexed by Γ with the same neutral stub, while preserving its position and structural identifier. Queries, reasoning text, previous summaries, and all non-intervened observations remain fixed. DCK first defines the unclipped closure-knockout influence

$$\widetilde{I}_{t,s} = \mathcal{J}_A(H_s; a^\star) - \mathcal{J}_A\big(\mathsf{K}_{\Gamma_{t,s}}(H_s); a^\star\big), \tag{8}$$

and uses the bounded target

$$I_{t,s} = \mathrm{clip}\Big(\widetilde{I}_{t,s}, -c, c\Big) \tag{9}$$

for distillation, where the clipping constant c is fixed before training.

A direct baseline masks only the root observation:

$$\widetilde{I}^{\mathrm{single}}_{t,s} = \mathcal{J}_A(H_s; a^\star) - \mathcal{J}_A\big(\mathsf{K}_{\{t\}}(H_s); a^\star\big), \qquad I^{\mathrm{single}}_{t,s} = \mathrm{clip}\left(\widetilde{I}^{\mathrm{single}}_{t,s}, -c, c\right). \tag{10}$$

The additional lineage-mediated influence captured by closure knockout is

$$\widetilde{\Delta}_{t,s} = \widetilde{I}_{t,s} - \widetilde{I}^{\mathrm{single}}_{t,s}. \tag{11}$$

Proposition 1 (Horizon consistency). Assume that entity normalization and the stop set are prefix-computable. Then $D_{\le s}(t)$, $\Gamma_{t,s}$, and the inputs used to predict $I_{t,s}$ depend only on information available by horizon s. The offline target additionally uses $a^\star$, but no query, observation, reasoning step, or summary generated after s enters its construction. This establishes absence of future-trajectory leakage, rather than absence of supervision.

Proposition 2 (Closure increment). The unclipped increment satisfies

$$\widetilde{\Delta}_{t,s} = \mathcal{J}_A\big(\mathsf{K}_{\{t\}}(H_s); a^\star\big) - \mathcal{J}_A\big(\mathsf{K}_{\Gamma_{t,s}}(H_s); a^\star\big). \tag{12}$$

This follows by expanding Eqs. equation 8 and equation 10; their factual-context terms cancel. The identity states precisely what closure adds after the root has already been masked. It does not imply $\widetilde{\Delta}_{t,s} \ge 0$, since descendants may be supportive or misleading.

3.4 Agent-Specific Influence Distillation

The gold answer is unavailable during deployment, so DCK distills $I_{t,s}$ into a lightweight predictor. When observation o_t first enters the actual working context, we cache the mean of the frozen agent's last-layer hidden states over the tokens corresponding only to the visible observation payload:

$$\mathbf{h}^A_t = \frac{1}{|\mathcal{S}^A_t|} \sum_{k \in \mathcal{S}^A_t} \mathbf{H}^A_{L,k}, \tag{13}$$

where $\mathcal{S}_t^A$ excludes message delimiters, tool metadata, query text, special tokens, and padding. The representation is computed once and reused at later horizons.

We augment it with seven horizon features,

$$\mathbf{f}_{t,s}^A = [t, s, s-t, |D_{\le s}(t) \cap \mathcal{A}_s|, |E_t|, \mathrm{Tok}_A(o_t), r_t]^\top, \tag{14}$$

where r_t indicates whether any normalized query in Q_t exactly repeats a query in an earlier Q_u. Let $\mathbf{z}_A(\cdot)$ denote feature standardization using statistics from the training split for agent A. An agent-specific predictor estimates

$$\widehat{I}_{t,s} = g_{\psi_A}\left([\mathbf{h}_t^A; \mathbf{z}_A(\mathbf{f}_{t,s}^A)]\right), \tag{15}$$

where g_{ψ_A} is a lightweight multilayer perceptron. It is trained with

$$\mathcal{L}_{\mathrm{inf}}(\psi_A) = \frac{1}{N_A} \sum_{(t,s)\in\mathcal{D}_A} \mathrm{Huber}\left(\widehat{I}_{t,s} - I_{t,s}\right). \tag{16}$$

Only ψ_A and the associated normalization statistics are learned. The search agent, tools, and summary models remain frozen. Each search-agent checkpoint owns a separate head because both the hidden-state space and the likelihood target are agent-specific. During inference, DCK requires neither $a^\star$ nor knockout forward passes.

3.5 Budget-Preserving Protected Compaction

Suppose host $\mathcal{H}$ triggers compaction at horizon s_c. DCK scores every addressable observation and chooses a fixed-capacity set

$$B_{s_c} = \mathrm{TopK}_K\left\{(t, \widehat{I}_{t,s_c}) : t \in \mathcal{A}_{s_c}\right\}. \tag{17}$$

Ties are resolved deterministically. Selection uses predicted influence, whereas the selected observations are serialized in their original chronological order; neither scores nor ranks are written into the agent prompt. The constants K and the per-observation retention cap are fixed globally and reported in the experimental setup.

Let $\rho(o_t)$ be a bounded verbatim serialization of o_t containing its stable identifier, compact source metadata, and a capped evidence payload. Let P_{s_c} denote the chronologically ordered sequence induced by B_{s_c}. Before summarization, protected payloads are replaced at their original positions by stable identifiers:

$$M_{\mathcal{H}}(H_{s_c}, B_{s_c}) = \mathrm{ReplaceProtectedWithIDs}(H_{s_c}, B_{s_c}). \tag{18}$$

Under a fixed reset-context budget L_{reset}, the host summary receives

$$L_{\mathrm{sum}} = L_{\mathrm{reset}} - \mathrm{Tok}_A\left(\bigoplus_{t\in P_{s_c}} \rho(o_t)\right) - L_{\mathrm{overhead}}, \tag{19}$$

then produces

$$S_{s_c}^{\mathcal{H}} = \mathcal{S}_{\mathcal{H}}\left(M_{\mathcal{H}}(H_{s_c}, B_{s_c}); L_{\mathrm{sum}}\right), \tag{20}$$

$$\widetilde{H}_{s_c}^{\mathcal{H}} = \mathrm{Compose}_{\mathcal{H}}\left(x, S_{s_c}^{\mathcal{H}}, \bigoplus_{t\in P_{s_c}} \rho(o_t)\right). \tag{21}$$

Thus, protected evidence consumes part of the same reset budget rather than enlarging the context available to DCK. At every later compaction event, the current addressable observations are rescored and a new protected set is selected. Protection is therefore dynamic and nonpersistent; an observation absorbed into an earlier summary cannot be recovered from the external archive.

Host instantiations. For ReSum, DCK runs after (τ_t, a_t, o_t) is appended and the trigger fires, but before the external summary tool, extending the reset state from (x, S_t) to (x, S_t, P_t). For SelfCompact, DCK is invoked after the rubric returns a compression decision and before the same-model summarization call. In both cases, the host trigger, summary generation procedure, and continuation logic remain unchanged. A head can be transferred across these hosts only when the underlying search-agent checkpoint is fixed; host-specific thresholds, prompts, budgets, model choices, and all optimization hyperparameters are specified in the experimental setup.

4 EXPERIMENT

4.1 EXPERIMENTAL SETUP

Evaluation Protocols. To evaluate the generality of protected compaction, we conduct evaluations across three benchmarks spanning different languages and search depths: the GAIA text-only validation subset Mialon et al. (2024), BrowseComp-ZH Zhou et al. (2025a), and BrowseComp Wei et al. (2025). Every method runs three cross-method paired rollouts per question; final answers are judged by Qwen2.5-72B-Instruct under the frozen ReSum judge prompts Wu et al. (2025c). We report Pass@1 (mean accuracy over the three paired rollouts) and Pass@3 (fraction of questions solved by at least one rollout), which reflect the method's expected stability and its potential peak performance, respectively; efficiency metrics follow §E. Paired comparisons use question-level cluster bootstrap with one-sided sign-flip tests, Holm-corrected over the two core comparisons (DCK vs. ReSum and DCK vs. Single Knockout on BrowseComp); the full protocol is specified in §D.

Baseline Methods. We benchmark DCK against hosts and baselines spanning complementary compaction regimes: (1) the ReSum host, which triggers summarization at 90% of the window and restarts from the question plus a summary within the reset budget L_{reset}, using GPT-OSS-120B as its default summary tool Wu et al. (2025c); (2) the SelfCompact host, which lets WebSailor-7B itself run rubric probes from the third tool round, permits compression once the working context passes a low threshold, enforces an unconditional backstop at a fixed upper threshold, and performs at most one summary per trajectory Li et al. (2026a); (3) ReAct, which performs no compaction and, upon hitting the window, is forced to emit a judged final answer; (4) Recent History, which shares the ReSum trigger but retains only the most recent 22K tokens without calling any summary model; and (5) Fixed-Interval Summary, which ignores the context state and restarts from a WebSailor-7B self-summary every ten tool calls. Two controlled add-ons share the trigger, summarizer, and budget of ReSum + DCK: *Random Protection* protects a random set of the same size and token budget, and *Single Knockout* differs only in its single-observation supervision labels.

Implementation Details. We leverage frozen search agents of three scales—WebSailor-3B, WebSailor-7B Li et al. (2025b), and WebSailor-30B-A3B—to assess DCK. All agents share curated prompts, Search and Visit tools with a common page evidence extractor, a 32K-token context window, a cap of 60 LLM calls per rollout, and identical decoding settings. Each DCK head is trained once per agent checkpoint from 300 questions sampled from OpenSeeker-v1-Data and disjoint from all evaluation benchmarks (270 English, 30 Chinese; question-level 90/10 split). Supervision snapshots come from the ReSum + GPT-OSS-120B host, and closure-knockout labels are clipped to a fixed range. The head is a lightweight two-layer MLP over the cached last-layer observation representation and the seven horizon features of §3.4. At deployment we fix $K{=}8$ protected observations per compaction event; protected blocks are paid for out of the shared L_{reset} budget rather than enlarging the reset, heads are never shared across scales, and the SelfCompact transfer reuses the 7B ReSum-trained head without relabeling or retraining. Additional details are provided in §C.

4.2 MAIN RESULTS

Table 1 reports the main comparison. DCK improves ReSum on every dataset and metric, with the largest relative gains on the BrowseComp variants, where repeated compaction is the norm. Both core paired comparisons remain significant after Holm correction, with unsolved-to-solved flips dominating regressions in both (§D). The controlled add-ons isolate the origin of the gains. Random Protection matches the protected set size and token budget but selects tokens uniformly at random; it moves ReSum only marginally, so the improvement cannot be explained by protected tokens alone. Single Knockout is identical to DCK except that its supervision ignores descendant lineage; it recovers much of the gap yet stays clearly below DCK throughout, confirming that descendant closure captures delayed influence that single-observation knockout misses. The selection-level diagnostic in §G reproduces this ordering at the ranking level, locating the gains in the supervision signal rather than deployment-time differences.

DCK also transfers to the rubric-driven SelfCompact host with the frozen 7B head, improving every dataset and metric without relabeling or retraining. Because host and summarizer change simultaneously across the two blocks of Table 1, we claim only within-host paired increments.

Table 1: Main results and controlled retention comparisons with WebSailor-7B. Pass@1 / Pass@3 over three paired rollouts; current-matched rows share questions, tools, and judge; † rows quoted from ReSum Table 1 for context, excluded from paired statistics; ‡ denotes our WebSailor-7B protocol adaptation; best current-matched result per column in **bold**.

Agent	Paradigm	GAIA P@1↑ / P@3↑	BrowseComp-ZH P@1↑ / P@3↑	BrowseComp P@1↑ / P@3↑
Claude-4†	ReAct	68.3 / –	29.1 / –	12.2 / –
OpenAI-o3†	ReAct	70.5 / –	58.1 / –	50.9 / –
Kimi-K2†	ReAct	57.7 / –	28.8 / –	14.1 / –
DeepSeek-v3.1†	ReAct	63.1 / –	49.2 / –	30.0 / –
WebSailor-7B	ReAct	32.0 / 44.7	12.8 / 25.3	5.6 / 10.4
WebSailor-7B	Recent History	32.7 / 47.6	14.7 / 27.7	5.3 / 9.3
WebSailor-7B	Fixed-Interval Summary‡	35.0 / 51.5	16.3 / 30.1	7.8 / 13.0
ReSum host (GPT-OSS-120B summarizer)				
WebSailor-7B	ReSum	41.8 / 60.2	18.8 / 35.0	10.3 / 16.8
WebSailor-7B	+ Random Protection	42.7 / 62.1	19.3 / 36.0	10.7 / 17.0
WebSailor-7B	+ Single Knockout	44.3 / 64.1	20.9 / 38.1	11.4 / 17.8
WebSailor-7B	**+ DCK (ours)**	**46.6 / 67.0**	**22.3 / 40.1**	**12.4 / 19.8**
SelfCompact host‡ (rubric-driven, frozen ReSum-trained head)				
WebSailor-7B	SelfCompact	38.2 / 57.3	18.5 / 32.9	9.2 / 14.7
WebSailor-7B	+ DCK (ours)	40.8 / 60.2	18.7 / 34.3	10.0 / 15.7

4.3 Cross-Summarizer Generalization

Table 2 fixes the 7B closure head and varies the summary tool. DCK improves every tool on every dataset and metric, including the self-summary setting where summarizer and agent coincide. Margins are largest for the strongest tool, indicating that protected evidence and generated summaries are complementary: protection matters most when precise details risk being paraphrased away. Without per-tool calibration, the learned retention signal is summarizer-agnostic at a fixed checkpoint.

4.4 Agent-Scale Generalization

Table 3 fixes GPT-OSS-120B and scales the search agent. The paired DCK increment is positive on all datasets from the smallest to the largest agent and is preserved on the harder benchmarks, where compaction is frequent. Because every head is agent-specific and trained on the agent's own trajectories, the retention signal distilled from closure-knockout influence remains informative across hidden-state spaces and scales.

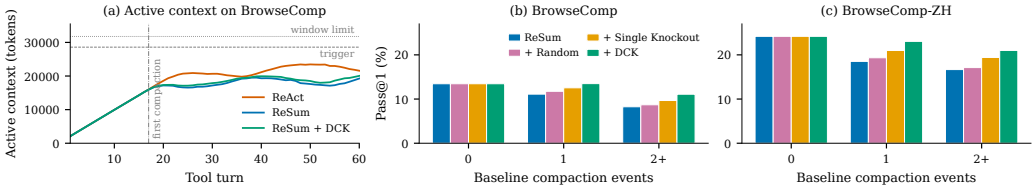


Figure 3: Context dynamics and DCK gains. (a) Working-context tokens per tool turn on BrowseComp: curves coincide before the first compaction, ReAct climbs toward the window limit, and DCK resets slightly above ReSum within budget. (b, c) Pass@1 by baseline compaction count on BrowseComp and BrowseComp-ZH: protection variants coincide on 0-compaction questions; the DCK margin concentrates on 1- and 2+-compaction slices, where lossy compaction occurs.

Table 2: Cross-summary-tool results with WebSailor-7B under the frozen ReSum trigger. All rows share the 7B closure head without per-tool relabeling or calibration; GPT-OSS-120B rows are reused from Table 1; better result per tool and dataset in **bold**.

Summary Tool	Method	GAIA P@1↑ / P@3↑	BrowseComp-ZH P@1↑ / P@3↑	BrowseComp P@1↑ / P@3↑
WebSailor-7B self-summary	ReSum	35.6 / 54.4	16.7 / 31.1	7.9 / 13.8
	+ DCK	**38.8 / 59.2**	**17.8 / 32.5**	**8.6 / 14.8**
Qwen3-30B-A3B-Thinking	ReSum	34.6 / 49.5	13.4 / 25.6	5.8 / 10.8
	+ DCK	**37.5 / 52.4**	**14.5 / 27.3**	**6.5 / 11.8**
GPT-OSS-120B	ReSum	41.8 / 60.2	18.8 / 35.0	10.3 / 16.8
	+ DCK	**46.6 / 67.0**	**22.3 / 40.1**	**12.4 / 19.8**
Qwen3-235B	ReSum	42.7 / 63.1	18.0 / 32.9	8.7 / 15.2
	+ DCK	**45.6 / 66.0**	**19.5 / 36.7**	**9.6 / 16.7**
DeepSeek-R1-671B	ReSum	40.1 / 58.3	17.0 / 31.5	10.1 / 16.1
	+ DCK	**43.0 / 62.1**	**18.2 / 33.6**	**11.2 / 17.3**

Table 3: Scaling DCK across frozen WebSailor agents with GPT-OSS-120B. Each agent uses its own head, never shared across scales; 7B rows are reused from Table 1; better result per agent and dataset in **bold**. The paired DCK increment stays positive from 3B to 30B-A3B on all datasets.

Search Agent	Method	GAIA P@1↑ / P@3↑	BrowseComp-ZH P@1↑ / P@3↑	BrowseComp P@1↑ / P@3↑
WebSailor-3B	ReSum	40.8 / 64.1	15.2 / 28.4	8.5 / 15.6
	+ DCK	**43.7 / 68.0**	**16.7 / 30.1**	**9.6 / 17.1**
WebSailor-7B	ReSum	41.8 / 60.2	18.8 / 35.0	10.3 / 16.8
	+ DCK	**46.6 / 67.0**	**22.3 / 40.1**	**12.4 / 19.8**
WebSailor-30B-A3B	ReSum	51.5 / 68.9	27.3 / 46.0	18.8 / 31.0
	+ DCK	**53.7 / 72.8**	**28.6 / 48.1**	**20.7 / 32.9**

4.5 Efficiency and Context Dynamics

Figure 3a traces working context per tool turn. Because the head runs only after the host trigger, the curves coincide until the first compaction, and protected blocks shift the reset level only mildly within budget. Figure 3b–c splits questions by baseline compaction count: protection variants are indistinguishable on never-compacting questions, and the DCK margin concentrates on questions with at least one compaction—precisely where lossy compaction occurs. On efficiency, compaction hosts issue more tool calls than ReAct because resets postpone the window constraint, while DCK adds only a lightweight scoring pass per compaction event and a modest token increase from carrying protected blocks: DCK converts its protected budget into accuracy rather than longer search (§E). The distilled heads also recover the oracle ranking: predicted influence correlates moderately with closure-knockout labels at all three scales, and selector-level agreement with the oracle top-8 set follows the end-task ordering (§G). Budget bookkeeping confirms exact conservation of the reset budget per event (§H). Case studies (§I) delineate the boundary: verbatim protection rescues details that summaries paraphrase away, the payload cap occasionally drops required mid-page evidence, and part of the ranking gain does not convert when the agent ignores protected evidence.

5 Conclusion

We presented Descendant Closure Knockout (DCK), a retention layer for lossy context compaction in long-horizon search agents. DCK masks an observation together with its entity-lineage descendants, measures the change in the frozen agent's gold-answer likelihood, distills the signal into a lightweight agent-specific head, and protects bounded verbatim evidence within the host's reset budget while leaving trigger and summarizer untouched. Across scales, summary tools, and hosts, protected compaction consistently improves its host, with gains concentrated on questions that compact at least once and attributed by controlled comparisons to closure-aware supervision.

AI USE STATEMENT

In this work, we used generative AI tools to assist with language editing and grammar polishing of the manuscript. We have not used generative AI tools for generating research ideas, conducting the core experiments, or producing the reported results, and no other disclosure-required tasks are applicable to this work. All AI-assisted text was carefully reviewed, verified, and revised by the authors, who take full responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI. All research ideas, methodological designs, experimental implementations, and result analyses were conducted solely by the authors.

ETHICS STATEMENT

This work does not raise notable ethical concerns. All experiments are conducted on publicly available benchmarks (GAIA, BrowseComp, and BrowseComp-ZH) with publicly released models, and no human subjects or personal data are involved. The proposed method operates as a context retention layer within the agent's own context management pipeline and does not access, store, or disclose any private user information. We are not aware of potential harmful insights, conflicts of interest, or fairness concerns arising from this work.

REPRODUCIBILITY STATEMENT

We summarize our efforts toward reproducibility. All benchmarks, agent configurations, compaction hosts, and baselines are specified in §4.1; detailed hyperparameters, decoding settings, and the DCK head training protocol are provided in the appendix (§C). Our evaluation protocol, including paired rollouts, judge prompts, and statistical testing procedures, is fully described in §D and §E. We release the source code and configuration files in §J.

REFERENCES

S. Bai, L. Bing, L. Lei, R. Li, X. Li, X. Lin, E. Min, L. Su, B. Wang, L. Wang, L. Wang, S. Wang, X. Wang, Y. Zhang, Z. Zhang, G. Chen, L. Chen, Z. Cheng, Y. Deng, Z. Huang, D. Ng, J. Ni, Q. Ren, X. Tang, B. L. Wang, H. Wang, N. Wang, C. Wei, Q. Wu, J. Xia, Y. Xiao, H. Xu, X. Xu, C. Xue, Z. Yang, Z. Yang, F. Ye, H. Ye, J. Yu, C. Zhang, W. Zhang, H. Zhao, and P. Zhu. Mirothinker-1.7 & H1: towards heavy-duty research agents via verification. *CoRR*, abs/2603.15726, 2026. doi: 10.48550/ARXIV.2603.15726. URL https://doi.org/10.48550/arXiv.2603.15726.

DeepSeek-AI, Aixin Liu, Aoxue Mei, Bangcai Lin, Bing Xue, Bingxuan Wang, Bingzheng Xu, Bochao Wu, Bowei Zhang, Chaofan Lin, Chen Dong, Chengda Lu, Chenggang Zhao, Chengqi Deng, Chenhao Xu, Chong Ruan, Damai Dai, Daya Guo, Dejian Yang, Deli Chen, Erhang Li, Fangqi Zhou, Fangyun Lin, Fucong Dai, Guangbo Hao, Guanting Chen, Guowei Li, H. Zhang, Hanwei Xu, Hao Li, Haofen Liang, Haoran Wei, Haowei Zhang, Haowen Luo, Haozhe Ji, Honghui Ding, Hongxuan Tang, Huanqi Cao, Huazuo Gao, Hui Qu, Hui Zeng, Jialiang Huang, Jiashi Li, Jiaxin Xu, Jiewen Hu, Jingchang Chen, Jingting Xiang, Jingyang Yuan, Jingyuan Cheng, Jinhua Zhu, Jun Ran, Junguang Jiang, Junjie Qiu, Junlong Li, Junxiao Song, Kai Dong, Kaige Gao, Kang Guan, Kexin Huang, Kexing Zhou, Kezhao Huang, Kuai Yu, Lean Wang, Lecong Zhang, Lei Wang, Liang Zhao, Liangsheng Yin, Lihua Guo, Lingxiao Luo, Linwang Ma, Litong Wang, Liyue Zhang, M. S. Di, M. Y Xu, Mingchuan Zhang, Minghua Zhang, Minghui Tang, Mingxu Zhou, Panpan Huang, Peixin Cong, Peiyi Wang, Qiancheng Wang, Qihao Zhu, Qingyang Li, Qinyu Chen, Qiushi Du, Ruiling Xu, Ruiqi Ge, Ruisong Zhang, Ruizhe Pan, Runji Wang, Runqiu Yin, Runxin Xu, Ruomeng Shen, Ruoyu Zhang, S. H. Liu, Shanghao Lu, Shangyan Zhou, Shanhuang Chen, Shaofei Cai, Shaoyuan Chen, Shengding Hu, Shengyu Liu, Shiqiang Hu, Shirong Ma, Shiyu Wang, Shuiping Yu, Shunfeng Zhou, Shuting Pan, Songyang Zhou, Tao Ni, Tao Yun, Tian Pei, Tian Ye, Tianyuan Yue, Wangding Zeng, Wen Liu, Wenfeng Liang, Wenjie Pang, Wenjing Luo, Wenjun Gao, Wentao Zhang, Xi Gao, Xiangwen Wang, Xiao Bi, Xiaodong Liu, Xiaohan Wang, Xiaokang Chen, Xiaokang Zhang, Xiaotao Nie, Xin Cheng, Xin Liu, Xin Xie, Xingchao Liu, Xingkai Yu, Xingyou Li, Xinyu Yang, Xinyuan Li, Xu Chen, Xuecheng Su, Xuehai Pan, Xuheng Lin, Xuwei Fu, Y. Q. Wang, Yang Zhang, Yanhong Xu, Yanru Ma, Yao Li, Yao Li, Yao Zhao, Yaofeng Sun, Yaohui Wang, Yi Qian, Yi Yu, Yichao Zhang, Yifan Ding, Yifan

Shi, Yiliang Xiong, Ying He, Ying Zhou, Yinmin Zhong, Yishi Piao, Yisong Wang, Yixiao Chen, Yixuan Tan, Yixuan Wei, Yiyang Ma, Yiyuan Liu, Yonglun Yang, Yongqiang Guo, Yongtong Wu, Yu Wu, Yuan Cheng, Yuan Ou, Yuanfan Xu, Yuduan Wang, Yue Gong, Yuhan Wu, Yuheng Zou, Yukun Li, Yunfan Xiong, Yuxiang Luo, Yuxiang You, Yuxuan Liu, Yuyang Zhou, Z. F. Wu, Z. Z. Ren, Zehua Zhao, Zehui Ren, Zhangli Sha, Zhe Fu, Zhean Xu, Zhenda Xie, Zhengyan Zhang, Zhewen Hao, Zhibin Gou, Zhicheng Ma, Zhigang Yan, Zhihong Shao, Zhixian Huang, Zhiyu Wu, Zhuoshu Li, Zhuping Zhang, Zian Xu, Zihao Wang, Zihui Gu, Zijia Zhu, Zilin Li, Zipeng Zhang, Ziwei Xie, Ziyi Gao, Zizheng Pan, Zongqing Yao, Bei Feng, Hui Li, J. L. Cai, Jiaqi Ni, Lei Xu, Meng Li, Ning Tian, R. J. Chen, R. L. Jin, S. S. Li, Shuang Zhou, Tianyu Sun, X. Q. Li, Xiangyue Jin, Xiaojin Shen, Xiaosha Chen, Xinnan Song, Xinyi Zhou, Y. X. Zhu, Yanping Huang, Yaohui Li, Yi Zheng, Yuchen Zhu, Yunxian Ma, Zhen Huang, Zhipeng Xu, Zhongyu Zhang, Dongjie Ji, Jian Liang, Jianzhong Guo, Jin Chen, Leyi Xia, Miaojun Wang, Mingming Li, Peng Zhang, Ruyi Chen, Shangmian Sun, Shaoqing Wu, Shengfeng Ye, T. Wang, W. L. Xiao, Wei An, Xianzu Wang, Xiaowen Sun, Xiaoxiang Wang, Ying Tang, Yukun Zha, Zekai Zhang, Zhe Ju, Zhen Zhang, and Zihua Qu. Deepseek-v3.2: Pushing the frontier of open large language models. 2025. URL https://arxiv.org/abs/2512.02556.

Jiaxuan Gao, Wei Fu, Minyang Xie, Shusheng Xu, Chuyi He, Zhiyu Mei, Banghua Zhu, and Yi Wu. Beyond ten turns: Unlocking long-horizon agentic search with large-scale asynchronous RL. *CoRR*, abs/2508.07976, 2025. doi: 10.48550/ARXIV.2508.07976. URL https://doi.org/10.48550/arXiv.2508.07976.

Bowen Jin, Hansi Zeng, Zhenrui Yue, Dong Wang, Hamed Zamani, and Jiawei Han. Search-r1: Training llms to reason and leverage search engines with reinforcement learning. *CoRR*, abs/2503.09516, 2025. doi: 10.48550/ARXIV.2503.09516. URL https://doi.org/10.48550/arXiv.2503.09516.

Baixuan Li, Bo Zhang, Dingchu Zhang, Fei Huang, Guangyu Li, Guoxin Chen, Huifeng Yin, Jialong Wu, Jingren Zhou, Kuan Li, Liangcai Su, Litu Ou, Liwen Zhang, Pengjun Xie, Rui Ye, Wenbiao Yin, Xinmiao Yu, Xinyu Wang, Xixi Wu, Xuanzhong Chen, Yida Zhao, Zhen Zhang, Zhengwei Tao, Zhongwang Zhang, Zile Qiao, Chenxi Wang, Donglei Yu, Gang Fu, Haiyang Shen, Jiayin Yang, Jun Lin, Junkai Zhang, Kui Zeng, Li Yang, Hailong Yin, Maojia Song, Ming Yan, Peng Xia, Qian Xiao, Rui Min, Ruixue Ding, Runnan Fang, Shaowei Chen, Shen Huang, Shihang Wang, Shihao Cai, Weizhou Shen, Xiaobin Wang, Xin Guan, Xinyu Geng, Yingcheng Shi, Yuning Wu, Zhuo Chen, Zijian Li, and Yong Jiang. Tongyi deepresearch technical report. *CoRR*, abs/2510.24701, 2025a. doi: 10.48550/ARXIV.2510.24701. URL https://doi.org/10.48550/arXiv.2510.24701.

Kuan Li, Zhongwang Zhang, Huifeng Yin, Liwen Zhang, Litu Ou, Jialong Wu, Wenbiao Yin, Baixuan Li, Zhengwei Tao, Xinyu Wang, Weizhou Shen, Junkai Zhang, Dingchu Zhang, Xixi Wu, Yong Jiang, Ming Yan, Pengjun Xie, Fei Huang, and Jingren Zhou. Websailor: Navigating superhuman reasoning for web agent. *CoRR*, abs/2507.02592, 2025b. doi: 10.48550/ARXIV.2507.02592. URL https://doi.org/10.48550/arXiv.2507.02592.

Tianjian Li, Jingyu Zhang, William Jurayj, Xi Wang, Chuanyang Jin, Mehrdad Farajtabar, Eric T. Nalisnick, and Daniel Khashabi. Self-compacting language model agents. *CoRR*, abs/2606.23525, 2026a. doi: 10.48550/ARXIV.2606.23525. URL https://doi.org/10.48550/arXiv.2606.23525.

Xiaoxi Li, Guanting Dong, Jiajie Jin, Yuyao Zhang, Yujia Zhou, Yutao Zhu, Peitian Zhang, and Zhicheng Dou. Search-o1: Agentic search-enhanced large reasoning models. In Christos Christodoulopoulos, Tanmoy Chakraborty, Carolyn Rose, and Violet Peng (eds.), *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing, EMNLP 2025, Suzhou, China, November 4-9, 2025*, pp. 5420–5438. Association for Computational Linguistics, 2025c. doi: 10.18653/V1/2025.EMNLP-MAIN.276. URL https://doi.org/10.18653/v1/2025.emnlp-main.276.

Zhuofeng Li, Dongfu Jiang, Xueguang Ma, Haoxiang Zhang, Ping Nie, Yuyu Zhang, Kai Zou, Jianwen Xie, Yu Zhang, and Wenhu Chen. Openresearcher: A fully open pipeline for long-horizon deep research trajectory synthesis. *CoRR*, abs/2603.20278, 2026b. doi: 10.48550/ARXIV.2603.20278. URL https://doi.org/10.48550/arXiv.2603.20278.

Junteng Liu, Yunji Li, Chi Zhang, Jingyang Li, Aili Chen, Ke Ji, Weiyu Cheng, Zijia Wu, Chengyu Du, Qidi Xu, Jiayuan Song, Zhengmao Zhu, Wenhu Chen, Pengyu Zhao, and Junxian He. Webexplorer: Explore and evolve for training long-horizon web agents. *CoRR*, abs/2509.06501, 2025. doi: 10.48550/ARXIV.2509.06501. URL https://doi.org/10.48550/arXiv.2509.06501.

Miao Lu, Weiwei Sun, Weihua Du, Zhan Ling, Xuesong Yao, Kang Liu, and Jiecao Chen. Scaling LLM multi-turn RL with end-to-end summarization-based context management. *CoRR*, abs/2510.06727, 2025a. doi: 10.48550/ARXIV.2510.06727. URL https://doi.org/10.48550/arXiv.2510.06727.

Rui Lu, Zhenyu Hou, Zihan Wang, Hanchen Zhang, Xiao Liu, Yujiang Li, Shi Feng, Jie Tang, and Yuxiao Dong. Deepdive: Advancing deep search agents with knowledge graphs and multi-turn RL. *CoRR*, abs/2509.10446, 2025b. doi: 10.48550/ARXIV.2509.10446. URL https://doi.org/10.48550/arXiv.2509.10446.

Yijun Lu, Rui Ye, Yuwen Du, Jiajun Wang, Songhua Liu, and Siheng Chen. Longseeker: Elastic context orchestration for long-horizon search agents. *CoRR*, abs/2605.05191, 2026. doi: 10.48550/ARXIV.2605.05191. URL https://doi.org/10.48550/arXiv.2605.05191.

Lang Mei, Xiaohan Yu, Chong Chen, Liyan Liu, Xiangnan Chen, Jinchao Ma, Chao Feng, Li Huang, Siyu Mo, Sichen Kang, Yunkun Xu, Zhihan Yang, Zhujun Xue, Jingren Zhang, Qing He, Yingdi Huang, Hao Jiang, Ziao Ma, Zewei Pan, Minhao Sun, Zhuo Tao, Jinzhao Xiao, Gangtao Xin, Huanyao Zhang, Wenjian Zhang, Jiangshan Zhang, Guojie Zhu, Jiaxin Mao, and Wentao Zhang. Searchart: Training long-horizon search agent with scalable synthetic and verified task. *CoRR*, abs/2607.24850, 2026. doi: 10.48550/ARXIV.2607.24850. URL https://doi.org/10.48550/arXiv.2607.24850.

Grégoire Mialon, Clémentine Fourrier, Thomas Wolf, Yann LeCun, and Thomas Scialom. GAIA: a benchmark for general AI assistants. In *The Twelfth International Conference on Learning Representations, ICLR 2024, Vienna, Austria, May 7-11, 2024*. OpenReview.net, 2024. URL https://openreview.net/forum?id=fibxvahvs3.

Reiichiro Nakano, Jacob Hilton, Suchir Balaji, Jeff Wu, Long Ouyang, Christina Kim, Christopher Hesse, Shantanu Jain, Vineet Kosaraju, William Saunders, Xu Jiang, Karl Cobbe, Tyna Eloundou, Gretchen Krueger, Kevin Button, Matthew Knight, Benjamin Chess, and John Schulman. Webgpt: Browser-assisted question-answering with human feedback. *CoRR*, abs/2112.09332, 2021. URL https://arxiv.org/abs/2112.09332.

Jiaqi Shao, Yufeng Miao, Wei Zhang, and Bing Luo. Foldact: Efficient and stable context folding for long-horizon search agents. *CoRR*, abs/2512.22733, 2025. doi: 10.48550/ARXIV.2512.22733. URL https://doi.org/10.48550/arXiv.2512.22733.

Huatong Song, Jinhao Jiang, Yingqian Min, Jie Chen, Zhipeng Chen, Wayne Xin Zhao, Lei Fang, and Ji-Rong Wen. R1-searcher: Incentivizing the search capability in llms via reinforcement learning. *CoRR*, abs/2503.05592, 2025. doi: 10.48550/ARXIV.2503.05592. URL https://doi.org/10.48550/arXiv.2503.05592.

Weiwei Sun, Miao Lu, Zhan Ling, Kang Liu, Xuesong Yao, Yiming Yang, and Jiecao Chen. Scaling long-horizon LLM agent via context-folding. *CoRR*, abs/2510.11967, 2025. doi: 10.48550/ARXIV.2510.11967. URL https://doi.org/10.48550/arXiv.2510.11967.

Zhengwei Tao, Jialong Wu, Wenbiao Yin, Junkai Zhang, Baixuan Li, Haiyang Shen, Kuan Li, Liwen Zhang, Xinyu Wang, Yong Jiang, Pengjun Xie, Fei Huang, and Jingren Zhou. Webshaper: Agentically data synthesizing via information-seeking formalization. *CoRR*, abs/2507.15061, 2025. doi: 10.48550/ARXIV.2507.15061. URL https://doi.org/10.48550/arXiv.2507.15061.

Jason Wei, Zhiqing Sun, Spencer Papay, Scott McKinney, Jeffrey Han, Isa Fulford, Hyung Won Chung, Alex Tachard Passos, William Fedus, and Amelia Glaese. Browsecomp: A simple yet challenging benchmark for browsing agents. *CoRR*, abs/2504.12516, 2025. doi: 10.48550/ARXIV.2504.12516. URL https://doi.org/10.48550/arXiv.2504.12516.

Jialong Wu, Baixuan Li, Runnan Fang, Wenbiao Yin, Liwen Zhang, Zhenglin Wang, Zhengwei Tao, Dingchu Zhang, Zekun Xi, Robert Tang, Yong Jiang, Pengjun Xie, Fei Huang, and Jingren Zhou. Webdancer: Towards autonomous information seeking agency. In Danielle Belgrave, Cheng Zhang, Laura N. Montoya, Hsuan-Tien Lin, Razvan Pascanu, Piotr Koniusz, Marzyeh Ghassemi, Nancy Chen, Iván Vladimir Meza Ruíz, and Arturo Loaiza-Bonilla (eds.), *Advances in Neural Information Processing Systems 38: Annual Conference on Neural Information Processing Systems 2025, NeurIPS 2025, San Diego, CA, USA, December 2-7, 2025 / Mexico City, Mexico, November 30 - December 5, 2025*, 2025a. URL http://papers.nips.cc/paper_files/paper/2025/hash/af043aee9cb137785d93195c6cf4cd96-Abstract-Conference.html.

Jialong Wu, Wenbiao Yin, Yong Jiang, Zhenglin Wang, Zekun Xi, Runnan Fang, Linhai Zhang, Yulan He, Deyu Zhou, Pengjun Xie, and Fei Huang. Webwalker: Benchmarking llms in web traversal. In Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar (eds.), *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2025, Vienna, Austria, July 27 - August 1, 2025*, pp. 10290–10305. Association for Computational Linguistics, 2025b. doi: 10.18653/V1/2025.ACL-LONG.508. URL https://doi.org/10.18653/v1/2025.acl-long.508.

Xixi Wu, Kuan Li, Yida Zhao, Liwen Zhang, Litu Ou, Huifeng Yin, Zhongwang Zhang, Yong Jiang, Pengjun Xie, Fei Huang, Minhao Cheng, Shuai Wang, Hong Cheng, and Jingren Zhou. Resum: Unlocking long-horizon search intelligence via context summarization. *CoRR*, abs/2509.13313, 2025c. doi: 10.48550/ARXIV.2509.13313. URL https://doi.org/10.48550/arXiv.2509.13313.

Yong Wu, Yanzhao Zheng, Tianze Xu, ZhenTao Zhang, YuanQiang Yu, JiHuai Zhu, Chao Ma, Binbin Lin, Baohua Dong, Hangcheng Zhu, Ruohui Huang, and Gang Yu. Contextbudget: Budget-aware context management for long-horizon search agents. *CoRR*, abs/2604.01664, 2026. doi: 10.48550/ARXIV.2604.01664. URL https://doi.org/10.48550/arXiv.2604.01664.

Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik R. Narasimhan, and Yuan Cao. React: Synergizing reasoning and acting in language models. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023. URL https://openreview.net/forum?id=WE_vluYUL-X.

Yilun Yao, Shan Huang, Elsie Dai, Zhewen Tan, Zhenyu Duan, Shousheng Jia, Yanbing Jiang, and Tong Yang. ARC: active and reflection-driven context management for long-horizon information seeking agents. *CoRR*, abs/2601.12030, 2026. doi: 10.48550/ARXIV.2601.12030. URL https://doi.org/10.48550/arXiv.2601.12030.

Rui Ye, Zhongwang Zhang, Kuan Li, Huifeng Yin, Zhengwei Tao, Yida Zhao, Liangcai Su, Liwen Zhang, Zile Qiao, Xinyu Wang, Pengjun Xie, Fei Huang, Siheng Chen, Jingren Zhou, and Yong Jiang. Agentfold: Long-horizon web agents with proactive context management. *CoRR*, abs/2510.24699, 2025. doi: 10.48550/ARXIV.2510.24699. URL https://doi.org/10.48550/arXiv.2510.24699.

Lu Yi, Runlin Lei, Liuyi Yao, Yuexiang Xie, Yuyang Li, Wenhao Zhang, Zhewei Wei, Yaliang Li, and Jian-Yun Nie. Learning agent-compatible context management for long-horizon tasks. *CoRR*, abs/2605.30785, 2026. doi: 10.48550/ARXIV.2605.30785. URL https://doi.org/10.48550/arXiv.2605.30785.

Hongli Yu, Tinghong Chen, Jiangtao Feng, Jiangjie Chen, Weinan Dai, Qiying Yu, Ya-Qin Zhang, Wei-Ying Ma, Jingjing Liu, Mingxuan Wang, and Hao Zhou. Memagent: Reshaping long-context LLM with multi-conv rl-based memory agent. *CoRR*, abs/2507.02259, 2025. doi: 10.48550/ARXIV.2507.02259. URL https://doi.org/10.48550/arXiv.2507.02259.

Qianhao Yuan, Jie Lou, Zichao Li, Jiawei Chen, Yaojie Lu, Hongyu Lin, Le Sun, Debing Zhang, and Xianpei Han. Memsearcher: Iterative memory integration for search agent via end-to-end reinforcement learning. In Maria Liakata, Viviane P. Moreira, Jiajun Zhang, and David Jurgens (eds.), *Findings of the Association for Computational Linguistics, ACL 2026, San Diego, California, United States, July 2-7, 2026*, pp. 14965–14977. Association for Computational Linguistics,

2026. doi: 10.18653/V1/2026.FINDINGS-ACL.736. URL https://doi.org/10.18653/v1/2026.findings-acl.736.

Haoxiang Zhang, Qixin Xu, Zhuofeng Li, Lei Zhang, Pengcheng Jiang, Yu Zhang, and Julian J. McAuley. Masking stale observations helps search agents - until it doesn't: A regime map and its mechanism. *CoRR*, abs/2606.00408, 2026. doi: 10.48550/ARXIV.2606.00408. URL https://doi.org/10.48550/arXiv.2606.00408.

Peilin Zhou, Bruce Leon, Xiang Ying, Can Zhang, Yifan Shao, Qichen Ye, Dading Chong, Zhiling Jin, Chenxuan Xie, Meng Cao, Yuxin Gu, Sixin Hong, Jing Ren, Jian Chen, Chao Liu, and Yining Hua. Browsecomp-zh: Benchmarking web browsing ability of large language models in chinese. *CoRR*, abs/2504.19314, 2025a. doi: 10.48550/ARXIV.2504.19314. URL https://doi.org/10.48550/arXiv.2504.19314.

Zijian Zhou, Ao Qu, Zhaoxuan Wu, Sunghwan Kim, Alok Prakash, Daniela Rus, Jinhua Zhao, Bryan Kian Hsiang Low, and Paul Pu Liang. MEM1: learning to synergize memory and reasoning for efficient long-horizon agents. *CoRR*, abs/2506.15841, 2025b. doi: 10.48550/ARXIV.2506.15841. URL https://doi.org/10.48550/arXiv.2506.15841.

A Appendix

A.1 The process of DCK algorithm

Algorithm 1 Offline Training of the DCK Influence Head

Input: Frozen agent π_{θ_A}; host snapshots $\mathcal{T}_A = \{(H_s, a^\star)\}$ collected without DCK; clipping bound c

Output: Agent specific influence head g_{ψ_A} and feature standardizer $\mathbf{z}_A$

$\mathcal{D}_A \leftarrow \emptyset$
for all $(H_s, a^\star) \in \mathcal{T}_A$ **do**
  Identify $\mathcal{A}_s$ and construct $E_{1:s}$ and $D_{\le s}(\cdot)$
  $\ell_s \leftarrow \mathcal{J}_A(H_s; a^\star)$
  for all $t \in \mathcal{A}_s$ **do**
    $\Gamma_{t,s} \leftarrow \{t\} \cup \big(D_{\le s}(t) \cap \mathcal{A}_s\big)$
    $I_{t,s} \leftarrow \mathrm{clip}\big(\ell_s - \mathcal{J}_A(\mathsf{K}_{\Gamma_{t,s}}(H_s); a^\star), -c, c\big)$
    $\mathcal{D}_A \leftarrow \mathcal{D}_A \cup \{(\mathbf{h}_t^A, \mathbf{f}_{t,s}^A, I_{t,s})\}$
  end for
end for
$\mathbf{z}_A \leftarrow \mathrm{FitStandardizer}(\{\mathbf{f} : (\mathbf{h}, \mathbf{f}, I) \in \mathcal{D}_A\})$
$\psi_A \leftarrow \arg\min_{\psi} \dfrac{1}{|\mathcal{D}_A|} \displaystyle\sum_{(\mathbf{h},\mathbf{f},I)\in\mathcal{D}_A} \mathrm{Huber}(g_\psi([\mathbf{h}; \mathbf{z}_A(\mathbf{f})]) - I)$
return $g_{\psi_A}, \mathbf{z}_A$

Algorithm 2 DCK Protected Compaction at Inference

Input: Working context H_s; trained g_{ψ_A} and $\mathbf{z}_A$; host $\mathcal{H} = (\chi_{\mathcal{H}}, \mathcal{S}_{\mathcal{H}}, \mathrm{Compose}_{\mathcal{H}})$; capacity K; reset budget L_{reset}

Output: Updated working context $\widetilde{H}_s^{\mathcal{H}}$

if $\chi_{\mathcal{H}}(H_s) = 0$ **then**
  return H_s
end if
Identify the current addressable set $\mathcal{A}_s$
for all $t \in \mathcal{A}_s$ **do**
  $\widehat{I}_{t,s} \leftarrow g_{\psi_A}([\mathbf{h}_t^A; \mathbf{z}_A(\mathbf{f}_{t,s}^A)])$
end for
$B_s \leftarrow \mathrm{TopK}_K\{(t, \widehat{I}_{t,s}) : t \in \mathcal{A}_s\}$
$P_s \leftarrow \mathrm{Chronological}(B_s)$
$L_{\mathrm{sum}} \leftarrow L_{\mathrm{reset}} - \mathrm{Tok}_A\left(\bigoplus_{t\in P_s} \rho(o_t)\right) - L_{\mathrm{overhead}}$
$S_s^{\mathcal{H}} \leftarrow \mathcal{S}_{\mathcal{H}}(M_{\mathcal{H}}(H_s, B_s); L_{\mathrm{sum}})$
$\widetilde{H}_s^{\mathcal{H}} \leftarrow \mathrm{Compose}_{\mathcal{H}}\left(x, S_s^{\mathcal{H}}, \bigoplus_{t\in P_s} \rho(o_t)\right)$
return $\widetilde{H}_s^{\mathcal{H}}$

B Proofs

B.1 Proof of Lemma 1

Fix a root t and horizon s. We prove by induction over $u = t+1, \ldots, s$ that the temporal recursive rule admits u if and only if u is reachable from t in G_s. For the base case $u = t+1$, admission holds exactly when $\eta(Q_u) \cap E_t \neq \emptyset$, which is equivalent to the edge $(t, u) \in \mathcal{E}_s$.

Assume the equivalence holds for every index smaller than u. The recursive rule admits u exactly when there exists $v \in \{t\} \cup D_{<u}(t)$ such that $\eta(Q_u) \cap E_v \neq \emptyset$. By definition, this condition is

equivalent to $(v, u) \in \mathcal{E}_s$. By the induction hypothesis, either $v = t$ or there is a path $t \rightsquigarrow v$, and hence $t \rightsquigarrow v \to u$. Conversely, if $t \rightsquigarrow u$, let (v, u) be the final edge on such a path. Then either $v = t$ or v has already been admitted before processing u, so the recursive rule admits u. Thus the recursive closure equals the graph reachability set. Reachability is unique, and every edge obeys $v < u$, so the graph is acyclic and a chronological scan is sufficient.

B.2 Proof of Proposition 1

Let $\mathcal{F}_s$ be the information available by horizon s, including the question, queries and observations through s, previous summaries and protected blocks, cached observation representations, and lineage metadata. The stop-entity set is fixed from the head-training split before any target trajectory is processed. Consequently, each first-introduced entity set E_t for $t \le s$ is a function of the prefix through t, and every edge in G_s is determined by $Q_{1:s}$ and $E_{1:s}$. Hence $D_{\le s}(t)$ and $\Gamma_{t,s}$ are $\mathcal{F}_s$-measurable.

The cached representation $\mathbf{h}_t^A$ is computed when o_t first enters the working context, and each component of $\mathbf{f}_{t,s}^A$ depends only on indices, addressable descendants, first-introduced entities, observation length, and earlier-query repetition through s. Therefore the predictor input is also $\mathcal{F}_s$-measurable. The factual and knockout contexts both terminate at horizon s. Computing $I_{t,s}$ additionally requires the offline gold answer $a^\star$, but no trajectory variable after s is used. Thus $I_{t,s}$ is measurable with respect to $\mathcal{F}_s \vee \sigma(a^\star)$ and contains no future-trajectory leakage.

B.3 Proof of Proposition 2

Expanding the two definitions gives

$$\widetilde{\Delta}_{t,s} = \widetilde{I}_{t,s} - \widetilde{I}_{t,s}^{\text{single}} \tag{22}$$

$$= \left[\mathcal{J}_A(H_s; a^\star) - \mathcal{J}_A\big(\mathsf{K}_{\Gamma_{t,s}}(H_s); a^\star\big)\right] \tag{23}$$

$$- \left[\mathcal{J}_A(H_s; a^\star) - \mathcal{J}_A\big(\mathsf{K}_{\{t\}}(H_s); a^\star\big)\right] \tag{24}$$

$$= \mathcal{J}_A\big(\mathsf{K}_{\{t\}}(H_s); a^\star\big) - \mathcal{J}_A\big(\mathsf{K}_{\Gamma_{t,s}}(H_s); a^\star\big)\,. \tag{25}$$

The factual-context terms cancel exactly. If no addressable descendant exists, then $\Gamma_{t,s} = \{t\}$ and the increment is zero. No sign follows in general because the additionally masked descendants may support or mislead the frozen agent.

B.4 Boundary on Environment-Level Total Influence

For completeness, let $H_s^{\text{env}(t)}$ denote the context produced by changing the root observation and allowing subsequent queries, observations, reasoning, and summaries to respond in the environment. Define

$$TE_{t,s} = \mathcal{J}_A(H_s; a^\star) - \mathcal{J}_A(H_s^{\text{env}(t)}; a^\star). \tag{26}$$

Then the residual

$$r_{t,s} = \mathcal{J}_A\big(\mathsf{K}_{\Gamma_{t,s}}(H_s); a^\star\big) - \mathcal{J}_A(H_s^{\text{env}(t)}; a^\star) \tag{27}$$

satisfies the definitional decomposition

$$TE_{t,s} = \widetilde{I}_{t,s} + r_{t,s}. \tag{28}$$

The residual can contain information retained in factual queries and reasoning, non-entity dependencies, content already absorbed into an earlier summary, and changes induced by environmental re-execution. The decomposition supplies a claim boundary rather than an identification result; DCK does not assume $r_{t,s} = 0$.

C Experimental Details and Hyperparameters

Inference configuration. All current-matched runs share the following settings. Search agents are served frozen with a 32K-token context window (L_{host}=31,768 tokens) and a cap of 60 LLM calls per rollout; decoding uses temperature 0.6 and top-p 0.95. The Search and Visit tools, the page

evidence extractor, and the judge (Qwen2.5-72B-Instruct with the frozen ReSum judge prompts and parser) are identical across methods. Each method runs three paired rollout iterations per question; runs are interleaved by question, iteration, and method, and raw tool returns are archived with UTC timestamps.

Host configurations. The ReSum host triggers summarization at 90% of the window (28,591 tokens) and restarts from the question plus a summary. The SelfCompact host starts rubric probes at the third tool round at a fixed interval, permits compression once the context exceeds 20% of the window, enforces an unconditional backstop at 30%, and performs at most one summary per trajectory. The Fixed-Interval Summary baseline restarts from a WebSailor-7B self-summary every 10 completed tool calls. Recent History uses the ReSum trigger but keeps only the most recent 22K tokens without calling any summary model.

DCK protection configuration. We fix $K{=}8$ protected observations per compaction event. The verbatim serialization $\rho(o_t)$ contains a stable observation identifier, compact source metadata, and an evidence payload capped at 512 tokens, truncated to head 384 plus tail 128 tokens when the raw payload exceeds the cap; the full protected block is capped at 704 tokens. The reset budget is $L_{\text{reset}}{=}12{,}288$ tokens for both the baseline and DCK, constructed as $1024 \cdot \lceil (L_{\text{overhead,max}} + L_{\text{protected,max}} + L_{\text{summary,min}})/1024 \rceil$ from the component budgets $2{,}560{+}5{,}632{+}4{,}096$. The window safety check $L_{\text{reset}} + L_{\text{turn,max}} = 15{,}464 \leq 0.9\, L_{\text{host}}$ holds by construction. Protected blocks are re-ranked at every compaction event (dynamic, non-persistent protection).

Head training configuration. Heads are trained from 300 questions sampled from OpenSeeker-v1-Data (270 English, 30 Chinese) with a question-level 90/10 split stratified by language. Labels are closure-knockout influences clipped to $[-5, 5]$. Each head is a two-layer MLP (input_dim $\rightarrow$ $512 \rightarrow 1$) over the cached observation representation concatenated with the seven standardized horizon features; input dimension equals the agent hidden size plus seven (7B: $3591 \rightarrow 1{,}839{,}617$ parameters; 3B and 30B-A3B: $2055 \rightarrow 1{,}053{,}185$). Training uses AdamW (learning rate 10^{-3}, betas $(0.9, 0.999)$, $\epsilon = 10^{-8}$, weight decay 0.01), batch size 256, three epochs, gradient-norm clip 1.0, and seed 42. The 7B single-knockout comparator head shares questions, trajectories, candidates, cached hidden states, features, capacity, optimizer, data order, and checkpoint rule with the closure head; only the supervision label differs.

D Additional Paired Statistics

Table 4: Question-level paired statistics on BrowseComp for WebSailor-7B. We report mean Pass@1 differences with 95% percentile cluster-bootstrap intervals, one-sided sign-flip p-values, and Holm-adjusted p-values over the two core comparisons; both core comparisons remain significant after Holm correction.

Comparison	ΔP@1 (pp)	95% CI	p / Holm p
DCK vs. ReSum	$+2.1$	$[1.3,\ 2.9]$	.0001 / .0002
DCK vs. Single Knockout	$+1.0$	$[0.2,\ 1.7]$	.006 / .006
Random vs. ReSum	$+0.4$	$[-0.2,\ 1.0]$	.083 / –
SelfCompact+DCK vs. SelfCompact	$+0.8$	$[0.2,\ 1.4]$	.007 / –

E Full Efficiency Results

Trajectory token counts follow the evaluation protocol and cover the full recorded trajectory, including history that is later compacted away and restart messages, so compaction hosts legitimately exceed the window-free ReAct budget in recorded tokens while the active working context stays bounded (Figure 3a). ReSum-family methods issue more tool calls than ReAct because compaction postpones the window constraint and exploration continues under the call cap. Within the ReSum family, DCK adds only the head scoring pass at each compaction event: unconditional tool calls remain at the host level (51.4 vs. 51.5 on BrowseComp, 52.4 vs. 53.1 on BrowseComp-ZH), and

Table 5: Per-question rollout decomposition of the DCK vs. ReSum and DCK vs. Single Knockout comparisons on BrowseComp: number of questions whose three paired rollouts net-improved, net-degraded, or stayed tied, together with the graded rollout-level differences and supplementary per-dataset mean Pass@1 differences. Improvements dominate regressions in both comparisons.

	DCK vs. ReSum	DCK vs. Single
Questions better / worse / tied	72 / 12 / 1182	26 / 11 / 1229
Rollouts $+3$ / $+2$ / $+1$	10 / 8 / 54	11 / 8 / 7
Rollouts -1 / -2 / -3	6 / 4 / 2	5 / 4 / 2
Supplementary ΔP@1, GAIA	$+4.9$	$+2.3$ (vs. Single)
Supplementary ΔP@1, BrowseComp-ZH	$+3.5$	$+1.4$ (vs. Single)
Supplementary ΔP@1, SelfCompact hosts (GAIA / ZH / BC)	$+2.6$ / $+0.2$ / $+0.8$	

Table 6: Efficiency of WebSailor-7B methods on BrowseComp and BrowseComp-ZH. Token counts use the evaluation tokenizer over the full recorded trajectory; unconditional averages are over all rollouts, conditional averages over judge-correct rollouts only, with the number of correct rollouts (N_{corr}, out of $3{,}798$ and 867 respectively). Within the ReSum family, DCK adds only a lightweight head-scoring pass per compaction event: unconditional tool calls stay at the host level, and unconditional token totals grow only modestly from carrying the protected blocks.

	BrowseComp				BrowseComp-ZH			
Method	Tokens	Calls	Corr. tokens	N_{corr}	Tokens	Calls	Corr. tokens	N_{corr}
ReAct	29571	37.7	28733	213	30497	38.6	26168	111
Recent History	52827	41.7	47347	202	28168	34.2	21957	127
Fixed-Interval Summary	75189	43.1	77608	295	57333	33.6	55587	141
ReSum	58517	51.5	49656	390	57578	53.1	51183	163
+ Random Protection	60645	50.9	51369	406	58487	53.1	51455	167
+ Single Knockout	61241	51.2	54347	433	57942	52.8	52518	181
+ DCK (ours)	61508	51.4	54667	470	57765	52.4	53634	193
SelfCompact	35443	38.4	23627	351	28695	29.5	22961	160
+ DCK (ours)	35761	34.3	25353	381	33136	29.9	28075	162

the modest growth in unconditional token totals ($+5.1\%$ and $+0.3\%$) reflects the protected verbatim blocks carried in each reset context.

F Compaction-Slice Details

Table 7: Pass@1 by number of baseline compaction events on GAIA (58 / 29 / 16 questions per slice). Protection variants coincide on the 0-compaction slice; the DCK margin concentrates on the 1- and 2+-compaction slices.

Method	0 comp.	1 comp.	2+ comp.
ReSum	34.5	50.6	52.1
ReSum + DCK	34.5	63.2	60.4

Slice sizes follow from the compaction event logs: 260 / 420 / 586 BrowseComp questions and 62 / 81 / 146 BrowseComp-ZH questions fall in the 0-, 1-, and 2+-compaction slices. On GAIA the slice composition differs in character: questions that never compact are the hardest early-answer cases, while compaction correlates with longer exploration, which is why the 0-compaction Pass@1 is lowest there.

G Head Diagnostics

Selector-level agreement with the oracle top-8 closure-influence set on the 7B validation split is 0.61 for the closure head, 0.53 for the single-knockout head, and 0.44 for a uniform random selector of

Table 8: Agent-specific head diagnostics. Valid scoring events are compaction events captured plus terminal-horizon snapshots minus score-context overflows; the single-knockout head is trained only for 7B. All heads select three epochs, and validation Huber/Spearman show a moderate fit to the closure-knockout labels at all three scales.

	3B	7B	30B-A3B
Head parameters	1.05M	1.84M	1.05M
Valid scoring events	653	676	669
Labeled rows (train / val)	8005 (7214 / 791)	8383 (7552 / 831)	8763 (7894 / 869)
Closure head val. Huber	0.521	0.486	0.462
Closure head val. Spearman	0.512	0.547	0.529
Clip saturation rate	0.031	0.038	0.041
Single head val. Spearman	–	0.489	–

the same capacity (hypergeometric expectation ≈ 0.45), mirroring the end-task ordering of Table 1. The head adds about 4 ms of latency per compaction event.

H Protection Statistics

Table 9: Per-compaction-event averages and realized reset-context sizes. Every realized reset context closes exactly under the shared L_{reset}=12,288-token cap; DCK reallocates budget from summary tokens to protected verbatim blocks rather than enlarging the reset.

	GAIA	BrowseComp-ZH	BrowseComp
Addressable candidates	14.2	16.5	17.8
Protected items ($\leq K$=8)	6.9	7.4	7.6
Protected tokens	3924	4210	4386
Summary tokens (DCK)	6150	6180	5940
Overflow events	0	0	0
Realized reset, ReSum baseline	9900	9850	10020
Realized reset, ReSum + DCK	12074	11940	11846
Realized reset, SelfCompact	4520 (baseline) / 8606 (+ DCK)		

I Qualitative Case Studies

Four cases delineate the boundary of the method. On one BrowseComp question, the final evidence chain required an exact registration date first seen at tool turn 6: the summarizer paraphrased the date incorrectly under the ReSum baseline, whereas DCK protected the source observation verbatim, and both rollouts that protected it answered correctly. Conversely, on another BrowseComp question, the required evidence lay in the middle of a long page return: the head-to-tail payload cap retained only the outer segments and lost the mid-page evidence, a failure of payload truncation rather than selection, since the head ranked the observation highly. On a BrowseComp-ZH question, the single-knockout head failed to protect an early observation whose entities seeded later queries, and the summarizer dropped the relation, whereas the closure head protected the root observation together with its lineage and solved the question. Finally, on a GAIA question, both protection variants retained the relevant evidence verbatim, but the agent failed to use it in the final answer, showing that protection guarantees survival of evidence, not its exploitation—delineating the boundary between retention and policy ability.

J Code Link

Fully anonymized code is available at: https://anonymous.4open.science/r/DCK-70F1/