M6: An Evidence-Bounded Dialogic Regime Controller for Memory-Augmented Conversational Retrieval

Working Draft

TOF MUCK, Independent research draft, France

Conversational systems that combine retrieval, memory, and large language model generation must coordinate not only what information is retrieved, but also how the dialogue should proceed across turns. In many practical systems, this interactional control problem is absorbed into prompting conventions, style presets, or loosely specified response heuristics. Such approaches are difficult to audit, difficult to stabilize, and difficult to evaluate independently from the rest of the pipeline. This paper presents M6, a modular controller for temporal dialogic coherence in a memory-augmented conversational architecture. M6 does not model user psychology, does not rank retrieval candidates, and does not directly generate final responses. Instead, it computes a bounded dialogic regime from a short present window, a previously validated local state, and evidence candidates anchored in conversation traces. The module combines threshold-bounded state updates, hysteresis, persistence constraints, evidence-gated transition validation, deterministic regime derivation, and fail-open fallback behavior. Its output is a compact control state that can condition downstream arbitration and response modules while remaining inspectable, reversible, and operationally safe. We argue that dialogic regime control should be treated as a first-class systems problem in conversational retrieval pipelines, distinct from retrieval ranking, affect modeling, and surface-level style prompting. We describe the design of M6, formalize its validation contract, and propose an evaluation program for studying coherence preservation, fallback robustness, and downstream impact.

CCS Concepts: • **Information systems** → **Retrieval models and ranking**; • **Computing methodologies** → *Artificial intelligence*; *Natural language processing*.

Additional Key Words and Phrases: conversational retrieval, dialogue systems, retrieval-augmented generation, dialogue policy, interaction control, fail-open systems, memory-augmented LLMs

1 INTRODUCTION

Recent conversational AI systems increasingly rely on multi-stage architectures that combine short-term conversational context, longer-term trace memory, retrieval components, arbitration logic, and large language model generation. In such systems, response quality depends not only on the relevance of retrieved material or the capability of the final generator, but also on the stability of the system's interactional posture across turns. A system may retrieve the right evidence and still respond under an ill-suited regime: too direct when justification is weak, too expansive when stabilization is needed, or too cautious when the interaction warrants a clear position.

In practice, this control layer is often under-specified. Many systems handle it indirectly through prompt wording, ad hoc tone presets, or implicit conventions distributed across modules. This makes the interactional logic hard to inspect and hard to test. It also creates a structural problem for modular architectures: when posture control is left implicit, downstream behavior becomes difficult to attribute. Failures can be mistakenly assigned to retrieval, prompting, or generation, even when the actual issue lies in unstable turn-to-turn regime selection.

This paper proposes a different framing. We present M6, a dialogic regime controller embedded in a larger memory-augmented conversational pipeline. M6 is not a style engine and not an affect detector. Its purpose is to maintain temporal dialogic coherence across turns by producing a validated intermediate control state. The module reads a bounded slice of the present, considers the previously validated regime, checks anchored conversational evidence, and derives a closed regime label under explicit transition constraints.

Author's address: Tof Muck, Independent research draft, France.

The central claim of this paper is that dialogic posture should be represented as an auditable state variable rather than as a hidden side effect of prompting. By isolating this function, we obtain four benefits. First, regime change becomes measurable. Second, invalid or weakly justified transitions can be blocked or degraded. Third, posture control can be separated cleanly from retrieval ranking. Fourth, the system can fail open: when M6 cannot produce a valid control state, the broader pipeline falls back to a neutral path instead of propagating speculative behavior.

The contribution of this work is primarily architectural and methodological. We do not claim a new ranking model, a new affective classifier, or a new general dialogue policy learner. Rather, we describe how a bounded, evidence-aware controller can regulate interactional regime inside a modular conversational retrieval system.

2 RELATED WORK

2.1 Conversational information seeking and conversational search

Conversational information seeking (CIS) has emerged as a major IR area concerned with multi-turn, language-based interactions between users and information systems. The monograph by Zamani et al. [15] provides the broadest synthesis of this space, spanning conversational search, question answering, recommendation, mixed initiative, and evaluation methodology. Within the benchmark tradition, the TREC Conversational Assistance Track (CAsT) helped establish conversational retrieval as a context-sensitive passage ranking problem in which performance depends on handling dialogue history, coreference, and turn-to-turn coherence rather than isolated queries alone [2]. More recent datasets such as InSCIt move beyond simple answer retrieval by explicitly modeling mixed-initiative information-seeking interactions in which the agent may answer directly, clarify, or provide relevant background grounded in evidence [12].

In the LLM era, a substantial part of the conversational search literature has focused on better context understanding for retrieval. Recent work has shown that large language models can help infer contextual search intent [9], improve conversational query rewriting [14], and support hybrid conversational access to structured and unstructured corpora [7]. Other work adapts LLMs directly for conversational dense retrieval, showing that retrieval models themselves can benefit from session-aware instruction tuning and contrastive objectives [8].

M6 is adjacent to this line of work, but addresses a different problem. It does not attempt to improve query reformulation, turn resolution, or dense retrieval quality. Instead, it assumes that conversational retrieval pipelines need an explicit control layer that regulates the interactional regime under which retrieval, arbitration, and response generation operate.

2.2 Retrieval-augmented generation and adaptive retrieval control

The introduction of retrieval-augmented generation (RAG) reframed knowledge-intensive NLP as the combination of parametric generation and non-parametric memory access [6]. Since then, many works have studied when retrieval should be invoked, how retrieved evidence should be used, and how generation can remain grounded in retrieved material.

Several recent papers are especially relevant to M6's positioning. Active Retrieval Augmented Generation (FLARE) allows retrieval decisions to unfold dynamically during generation rather than only once at input time [4]. Self-RAG goes further by coupling adaptive retrieval with critique and self-reflection tokens, making retrieval usage itself part of the model's controllable inference process [1]. At the same time, groundedness studies show that retrieval does not automatically guarantee faithful use of evidence, especially in long-form generation [10]. Complementing this, Xie et al. [13] demonstrate that search-augmented LLMs can over-search, invoking retrieval unnecessarily and thereby harming efficiency, abstention, and sometimes factuality, especially in multi-turn settings.

M6 differs from these approaches in both scope and mechanism. It is not a retrieve-on-demand generator, not a self-critique decoder, and not a grounding verifier. Its role is upstream and orthogonal: it computes a bounded control state that can influence downstream modules while remaining separate from retrieval ranking and generation internals.

2.3 Dialogue state, policy, and controllable interaction

There is also a clear relationship between M6 and the dialogue systems literature, especially work on dialogue state tracking and controllable dialogue policy. Dialogue state tracking aims to estimate the current belief state of a conversation from its interaction history, often in service-oriented or task-oriented settings. Representative approaches include reading-comprehension-based state tracking [3] and schema-driven prompting with language models [5]. In parallel, controllable dialogue work has explored explicit or latent dialogue acts as intermediate policy variables for shaping system behavior, for example through latent act models such as DiactTOD [11].

M6 is related to these traditions, but it is not equivalent to either of them. It does not track slot-value belief states, user goals, or database constraints, and it is not a general end-to-end dialogue manager. Nor does it directly emit surface dialogue acts for response realization. Instead, it maintains a narrower intermediate state concerned with interactional regime: justification demand, admissible assertiveness, stabilization pressure, and related temporal control signals.

3 PROBLEM SETTING AND SCOPE

We study the following systems problem: how can a memory-augmented conversational architecture maintain coherent interactional behavior across turns without introducing opaque style heuristics, retrieval contamination, or ungrounded psychological inference?

This problem appears when several conditions hold simultaneously: the system is multi-turn rather than single-shot; retrieval and generation are separated into distinct modules; memory can be injected or withheld dynamically; downstream modules require stable control signals; and uncontrolled posture shifts can damage interpretability or trust.

We define *dialogic regime control* as the problem of selecting and stabilizing the interactional conditions under which the next turn should be processed or produced. In our setting, this includes regulating how much explicit justification is required, how assertive the system may be, and whether the next turn should preserve neutrality, adopt caution, stabilize tension, or assume a more directive but justified stance.

M6 is intentionally narrow in scope. It is not a retrieval ranker, not a user psychology model, and not a free-form style engine. Its purpose is to produce an auditable intermediate control signal while preserving architectural separation between retrieval, arbitration, and final generation.

4 METHOD

This section formalizes M6 as a bounded state-transition controller. The goal is not to learn a latent user model, but to compute a validated dialogic regime state from local conversational evidence under explicit operational constraints.

4.1 Turn-level inputs and state

For each turn t, let W_t be the bounded present window, N_t the canonical "Now" representation, x_{t-1} the previous validated threshold state when available, P the active threshold profile, and E_t the set of evidence candidates anchored in conversation traces.

M6 maintains a threshold vector

$$x_t \in [0,1]^{10}, \tag{1}$$

whose coordinates correspond to ten operational dimensions: conflictuality, admissible assertiveness, disclosure, interpretive suspension, conceptual density, inferential speed, explicitation requirement, regime stability, exploratory openness, and justification demand.

Let $A(P) \subseteq \{1,\ldots,10\}$ denote the active thresholds for profile P. In the release profile, all ten thresholds are active. At runtime, the model first produces a raw proposal

$$\hat{x}_t \in [0,1]^{10}, \tag{2}$$

but this proposal is never accepted directly.

4.2 Bounded state update

For each active threshold $i \in A(P)$, M6 computes a bounded update from the previous validated value $x_{t-1,i}$ to the proposed value $\hat{x}_{t,i}$. Let

$$\Delta_i = \hat{x}_{t,i} - x_{t-1,i}, \tag{3}$$

and let $u_i > 0$ and $d_i > 0$ denote the maximum upward and downward movements permitted for threshold i. M6 first clips the proposed change:

$$\Delta_i' = \min\left(u_i, \max\left(-d_i, \Delta_i\right)\right), \tag{4}$$

and forms the intermediate value

$$z_i = x_{t-1,i} + \Delta_i'. \tag{5}$$

The result is then quantized to two decimal places:

$$q(z_i) = \frac{\mathrm{round}(100z_i)}{100}. \tag{6}$$

Finally, a hold rule removes low-amplitude jitter. Let $\epsilon_i(P)$ be the hold threshold for coordinate i under profile P. Then

$$x_{t,i} = \begin{cases} x_{t-1,i} & \text{if } |q(z_i) - x_{t-1,i}| < \epsilon_i(P), \\ q(z_i) & \text{otherwise.} \end{cases} \tag{7}$$

For inactive thresholds $i \notin A(P)$, M6 leaves the previous value unchanged.

4.3 Transition status and stability

For each active threshold, M6 computes a transition status

$$\tau_{t,i} \in \{\mathrm{OK}, \mathrm{CHALLENGE}, \mathrm{BLOCK}\}. \tag{8}$$

The base rule is simple: OK if the proposed motion is admitted without clipping, Challenge if the raw proposal exceeds the allowed delta and must be clipped, and Block if a later validation gate rejects the transition.

To summarize global movement, M6 computes a stability index from normalized threshold motion. For each active coordinate, let

$$m_i = \frac{|x_{t,i} - x_{t-1,i}|}{a_i}, \tag{9}$$

where $a_i = u_i$ for upward motion and $a_i = d_i$ for downward motion, clipped to at most 1.0. The stability index is then

$$\sigma_t = q\left(1 - \frac{1}{|A(P)|} \sum_{i \in A(P)} m_i\right). \tag{10}$$

4.4 Deterministic regime derivation

M6 maps the validated threshold vector to a closed regime set

$$R = \{\text{stabilisation}, \text{tension\_controle}, \text{prudent\_explicatif}, \text{directif\_justifie}, \text{neutre\_explicatif}\}. \tag{11}$$

Let $g_P(x_t)$ denote the profile-specific deterministic regime mapping. In the release profile, the first-match priority order is:

(1) STABILISATION if $S8 < 0.25$;
(2) TENSION_CONTROLE if $S1 \geq 0.65 \wedge S7 \geq 0.60 \wedge S10 \geq 0.60$;
(3) DIRECTIF_JUSTIFIE if $S1 \leq 0.40 \wedge S2 \geq 0.60 \wedge S10 \leq 0.60$;
(4) PRUDENT_EXPLICATIF if one of the prudence rules is activated;
(5) NEUTRE_EXPLICATIF otherwise.

Under the coupled release profile, prudence can be triggered either by standalone threshold conditions or by coupled conditions, including $S7 \geq 0.65$, $S10 \geq 0.70$, $S3 \leq 0.30$, $S4 \geq 0.70$, $S5 \geq 0.70$, $S6 \geq 0.70$, $(S5 \geq 0.62 \wedge S6 \geq 0.62)$, and $(S3 \leq 0.35 \wedge S9 \leq 0.35)$.

4.5 Hysteresis and evidence-gated switching

M6 applies hysteresis to avoid excessive regime switching. Let r_{t-1} be the previous effective regime and let

$$r_t^* = g_P(x_t, r_{t-1}) \tag{12}$$

be the candidate regime after applying profile-specific hysteresis rules.

Each evidence item has the form $e = (id, type, anchor)$ and must resolve to a trace in the active conversation. Let $\text{Strong}(E_t, age_t)$ denote the predicate that at least one validated evidence reference is strong enough to authorize a regime change at age age_t. Immediate-strong evidence is sufficient on its own, whereas assisted-strong evidence becomes sufficient only if $age_t \geq 2$.

Let $n_{\min}$ be the minimum persistence requirement as a function of regime stability and, in the coupled profile, of the stability index. The switch gate is then

$$\text{AllowSwitch}_t = [r_t^* = r_{t-1}] \vee [age_t \geq n_{\min}] \vee [\text{Strong}(E_t, age_t)]. \tag{13}$$

The effective regime becomes

$$r_t = \begin{cases} r_t^* & \text{if AllowSwitch}_t, \\ r_{t-1} & \text{otherwise.} \end{cases} \tag{14}$$

4.6 Downstream control signal

The final M6 state is not free-form text but a structured control object

$$z_t = (\hat{x}_t, x_t, \tau_t, r_t^*, r_t, \sigma_t, E_t', d_t), \tag{15}$$

where E_t' is the validated evidence subset and d_t the downstream directive bundle. The directive bundle can be written as

$$d_t = h(r_t), \tag{16}$$

where h is a deterministic regime-to-directives mapping. In the current release, the strongest downstream effect is on the arbitration surface, where M6 controls quantities such as assertiveness caps and explicit-reasoning requirements.

5 IMPLEMENTATION AND REPRODUCIBILITY

The current M6 implementation is designed as a production-adjacent systems component rather than a notebook-only experimental prototype. The runtime path is deliberately narrow: the module receives a bounded present window, a previous validated state when available, and a small set of evidence candidates derived from trace memory. The default implementation uses a provider abstraction for structured model calls, a deterministic post-processing layer, and a local persistence store for validated state snapshots and failure traces.

Three implementation choices are important for reproducibility. First, the module is intentionally configured for low-variance output generation: the release runtime uses deterministic decoding settings, bounded token budget, short timeout, and low retry count. Second, every accepted or rejected turn produces structured telemetry, including thresholds, transition statuses, evidence identifiers, effective regime, fallback source, and failure code when relevant. Third, the controller's local persistence layer stores both valid states and transition failures, allowing exact replay of acceptance and fail-open behavior at the conversation level.

From an experimental standpoint, this means that M6 can be evaluated in at least two ways: *online* as a live control module in a running architecture, or *offline* by replaying logged conversations through the same validation and transition logic. The latter is particularly useful for paired-condition experiments because it keeps retrieval, arbitration, and generation inputs fixed while varying only the regime-control condition.

6 VALIDATION AND FAIL-OPEN

M6 uses a two-stage acceptance procedure. The first stage validates whether a raw model output is structurally admissible. The second determines whether the resulting state may be persisted and exposed downstream, or whether the pipeline must degrade to a neutral fallback.

6.1 Structured output contract

Let y_t denote the raw model output produced at turn t. M6 requires y_t to satisfy a strict JSON contract with a fixed schema identifier, a fixed set of root keys, exact turn metadata consistency, a threshold profile identifier matching the active runtime policy, finite threshold values in $[0, 1]$, and evidence references resolvable against trace memory.

We define a structural decoder

$$D(y_t) \to (\hat{x}_t, E_t, meta_t) \quad \text{or} \quad \perp, \tag{17}$$

where $\perp$ denotes parse failure. Structural failure occurs when the output is not valid JSON, unknown root keys are present, mandatory fields are missing, field types are invalid, or schema versioning constraints are violated.

6.2 Header, evidence, and persistability

Let

$$meta_t = (turn\_id, conversation\_id, now\_seq, profile\_id) \tag{18}$$

and let

$$ctx_t = (turn_t, conv_t, seq_t, P) \tag{19}$$

be the runtime context. Header validation succeeds iff $turn\_id = turn_t$, $conversation\_id = conv_t$, $now\_seq = seq_t$, and $profile\_id = P$.

Evidence validation further requires that every proposed evidence identifier is listed, every evidence type belongs to the closed vocabulary, every anchor points to the active conversation, every evidence identifier matches its anchor deterministically, and every anchor resolves in the conversation store. Let $E'_t \subseteq E_t$ be the subset that satisfies all these constraints.

Let I_t be the set of validation issues generated at turn t. M6 distinguishes hard-invalid issues, which make the candidate state non-persistable, from guardrail issues, which allow persistence but mark the output as degraded. Formally,

$$\mathrm{Persistable}(I_t) \in \{\mathrm{true}, \mathrm{false}\}. \tag{20}$$

Then $\mathrm{Persistable}(I_t)$ is true if I_t is empty or if every issue is guardrail-persistable, and false otherwise.

6.3 Acceptance and failure taxonomy

Let

$$V(y_t, ctx_t, x_{t-1}) = (z_t, I_t) \tag{21}$$

be the full validator. M6 persists z_t iff $\mathrm{Persistable}(I_t)$ is true; otherwise the state is rejected and a validation failure is recorded.

Operationally, M6 distinguishes parse failure, validation failure, evidence invalidity or unresolvability, persistence failure, runtime failure, and circuit-open bypass. Let ϕ_t be the failure code emitted at turn t, with $\phi_t = \varnothing$ when no failure occurs. This yields a turn-level failure trace that can be aggregated into fail-open rate, parse-failure rate, validation-failure rate, circuit-open ratio, and evidence-rejection statistics.

6.4 Fail-open state machine

M6 uses a local circuit state

$$c_t \in \{\mathrm{closed}, \mathrm{half\_open}, \mathrm{open}\}. \tag{22}$$

Let f_t be the number of consecutive failures associated with the active conversation. The release policy uses an open threshold $K = 3$, a half-open retry interval $T_{\mathrm{retry}} = 60$ seconds, and a release alert threshold $\rho_{\max} = 0.20$ for the circuit-open ratio.

The state machine is:

(1) If $c_{t-1} = \mathrm{closed}$ and evaluation succeeds, keep $c_t = \mathrm{closed}$.
(2) If failures accumulate until $f_t \geq K$, transition to $c_t = \mathrm{open}$.
(3) If $c_{t-1} = \mathrm{open}$, skip live evaluation unless at least T_{retry} seconds have elapsed.
(4) If the retry window has elapsed, transition temporarily to $c_t = \mathrm{half\_open}$ and allow one fresh evaluation.
(5) If the half-open evaluation succeeds, reset failures and return to $c_t = \mathrm{closed}$.

(6) If the half-open evaluation fails, return to c_t = open.

When live evaluation fails or is bypassed by an open circuit, the pipeline continues with a fallback state instead of aborting the full turn. The downstream directive bundle is neutralized for the current turn, and the broader architecture continues without a fresh regime intervention. This is the core fail-open principle of M6: the module may abstain from control, but it may not force the rest of the system into undefined behavior.

7 EXPERIMENTAL DESIGN

This section defines an evaluation protocol for testing whether M6 improves conversational control without introducing unacceptable operational cost or unjustified regime changes.

7.1 Research questions and conditions

We propose four research questions:

- **RQ1:** Does M6 reduce unstable regime switching in multi-turn interaction?
- **RQ2:** Does evidence-gated validation improve justificatory discipline?
- **RQ3:** Does M6 improve downstream conversational behavior?
- **RQ4:** Is M6 operationally safe enough for online use?

We recommend evaluating at least three primary conditions:

- **B0** No-M6 baseline: the pipeline runs without M6 and downstream modules receive no regime control signal.
- **B1** Neutral-control baseline: the pipeline includes the M6 interface, but all downstream directives are neutral.
- **M6** Active controller: full M6 validation, regime derivation, and fail-open policy are enabled.

Two ablation conditions are also useful: **M6-no-evidence-gate** and **M6-no-hysteresis**.

7.2 Evaluation corpora and units of analysis

The evaluation should combine three dataset types. First, replayable internal conversation traces sampled from the target system's real usage logs after redaction and filtering. Second, a curated stress-test suite containing synthetic or hand-selected conversations designed to trigger abrupt tone escalation, weak evidence with large threshold jumps, repeated malformed outputs, rapid oscillation between prudence and directiveness, and long neutral stretches with low signal. Third, a smaller annotation set for downstream human evaluation, where multiple system outputs can be compared side by side under controlled prompts and shared conversation history.

We recommend two complementary units of analysis: the *turn* as the atomic unit for threshold updates, validation outcomes, and latency, and the *conversation* as the unit for regime stability, fallback behavior, and end-to-end interaction quality.

7.3 Annotation protocol

Human annotation is needed for at least two constructs that cannot be reduced to runtime logs alone: regime appropriateness for the local turn, and downstream response coherence and justified stance.

For each annotated turn or conversation segment, annotators should receive the conversation excerpt, the system condition identifier hidden or randomized, the resulting response, and, for expert analysis, optionally the M6 regime and evidence summary. Recommended annotation axes include coherence with the recent conversational scene, adequacy of caution versus directness, justification quality, degree of over-interpretation, and need for stabilization.

7.4 Metrics

We divide metrics into four families.

State-dynamics metrics. Regime flap rate, average threshold movement per turn, proportion of Block, Challenge, and Ok transitions, and average stability index.

Evidence-discipline metrics. Proportion of challenged transitions lacking strong evidence, evidence rejection rate, anchor-resolution failure rate, and precision of strong-evidence authorization on annotated cases.

Operational metrics. Parse failure rate, validation failure rate, persistence failure rate, fail-open rate, circuit-open ratio, and mean and tail latency added by M6.

Downstream quality metrics. Expert-rated coherence, expert-rated justified directiveness, rate of over-interpretive responses, pairwise preference between conditions, and scenario-specific response adequacy.

7.5 Procedure and analysis

For replay evaluation, each conversation is processed under all compared conditions while keeping every non-M6 component fixed. This yields paired observations at both turn and conversation levels. For downstream response comparison, the same prompts and conversation histories should be routed through the same architecture with only the M6 condition varied. Where stochastic generation cannot be fully eliminated, fixed seeds or repeated sampling should be used, and the analysis should report within-condition variance.

For stress testing, each failure scenario should be run repeatedly in order to estimate whether the failure is caught by validation, whether the fail-open path activates correctly, whether the circuit transitions as intended, and whether recovery occurs after the retry interval.

For turn-level paired comparisons, non-parametric paired tests or bootstrap confidence intervals are appropriate when distributional assumptions are weak. For conversation-level human ratings, mixed-effects models are preferable if multiple annotators and repeated conversations are involved. Primary reported outcomes should include effect sizes and confidence intervals, not only significance values.

7.6 Error analysis and deployment criteria

Quantitative results should be complemented by targeted error analysis, including false-positive prudence, unjustified directiveness that slipped past validation, unnecessary blocking under weak but legitimate transition cues, repeated fail-open cascades, and mismatches between internal regime stability and human judgments of coherence.

If M6 is evaluated for online use rather than only for publication, explicit acceptance criteria should be defined in advance. A reasonable deployment gate would require a practically meaningful reduction in regime flap rate, no substantial degradation in downstream user-rated quality, low persistent circuit-open ratio, bounded tail latency, and a fail-open path that preserves system usability even under repeated local M6 failures.

8 DISCUSSION

The experimental contribution of M6 is less about raw model capability than about architectural discipline. The core question is whether conversational retrieval systems benefit from an explicit, validated control layer between local context and downstream decision surfaces. If the answer is yes, then regime control should be studied as its own systems problem, rather than treated as an accidental by-product of prompt phrasing or generation-time sampling.

M6 also creates a useful separation of concerns for empirical analysis. Because it does not directly rewrite retrieval ranking, performance changes can be attributed more cleanly than in monolithic systems where retrieval, policy, and generation effects are entangled. If active M6 improves coherence while leaving retrieval ranking untouched, that would support the claim that interactional control is not reducible to better retrieval alone. Conversely, if the controller over-blocks, over-stabilizes, or falls open too frequently, that would reveal concrete limits of explicit regime governance.

Finally, M6 invites a more operational interpretation of "dialogue quality." Instead of reducing quality to answer relevance or fluency, it foregrounds whether a system can justify its posture, avoid ungrounded shifts, and recover safely from local control failure. These are empirical questions, and they are well suited to the kinds of paired, replay-based, and human-evaluated experiments defined in this paper.

9 LIMITATIONS

This work has several limitations. First, M6 is a modular control proposal implemented within one specific architecture. Its generality remains to be demonstrated across substantially different systems. Second, the strongest present operational effect appears in downstream arbitration, while the bridge from internal regime semantics to final response shaping remains narrower than the conceptual design would ideally support. Third, the threshold vocabulary is hand-crafted rather than learned end-to-end. This is an advantage for auditability but may reduce adaptability. Fourth, the evidence taxonomy is deliberately conservative. It may under-authorize legitimate shifts in cases where relevant conversational cues are not captured by the trace format.

Finally, the present draft describes a design and implementation rationale more strongly than a completed empirical program. Any submission claiming broad effectiveness would require the experiments outlined above.

10 CONCLUSION

M6 proposes a different way of thinking about control in conversational retrieval systems. Instead of treating dialogic posture as a hidden property of prompts or a vague stylistic overlay, it treats regime selection as a bounded and inspectable systems problem. The module combines local state, evidence-aware validation, deterministic regime derivation, and fail-open behavior to regulate interaction across turns without contaminating retrieval ranking or overclaiming psychological access.

The broader significance of this work is methodological. If conversational systems are to become more reliable, it may not be enough to improve retrieval or generation in isolation. We may also need explicit control modules that govern how those capabilities are deployed in context, under what regime, and with what safety guarantees.

ACKNOWLEDGMENTS

This manuscript was developed with assistance from generative AI tools used for drafting support, reformulation, structural editing, and exploratory synthesis. All architectural decisions, technical claims, validation choices, and final manuscript revisions were reviewed and approved by the human author, who assumes full responsibility for the content of this work.

REFERENCES

[1] Akari Asai, Zeqiu Wu, Yizhong Wang, Avirup Sil, and Hannaneh Hajishirzi. 2024. Self-RAG: Learning to Retrieve, Generate, and Critique through Self-Reflection. In *The Twelfth International Conference on Learning Representations (ICLR 2024)*. https://openreview.net/forum?id=hSyW5go0v8 OpenReview poster.

[2] Jeffrey Dalton, Chenyan Xiong, and Jamie Callan. 2019. CAsT 2019: The Conversational Assistance Track Overview. In *The Twenty-Eighth Text REtrieval Conference (TREC 2019) Proceedings*. https://trec.nist.gov/pubs/trec28/papers/OVERVIEW.CAsT.pdf

[3] Shuyang Gao, Abhishek Sethi, Sanchit Agarwal, Tagyoung Chung, and Dilek Hakkani-Tür. 2019. Dialog State Tracking: A Neural Reading Comprehension Approach. In *Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue*. Association for Computational Linguistics, Stockholm, Sweden, 264–273. https://doi.org/10.18653/v1/W19-5932

[4] Zhengbao Jiang, Frank F. Xu, Luyu Gao, Zhiqing Sun, Qian Liu, Jane Dwivedi-Yu, Yiming Yang, Jamie Callan, and Graham Neubig. 2023. Active Retrieval Augmented Generation. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Singapore, 7969–7992. https://doi.org/10.18653/v1/2023.emnlp-main.495

[5] Chia-Hsuan Lee, Hao Cheng, and Mari Ostendorf. 2021. Dialogue State Tracking with a Language Model using Schema-Driven Prompting. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Online and Punta Cana, Dominican Republic, 4937–4949. https://doi.org/10.18653/v1/2021.emnlp-main.404

[6] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. 2020. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. In *Advances in Neural Information Processing Systems*, Vol. 33. Curran Associates, Inc., 9459–9474. https://papers.nips.cc/paper_files/paper/2020/hash/6b493230205f780e1bc26945df7481e5-Abstract.html

[7] Shicheng Liu, Jialiang Xu, Wesley Tjangnaka, Sina Semnani, Chen Yu, and Monica Lam. 2024. SUQL: Conversational Search over Structured and Unstructured Data with Large Language Models. In *Findings of the Association for Computational Linguistics: NAACL 2024*. Association for Computational Linguistics, Mexico City, Mexico, 4535–4555. https://doi.org/10.18653/v1/2024.findings-naacl.283

[8] Kelong Mao, Chenlong Deng, Haonan Chen, Fengran Mo, Zheng Liu, Tetsuya Sakai, and Zhicheng Dou. 2024. ChatRetriever: Adapting Large Language Models for Generalized and Robust Conversational Dense Retrieval. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Miami, Florida, USA, 1227–1240. https://doi.org/10.18653/v1/2024.emnlp-main.71

[9] Kelong Mao, Zhicheng Dou, Fengran Mo, Jiewen Hou, Haonan Chen, and Hongjin Qian. 2023. Large Language Models Know Your Contextual Search Intent: A Prompting Framework for Conversational Search. In *Findings of the Association for Computational Linguistics: EMNLP 2023*. Association for Computational Linguistics, Singapore, 1211–1225. https://doi.org/10.18653/v1/2023.findings-emnlp.86

[10] Alessandro Stolfo. 2024. Groundedness in Retrieval-augmented Long-form Generation: An Empirical Study. In *Findings of the Association for Computational Linguistics: NAACL 2024*. Association for Computational Linguistics, Mexico City, Mexico, 1537–1552. https://doi.org/10.18653/v1/2024.findings-naacl.100

[11] Qingyang Wu, James Gung, Raphael Shu, and Yi Zhang. 2023. DiactTOD: Learning Generalizable Latent Dialogue Acts for Controllable Task-Oriented Dialogue Systems. In *Proceedings of the 24th Annual Meeting of the Special Interest Group on Discourse and Dialogue*. Association for Computational Linguistics, Prague, Czechia, 255–267. https://doi.org/10.18653/v1/2023.sigdial-1.24

[12] Zeqiu Wu, Ryu Parish, Hao Cheng, Sewon Min, Prithviraj Ammanabrolu, Mari Ostendorf, and Hannaneh Hajishirzi. 2023. InSCIt: Information-Seeking Conversations with Mixed-Initiative Interactions. *Transactions of the Association for Computational Linguistics* 11 (2023), 453–468. https://doi.org/10.1162/tacl_a_00559

[13] Roy Xie, Deepak Gopinath, David Qiu, Dong Lin, Haitian Sun, Saloni Potdar, and Bhuwan Dhingra. 2026. Over-Searching in Search-Augmented Large Language Models. In *Proceedings of the 19th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Rabat, Morocco, 7714–7739. https://doi.org/10.18653/v1/2026.eacl-long.361

[14] Fanghua Ye, Meng Fang, Shenghui Li, and Emine Yilmaz. 2023. Enhancing Conversational Search: Large Language Model-Aided Informative Query Rewriting. In *Findings of the Association for Computational Linguistics: EMNLP 2023*. Association for Computational Linguistics, Singapore, 5985–6006. https://doi.org/10.18653/v1/2023.findings-emnlp.398

[15] Hamed Zamani, Johanne R. Trippas, Jeff Dalton, and Filip Radlinski. 2023. Conversational Information Seeking. *Foundations and Trends® in Information Retrieval* 17, 3-4 (2023), 244–456. https://doi.org/10.1561/1500000081