

# Beyond Late Fusion: Reciprocal Critique and Revocable Decoding for Multimodal Generative Recommendation

Anonymous Author(s)

## Abstract

Multimodal generative recommenders learn textual and visual semantic identifiers (SIDs) with cross-conditioned objectives, yet inference typically decodes modality-specific candidates and merges completed lists by late fusion; the opposite view therefore cannot challenge an individual decision. We characterize this cross-modal train-inference utilization gap and find that nearly half of the misses remaining after direct-route fusion fall outside the union of four width-20 direct and reciprocal beams, beyond the reach of fixed-set reranking. We introduce **RECAP** (Reciprocal Evidence Critique with Actionable Prefix Repair), a 6.1M-parameter supervised correction plug-in for frozen backbones. RECAP uses cross-conditioned beliefs to localize the first likely divergence, abstains without a credible legal alternative, and otherwise revokes the dependent suffix, repairs it under refreshed evidence, and returns directly rescored proposals to the original beam budget before unchanged fusion. Across MQL4GRec, MACRec, and SynGR on three Amazon categories, reciprocal critique improves first-error localization over self critique in all nine settings. RECAP yields consistent 4.4%–5.7% relative NDCG@10 gains while retaining over 98.9% of existing Top-10 hits. Under measured cost, it also outperforms latency-matched wider native search, within-budget process-guided search, and a more expensive train-matched reranker.

**Anonymous code:** <https://github.com/recap-wsdm27-review/recap>

## CCS Concepts

• **Information systems** → **Recommender systems**; • **Computing methodologies** → *Learning latent representations*.

## Keywords

generative recommendation, multimodal recommendation, semantic identifiers, corrective decoding, constrained decoding

## 1 Introduction

Generative recommendation (GR) formulates retrieval as conditional sequence generation. Instead of scoring every item independently, the model emits a short semantic identifier (SID) that resolves to a catalog item [36]. Hierarchical SIDs are particularly attractive: coarse codes select a semantic region and later codes progressively identify a leaf. This structure combines compact output vocabularies with constrained full-catalog decoding, and has become a central interface for modern generative recommenders.

Multimodal GR extends this idea by representing each item in textual and visual SID spaces. Text captures attributes, function, and language-level semantics; images expose appearance, shape, and visual style. Several recent models consequently learn much more than two independent next-item tasks. MQL4GRec trains cross-modal generation and alignment [54]; MACRec adds reciprocal item- and sequence-level objectives around cross-modal quantization [55]; and SynGR explicitly regularizes cross-modal synergy [5].

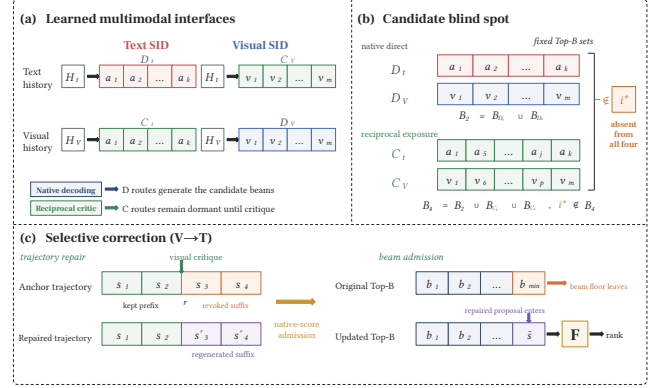


**Figure 1: Capability gap and repair. (a) Native inference uses learned direct routes but leaves reciprocal routes dormant as critics. (b) A target may lie outside all four fixed Top-B sets. (c) RECAP preserves the trusted prefix, repairs the suffix, and reinserts a native-scored proposal before unchanged fusion.**

These objectives establish that the opposite modality contains useful predictive structure.

In these models, however, inference remains markedly less interactive. A textual history directly generates textual candidates, a visual history directly generates visual candidates, and completed lists are combined by late fusion. Reciprocal prediction routes mainly serve as training supervision. They do not inspect a realized direct trajectory and say: *the first token is plausible, but the second token enters a branch that the other view strongly contradicts*. Fusion can reorder candidates already generated, but it cannot recover an item outside every supplied candidate list. This mismatch motivates our central question: *can reciprocal cross-modal knowledge act as a conservative inference-time critic that decides whether, where, and how a hierarchical SID trajectory should be revised?*

This problem differs from ordinary uncertainty estimation—direct confidence can remain high after an early wrong branch, and later codes are evaluated under the corrupted prefix—and from four-way ensembling, since any static union still inherits finite-beam pruning. The opportunity lies in using the other view as *structured counter-evidence* in the same output SID space, locating the earliest disputed level, and regenerating the suffix under a new legal prefix.

Figure 1 makes the gap and repair target explicit. Panel (a) distinguishes the learned direct routes  $D_T, D_V$ , which generate the native candidate beams, from the reciprocal routes  $C_T, C_V$ , which are cross-conditioned by the opposite view but remain dormant as critics during native decoding. Panel (b) then exposes the fixed-set limit: the two direct beams form  $B_2$ , explicit reciprocal queries enlarge this to  $B_4$ , yet a target outside  $B_4$  remains unreachable to any reranker over those candidates. Panel (c) shows how RECAP changes candidate membership rather than merely rescored it. In the  $V \rightarrow T$

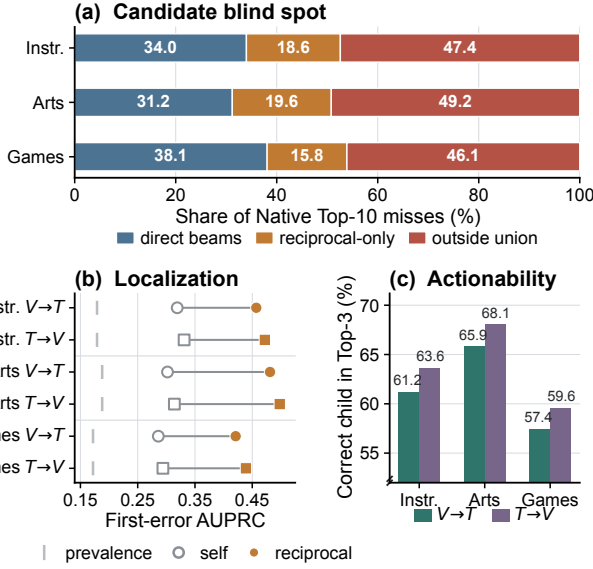


**Figure 2: SynGR motivation diagnostics. (a) Provenance of native misses; (b) direction-specific first-error AUPRC; (c) correct-child Top-3 actionability (y-axis starts at 52%). All beams use width 20.**

direction, reciprocal evidence critiques a realized textual trajectory, preserves its trusted prefix, revokes and regenerates the dependent suffix, and returns a directly scored proposal to the original Top- $B$  before unchanged fusion;  $T \rightarrow V$  is symmetric.

The resulting method, RECAP, is by design a lightweight supervised correction plug-in on a frozen backbone rather than a new multimodal encoder, SID tokenizer, or final reranker. This separation lets us ask a clean empirical question: does a capability already learned by the backbone become more useful when moved from parameter shaping to candidate-specific critique? Detailed SynGR diagnostics establish the candidate blind spot and repair opportunity; cross-backbone experiments then test whether the critic and its recommendation effect generalize beyond that diagnostic setting. Our evidence addresses four questions. First, a large fraction of fused misses remains outside even a four-way beam union. Second, reciprocal evidence detects first errors better than self-confidence. Third, it often ranks the correct repair direction highly and exposes a legal alternative subtree. Fourth, selective repair improves ranking without sacrificing the majority of already correct predictions.

Our contributions are threefold:

- We diagnose a *cross-modal train–inference utilization gap* in detail on SynGR and test critic reliability across three multimodal backbones, explicitly separating native cross-conditioned tasks from SynGR’s frozen wrapper.
- We introduce RECAP, a supervised frozen-backbone plug-in for cross-view critique, first-error localization, abstention, hierarchical revocation, and refreshed repair.
- We establish the mechanism through backbone-level gains, candidate-union and first-error diagnostics, simpler repair controls, preservation and harm accounting, and matched efficiency measurements.

## 2 Related Work

### 2.1 Sequential and Generative Recommendation

Sequential recommenders learn from ordered histories [19, 41]; generative recommendation instead retrieves by emitting textual or learned identifiers [9, 43, 45]. Tree retrieval and hierarchical SIDs organize this decision coarse-to-fine [36, 59, 60], while recent work improves identifier construction, collaborative alignment, tokenization, and multimodal fusion [16, 23, 35, 42, 44, 46, 52, 57]. RECAP assumes such a trained SID space and addresses a different failure: an early wrong branch confines every later decision to the wrong subtree of the identifier hierarchy.

### 2.2 Multimodal Generative Recommendation

Multimodal recommenders incorporate visual and textual content through representation learning, graph propagation, or self-supervised alignment [6, 24, 47, 48, 58]. Multimodal GR brings these signals into the generative interface: VIP5 conditions language-based recommendation on images [10]; MQL4GRec learns cross-modal generation tasks [54]; MACRec adds reciprocal item- and sequence-level generation [55]; and SynGR targets synergy through saliency-aware masking and regularization [5]. These methods mainly use cross-modal objectives to shape representations or completed candidate lists. RECAP instead evaluates direct and reciprocal beliefs over the same legal children of a realized SID trajectory, revises candidate membership, and reuses the native fusion rule unchanged.

### 2.3 Corrective Decoding and Selective Repair for Semantic IDs

Hierarchical SID generation introduces a structural dependency: once an early token selects the wrong branch, later tokens are generated under a corrupted prefix. Recent GR work therefore intervenes in decoding itself. PROMISE scores intermediate prefixes with a process reward model; GRC learns generation–reflection–correction trajectories; masked-diffusion GR relaxes a fixed autoregressive order; and HCGRec uses target-prefix hints during reinforcement-learning post-training [12, 32, 38, 51, 56]. More generally, beam search is not a neutral approximation: beam-aware training, stopping, rescoring, and search-error analyses show that finite beams may prune useful prefixes or prefer flawed completed sequences [17, 31, 39, 49, 53]. Constrained decoding instead preserves validity or specified conditions through grid allocation, dynamic beams, multimodal constraints, or logical lookahead [1, 15, 27, 28, 34].

Completed sequences also need not remain immutable. Deliberation, iterative refinement, insertion, and learned editing revise an initial trajectory [20, 29, 30, 40, 50]; diffusion language models replace one-pass commitment with repeated denoising [2, 14, 22, 26, 37]; and learned critics or process verifiers provide feedback for selection and revision [11, 21, 25]. Because confidence does not establish correctness, reject-option and selective-classification work evaluates whether to act through calibration, trust, and risk–coverage trade-offs [3, 4, 7, 8, 13, 18, 61].

Unlike self/process verification in PROMISE and GRC, RECAP audits a realized direct trajectory using the paired modality as same-output-space counter-evidence, while keeping the generator,

candidate budget, and fusion frozen. It reopens a legal subtree only when the first divergence is actionable; otherwise it abstains.

### 3 Problem Setup

#### 3.1 Multimodal Recommendation in Two SID Spaces

Let  $\mathcal{I}$  be an item catalog and  $i^* \in \mathcal{I}$  a user's next item. Each item has independently constructed textual and visual SIDs,

$$\mathbf{y}_i^T = (y_{i,1}^T, \dots, y_{i,L_T}^T), \quad \mathbf{y}_i^V = (y_{i,1}^V, \dots, y_{i,L_V}^V). \quad (1)$$

The two spaces may differ in depth, branching factor, and code semantics. We assume no position-wise alignment; their only necessary correspondence is item identity. Native disambiguation tokens are included in  $\mathbf{y}_i^m$ , so the SID-to-item mapping used for supervision is unique under each backbone's collision policy. For modality-specific histories  $H_T, H_V$ , the direct and reciprocal routes are

$$\begin{aligned} D_T &= p_\theta(\mathbf{y}^T | H_T), & D_V &= p_\theta(\mathbf{y}^V | H_V), \\ C_T &= q_{V \rightarrow T}(\mathbf{y}^T | H_V), & C_V &= q_{T \rightarrow V}(\mathbf{y}^V | H_T). \end{aligned} \quad (2)$$

The subscript denotes the output SID space. Thus  $D_T$  and  $C_T$  are comparable distributions over textual codes despite conditioning on different modality histories; likewise for  $D_V, C_V$ .

Native inference beam-decodes  $D_T, D_V$ , maps their SIDs to item candidates, and applies a fixed fusion rule  $F$ . We do not replace  $F$ , increase beam width, or add a final ranking score. Given a direct candidate, our task is instead to decide whether its hierarchical path is credible, locate the first likely error, and repair it only when reciprocal evidence identifies an actionable legal alternative.

#### 3.2 Legal-Set Normalization

For modality  $m \in \{T, V\}$ , let  $\mathcal{A}_m(\mathbf{y}_{<\ell})$  be legal children after prefix  $\mathbf{y}_{<\ell}$ . Comparing raw logits is misleading when illegal codes or output masks differ. We normalize both heads within the same legal set:

$$d_{m,\ell}(u) = \frac{\exp a_{m,\ell}^D(u)}{\sum_{v \in \mathcal{A}_m(\mathbf{y}_{<\ell})} \exp a_{m,\ell}^D(v)}, \quad (3)$$

$$c_{m,\ell}(u) = \frac{\exp a_{m,\ell}^C(u)}{\sum_{v \in \mathcal{A}_m(\mathbf{y}_{<\ell})} \exp a_{m,\ell}^C(v)}. \quad (4)$$

Both are evaluated under the direct candidate's current prefix. The reciprocal route therefore critiques a decision in the same output space; RECAP never compares a text token with a visual token.

#### 3.3 Static Candidate Sets and Revision Opportunity

Let  $\mathcal{B}_{D_T}$  and  $\mathcal{B}_{D_V}$  denote item sets produced by the two direct beams, and let  $\mathcal{B}_{C_T}, \mathcal{B}_{C_V}$  denote sets exposed by explicitly querying reciprocal routes with the same width. We define

$$\mathcal{B}_2 = \mathcal{B}_{D_T} \cup \mathcal{B}_{D_V}, \quad \mathcal{B}_4 = \mathcal{B}_2 \cup \mathcal{B}_{C_T} \cup \mathcal{B}_{C_V}. \quad (5)$$

For a target missed by native fused Top-10, membership in  $\mathcal{B}_2$  identifies a ranking/fusion residual; membership in  $\mathcal{B}_4 \setminus \mathcal{B}_2$  identifies an exposure opportunity available from dormant reciprocal routes.

No static reranker can recover the difficult cohort  $\mathcal{M}_{\text{blind}} = \{u : i_u^* \notin \mathcal{B}_4(u)\}$ ; revision must change the candidate set itself.

Suppose candidate  $\hat{\mathbf{y}}$  first diverges from the target at level  $r$ . Reopening from  $r$  defines the legal region

$$\mathcal{R}(\hat{\mathbf{y}}, r) = \{i \in \mathcal{I} : \mathbf{y}_{i,<r} = \hat{\mathbf{y}}_{<r}\}. \quad (6)$$

It is possible that  $i^* \in \mathcal{R}(\hat{\mathbf{y}}, r) \setminus \mathcal{B}_4$ : the trusted prefix is correct, but all finite beams chose the wrong child or pruned the correct continuation. This establishes the capability difference between revision and reranking. It does not imply that every item in  $\mathcal{M}_{\text{blind}}$  is repairable; localization, direction selection, and finite-budget search each introduce an observable loss that we measure separately.

## 4 Method

### 4.1 Overview

RECAP operates after a frozen multimodal generator has produced its native direct beams. It uses a reciprocal route in the same output SID space to critique the top direct trajectory, localizes the earliest actionable divergence, and reopens only the dependent suffix. A lightweight repairer then regenerates legal continuations under refreshed cross-modal evidence; accepted proposals are scored by the unchanged native function and reinserted before the original fusion. Training therefore learns only the locator and repairer from cached rollouts, while inference preserves the backbone, trie, beam budget, and final ranking protocol. Figure 3 summarizes this critique–revoke–repair path and its lightweight post-training interface.

### 4.2 Reciprocal Critique and Actionability

The base decoder first generates  $B$  candidates  $\{\hat{\mathbf{y}}_m^{(b)}\}_{b=1}^B$  in each SID space. RECAP designates the highest-scoring trajectory in each modality as the *revision anchor* and audits all its levels; the other  $B - 1$  hypotheses remain untouched. This avoids assigning every diverse beam hypothesis an independent binary correctness label. The evidence vector contains direct and reciprocal support for the committed token, top-two margins, normalized entropies, Jensen–Shannon disagreement, and shared-alternative agreement. A lightweight locator then maps the feature sequence to

$$\begin{aligned} \mathbf{x}_{m,\ell} &= f(d_{m,\ell}, c_{m,\ell}, \hat{\mathbf{y}}_{m,\ell}, \mathbf{y}_{m,<\ell}), \\ \boldsymbol{\pi}_m &= g_\phi(\mathbf{x}_{m,1:L_m}) \in \Delta^{L_m+1}, \end{aligned} \quad (7)$$

over  $\{1, \dots, L_m, \text{No-ERROR}\}$ . For a frozen rollout  $\hat{\mathbf{y}}$  and target  $\mathbf{y}^*$ , supervision is

$$z(\hat{\mathbf{y}}, \mathbf{y}^*) = \begin{cases} \min\{\ell : \hat{\mathbf{y}}_\ell \neq \mathbf{y}_\ell^*\}, & \hat{\mathbf{y}} \neq \mathbf{y}^*, \\ \text{No-ERROR}, & \text{otherwise.} \end{cases} \quad (8)$$

This mutually exclusive label respects hierarchy: beliefs after the first wrong ancestor are stale rather than independent negative examples. Before feature construction, we fit one temperature per route on validation NLL. The locator also receives rank and within-level quantile features, which are invariant to logit scale. Its two layers add learned depth and route-type embeddings to  $\mathbf{x}_{m,\ell}$ , mask padded levels, and score every level together with a learned No-ERROR query; parameters are shared across audit directions within a backbone. Here No-ERROR is a sequence-correctness supervision label, not the final KEEP decision; an erroneous but non-actionable

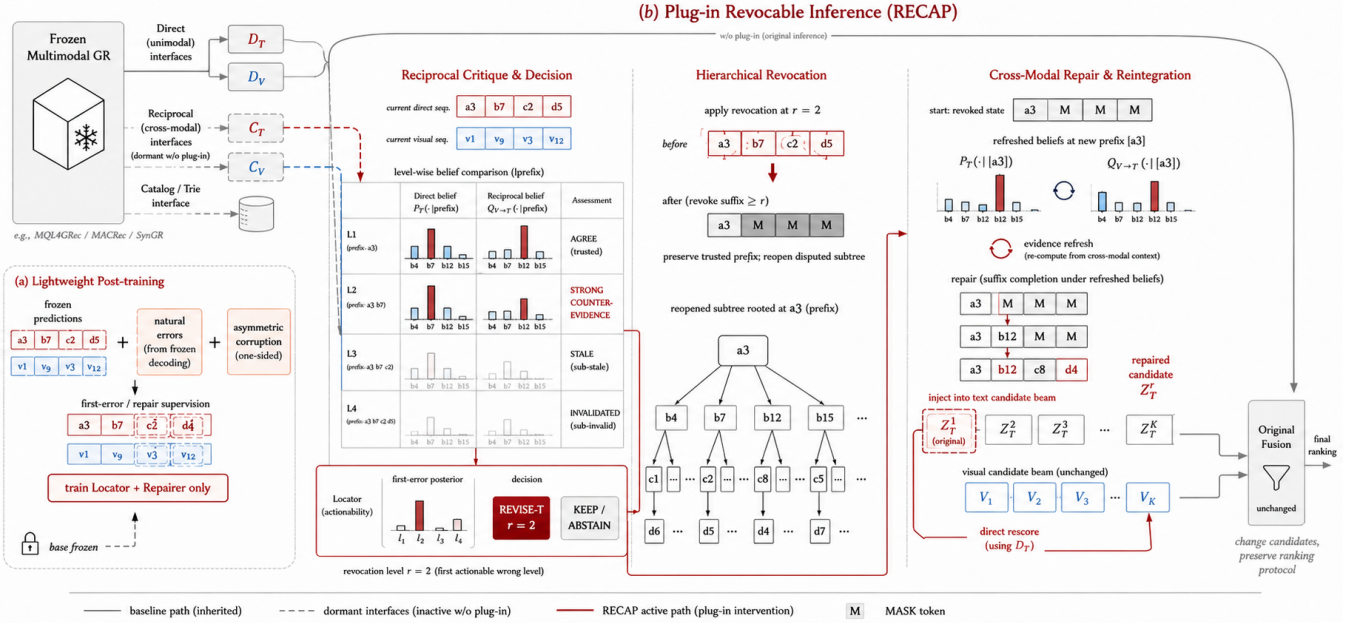


**Figure 3: RECAP training and inference pipeline. A frozen backbone exposes direct and reciprocal interfaces; lightweight post-training learns a first-error locator and suffix repairer. Actionable critique revokes the disputed suffix, refreshes prefix-conditioned evidence, and returns directly rescored proposals to the fixed-width beam and unchanged fusion.**

anchor is preserved by the separate gate below. We train on natural top-trajectory errors from the training split so that the locator observes the structured suffixes encountered at deployment. One-sided hierarchical corruption augments rare depths while keeping the opposite modality intact: after replacing a legal child at sampled depth, we draw all descendants from its new subtree. Natural and corrupted examples are mixed 3:1, with a natural-only ablation. Here and throughout the experiments, “first error” denotes the first SID divergence from the single held-out next item under offline evaluation; an alternative branch is not thereby proven semantically invalid.

Detectability does not imply repairability. A reciprocal route may disagree yet remain diffuse, or both routes may be wrong. For proposed level  $\ell$ , we form a tempered product score and its best-alternative margin:

$$s_{m,\ell}(u) = \lambda \log d_{m,\ell}(u) + (1 - \lambda) \log c_{m,\ell}(u),$$

$$A_{m,\ell} = \max_{u \in \mathcal{A}_m(\hat{y}_{<\ell}) \setminus \{\hat{y}_\ell\}} s_{m,\ell}(u) - s_{m,\ell}(\hat{y}_\ell). \quad (9)$$

Let  $\hat{z} = \arg \max_z \pi_m(z)$ . The plug-in abstains if  $\hat{z} = \text{No-ERROR}$ ,  $\pi_m(\hat{z}) < \tau_{\text{err}}$ ,  $A_{m,\hat{z}} < \tau_{\text{act}}$ , or normalized reciprocal entropy  $\bar{H}(c) = H(c) / \log |\mathcal{A}_m(\hat{y}_{<\hat{z}})|$  exceeds  $\tau_H$ ; a singleton legal set also forces abstention. Thresholds and  $\lambda$  are selected on validation and then frozen. Headline results use one shared default, and Table 4 separately reports the Always-revise stress test. The two components have distinct roles:  $g_\phi$  estimates where the path first failed, whereas  $A$  tests whether available evidence supplies a credible direction. The gate treats KEEP/ABSTAIN as a first-class decision: consensus enters only as soft evidence aggregation, and shared-alternative agreement serves only as a locator feature.

### 4.3 Revocation, Repair, and Reintegration

If the gate accepts  $\hat{z} = r$ , RECAP preserves  $\hat{y}_{<r}$  and revokes  $\hat{y}_{r:L_m}$ :

$$(\hat{y}_{<r}, \hat{y}_{r:L_m}) \rightarrow (\hat{y}_{<r}, [\text{MASK}]_{r:L_m}). \quad (10)$$

The entire suffix is removed because a residual code is meaningful only under its ancestors. Replacing token  $r$  while retaining descendants from the old branch can produce an illegal or semantically incoherent path. Revocation is consequently a catalog operation: retain evidence judged reliable, reopen the disputed subtree, and search only within its legal descendants.

After reopening the subtree, we study three repair operators under the same gate and beam budget. A parameter-free product-of-experts (PoE) chooses each legal suffix token using

$$p_{\text{PoE}}(u \mid y_{<k}) \propto d_{m,k}(u)^\lambda c_{m,k}(u)^{1-\lambda}. \quad (11)$$

A learned one-shot repairer predicts the revoked suffix as

$$p_\psi(y_{r:L} \mid y_{<r}, \mathbf{d}, \mathbf{c}, m) = \prod_{k=r}^L p_\psi(y_k \mid y_{<k}, d_{m,k}, c_{m,k}, m). \quad (12)$$

Each decoder input concatenates the preserved-prefix state, frozen direct and reciprocal route summaries, a depth embedding, and a modality embedding; a trie mask restricts every factor to  $\mathcal{A}_m(y_{<k})$ . The two-layer decoder is shared across audit directions within a backbone and is trained on the same 3:1 mixture of natural and one-sided corrupted suffixes. The full variant runs at most  $R = 2$  refinement rounds. Round one completes the masked suffix left-to-right. We then recompute the complete feature sequence under the proposal and rerun  $g_\phi$ . If it predicts a gated level  $r' \geq r$ , the next round reopens from  $r'$  and regenerates that sub-suffix; otherwise

refinement stops. It also stops when the proposal is unchanged or the budget is exhausted. The highest-direct-score legal proposal from these rounds is passed to reintegration; it need not outrank the audited anchor.

Evidence refresh is the reason to permit multiple rounds. If  $\hat{y}_2$  changes, then a previous belief  $c_3(\cdot \mid \hat{y}_1, \hat{y}_2, H_m)$  no longer evaluates the new prefix. Reusing it would contradict the critic’s own prefix-conditioned semantics. Iterative repair may be implemented by masked denoising, but RECAP does not require a diffusion schedule. Its empirical obligation is narrower: refreshed refinement should improve upon matched PoE and learned one-shot controls.

For an accepted repair, the proposal receives only the backbone’s original direct scoring function,  $S_m^{\text{base}}(\mathbf{y})$ , including native length normalization. We merge it with the untouched  $B$ -candidate direct beam, deduplicate by item ID, and retain the top  $B$  direct scores. Thus a repair must beat the beam floor, not the already top-ranked anchor. Critic confidence, locator probability, and actionability never enter ranking; invalid or low-scoring proposals leave the beam unchanged. Textual and visual lists then enter the original  $F$ . RECAP changes candidate membership while preserving both final slot count and ranking protocol.

#### 4.4 Training and End-to-End Inference

The backbone is frozen during plug-in training. We cache direct rollouts and legal belief summaries, then minimize

$$\mathcal{L}_{\text{RECAP}} = \mathcal{L}_{\text{loc}} + \eta \mathcal{L}_{\text{rep}}, \quad \mathcal{L}_{\text{loc}} = -\log \pi_m(z), \quad (13)$$

where  $\mathcal{L}_{\text{rep}}$  is suffix cross-entropy from the true first error, with  $\eta = 0$  for PoE. No test examples are used for threshold or variant selection.

The locator and learned repairer are new plug-in capacity trained with target-side first-error and suffix supervision; only the backbone and reciprocal routes are inherited. We isolate route evidence from added capacity with an identical self-only locator/repairer, shuffled reciprocal features, parameter-free PoE, and one-shot versus iterative controls.

Let  $\mathcal{F}$  be one batched model evaluation. Auditing adds reciprocal evaluation along at most  $L$  anchor prefixes per modality, batched in practice; the locator itself costs  $O(Lh)$ . If fraction  $\rho$  passes the gate, iterative repair adds at most  $O(\rho R\mathcal{F})$ . Final fusion and full-catalog lookup retain the backbone’s complexity. This makes selectivity computationally as well as statistically important.

For each SID space, native beam decoding first produces  $B$  candidates and its top trajectory becomes the revision anchor. Direct and reciprocal beliefs are evaluated on that anchor, the locator proposes the earliest disputed level, and the actionability gate either abstains or reopens the corresponding legal subtree. Accepted suffixes are repaired for at most  $R$  rounds with refreshed prefix-conditioned evidence. The best legal proposal under the backbone’s direct score is merged into the original beam, deduplicated, and truncated back to  $B$  before unchanged fusion. No target-side image, future interaction, oracle label, critic score, or additional final slot enters inference. MQL4GRec and MACRec reuse native cross-conditioned tasks; SynGR uses the frozen wrapper in Section 5.1.

**Table 1: Amazon 2018 datasets after shared 5-core filtering.**

| Dataset     | Users  | Items  | Interactions |
|-------------|--------|--------|--------------|
| Instruments | 17,112 | 6,250  | 136,226      |
| Arts        | 22,171 | 9,416  | 174,079      |
| Games       | 42,259 | 13,839 | 373,514      |

## 5 Experiments

### 5.1 Experimental Setup

We evaluate four aspects of RECAP: reciprocal-critic actionability, effectiveness across frozen backbones and datasets, comparison with measured-cost alternatives, and the mechanisms governing selective revision.

*Data and metrics.* We use the Amazon 2018 data [33] under the shared 5-core, leave-one-out, full-catalog protocol [5, 54, 55]. Table 1 reports the resulting scale. We measure Hit Rate (H) and NDCG (N) at 5 and 10 with each backbone’s collision handling and constrained decoder.

*Backbones and training.* We instantiate RECAP on MQL4GRec, MACRec, and SynGR. MQL4GRec and MACRec expose bidirectional cross-conditioned tasks directly. For SynGR, we freeze its joint model, split the SID segments, and query its published view-conditioned segment objectives under teacher forcing on the *realized candidate prefix*, never the benchmark target prefix. Direct and reciprocal queries are normalized over the same legal child set; target SIDs supervise training labels only and are unavailable to test-time queries. The wrapper supplies read-only, same-output-space beliefs: it changes neither the legal candidates nor the backbone’s direct score or fusion, and its computation is included in measured end-to-end latency. A two-layer locator and repair decoder add 6.1M parameters. They are trained for eight epochs on cached training-split rollouts with AdamW (learning rate  $2 \times 10^{-4}$ , batch 512); validation selects gates, evidence mixing, stopping, and all comparison-model hyperparameters. The natural/corrupted training ratio is 3:1 unless stated otherwise.

*Fair comparisons and statistics.* All methods use identical histories, SIDs, tries, checkpoints, final slot counts, and fusion. Four-way ensembling calibrates route scores on validation. Its 6.3M-parameter reranker is trained on the same training-split rollouts as RECAP, uses only scores and presence bits from the fixed four-beam union, and uses validation only for tuning and early stopping. Our PROMISE-style control trains a self/process prefix verifier on those same rollouts and guides legal trie search without reciprocal features; its width is selected under the same measured latency budget. Wider native beams use the frozen grid  $\mathcal{G} = \{20, 30, 40, 50, 60\}$ . For setting  $d$ ,  $B_d^*$  is the largest grid width whose measured latency does not exceed RECAP; it is selected from latency alone. Latency is measured end to end on one A100-80GB at batch 256 after 200 warm-up and 1,000 timed batches, excluding preprocessing and cached-trie construction uniformly. We run five paired seeds. Primary uncertainty uses a hierarchical bootstrap that resamples seeds and then paired users within each sampled seed; published values are contextual references and never enter paired comparisons.

**Reproducibility.** We hash checkpoints, SIDs, tries, splits, fusion, and plug-in artifacts; join cached routes by stable sample ID; and store legal-set distributions needed to recompute ranks, entropy, and actionability. Model selection uses validation only, while paired seed and user outcomes are retained for recomputation. The anonymous repository linked in the abstract contains code, configurations, and table-generation utilities.

## 5.2 Reciprocal Critique Diagnostics

Figure 2(a) separates native Top-10 misses by candidate provenance. Direct beams already contain 31.2–38.1% of targets, reciprocal beams expose another 15.8–19.6%, and 46.1–49.2% remain outside the complete four-way width-20 union. A fixed-union scorer cannot recover the last group. For first-error detection, each (user, audit direction, SID level) pair is one binary instance: the oracle first-divergence level is positive, all other evaluated levels are negative, and trajectories without a divergence contribute only negatives. Figure 2(b) reports both audit directions: reciprocal and self AUPRC span 0.421–0.498 and 0.286–0.331, respectively, while the dataset-level prevalence baselines span 0.172–0.188. Macro-averaging the two directions within each dataset gives 0.43–0.49 for reciprocal critique and 0.29–0.33 for self critique. Among true first-divergence cases, Figure 2(c) reports whether the oracle legal child is in the reciprocal Top-3; the two directions span 57.4–68.1%. Both audit directions improve over their self baselines on every dataset. Together these results establish missing candidates, a usable error signal, and a plausible repair direction rather than disagreement alone. Figure 4 traces an Arts target absent from all four route beams (each  $B = 20$ ). At L2, direct evidence favors the original decision (.56 versus .17), whereas reciprocal evidence favors the alternative (.43 versus .10); the locator, actionability, and entropy gates therefore trigger repair at the oracle level ( $\hat{r} = r^* = 2$ ). Two repair rounds produce the accepted proposal, whose native direct score  $-8.42$  exceeds the unchanged beam floor  $-8.57$ ; it enters at direct rank 13 and reaches fused rank 7 under unchanged fusion.

## 5.3 Overall Performance

Table 2 separates source-tagged published context from the paired evidence used for every claim. The gain rows make the plug-in effect explicit across all four metrics without treating the published block as protocol-matched. Relative NDCG@10 gains range from 4.4% to 5.7% (5.0% on average), while absolute H@10 gains range from 0.0053 to 0.0080.

Figure 5 makes intermediate reliability and final uncertainty visible in the same nine-setting order. Reciprocal AUPRC exceeds its same-capacity self baseline in every row, including native reciprocal tasks and the frozen SynGR wrapper. Hierarchical seed/user bootstrap intervals for paired NDCG@10 remain on the positive side of zero. Rows are aligned only to ease joint inspection; we do not claim that a larger AUPRC improvement causes a larger recommendation gain.

## 5.4 Measured-Cost Comparisons

Figure 6(a) and Table 3 compare alternative uses of inference budget. On Arts, widening native search from B20 to latency-matched B40 raises NDCG@10 from 0.1010 to 0.1032; the first over-budget point,

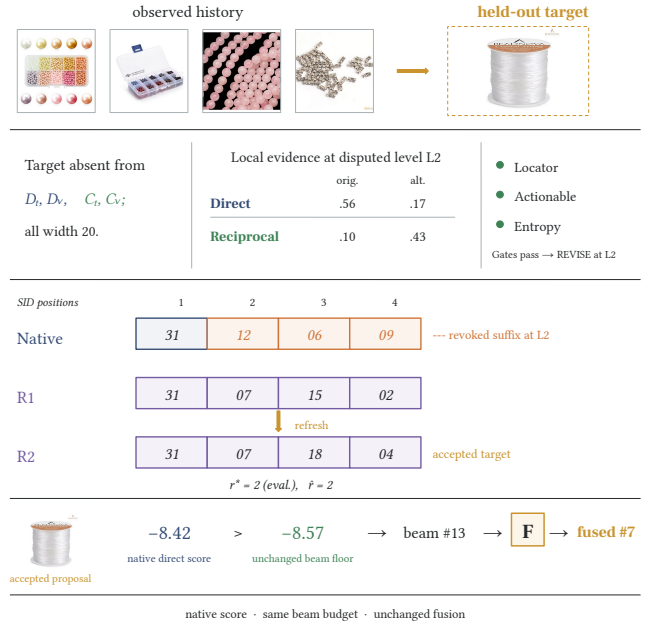


**Figure 4: Repair trace on Arts. The target is absent from all four width-20 route beams. Reciprocal evidence overturns the direct preference at L2, the gates trigger suffix repair, and two repair rounds yield a proposal whose native score exceeds the unchanged beam floor. It enters at direct rank 13 and reaches fused rank 7 under unchanged fusion.**

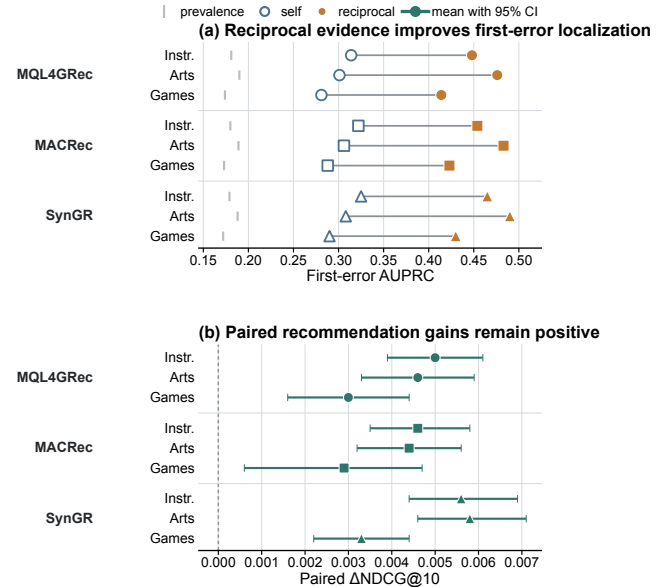


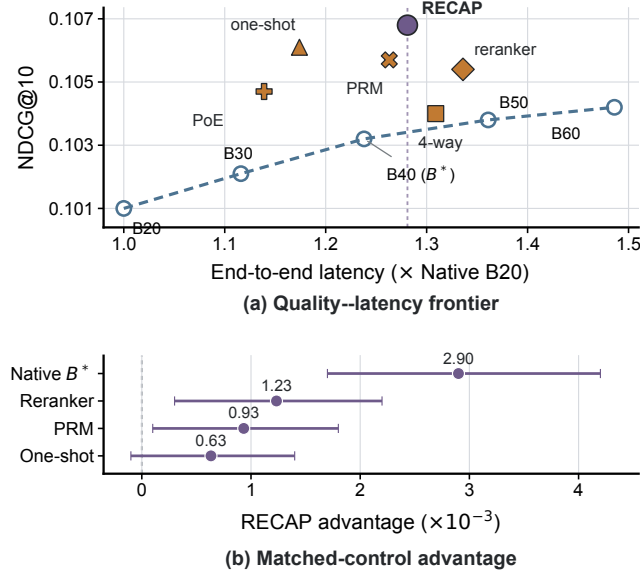
**Figure 5: Cross-setting reliability: (a) first-error AUPRC; (b) paired NDCG@10 gains with 95% bootstrap intervals.**

B50, reaches 0.1038. Four-way ensembling and a train-matched reranker reach 0.1040 and 0.1054. Parameter-free PoE reaches 0.1047, PRM-guided prefix search 0.1057, and one-shot repair

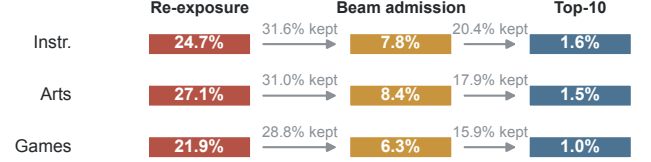
**Table 2: Full-catalog performance. Published values are transcribed verbatim from MACRec Table 2 as non-protocol-matched context; only paired Native/+RECAP rows support claims.**

| Method                                                              | Instruments  |              |              |              | Arts         |              |              |              | Games        |              |              |              |
|---------------------------------------------------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
|                                                                     | H@5          | H@10         | N@5          | N@10         | H@5          | H@10         | N@5          | N@10         | H@5          | H@10         | N@5          | N@10         |
| <i>Published context from MACRec Table 2; not protocol-matched</i>  |              |              |              |              |              |              |              |              |              |              |              |              |
| BERT4Rec                                                            | .0856        | .1081        | .0667        | .0739        | .0697        | .0922        | .0502        | .0575        | .0426        | .0725        | .0270        | .0366        |
| SASRec                                                              | .0946        | .1233        | .0654        | .0746        | .0951        | .1250        | .0610        | .0706        | .0587        | .0985        | .0333        | .0461        |
| FDSA                                                                | .0987        | .1249        | .0775        | .0859        | .0832        | .1190        | .0583        | .0695        | .0614        | .0988        | .0389        | .0509        |
| S3-Rec                                                              | .0937        | .1123        | .0693        | .0743        | .0739        | .1030        | .0511        | .0630        | .0527        | .0903        | .0351        | .0468        |
| MISSRec                                                             | .1089        | .1361        | .0797        | .0880        | .1021        | .1321        | .0699        | .0815        | .0674        | .1048        | .0385        | .0499        |
| P5-CID                                                              | .0839        | .1119        | .0678        | .0704        | .0713        | .0994        | .0607        | .0662        | .0532        | .0824        | .0331        | .0454        |
| VIP5                                                                | .0892        | .1071        | .0815        | .0872        | .0704        | .0959        | .0586        | .0635        | .0480        | .0758        | .0328        | .0418        |
| TIGER                                                               | .1007        | .1221        | .0882        | .0950        | .0894        | .1167        | .0718        | .0806        | .0523        | .0857        | .0345        | .0453        |
| <i>Paired frozen-backbone comparisons under our common protocol</i> |              |              |              |              |              |              |              |              |              |              |              |              |
| MQL4GRec Native                                                     | .1115        | .1375        | .0977        | .1060        | .1037        | .1327        | .0857        | .0950        | .0637        | .1033        | .0421        | .0548        |
| MQL4GRec +RECAP                                                     | <b>.1154</b> | <b>.1438</b> | <b>.1014</b> | <b>.1110</b> | <b>.1070</b> | <b>.1385</b> | <b>.0890</b> | <b>.0996</b> | <b>.0663</b> | <b>.1087</b> | <b>.0440</b> | <b>.0578</b> |
| Gain (%)                                                            | +3.5%        | +4.6%        | +3.8%        | +4.7%        | +3.2%        | +4.4%        | +3.9%        | +4.8%        | +4.1%        | +5.2%        | +4.5%        | +5.5%        |
| MACRec Native                                                       | .1110        | .1363        | .0965        | .1046        | .1046        | .1329        | .0868        | .0953        | .0671        | .1078        | .0435        | .0565        |
| MACRec +RECAP                                                       | <b>.1148</b> | <b>.1422</b> | <b>.1001</b> | <b>.1092</b> | <b>.1078</b> | <b>.1386</b> | <b>.0899</b> | <b>.0997</b> | <b>.0697</b> | <b>.1131</b> | <b>.0454</b> | <b>.0594</b> |
| Gain (%)                                                            | +3.4%        | +4.3%        | +3.7%        | +4.4%        | +3.1%        | +4.3%        | +3.6%        | +4.6%        | +3.9%        | +4.9%        | +4.4%        | +5.1%        |
| SynGR Native                                                        | .1321        | .1768        | .1045        | .1189        | .1145        | .1449        | .0919        | .1010        | .0702        | .1092        | .0471        | .0596        |
| SynGR +RECAP                                                        | <b>.1371</b> | <b>.1848</b> | <b>.1088</b> | <b>.1245</b> | <b>.1192</b> | <b>.1526</b> | <b>.0961</b> | <b>.1068</b> | <b>.0732</b> | <b>.1151</b> | <b>.0494</b> | <b>.0629</b> |
| Gain (%)                                                            | +3.8%        | +4.5%        | +4.1%        | +4.7%        | +4.1%        | +5.3%        | +4.6%        | +5.7%        | +4.3%        | +5.4%        | +4.9%        | +5.5%        |

Paired uncertainty: seed-level SD of  $\Delta N@10$  is at most 0.0019, and 44/45 seed deltas are positive.

**Figure 6: Quality-latency frontier and measured-cost comparisons. (a) SynGR on Arts. (b) RECAP advantage over four controls, equally averaged across three datasets (95% paired intervals).**

0.1061, versus 0.1068 for refreshed RECAP. This progression isolates reciprocal evidence, learned prefix verification, learned repair, and refresh.

**Figure 7: Blind-cohort candidate attrition. Legal re-exposure is progressively filtered by native beam admission and unchanged fusion before fused Top-10.**

The latency-only choice of  $B_d^*$  is directly checkable on every SynGR dataset: Instruments and Arts use  $B_{40}$ , while Games uses  $B_{50}$ ; the next wider native beam exceeds RECAP latency in every case. Figure 6(b) makes the equal-dataset paired comparisons explicit. Intervals remain above zero against  $B_d^*$ , the train-matched reranker, and PRM-guided search. The smaller advantage over one-shot repair slightly crosses zero, so we treat refresh as a consistent ranking pattern with limited resolved margin rather than a separately established significant gain.

Candidate provenance explains the gap. Let  $\mathcal{M}_{20}$  be Native- $B_{20}$  misses and  $\mathcal{M}_{\text{blind}} \subset \mathcal{M}_{20}$  exclude targets in all four width-20 beams. Within  $\mathcal{M}_{\text{blind}}$ , Figure 7 shows that RECAP re-exposes 21.9–27.1% of targets; 6.3–8.4% survive native rescoring and beam truncation, and 1.0–1.6% enter fused Top-10. Conversely, 60.4–68.9% of all successful rescues originate in  $\mathcal{M}_{\text{blind}}$ . Most re-exposed targets are still filtered by unchanged scoring and fusion, so re-exposure alone rarely converts to a hit.

**Table 3: Measured-cost controls. H@10/N@10 are absolute metrics; end-to-end latency is normalized to Native B = 20. Latency I/A/G lists Instruments/Arts/Games. \* marks the latency-selected  $B_d^*$ .**

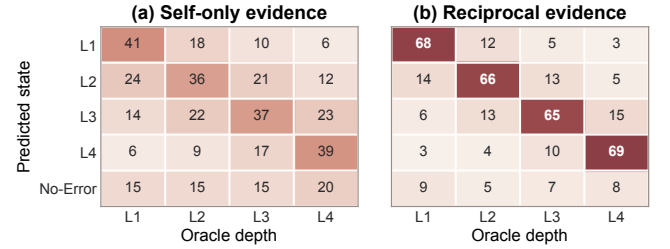
| Decoder                                                  | Instruments  |              | Arts         |              | Games        |              | Latency I/A/G     |
|----------------------------------------------------------|--------------|--------------|--------------|--------------|--------------|--------------|-------------------|
|                                                          | H@10         | N@10         | H@10         | N@10         | H@10         | N@10         | × native          |
| <i>SynGR beam-width and decoder frontier</i>             |              |              |              |              |              |              |                   |
| Native B20                                               | .1768        | .1189        | .1449        | .1010        | .1092        | .0596        | 1.000/1.000/1.000 |
| Native B30                                               | .1792        | .1201        | .1468        | .1021        | .1105        | .0604        | 1.108/1.116/1.086 |
| Native B40                                               | .1806*       | .1210*       | .1482*       | .1032*       | .1116        | .0609        | 1.225/1.238/1.171 |
| Native B50                                               | .1814        | .1216        | .1490        | .1038        | .1121*       | .0613*       | 1.348/1.361/1.248 |
| Native B60                                               | .1819        | .1220        | .1495        | .1042        | .1125        | .0615        | 1.472/1.486/1.329 |
| Four-way ensemble                                        | .1807        | .1214        | .1490        | .1040        | .1111        | .0607        | 1.301/1.309/1.294 |
| Train-matched reranker                                   | .1826        | .1231        | .1508        | .1054        | .1136        | .0620        | 1.329/1.336/1.321 |
| PoE repair                                               | .1817        | .1224        | .1498        | .1047        | .1128        | .0617        | 1.128/1.139/1.124 |
| PRM-guided prefix search                                 | .1831        | .1235        | .1513        | .1057        | .1137        | .0622        | 1.252/1.263/1.246 |
| One-shot repair                                          | .1836        | .1238        | .1519        | .1061        | .1143        | .0624        | 1.164/1.174/1.158 |
| RECAP                                                    | <b>.1848</b> | <b>.1245</b> | <b>.1526</b> | <b>.1068</b> | <b>.1151</b> | <b>.0629</b> | 1.270/1.281/1.270 |
| <i>Backbone transfer of the measured-cost comparison</i> |              |              |              |              |              |              |                   |
| MQL4GRec Native B20                                      | .1375        | .1060        | .1327        | .0950        | .1033        | .0548        | 1.000/1.000/1.000 |
| MQL4GRec Native B*                                       | .1408        | .1081        | .1355        | .0970        | .1061        | .0561        | 1.234/1.241/1.238 |
| MQL4GRec +RECAP                                          | .1438        | .1110        | .1385        | .0996        | .1087        | .0578        | 1.268/1.274/1.263 |
| MACRec Native B20                                        | .1363        | .1046        | .1329        | .0953        | .1078        | .0565        | 1.000/1.000/1.000 |
| MACRec Native B*                                         | .1394        | .1066        | .1357        | .0971        | .1105        | .0577        | 1.244/1.252/1.245 |
| MACRec +RECAP                                            | .1422        | .1092        | .1386        | .0997        | .1131        | .0594        | 1.279/1.286/1.273 |

## 5.5 Mechanism and Selectivity

Table 4 separates the method’s dependencies. Parameter-free PoE recovers 0.0037 of the full 0.0058 NDCG@10 gain, showing that reciprocal route evidence is useful before adding learned suffix capacity. Natural-only training remains effective but loses 0.0013 NDCG, so structured corruption broadens supervision without replacing natural errors. An identical-capacity self-only repairer retains 31% of the full NDCG gain, while shuffled reciprocal evidence nearly returns to native. Replacing first-error localization with global minimum confidence or changing only one token also degrades quality. Always-revise disables every KEEP condition and forces one revision for every user; it increases harm from 0.15% to 0.67% despite finding more rescues.

Figure 8 exposes the layer-wise structure behind the scalar AUPRC results. Conditioning each column on an oracle first-divergence depth, reciprocal evidence concentrates substantially more mass on the correct level at every depth and reduces erroneous No-Error outputs. The comparison includes all locator states rather than conditioning on a successful localization.

The four operating controls have distinct roles. The mixture  $\lambda$  determines how strongly reciprocal evidence can challenge the realized route;  $\tau_{\text{err}}$  and  $\tau_{\text{act}}$  require both a credible error and a viable alternative; and  $\tau_H$  rejects diffuse repair distributions. We therefore vary one control at a time around a single validation-frozen profile rather than selecting a separate optimum for every dataset or metric. The paired accounting is user-level: Trigger records intervention frequency, whereas Rescue and Harm record the two directions of Top-10 conversion. Their difference must equal the H@10 change, and Keep measures retention conditional on a native hit. Finally,

**Figure 8: First-error localization on SynGR/Arts. Columns condition on oracle depth and sum to 100%; rows include all locator outputs.**

the oracle rows release localization, direction, and finite-search constraints in sequence. They are diagnostic ceilings, not additional methods, and identify where a larger repair budget would help only after the correct branch has been reopened.

The full hyperparameter view in Figure 9 separates four controls rather than presenting a single favorable slice. Shared defaults remain near a robust region, but the nearby HR/NDCG optima, slopes, and plateaus differ across datasets. Repair-width gains also saturate near the shared default  $K_{\text{rep}} = 20$ : widening to 40 changes NDCG@10 by at most 0.0001 on all three SynGR datasets.

Table 5 reports the full intervention accounting across nine settings. Full RECAP triggers on 17.4–24.7% of users, converts 0.61–0.96% of all users from miss to hit, and harms only 0.08–0.16%. It consequently retains 98.95–99.26% of native Top-10 hits. Figure 10 makes the operating mechanism visible: executed revisions concentrate at intermediate depths but exhibit dataset-specific profiles,

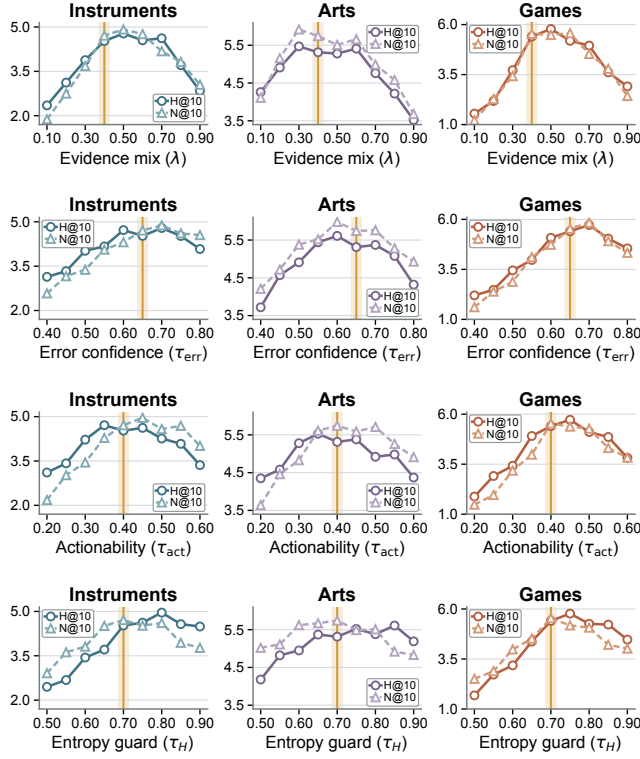


Figure 9: One-factor sensitivity around shared validation-frozen defaults. Curves show relative  $H@10/NDCG@10$  gains; amber bands mark defaults; vertical ranges are fixed within dataset columns.

Table 4: Mechanism analysis on SynGR/Arts. Trigger/Rescue/Harm are user percentages; oracle rows report finite-search headroom.

| Variant                      | H@10         | N@10         | Trigger | Rescue | Harm |
|------------------------------|--------------|--------------|---------|--------|------|
| Native                       | .1449        | .1010        | 0.0     | 0.00   | 0.00 |
| Full RECAP                   | <b>.1526</b> | <b>.1068</b> | 24.7    | .92    | .15  |
| Natural-only training        | .1510        | .1055        | 23.8    | .80    | .19  |
| Self-only evidence           | .1478        | .1028        | 19.4    | .47    | .18  |
| Shuffled reciprocal evidence | .1461        | .1018        | 16.8    | .29    | .17  |
| PoE repair                   | .1498        | .1047        | 24.7    | .70    | .21  |
| Lowest-confidence locator    | .1485        | .1035        | 22.4    | .57    | .21  |
| Token-only replacement       | .1492        | .1040        | 24.5    | .66    | .23  |
| Always-revise                | .1500        | .1043        | 100.0   | 1.18   | .67  |
| Oracle first-error level     | .1549        | .1086        | –       | –      | –    |
| + correct child              | .1572        | .1105        | –       | –      | –    |
| + wider repair budget        | .1608        | .1135        | –       | –      | –    |

while net  $H@10$  utility peaks near the default gate and declines under broader intervention. Together with Always-revise in Table 4, this shows why selection matters: the method combines an informative critic with a credible KEEP decision rather than treating every disagreement as an error.

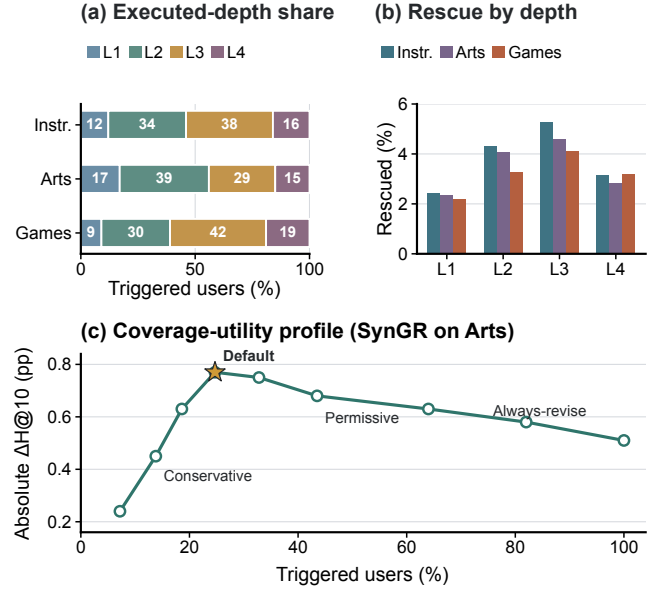


Figure 10: Selective-revision diagnostics on SynGR: executed-depth share, rescue by depth, and Arts coverage-utility (star: default).

Table 5: Paired intervention accounting (%). Trigger/Rescue/Harm use all users; Keep conditions on native hits; Rescue – Harm equals  $\Delta H@10$  in points.

| Backbone | Dataset | Trigger | Rescue | Harm | Keep  |
|----------|---------|---------|--------|------|-------|
| MQL4GRec | Instr.  | 20.5    | .75    | .12  | 99.13 |
|          | Arts    | 22.8    | .72    | .14  | 98.95 |
|          | Games   | 17.4    | .62    | .08  | 99.23 |
| MACRec   | Instr.  | 21.1    | .71    | .12  | 99.12 |
|          | Arts    | 23.3    | .70    | .13  | 99.02 |
|          | Games   | 18.1    | .61    | .08  | 99.26 |
| SynGR    | Instr.  | 22.6    | .96    | .16  | 99.10 |
|          | Arts    | 24.7    | .92    | .15  | 98.96 |
|          | Games   | 19.4    | .68    | .09  | 99.18 |

## 6 Conclusion

Reciprocal routes can critique realized SID commitments rather than merely shape representations or add fusion scores. RECAP turns this signal into selective hierarchical correction: first-error localization, prefix-preserving revocation, refreshed repair, and native-score reintegration under unchanged fusion. It improves three frozen backbones on three datasets, and its advantage persists under matched latency and against a train-matched reranker. Paired accounting shows that these gains come with nearly all native hits retained. Multimodal inference can therefore let one view request reconsideration of a specific decision rather than only vote on completed lists.

## Ethical Considerations

This work uses public interaction benchmarks and introduces no new user data. Selective repair can inherit modality, popularity, and exposure biases and alter which users, providers, or catalog regions receive additional opportunities. Deployment therefore requires subgroup and provider audits and exposure monitoring; abstention does not by itself guarantee fairness or robustness.

## References

- [1] Peter Anderson, Basura Fernando, Mark Johnson, and Stephen Gould. 2017. Guided Open Vocabulary Image Captioning with Constrained Beam Search. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. 936–945. doi:10.18653/v1/D17-1098
- [2] Jacob Austin, Daniel D. Johnson, Jonathan Ho, Daniel Tarlow, and Rianne van den Berg. 2021. Structured Denoising Diffusion Models in Discrete State-Spaces. In *Advances in Neural Information Processing Systems*, Vol. 34.
- [3] Peter L. Bartlett and Marten H. Wegkamp. 2008. Classification with a Reject Option Using a Hinge Loss. *Journal of Machine Learning Research* 9 (2008), 1823–1840.
- [4] Luis Felipe Cattelan and Danilo Silva. 2024. How to Fix a Broken Confidence Estimator: Evaluating Post-hoc Methods for Selective Classification with Deep Neural Networks. In *Proceedings of the Fortieth Conference on Uncertainty in Artificial Intelligence (Proceedings of Machine Learning Research, Vol. 244)*. PMLR, 547–584.
- [5] Wei Chen, Xingyu Guo, Shuang Li, Fuwei Zhang, Meng Yuan, Jing Fan, Zhao Zhang, Deqing Wang, and Fuzhen Zhuang. 2026. SynGR: Unleashing the Potential of Cross-Modal Synergy for Generative Recommendation. *CoRR* abs/2605.18920 (2026). <https://arxiv.org/abs/2605.18920>
- [6] Yashar Deldjoo, Markus Schedl, Balázs Hidasi, Yinwei Wei, and Xiangnan He. 2022. Multimedia Recommender Systems: Algorithms and Challenges. In *Recommender Systems Handbook*. Springer, 973–1014. doi:10.1007/978-1-0716-2197-4\_25
- [7] Yonatan Geifman and Ran El-Yaniv. 2017. Selective Classification for Deep Neural Networks. In *Advances in Neural Information Processing Systems*, Vol. 30. 4878–4887.
- [8] Yonatan Geifman and Ran El-Yaniv. 2019. SelectiveNet: A Deep Neural Network with an Integrated Reject Option. In *Proceedings of the 36th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 97)*. 2151–2159.
- [9] Shijie Geng, Shuchang Liu, Zuohui Fu, Yingqiang Ge, and Yongfeng Zhang. 2022. Recommendation as Language Processing (RLP): A Unified Pretrain, Personalized Prompt and Predict Paradigm (P5). In *Proceedings of the 16th ACM Conference on Recommender Systems*. 299–315. doi:10.1145/3523227.3546767
- [10] Shijie Geng, Juntao Tan, Shuchang Liu, Zuohui Fu, and Yongfeng Zhang. 2023. VIP5: Towards Multimodal Foundation Models for Recommendation. In *Findings of the Association for Computational Linguistics: EMNLP 2023*. Association for Computational Linguistics, 9606–9620. doi:10.18653/v1/2023.findings-emnlp.644
- [11] Zhibin Gou, Zhihong Shao, Yeyun Gong, Yelong Shen, Yujiu Yang, Nan Duan, and Weizhu Chen. 2024. CRITIC: Large Language Models Can Self-Correct with Tool-Interactive Critiquing. In *International Conference on Learning Representations*.
- [12] Chengcheng Guo, Kuo Cai, Yu Zhou, Qiang Luo, Ruiming Tang, Han Li, Kun Gai, and Guorui Zhou. 2026. PROMISE: Process Reward Models Unlock Test-Time Scaling Laws in Generative Recommendations. *CoRR* abs/2601.04674 (2026). <https://arxiv.org/abs/2601.04674>
- [13] Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q. Weinberger. 2017. On Calibration of Modern Neural Networks. In *Proceedings of the 34th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 70)*. 1321–1330.
- [14] Zhengfu He, Tianxiang Sun, Qiong Tang, Kuanning Wang, Xuanjing Huang, and Xipeng Qiu. 2023. DiffusionBERT: Improving Generative Masked Language Models with Diffusion Models. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*. 4521–4534. doi:10.18653/v1/2023.acl-long.248
- [15] Chris Hokamp and Qun Liu. 2017. Lexically Constrained Decoding for Sequence Generation Using Grid Beam Search. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*. 1535–1546. doi:10.18653/v1/P17-1141
- [16] Wenyue Hua, Shuyuan Xu, Yingqiang Ge, and Yongfeng Zhang. 2023. How to Index Item IDs for Recommendation Foundation Models. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval in the Asia Pacific Region*. 195–204. doi:10.1145/3624918.3625339
- [17] Liang Huang, Kai Zhao, and Mingbo Ma. 2017. When to Finish? Optimal Beam Search for Neural Text Generation (modulo beam size). In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. 2134–2139. doi:10.18653/v1/D17-1227
- [18] Heinrich Jiang, Been Kim, Melody Y. Guan, and Maya R. Gupta. 2018. To Trust or Not to Trust a Classifier. In *Advances in Neural Information Processing Systems*, Vol. 31. 5546–5557.
- [19] Wang-Cheng Kang and Julian J. McAuley. 2018. Self-Attentive Sequential Recommendation. In *2018 IEEE International Conference on Data Mining*. 197–206. doi:10.1109/ICDM.2018.00035
- [20] Jason Lee, Elman Mansimov, and Kyunghyun Cho. 2018. Deterministic Non-Autoregressive Neural Sequence Modeling by Iterative Refinement. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. 1173–1182. doi:10.18653/v1/D18-1149
- [21] José Lezama, Huiwen Chang, Lu Jiang, and Irfan Essa. 2022. Improved Masked Image Generation with Token-Critic. In *European Conference on Computer Vision*. 70–86.
- [22] Xiang Lisa Li, John Thickstun, Ishaan Gulrajani, Percy Liang, and Tatsunori B. Hashimoto. 2022. Diffusion-LM Improves Controllable Text Generation. In *Advances in Neural Information Processing Systems*, Vol. 35. 4328–4343.
- [23] Yongqi Li, Nan Yang, Liang Wang, Furu Wei, and Wenjie Li. 2023. Multiview Identifiers Enhanced Generative Retrieval. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*. 6636–6648. doi:10.18653/v1/2023.acl-long.366
- [24] Yang Li, Qi'ao Zhao, Chen Lin, Jinsong Su, and Zhilin Zhang. 2024. Who To Align With: Feedback-Oriented Multi-Modal Alignment in Recommendation Systems. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 667–676. doi:10.1145/3626772.3657829
- [25] Hunter Lightman, Vineet Kosaraju, Yuri Burda, Harrison Edwards, Bowen Baker, Teddy Lee, Jan Leike, John Schulman, Ilya Sutskever, and Karl Cobbe. 2024. Let's Verify Step by Step. In *International Conference on Learning Representations*.
- [26] Aaron Lou, Chenlin Meng, and Stefano Ermon. 2024. Discrete Diffusion Modeling by Estimating the Ratios of the Data Distribution. In *Proceedings of the 41st International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 235)*. 32819–32848.
- [27] Ximing Lu, Sean Welleck, Peter West, Liwei Jiang, Jungo Kasai, Daniel Khashabi, Ronan Le Bras, Lianhui Qin, Youngjae Yu, Rowan Zellers, Noah A. Smith, and Yejin Choi. 2022. NeuroLogic A'esque Decoding: Constrained Text Generation with Lookahead Heuristics. In *Proceedings of NAACL-HLT*. 780–799. doi:10.18653/v1/2022.naacl-main.57
- [28] Ximing Lu, Peter West, Rowan Zellers, Ronan Le Bras, Chandra Bhagavatula, and Yejin Choi. 2021. NeuroLogic Decoding: (Un)supervised Neural Text Generation with Predicate Logic Constraints. In *Proceedings of NAACL-HLT*. 4288–4299. doi:10.18653/v1/2021.naacl-main.339
- [29] Jonathan Mallinson, Aliaksei Severyn, Eric Malmi, and Guillermo Garrido. 2020. FELIX: Flexible Text Editing Through Tagging and Insertion. In *Findings of the Association for Computational Linguistics: EMNLP 2020*. 1244–1255. doi:10.18653/v1/2020.findings-emnlp.111
- [30] Eric Malmi, Sebastian Krause, Sascha Rothe, Daniil Mirylenka, and Aliaksei Severyn. 2019. Encode, Tag, Realize: High-Precision Text Editing. In *Proceedings of EMNLP-IJCNLP*. 5053–5064. doi:10.18653/v1/D19-1510
- [31] Clara Meister, Ryan Cotterell, and Tim Vieira. 2020. If Beam Search Is the Answer, What Was the Question?. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*. 2173–2185. doi:10.18653/v1/2020.emnlp-main.170
- [32] Lingyu Mu, Hao Deng, Haibo Xing, Jinxin Hu, Yu Zhang, Xiaoyi Zeng, and Jing Zhang. 2026. Masked Diffusion Generative Recommendation. *CoRR* abs/2601.19501 (2026). <https://arxiv.org/abs/2601.19501>
- [33] Jianmo Ni, Jiacheng Li, and Julian McAuley. 2019. Justifying Recommendations using Distantly-Labeled Reviews and Fine-Grained Aspects. In *Proceedings of EMNLP-IJCNLP*. 188–197. doi:10.18653/v1/D19-1018
- [34] Matt Post and David Vilar. 2018. Fast Lexically Constrained Decoding with Dynamic Beam Allocation for Neural Machine Translation. In *Proceedings of NAACL-HLT*. 1314–1324. doi:10.18653/v1/N18-1119
- [35] Haohao Qu, Wenqi Fan, Zihuai Zhao, and Qing Li. 2025. TokenRec: Learning to Tokenize ID for LLM-Based Generative Recommendations. *IEEE Transactions on Knowledge and Data Engineering* 37, 10 (2025), 6216–6231. doi:10.1109/TKDE.2025.3599265
- [36] Shashank Rajput, Nikhil Mehta, Anima Singh, Raghunandan H. Keshavan, Trung Vu, Lukasz Heldt, Lichan Hong, Yi Tay, Vinh Q. Tran, Jonah Samost, Maciej Kula, Ed H. Chi, and Maheswaran Sathiamoorthy. 2023. Recommender Systems with Generative Retrieval. In *Advances in Neural Information Processing Systems*, Vol. 36.
- [37] Subham Sekhar Sahoo, Marianne Arriola, Yair Schiff, Aaron Gokaslan, Edgar Marroquin, Justin T. Chiu, Alexander Rush, and Volodymyr Kuleshov. 2024. Simple and Effective Masked Diffusion Language Models. In *Advances in Neural Information Processing Systems*.
- [38] Kulin Shah, Bhuvesh Kumar, Neil Shah, and Liam Collins. 2025. Masked Diffusion for Generative Recommendation. *CoRR* abs/2511.23021 (2025). <https://arxiv.org/abs/2511.23021>

- [39] Felix Stahlberg and Bill Byrne. 2019. On NMT Search Errors and Model Errors: Cat Got Your Tongue?. In *Proceedings of EMNLP-IJCNLP*. 3354–3360. doi:10.18653/v1/D19-1331
- [40] Mitchell Stern, William Chan, Jamie Kiros, and Jakob Uszkoreit. 2019. Insertion Transformer: Flexible Sequence Generation via Insertion Operations. In *Proceedings of the 36th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 97)*. 5976–5985.
- [41] Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. 2019. BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations from Transformer. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. 1441–1450. doi:10.1145/3357384.3357895
- [42] Weiwei Sun, Lingyong Yan, Zheng Chen, Shuaiqiang Wang, Haichao Zhu, Pengjie Ren, Zhumin Chen, Dawei Yin, Maarten de Rijke, and Zhaochun Ren. 2023. Learning to Tokenize for Generative Retrieval. In *Advances in Neural Information Processing Systems*, Vol. 36. 7341–7360.
- [43] Yi Tay, Vinh Q. Tran, Mostafa Dehghani, Jianmo Ni, Dara Bahri, Harsh Mehta, Zhen Qin, Kai Hui, Zhe Zhao, Jai Prakash Gupta, Tal Schuster, William W. Cohen, and Donald Metzler. 2022. Transformer Memory as a Differentiable Search Index. In *Advances in Neural Information Processing Systems*, Vol. 35. 21831–21843.
- [44] Wenjie Wang, Honghui Bao, Xinyu Lin, Jizhi Zhang, Yongqi Li, Fuli Feng, See-Kiong Ng, and Tat-Seng Chua. 2024. Learnable Item Tokenization for Generative Recommendation. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*. ACM, 2400–2409. doi:10.1145/3627673.3679569
- [45] Yujing Wang, Yingyan Hou, Haonan Wang, Ziming Miao, Shibin Wu, Qi Chen, Yuqing Xia, Chengmin Chi, Guoshuai Zhao, Zheng Liu, Xing Xie, Hao Sun, Weiwei Deng, Qi Zhang, and Mao Yang. 2022. A Neural Corpus Indexer for Document Retrieval. In *Advances in Neural Information Processing Systems*, Vol. 35. 25600–25614.
- [46] Ye Wang, Jiahao Xun, Mingjie Hong, Jieming Zhu, Tao Jin, Wang Lin, Haoyuan Li, Linjun Li, Yan Xia, Zhou Zhao, and Zhenhua Dong. 2024. EAGER: Two-Stream Generative Recommender with Behavior-Semantic Collaboration. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 3245–3254. doi:10.1145/3637528.3671885
- [47] Wei Wei, Chao Huang, Lianghao Xia, and Chuxu Zhang. 2023. Multi-Modal Self-Supervised Learning for Recommendation. In *Proceedings of the ACM Web Conference 2023*. 790–800. doi:10.1145/3543507.3583206
- [48] Yinwei Wei, Xiang Wang, Liqiang Nie, Xiangnan He, Richang Hong, and Tat-Seng Chua. 2019. MMGCN: Multi-modal Graph Convolution Network for Personalized Recommendation of Micro-video. In *Proceedings of the 27th ACM International Conference on Multimedia*. 1437–1445. doi:10.1145/3343031.3351034
- [49] Sam Wiseman and Alexander M. Rush. 2016. Sequence-to-Sequence Learning as Beam-Search Optimization. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. 1296–1306. doi:10.18653/v1/D16-1137
- [50] Yingce Xia, Fei Tian, Lijun Wu, Jianxin Lin, Tao Qin, Nenghai Yu, and Tie-Yan Liu. 2017. Deliberation Networks: Sequence Generation Beyond One-Pass Decoding. In *Advances in Neural Information Processing Systems*, Vol. 30. 1784–1794.
- [51] Haibo Xing, Hao Deng, Lingyu Mu, Jinxin Hu, Yu Zhang, Xiaoyi Zeng, and Jing Zhang. 2026. Learning to Reflect and Correct: Towards Better Decoding Trajectories for Large-Scale Generative Recommendation. *CoRR* abs/2602.23639 (2026). <https://arxiv.org/abs/2602.23639>
- [52] Liu Yang, Fabian Paischer, Kaveh Hassani, Jiacheng Li, Shuai Shao, Zhang Gabriel Li, Yun He, Xue Feng, Nima Noorshams, Sem Park, Bo Long, Robert D. Nowak, Xiaoli Gao, and Hamid Eghbalzadeh. 2025. Unifying Generative and Dense Retrieval for Sequential Recommendation. *Transactions on Machine Learning Research* (2025).
- [53] Yilin Yang, Liang Huang, and Mingbo Ma. 2018. Breaking the Beam Search Curse: A Study of (Re-)Scoring Methods and Stopping Criteria for Neural Machine Translation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. 3054–3059. doi:10.18653/v1/D18-1342
- [54] Jianyang Zhai, Zi-Feng Mai, Chang-Dong Wang, Feidiao Yang, Xiawu Zheng, Hui Li, and Yonghong Tian. 2025. Multimodal Quantitative Language for Generative Recommendation. In *International Conference on Learning Representations*.
- [55] Fuwei Zhang, Xiaoyu Liu, Dongbo Xi, Jishen Yin, Huan Chen, Peng Yan, Fuzhen Zhuang, and Zhao Zhang. 2026. Multi-Aspect Cross-modal Quantization for Generative Recommendation. In *Proceedings of the Fortieth AAAI Conference on Artificial Intelligence*, Vol. 40. 16271–16279. doi:10.1609/aaai.v40i19.38664
- [56] Kangning Zhang, Haotian Fang, Xukun Luo, Hao Yin, Yang Gao, Peng Yan, Weiwen Liu, Weinan Zhang, and Yong Yu. 2026. Learning from Unreachable Rewards: Hint-Conditioned Reinforcement Learning for Generative Recommendation. In *Proceedings of the 35th ACM International Conference on Information and Knowledge Management*. doi:10.1145/3799682.3840842
- [57] Bowen Zheng, Yupeng Hou, Hongyu Lu, Yu Chen, Wayne Xin Zhao, Ming Chen, and Ji-Rong Wen. 2024. Adapting Large Language Models by Integrating Collaborative Semantics for Recommendation. In *2024 IEEE 40th International Conference on Data Engineering*. IEEE, 1435–1448. doi:10.1109/ICDE60146.2024.00118
- [58] Xin Zhou and Zhiqi Shen. 2023. A Tale of Two Graphs: Freezing and Denoising Graph Structures for Multimodal Recommendation. In *Proceedings of the 31st ACM International Conference on Multimedia*. 935–943. doi:10.1145/3581783.3611948
- [59] Han Zhu, Daqing Chang, Zirui Xu, Pengye Zhang, Xiang Li, Jie He, Han Li, Jian Xu, and Kun Gai. 2019. Joint Optimization of Tree-based Index and Deep Model for Recommender Systems. In *Advances in Neural Information Processing Systems*, Vol. 32. 3973–3982.
- [60] Han Zhu, Xiang Li, Pengye Zhang, Guozheng Li, Jie He, Han Li, and Kun Gai. 2018. Learning Tree-based Deep Model for Recommender Systems. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 1079–1088. doi:10.1145/3219819.3219826
- [61] Liu Ziyin, Zhikang Wang, Paul Pu Liang, Ruslan Salakhutdinov, Louis-Philippe Morency, and Masahito Ueda. 2019. Deep Gamblers: Learning to Abstain with Portfolio Theory. In *Advances in Neural Information Processing Systems*, Vol. 32. 10622–10632.