ARC-AGI-3를 위한 비 LLM 그래프 탐색 에이전트: 새로움 기반 탐색과 사례 기반 기억 활용

모하마드 사다티　로라 마스카렐　김선동
광주과학기술원 AI융합학과
mohammadsaadati@gm.gist.ac.kr, {lmascarell, sundong}@gist.ac.kr

A Non-LLM Graph-Search Agent for ARC-AGI-3 with Novelty-Guided Exploration and Case-Based Memory

Mohammad Saadati　Laura Mascarell　Sundong Kim
Department of AI Convergence, GIST

Abstract

The ARC-AGI-3 benchmark (Abstraction and Reasoning Corpus for Artificial General Intelligence) evaluates agentic intelligence through novel turn-based environments in which agents must explore, infer goals, and plan without instructions. Frontier AI systems score below 1% while humans solve 100%, and the accompanying ARC Prize 2026 Kaggle competition forbids internet access at evaluation, disqualifying the large language model (LLM)-at-inference approaches that currently lead the benchmark. We treat this constraint as motivating a distinct problem formulation, and we propose a non-LLM-at-inference foundation comprising three components. The first is a state representation that separates gameplay content from user-interface (UI) overlay to enable correct cycle detection. The second is an Iterative Deepening A* (IDA*) search equipped with admissible heuristics that operate without prior knowledge of the goal. The third is an external case-based memory for within-environment cross-level transfer that requires no weight updates. We present this design as a proof of concept, demonstrate its behavior on a representative environment against uninformed-search and ablated variants, and discuss what the design supports, what it does not, and how it serves as a foundation for further offline-deployable work on the benchmark.

1. Introduction[1]

The ARC-AGI benchmark series [1] measures fluid intelligence, and ARC-AGI-3 [2] departs from its static input-output format: the agent must interact with a turn-based environment, infer its mechanics and winning conditions, and carry knowledge across increasingly difficult levels. Humans solve all environments with full reliability, while frontier AI systems score below 1%.

The ARC Prize 2026 Kaggle competition [3] imposes a strict deployment constraint: submissions run without internet access, within bounded compute, and must be released open-source. This disqualifies current top approaches such as Duke Hill-climbing and Arcgentica [2] that depend on frontier-model API calls at play time.

We treat this constraint as a forcing function that motivates the relevant problem. An agent without LLM access must carry its reasoning in whatever structure was compiled offline. We argue that principled graph-based search over a redefined state space, with goal-agnostic admissible heuristics and a case-based memory, is a workable foundation. LLMs may be used during development but not at play. As a proof of concept, we demonstrate on a representative environment that the novelty-guided agent sustains state discovery at a substantially higher rate than an uninformed breadth-first variant and a novelty-ablated variant, while a frame-change-greedy baseline plateaus.

2. Problem Setting

An ARC-AGI-3 environment is a sequence of levels with unannounced win conditions. At each turn, the agent receives a 64×64 frame from a 16-color palette and selects from five key actions, an Undo, and a coordinate-click over 4,096 cells, a branching factor near 4,102. No instructions are given: all must be inferred from interaction. Benchmark scoring rewards solving each level in few actions relative to a human baseline, so action efficiency is what matters.

[1] 본 연구는 IITP 디지털혁신기술 국제공동연구(RS-2024-00445087), 과학기술원 InnoCORE 사업(26-InnoCORE-01)에 의해 수행되었습니다.

Kaggle forbids internet access and requires open-source release, so agents relying on LLM calls during play are ineligible. We restrict to within-environment cross-level transfer. Cross-environment transfer is beyond this paper.

3. Approach

Figure 1 summarizes the agent architecture. The agent masks the UI overlay from each frame to derive a state, searches the resulting state graph with Iterative Deepening A* (IDA*) and goal-agnostic heuristics, reduces the action space and shapes search with intrinsic rewards, and reuses action sub-sequences via case-based memory.

3.1 Separating gameplay from interface

Every frame an ARC-AGI-3 agent observes carries two distinct kinds of information. The *playground region* encodes the game situation itself, while the *user-interface overlay* encodes quantities such as remaining action budget, current level, or life count, independent of the game situation.

For a graph-based agent relying on revisit detection, this distinction is decisive. If the raw frame is taken as state, every turn produces a formally distinct state and cycle detection fails silently. We therefore define the state as the frame with the UI overlay masked out, so that two states are equal if and only if their underlying game situations are equal. An action-budget detector identifies the overlay region using per-frame color and connected-component statistics, without environment-specific tuning. This redefinition is a precondition for the cycle detection that the search components below depend on.

3.2 Iterative deepening search on the state graph

The agent constructs a directed graph $G = (V, E)$ whose vertices are observed states and whose edges are labeled with actions. Exploration proceeds by IDA* with successively increasing depth bounds. Under uniform action cost, iterative deepening returns shortest-action-count solutions, and IDA* preserves optimality under any admissible heuristic.

In ARC-AGI-3, the goal must be discovered through play, so heuristics reflect exploration rather than the objective. We use three: the trivial heuristic $h_0(s) \equiv 0$ is admissible and reduces IDA* to iterative-deepening breadth-first search (BFS). The *frontier-distance* heuristic biases search toward unexplored territory:

$$h_{\text{frontier}}(s) = \min \{ d_G(s, s') : s' \in V_{\text{frontier}} \}$$

where d_G is shortest-path distance in G and $V_{frontier}$ is the set of unexpanded vertices. The *pixel-novelty* heuristic rewards states differing from those visited:

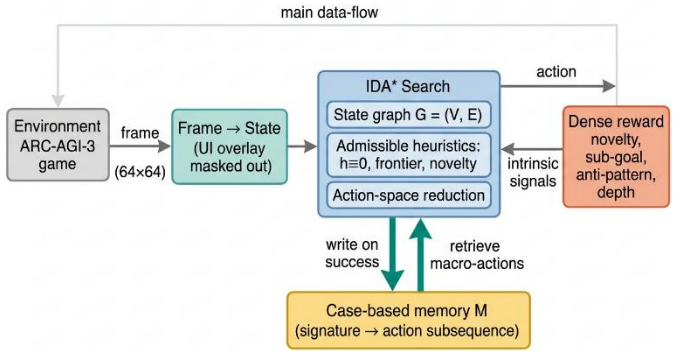

Figure 1. Agent architecture: IDA* searches over masked-state graphs, guided by novelty-based heuristics, with macro-actions retrievable from external case-based memory once it populates from level 1 onward.

$$h_{\text{novelty}}(s) = \alpha \cdot \min \{ \text{sim}(s, s') : s' \in V_{\text{visited}} \}$$

where $sim(\cdot,\cdot)$ is pixel-wise similarity (the fraction of matching grid cells) and $\alpha > 0$ is a small scaling constant. Formal admissibility is not guaranteed in this goal-discovery setting. The heuristic acts as a search bias.

3.3 Reducing the action space

Three regularities bring the branching factor into a tractable range. The *pixel-grouping* exploits 4-connected click semantics: clicks on a 4-connected same-colored region are equivalent, so we collapse each region to a representative action. The *no-effect pruning* refuses to re-expand actions whose expansion produced no playground change. The *foreground-background segmentation* heuristic identifies the dominant border color as background and treats other colors as foreground.

3.4 Failure handling and dense reward

Game-over states are not announced. The agent recognizes them from a vanished life indicator, abrupt playground reset, or prolonged non-progression, and records the terminal action sequence. Subsequent planning penalizes reproductions as anti-patterns.

Extrinsic reward in ARC-AGI-3 is sparse, so we drive search between reward events using four intrinsic signals. The *novelty* signal (dominant) rewards actions producing states not already in the graph. The *sub-goal progress* signal rewards any playground change, subordinated to novelty to avoid local-minimum pathologies. The *anti-pattern* signal contributes a large negative reward for matches to stored game-over trajectories. Finally, the *exploration-depth* signal adds a bonus proportional to graph depth. The four signals combine as a weighted sum that biases IDA* node expansion. Weights (1.0 novelty, 0.2 sub-goal, -10 anti-pattern, 0.05 depth) are fixed by design choice, with pixel-novelty scaling $\alpha = 0.1$, a pixel-similarity match threshold of 0.98, IDA* depth bounds from 1 to 12 in steps of 1, and a per-level budget of 800 actions. Principled selection of these values is future work.

3.5 Case-based external memory

The components above permit single-level exploration but no transfer. Within an environment, consecutive levels share object semantics, action affordances, and mechanics, so starting from scratch wastes structure already discovered. An external case-based memory *M* stores reusable action sub-sequences.

Each entry in *M* has the form *(signature, action-subsequence, outcome, usage-statistics)*, where the signature hashes the frame's connected-component color histogram and bounding-box structure, a coarse linear-time descriptor matching frames with shared structure. Memory is written by segmenting solved-level trajectories at sub-goal events. Each IDA* expansion queries *M*, and the top-*k* matches become macro-actions. Successful reuse promotes entries, repeated failures prune them. No gradients or weights are updated. In a new environment *M* starts empty and populates from level 1 onward. In the taxonomy of Hu et al. [4] this is a token-level, flat, experiential, case-based memory, and *M* grows indefinitely, satisfying continual-learning conditions [5].

4. Proof-of-Concept Demonstration

We demonstrate the agent on ft09, a representative public environment that exercises novelty, pattern matching, and anti-pattern avoidance, under a Kaggle-like regime: offline execution with bounded compute and a fixed per-level action budget. We compare against a *uniform random policy*, a *frame-change-greedy* baseline that picks any frame-changing action at random. To isolate the search machinery and novelty guidance, we add two comparisons: an uninformed breadth-first search (the trivial heuristic $h_0 \equiv 0$, retaining iterative deepening but no exploration bias), and a no-novelty ablation of our agent (action-space reduction and the sub-goal reward retained, novelty signal disabled). Because the case-based memory populates from level 1 onward, it is active within the demonstrated run; fuller multi-level evaluation is left to future work. Figure 2 shows that the frame-change-greedy and uniform-random agents plateau, while the breadth-first and no-novelty agents keep discovering at lower rates than the novelty-guided agent. The ablation gap isolates the contribution of novelty-guided exploration, and the breadth-first vs. frame-change-greedy gap shows what search adds over a reactive policy.

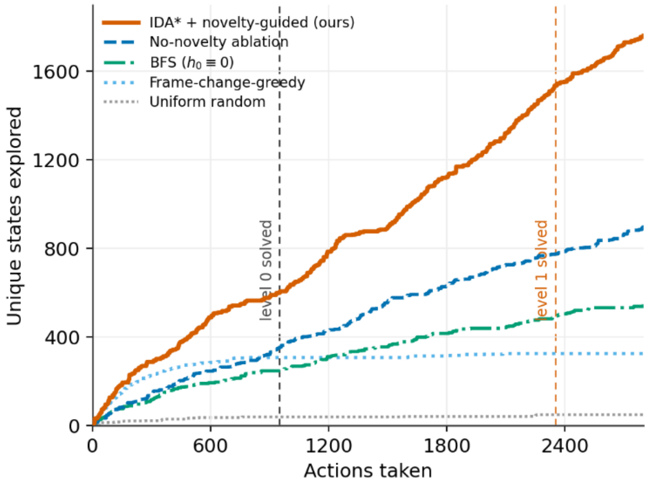

Figure 2. Unique states explored vs. actions taken. The IDA* + novelty-guided agent (ours) sustains the highest discovery rate. The frame-change-greedy and uniform-random baselines plateau, while the breadth-first search ($h_0 \equiv 0$) and no-novelty ablation keep discovering at lower rates. Vertical lines mark where the novelty-guided agent solved level 0 and level 1.

5. Discussion, Limitations, and Future Work

This work is a proof-of-concept demonstration, not a measure of readiness for the Kaggle competition: the demonstrated environment is materially easier than the private scoring sets and non-representative of their mechanics. We do not claim to solve ARC-AGI-3, but show that the design is viable under the Kaggle constraint — running offline, solving at least level 0, and accumulating memory entries within a run. Stronger claims (competitive action efficiency on held-out sets, cross-environment generalization) remain to be established.

Limitations: cross-environment transfer is outside the memory's scope; goal-condition inference is treated as an external module; dense-reward weights are set by design rather than ablation; and the anti-pattern mechanism uses exact matching. Component-level ablations are the immediate next step, alongside object-level signatures and theory-based world-model construction [6].

References

[1] F. Chollet, "On the Measure of Intelligence," arXiv:1911.01547, 2019.
[2] ARC Prize Foundation, "ARC-AGI-3: A New Challenge for Frontier Agentic Intelligence," arXiv:2603.24621, March 2026.
[3] ARC Prize Foundation, "ARC Prize 2026 Competition," https://www.kaggle.com/competitions/arc-prize-2026-arc-agi-3/.
[4] Y. Hu et al., "Memory in the Age of AI Agents: A Survey," arXiv:2512.13564, 2025.
[5] D. Abel et al., "A Definition of Continual Reinforcement Learning," in Advances in Neural Information Processing Systems, vol. 36, 2023.
[6] Z. Ahmed et al., "Synthesizing World Models for Bilevel Planning," Trans. Mach. Learn. Res., 2025.