

Supplementary Information

Failure-Aware Adaptive Inference for Reliable Long-Horizon LLM Agents

Youla Yang

1 Supplementary Methods

1.1 Detailed Failure Memory Retrieval

FAAI maintains an online memory storing previous execution experiences:

$$e_i = (s_i, a_i, o_{i+1}, r_i, f_i), \quad (1)$$

where s_i denotes the historical state, a_i denotes the executed action, o_{i+1} represents the resulting observation, r_i is the obtained feedback, and f_i indicates whether the execution resulted in failure.

Given a candidate decision $(s_t, \tilde{a}_t)$, FAAI retrieves similar historical failure experiences. The retrieval score is computed as:

$$q_i = \text{sim}(\phi(s_t, \tilde{a}_t), \phi(s_i, a_i)), \quad (2)$$

where $\phi(\cdot)$ denotes the representation function.

The retrieved memory is defined as:

$$\mathcal{M}_t = \text{TopK}_{e_i \in \mathcal{M}}(q_i). \quad (3)$$

Unlike conventional memory mechanisms that preserve historical information uniformly, FAAI prioritizes experiences associated with previous failures because they provide direct evidence of potentially harmful decisions.

1.2 Failure Risk Estimation Details

The execution risk score ρ_t estimates the probability that the candidate action may lead to failure:

$$\rho_t = \mathcal{R}(s_t, \tilde{a}_t, \mathcal{M}_t). \quad (4)$$

The risk estimator considers three factors:

$$\rho_t = \alpha d_t + \beta f_t + \gamma c_t, \quad (5)$$

where:

- d_t represents the similarity between the current decision and previous failed decisions;
- f_t represents the historical failure frequency of retrieved experiences;
- c_t represents contextual consistency between the current state and retrieved failure cases.

The risk score is normalized into $[0, 1]$. No additional training or parameter optimization is required.

1.3 Adaptive Memory Management

To maintain bounded memory size, each experience is assigned a utility score:

$$U_i = \alpha R_i + \beta I_i + \eta F_i - \lambda C_i, \quad (6)$$

where:

- R_i denotes task relevance;
- I_i denotes information value;
- F_i denotes failure importance;
- C_i denotes storage cost.

When memory capacity exceeds $M_{\max}$, low-utility experiences are removed:

$$\mathcal{M} \leftarrow \text{Top}_{M_{\max}}(\mathcal{M} \cup e_t; U). \quad (7)$$

This mechanism enables FAAI to preserve informative failure patterns while preventing unlimited memory growth.

Algorithm 1 Failure-Aware Adaptive Inference

```
1: Initialize failure memory  $\mathcal{M}$ 
2: for each timestep  $t$  do
3:   Observe current state  $s_t$ 
4:   Generate candidate action  $\tilde{a}_t \sim \pi_\theta(a|s_t, h_t)$ 
5:   Retrieve related failure experiences  $\mathcal{M}_t$ 
6:   Estimate execution risk  $\rho_t$ 
7:   if  $\rho_t > \tau$  then
8:     Activate self-correction:  $a_t = \mathcal{C}(s_t, \tilde{a}_t, \mathcal{M}_t)$ 
9:   else
10:    Execute candidate action:  $a_t = \tilde{a}_t$ 
11:   end if
12:   Observe feedback and update memory
13: end for
```

1.4 FAAI Algorithm

2 Additional Experimental Details

2.1 Implementation Settings

All experiments use the same pretrained language model, environment configuration, action space, and decoding parameters.

FAAI operates entirely at inference time and does not require:

- parameter updates;
- task-specific fine-tuning;
- additional verifier training.

The evaluation considers four interaction horizons:

$$T \in \{5, 10, 20, 40\}.$$

For each horizon, every method is evaluated on 24 independent episodes.

2.2 Prompt Configuration

The same base instruction template is used for all compared methods. The only difference is the inference strategy applied during execution.

For FAAI, the additional context contains retrieved failure examples:

Previous failure cases: (state, action, outcome)

Analyze whether the candidate action may cause similar failure.

If necessary, revise the action before execution.

3 Additional Results

3.1 Horizon-wise Performance

Table 1: Success rate (%) under different interaction horizons.

Method	T=5	T=10	T=20	T=40
Direct LLM	4.17	0.00	8.33	0.00
CoT	20.83	8.33	12.50	4.17
ReAct	0.00	0.00	0.00	0.00
Reflection	20.83	12.50	8.33	4.17
FAAI	20.83	45.83	29.17	8.33

3.2 Failure Correction Analysis

FAAI correction behavior is analyzed using four metrics:

- correction activation rate;
- action modification rate;
- useful correction rate;
- harmful correction rate.

Across all evaluated trajectories, FAAI activates correction for 65.46% of decisions. Among triggered corrections, 76.35% modify the original action and 59.59% are classified as useful corrections.

4 Case Studies

4.1 Example of Successful Correction

A candidate action is identified as risky because similar historical actions previously resulted in execution failure.

FAAI retrieves the corresponding failure experience and activates self-correction. The revised action avoids the previous failure pattern and completes the task successfully.

4.2 Example of Failure Without Intervention

Without failure-aware intervention, the agent executes an action that appears locally reasonable but causes an irreversible state transition. Subsequent decisions are generated from the degraded state, resulting in failure propagation.