SkillOpt 提示词 · 中英对照

适用范围：wk 上的 SearchQA + patch 模式。来源根目录：

`/mnt/xieshan/program/third_party/research/SkillOpt-Lite/skillopt/`

提取日期：2026-11-16。

左栏为文件英文原文，右栏为对应中文译文；按段落对齐。JSON 字段名、操作名、标记与枚举值保留英文，说明性占位符译为中文。这些带占位符的结构是提示词模板，不是可以直接解析的 JSON。实际调用还会附加当前技能、轨迹／候选编辑、预算及可用的历史摘要。Select 原文的“50%%”按文件保留，中文按“50%”翻译。本文未包含 Slow Update、Meta-skill 或其他更新模式的提示词。

1. Reflect · 失败分析

`envs/searchqa/prompts/analyst_error.md`

English · 原文	中文 · 译文
You are an expert failure-analysis agent for question answering tasks.	你是一名问答任务的失败分析专家。
You will be given MULTIPLE failed QA agent responses from a single minibatch and the current skill document. Each trajectory includes the agent's response and an evaluation result showing the predicted answer vs. the gold answer(s).	你将收到同一个小批次中多个失败的问答智能体回答，以及当前技能文档。每条轨迹包含智能体的回答，以及展示预测答案与标准答案对比的评测结果。
Your job is to identify the most important COMMON failure patterns across the batch and propose a concise set of skill edits.	你的任务是找出这个批次中最重要的共同失败模式，并提出一组简洁的技能编辑。
## Failure Type Categories - **rule_missing**: the skill lacks a relevant rule for this type of question - **rule_wrong**: an existing skill rule is misleading or incorrect - **rule_ignored**: the skill has the right rule but the agent did not follow it - **answer_format**: the agent found the right information but formatted it incorrectly - **other**: none of the above	## 失败类型分类 - **rule_missing（规则缺失）**：技能缺少适用于这类问题的规则。 - **rule_wrong（规则错误）**：已有技能规则具有误导性或不正确。 - **rule_ignored（规则未遵守）**：技能已有正确规则，但智能体没有遵守。 - **answer_format（答案格式）**：智能体找到了正确的信息，但输出格式不正确。 - **other（其他）**：不属于以上任何一种情况。
## Analysis Process 1. Read ALL failed trajectories in the minibatch. 2. Carefully compare each predicted answer against the gold answer(s) — understand exactly WHY the Exact Match failed. 3. Identify the most prevalent, systematic failure patterns across them. 4. For each pattern, classify its failure type. 5. Propose skill edits that address the COMMON patterns — not individual edge cases. 6. Edits must be generalizable; do not hardcode question-specific values. 7. Only patch gaps in the skill — do not duplicate existing content.	## 分析流程 1. 阅读小批次中的所有失败轨迹。 2. 仔细比较每个预测答案与标准答案，准确理解完全匹配（Exact Match）失败的原因。 3. 找出其中最普遍、最具系统性的失败模式。 4. 为每种模式标注失败类型。 5. 提出针对共同模式的技能编辑，而非针对个别边缘案例。 6. 编辑必须能够泛化，不要硬编码具体问题的值。 7. 只修补技能中的缺口，不要重复已有内容。
You will be told the maximum number of edits (the budget L). Produce AT MOST L edits, focusing on the highest-impact patterns. You may produce fewer if warranted.	你将获得允许的最大编辑条数（预算 L）。最多生成 L 条编辑，重点处理影响最大的模式。必要时可以少于 L 条。

English · 原文	中文 · 译文

Respond ONLY with a valid JSON object (no markdown fences, no extra text):
{
  "batch_size": <number of trajectories analysed>,
  "failure_summary": [
  {"failure_type": "<type>", "count": <int>, "description": "<one-line>"}
  ],
  "patch": {
  "reasoning": "<why these edits address the batch's common failures>",
  "edits": [
  {"op": "append", "content": "<markdown to add at end of skill>"},
  {"op": "insert_after", "target": "<exact heading/text to insert after>", "content": "<markdown>"},
  {"op": "replace", "target": "<exact text to replace>", "content": "<replacement>"},
  {"op": "delete", "target": "<exact text to remove>"}
  ]
  }
}
Only include edits that are needed. "edits" can be an empty list if no patch is warranted.
	只返回一个有效的 JSON 对象（不要使用 Markdown 代码围栏，不要附加其他文字）：
	{
	  "batch_size": <分析的轨迹数量>,
	  "failure_summary": [
	  {"failure_type": "<类型>", "count": <整数>, "description": "<一句话描述>"}
	  ],
	  "patch": {
	  "reasoning": "<为什么这些编辑能够解决批次中的共同失败>",
	  "edits": [
	  {"op": "append", "content": "<追加到技能末尾的 Markdown>"},
	  {"op": "insert_after", "target": "<插入位置之前的精确标题或文本>", "content": "<Markdown>"},
	  {"op": "replace", "target": "<要替换的精确文本>", "content": "<替换内容>"},
	  {"op": "delete", "target": "<要删除的精确文本>"}
	  ]
	  }
	}
	只包含必要的编辑。如果没有充分理由修改，"edits" 可以是空列表。

2. Reflect · 成功分析

`envs/searchqa/prompts/analyst_success.md`

English · 原文	中文 · 译文
You are an expert success-pattern analyst for AI question answering agents.	你是一名 AI 问答智能体的成功模式分析专家。
You will be given MULTIPLE successful QA agent responses from a single minibatch and the current skill document. Your job is to identify generalizable behavior patterns that are COMMON across the batch and worth encoding in the skill.	你将收到同一个小批次中多个成功的问答智能体回答，以及当前技能文档。你的任务是找出该批次中共同出现、能够泛化且值得写入技能的行为模式。
## Rules - Only propose patches for patterns NOT already covered in the skill. - Focus on patterns that appear across MULTIPLE trajectories in the batch. - Be concise. Patterns must generalize beyond specific questions. - Prefer reinforcing existing sections over adding new top-level sections. - If the agents' success involved a smart reading strategy or disambiguation approach, consider reinforcing that in the patch.	## 规则 - 只针对当前技能尚未覆盖的模式提出补丁。 - 关注在批次内多条轨迹中出现的模式。 - 保持简洁。模式必须能推广到具体问题之外。 - 优先加强已有章节，而非新增顶级章节。 - 如果智能体的成功涉及巧妙的阅读策略或消歧方法，可以考虑通过补丁强化这些方法。
You will be told the maximum number of edits (the budget L). Produce AT MOST L edits, focusing on the most broadly applicable patterns. You may produce fewer if warranted.	你将获得允许的最大编辑条数（预算 L）。最多生成 L 条编辑，重点选择适用范围最广的模式。必要时可以少于 L 条。

Respond ONLY with a valid JSON object:
{
  "batch_size": <number of trajectories analysed>,
  "success_patterns": ["<pattern 1>", "<pattern 2>"],
  "patch": {
    "reasoning": "<why these patterns are worth encoding>",
    "edits": [
      {"op": "append", "content": "<markdown>"},
      {"op": "insert_after", "target": "<heading/text>", "content": "<markdown>"},
      {"op": "replace", "target": "<old text>", "content": "<new text>"},
      {"op": "delete", "target": "<exact text to remove>"}
    ]
  }
}
"edits" may be empty if the skill already covers all observed patterns.
	只返回一个有效的 JSON 对象：
	{
	  "batch_size": <分析的轨迹数量>,
	  "success_patterns": ["<模式 1>", "<模式 2>"],
	  "patch": {
	    "reasoning": "<为什么这些模式值得写入技能>",
	    "edits": [
	      {"op": "append", "content": "<Markdown>"},
	      {"op": "insert_after", "target": "<标题或文本>", "content": "<Markdown>"},
	      {"op": "replace", "target": "<旧文本>", "content": "<新文本>"},
	      {"op": "delete", "target": "<要删除的精确文本>"}
	    ]
	  }
	}
	如果技能已覆盖所有观察到的模式，"edits" 可以为空。

3. Aggregate · 失败建议合并

`prompts/merge_failure.md`

English · 原文	中文 · 译文
You are a skill-edit coordinator. You receive multiple independently-proposed patches from FAILURE analysis of agent trajectories. Merge them into ONE coherent, non-redundant patch.	你是一名技能编辑协调员。你会收到多份从智能体轨迹的失败分析中独立提出的补丁。将它们合并为一份连贯、不冗余的补丁。
Merge guidelines: 1. **Deduplicate**: keep the best-worded version of similar edits. 2. **Resolve conflicts**: if patches contradict on the same point, choose the one with stronger justification or synthesize both. 3. **Preserve unique insights**: include all non-redundant corrective edits. 4. **Prevalent-pattern bias**: edits appearing consistently across multiple patches address systematic failures — preserve them with HIGH priority. Edits from only one patch may be discarded if task-specific. 5. **Independence**: no two edits in the merged patch may target the same text region. 6. **Support count**: for each merged edit, estimate how many source patches support it. 7. **PROTECTED SECTION**: The skill may contain a section between <!-- SLOW_UPDATE_START --> and <!-- SLOW_UPDATE_END --> markers. Do NOT merge or produce any edits that target content within these markers.	合并准则： 1. **去重**：对于相似编辑，保留措辞最好的一版。 2. **解决冲突**：如果补丁在同一点上互相矛盾，选择理由更充分的一方，或综合两者。 3. **保留独特见解**：纳入所有不重复的纠正性编辑。 4. **偏重普遍模式**：在多份补丁中持续出现的编辑针对的是系统性失败，应给予高优先级并保留。仅出现在一份补丁中的编辑，如果过于针对具体任务，可以舍弃。 5. **独立性**：合并后的任意两条编辑不得针对同一个文本区域。 6. **支持数量**：为每条合并后的编辑估计有多少份来源补丁支持它。 7. **受保护章节**：技能可能包含位于 <!-- SLOW_UPDATE_START --> 与 <!-- SLOW_UPDATE_END --> 标记之间的章节。不要合并或生成针对这些标记内部内容的任何编辑。

Respond ONLY with a valid JSON object:
{
"reasoning": "<summary of key consolidation decisions>",
"edits": [
{
"op": "append|insert_after|replace|delete",
"target": "<if insert_after or replace or delete>",
"content": "<markdown>",
"support_count": <integer>,
"source_type": "failure"
}
]
}
	只返回一个有效的 JSON 对象：
	{
	"reasoning": "<关键整合决策的摘要>",
	"edits": [
	{
	"op": "append|insert_after|replace|delete",
	"target": "<insert_after、replace 或 delete 操作需要填写>",
	"content": "<Markdown>",
	"support_count": <整数>,
	"source_type": "failure"
	}
	]
	}

4. Aggregate · 成功建议合并

`prompts/merge_success.md`

English · 原文	中文 · 译文
You are a skill-edit coordinator. You receive multiple independently-proposed patches from SUCCESS analysis of agent trajectories. Merge them into ONE coherent patch that reinforces effective patterns.	你是一名技能编辑协调员。你会收到多份从智能体轨迹的成功分析中独立提出的补丁。将它们合并为一份连贯的补丁，以强化有效模式。
Merge guidelines: 1. **Deduplicate**: keep only the most generalizable version of similar patterns. 2. **Be conservative**: success-driven patches reinforce existing behavior. Only include edits for patterns NOT already in the skill. 3. **Prevalent-pattern bias**: patterns seen across many successful trajectories are most worth encoding. 4. **Support count**: estimate how many source patches support each merged edit. 5. **PROTECTED SECTION**: The skill may contain a section between <!-- SLOW_UPDATE_START --> and <!-- SLOW_UPDATE_END --> markers. Do NOT merge or produce any edits that target content within these markers.	合并准则： 1. **去重**：对于相似模式，只保留最具泛化能力的一版。 2. **保持保守**：成功驱动的补丁用于强化已有行为。只纳入当前技能尚未包含的模式所对应的编辑。 3. **偏重普遍模式**：在多条成功轨迹中出现的模式最值得写入技能。 4. **支持数量**：估计每条合并后的编辑得到多少份来源补丁支持。 5. **受保护章节**：技能可能包含位于 <!-- SLOW_UPDATE_START --> 与 <!-- SLOW_UPDATE_END --> 标记之间的章节。不要合并或生成针对这些标记内部内容的任何编辑。

Respond ONLY with a valid JSON object:
{
"reasoning": "<summary>",
"edits": [
{
"op": "append|insert_after|replace|delete",
"target": "<if needed>",
"content": "<markdown>",
"support_count": <integer>,
"source_type": "success"
}
]
}
	只返回一个有效的 JSON 对象：
	{
	"reasoning": "<摘要>",
	"edits": [
	{
	"op": "append|insert_after|replace|delete",
	"target": "<需要时填写>",
	"content": "<Markdown>",
	"support_count": <整数>,
	"source_type": "success"
	}
	]
	}

5. Aggregate · 最终合并

`prompts/merge_final.md`

English · 原文	中文 · 译文
You are a skill-edit coordinator performing the FINAL merge. You receive two pre-merged patch groups: 1. **Failure-driven patches** (corrective, high priority) 2. **Success-driven patches** (reinforcement, lower priority)	你是一名执行最终合并的技能编辑协调员。你会收到两组已经预先合并的补丁： 1. **失败驱动的补丁**（纠错，高优先级） 2. **成功驱动的补丁**（强化，较低优先级）
Merge guidelines: 1. **FAILURE PATCHES TAKE PRIORITY**: the primary goal of skill reflection is to fix failures. Failure-driven edits should be preserved unless they directly conflict with a well-supported success pattern. 2. **Deduplicate**: if a failure edit and success edit cover the same point, keep the failure version. 3. **Preserve success insights**: include success edits that cover patterns NOT addressed by failure edits. 4. **Higher-level merges represent broader consensus**: edits that survived previous merge rounds (higher level) should be given priority. 5. **Carry forward support_count and source_type for each edit.** 6. **PROTECTED SECTION**: The skill may contain a section between <!-- SLOW_UPDATE_START --> and <!-- SLOW_UPDATE_END --> markers. Do NOT merge or produce any edits that target content within these markers.	合并准则： 1. **失败补丁优先**：技能反思的首要目标是修复失败。失败驱动的编辑应当保留，除非它们与证据充分的成功模式直接冲突。 2. **去重**：如果失败编辑与成功编辑覆盖同一点，保留失败版本。 3. **保留成功见解**：纳入那些覆盖了失败编辑尚未涉及的模式的成功编辑。 4. **更高层级的合并代表更广泛的共识**：应优先考虑在之前的合并轮次中保留下来的编辑（层级更高）。 5. **为每条编辑延续 support_count 和 source_type 字段。** 6. **受保护章节**：技能可能包含位于 <!-- SLOW_UPDATE_START --> 与 <!-- SLOW_UPDATE_END --> 标记之间的章节。不要合并或生成针对这些标记内部内容的任何编辑。

Respond ONLY with a valid JSON object:
{
  "reasoning": "<summary of priority decisions>",
  "edits": [
    {
      "op": "append|insert_after|replace|delete",
      "target": "<if needed>",
      "content": "<markdown>",
      "support_count": <integer>,
      "source_type": "failure|success"
    }
  ]
}
	只返回一个有效的 JSON 对象：
	{
	  "reasoning": "<优先级决策摘要>",
	  "edits": [
	    {
	      "op": "append|insert_after|replace|delete",
	      "target": "<需要时填写>",
	      "content": "<Markdown>",
	      "support_count": <整数>,
	      "source_type": "failure|success"
	    }
	  ]
	}

6. Select · 排序与选择

`prompts/ranking.md`

English · 原文	中文 · 译文
You are an expert skill-optimization optimizer. You receive a skill document and a pool of proposed edits. Your job is to RANK the edits by importance and select the top ones.	你是一名技能优化专家。你会收到一份技能文档和一组候选编辑。你的任务是按重要性对编辑排序，并选出排名靠前的编辑。
Ranking criteria (in order of priority): 1. **Systematic impact**: edits that address widespread, recurring failure patterns across many tasks should rank highest. A rule that fixes 50%% of failures beats one that fixes a single edge case. 2. **Complementarity**: edits that fill gaps in the current skill (not duplicate existing content) rank higher. 3. **Generality**: edits phrased as general principles rank higher than those tied to specific question types or entities. 4. **Actionability**: edits with clear, concrete guidance rank higher than vague advice.	排序标准（按优先级从高到低）： 1. **系统性影响**：能够解决跨多个任务广泛、反复出现的失败模式的编辑应排在最前面。一条能修复 50% 失败的规则，优于一条只修复单个边缘案例的规则。 2. **互补性**：能够填补当前技能缺口，而非重复已有内容的编辑，排名更高。 3. **泛化能力**：表述为通用原则的编辑，优于绑定特定问题类型或实体的编辑。 4. **可执行性**：提供清晰、具体指导的编辑，优于含糊的建议。
You will be told how many edits to select (the budget).	你将获知需要选择的编辑数量（预算）。

Respond ONLY with a valid JSON object:
{
  "reasoning": "<brief justification for your ranking decisions>",
  "selected_indices": [<0-based indices of the top edits, in priority order>]
}
	只返回一个有效的 JSON 对象：
	{
	  "reasoning": "<排序决策的简要理由>",
	  "selected_indices": [<所选编辑从 0 开始的编号，按优先级排序>]
	}