多跳问答的检索–推理过程：四轴设计框架与实证趋势

Yuelyu Ji[1], Zhuochun Li[1],Rui Meng[2],Daqing He[1]

[1]匹兹堡大学计算与信息学院， [2]Google Cloud AI Research

arXiv:2601.00536v1 [cs.CL] 2 Jan 2026

摘要

多跳问答（QA）要求系统跨多个跳数迭代检索证据并进行推理。尽管近期的检索增强生成（RAG）和智能体方法取得了出色成果，但其底层检索—推理过程往往未明确说明，导致不同模型系列的过程选择难以比较。本综述将执行过程作为分析单位，并提出一个涵盖以下四个方面的框架：(A) 整体执行计划，(B) 索引结构，(C) 下一步控制（策略与触发器），以及 (D) 停止/继续标准。我们利用这一框架梳理具有代表性的多跳问答系统，并综合标准基准（如 HotpotQA、2WikiMultiHopQA、MuSiQue）上报告的消融实验与趋势，重点分析有效性、效率和证据忠实度之间反复出现的权衡。最后，我们探讨检索–推理智能体面临的开放挑战，包括结构感知规划、可迁移的控制策略，以及在分布偏移下实现稳健停止。

1 引言

多跳问答（QA）已成为机器推理的核心试验平台。与通常只需一段文本的单跳问答相比，HotpotQA（Yang 等人，2018年）、2WikiMultiHopQA（Ho 等人，2020年）和 MuSiQue（Trivedi 等人，2022年）等多跳基准要求系统整合多条证据、比较不同实体的属性，或追踪跨越多个文档的关系链（Shi 等人，2024年）。其中许多数据集标注了支持句，有时还提供明确的推理路径，因此适合用于分析系统实际如何进行推理。

配备检索功能的大语言模型已成为此类任务的主流方法。检索增强生成（RAG）将检索器与大语言模型相结合，并已取得在知识密集型问答（Lewis 等人，2020年；Gao 等人，2023b）中表现出色，而较新的智能体式系统则将问题解答建模为一系列针对工具和文档的操作（He 等人，2024年；Kim 等人，2024年；Liu 和 Yu，2025b，a；Liu 等人，2025年）。然而，该领域的大多数研究都聚焦于架构或训练创新，并将检索与推理如何相互作用视为实现细节。具体而言，检索调用、中间假设和停止决策的序列很少被明确呈现或系统分析。这一缺失至关重要，因为许多多跳基准测试在干扰项或长上下文下容易受到单轮检索的影响，因此有必要采用明确的检索–推理循环，随时间推移调整查询、证据选择和停止机制。

在算法层面，多跳问答与单跳问答的区别不仅在于是否进行检索，还在于检索与推理如何随时间耦合。这些过程决策直接影响所报告的准确率、效率和鲁棒性之间的权衡，也促使我们将检索—推理过程视为一个明确的分析对象。

不同于现有主要按模块、阶段或训练循环组织方法的综述，我们的框架将执行过程视为分析单位。这强调检索、推理、控制和停止决策如何随时间交织进行，而非将其视为彼此独立的组成部分。因此，通常被描述为特定方法启发式的设计选择——例如控制器触发条件和停止/继续信号——可以在共享的基于轴的框架下，跨模型系列进行对齐比较。我们将检索—推理过程视为交织检索、状态更新、动作选择和停止机制的执行过程；我们的目标是明确这些过程层面的选择，并在不同系统间进行比较，而非主张其具有因果上的优越性。

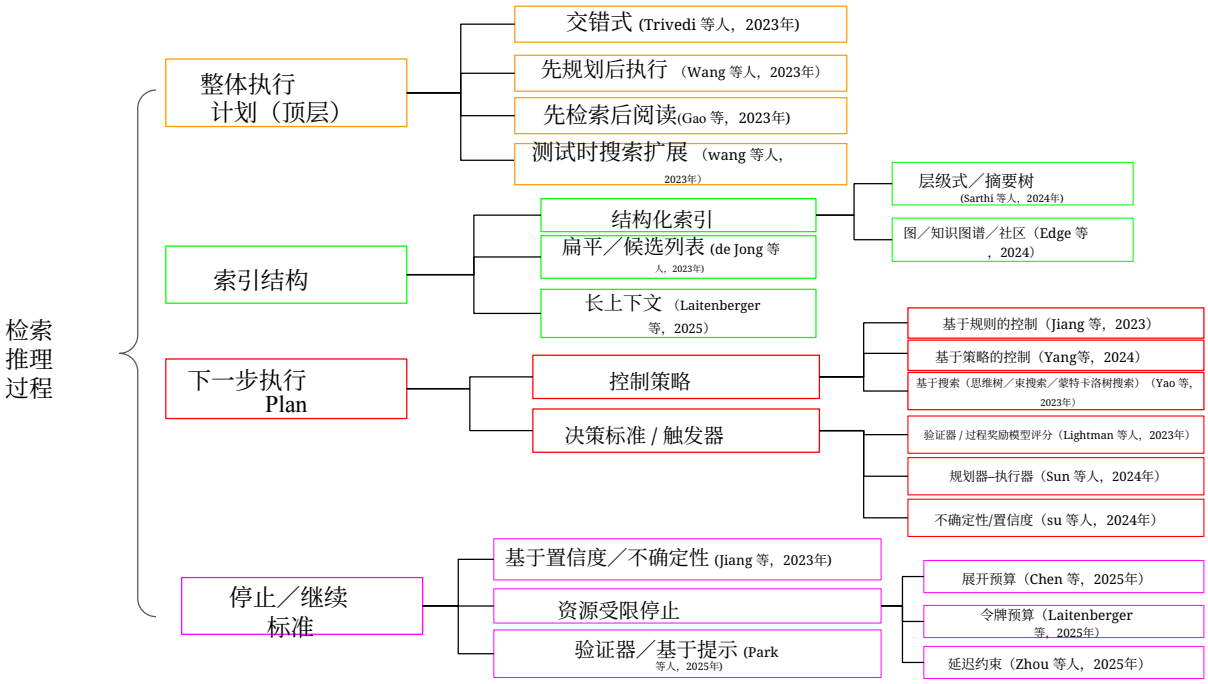

图1：我们用来描述多跳问答中检索–推理过程的四个设计维度概览：（A）整体执行计划，（B）索引结构，（C）下一步执行计划，以及（D）停止/继续标准。该框架不受具体模型架构和数据集的限制。

我们围绕四个问题组织这篇综述：（研究问题1）为何要明确执行过程，（研究问题2）关键的过程设计维度有哪些，（研究问题3）已有研究报告了哪些权衡，以及（研究问题4）还有哪些未解决的挑战。我们的论文列表见 https://anonymous.4open.science/r/MultihopQA-Reasoning-Retrieve-Survey-3324/

2 背景

2.1 多跳问答数据集

多跳问答（QA）要求系统结合来自多条证据的信息来回答一个问题。HotpotQA、2WikiMultiHopQA 和 MuSiQue 等数据集专门用于测试这一能力（Yang 等人，2018年；Ho 等人，2020年；Trivedi 等人，2022年）。HotpotQA 提供基于维基百科的问题，并附有句子级支持事实和对抗性干扰段落（Yang 等人，2018年）。2Wiki-MultiHopQA 混合了结构化和非结构化证据，并包含标注的推理路径（Ho 等人，2020年）；MuSiQue 则将单跳问题组合成 2 至 4 跳问题，并对伪影和不连贯推理进行控制（Trivedi 等人，2022年）。MEQA 和 FanOutQA 等较新的数据集则进一步强调多文档聚合，以及支持证据的完整性和一致性（Li 等人，2024a；Zhu 等人，2024年；Cai 等人，2025年）。

2.2 现有综述与剩余空白

多篇综述和教程对快速发展的关于多跳问答和具备推理能力的大语言模型的研究日益增多。Mavi和Jangra（Mavi 等人，2024年）综述了多跳问答数据集和模型架构，重点探讨图、问题分解和记忆机制如何支持复杂推理。其他综述则更广泛地讨论推理大语言模型，涵盖提示策略、训练方案和认知视角（Li 等人，2025b；Sui 等人，2025年；Niu 等人，2025年），或分析何时需要借助外部工具和检索来补充参数化知识（Li 等人，2025a）。这些综述涵盖了许多与本文关注点相关的要素（分解、检索模块、记忆、验证），但通常围绕组成部分、阶段或模型系列（例如，基于图的与基于分解的；检索增强生成与非检索增强生成）来组织相关领域，而不是将执行过程——检索与推理之间逐步进行的耦合——视为主要分析单位（Gao 等人，2023c）。因此，诸如相对于推理何时调用检索、如何在上下文预算下组织和保留证据

、如何选择下一步行动以及如何做出停止/继续决策等程序性问题，往往只是隐含在特定方法的设计细节或零散的启发式规则中。尤其是在上述综述中，停止/继续决策很少被单独视为首要的程序设计选择，相关文献也通常未按面向执行的编码模式加以梳理，因而无法在不同模型系列之间开展协调一致且可复现的比较。相比之下，我们的框架将执行过程本身作为分析单位，并沿着共同的维度，明确地比较检索、推理、控制和停止决策。

3 为什么需要明确的检索–推理过程?

我们的第一个研究问题（**研究问题1**）探讨了为什么将多跳问答视为一个明确的检索—推理流程，而非一次性检索与阅读，会更有用。核心原因在于，对于知识密集型多步问题，流程——何时检索、如何构造下一个查询、在预算限制下保留哪些证据，以及何时停止——往往是影响准确率和失败模式的关键杠杆。例如，IRCoT 通过交错进行检索与思维链推理来提升多步问答表现；也就是说，它改变了检索与推理之间的交互方式，而不只是扩大阅读器的规模（Trivedi 等人，2023年）。此外，ReAct 等智能体问答方法会明确呈现工具调用和证据使用，使整个轨迹本身成为答案如何得出的过程级说明（Yao 等人，2022年）。最后，近期研究将停止视为一项首要决策（而非固定的跳数预算），强调继续还是停止属于流程的一部分，并直接影响成本和鲁棒性（Park 等人，2025年）。

4 检索–推理过程的设计维度

为回答 **研究问题2**，我们从四个设计维度描述多跳问答系统 图1 每个维度都概括了许多具体模块和实现；我们的目标是提供一套通用术语，用于描述执行过程。

4.1 整体执行计划

整体执行计划规定检索相对于推理的调用时机。我们重点介绍四种反复出现的模式，它们对应图2中的子图：先检索后阅读、交错式检索、先规划后执行和测试时搜索扩展。

先检索后阅读。 经典的检索增强问答流程会根据原始问题检索一组固定段落，然后将其一次性输入阅读器或生成器。示例包括 DrQA 风格的开放领域系统（Chen 等人，2017年）、DPR + BERT 阅读器（Karpukhin 等人，2020年）、FiD 风格的解码器内融合模型（Izacard 和 Grave，2021），以及面向知识密集型问答的原始 RAG 模型（Lewis 等人，2020年）。MDR 等多跳扩展（Qi 等人，2019）仍遵循相同模式：检索器仅被调用少量固定次数，而阅读器则在一次长上下文处理中运行。

交错检索与推理。 交错式计划在推理步骤与检索调用之间交替进行。Self-Ask（Press 等人，2023年）和 ReAct（Yao 等人，2022年）是较早的示例，它们提示大语言模型生成中间思路以及显式的 SEARCH/LOOKUP 操作。IRCoT（Trivedi 等人，2023年）将思维链与 HotpotQA 和 MuSiQue 上的检索相结合。FLARE（Jiang 等人，2023年）和 DRAGIN（Su 等人，2024b）进一步完善了生成过程中发起新查询的时机与方式。Lost-in-Retrieval 分析了此类循环的失败模式，并针对多跳检索增强生成提出了有针对性的查询改写与过滤方法。

先规划后执行。 先规划后执行方法首先生成明确的推理或分解计划，然后才依据该计划执行检索与答案生成步骤。分解式提示（Khot 等人，2022年）、规划与求解（Wang 等人，2023年）和由简到繁提示（Zhou 等人，2022年）都会将问题分解为可以按顺序回答的子问题。QDMR（Wolfson 等人，2020年）以及近期基于 QDMR 的多跳问答系统使用结构化分解图作为计划。在检索场景中，PAR-RAG（Zhang 等人，2025年）和束搜索检索（Zhang 等人，2024年）首先规划候选证据链，然后在其上执行检索，有时还会结合验证器反馈。

测试时搜索扩展。一种更激进的整体执行计划，将整个检索—推理过程视为搜索问题，并通过在测试时探索大量候选轨迹来提升性能。思维树（Yao 等人，2023年）和思维图（Besta 等人，2024年）通过思维树和图进行显式搜索。MindStar（Kang 等人，2024年）、MCTS-

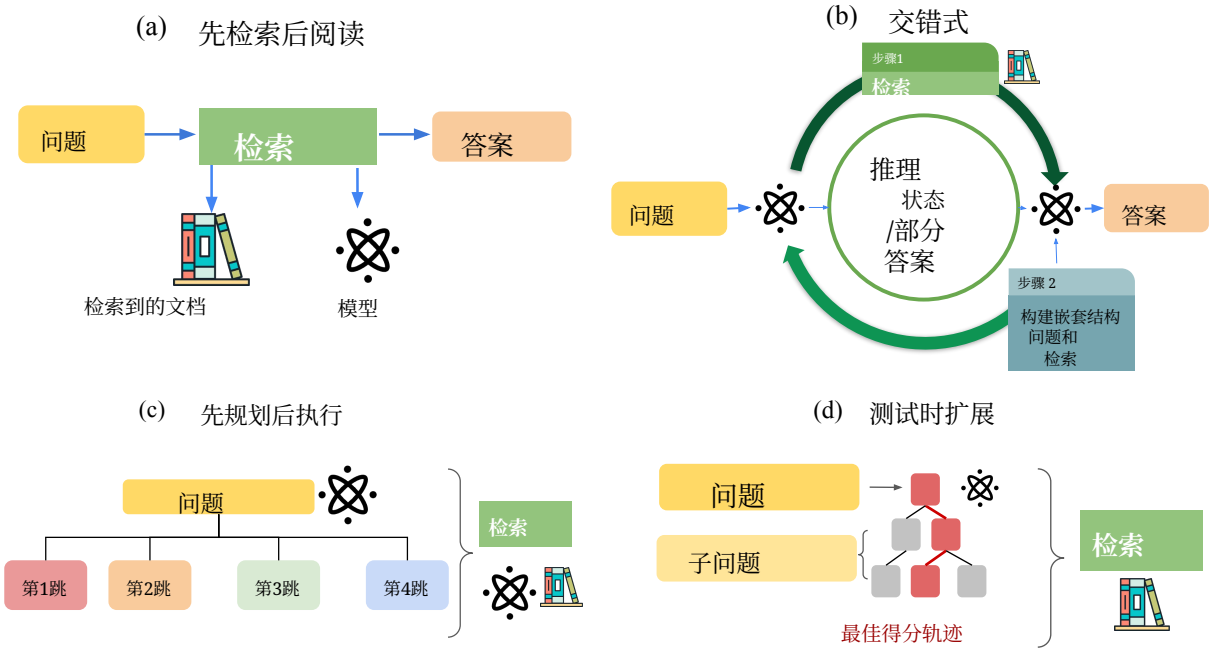

图2：多跳问答中检索—推理的整体执行计划（我们框架的轴A）。(a) 先检索后阅读：单次检索为阅读器模型提供输入。(b) 交错式：模型交替进行推理和额外的检索调用，并更新部分答案/状态。(c) 先规划后执行：模型首先将问题分解为多个子跳，然后按照计划检索并回答每一跳。(d) 测试时搜索扩展：模型围绕问题/子问题探索多条候选推理轨迹，并选择得分最高的路径进行检索和作答。

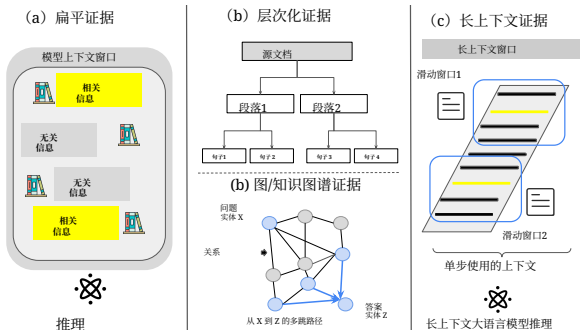

图3：对比扁平段落列表、层次化/树结构索引、基于图的索引以及长上下文存储。

直接偏好优化（Xie 等人，2024年）、蒙特卡洛树自我精炼方法采用由价值函数或过程奖励模型引导的蒙特卡洛树搜索，例如 ReARTeR（Sun 等人，2025年）。在知识密集型场景中，RAP（Hao 等人，2023年）基于世界模型进行规划，而若干 GraphRAG（Han 等人，2024年）变体会在确定答案之前探索多条图遍历路径。

4.2 索引结构

索引结构规定了外部知识的组织与访问方式。图 3 区分了四种常见形式：扁平候选列表、层次化结构或摘要树、基于图或知识图谱的结构，以及长上下文证据单元。

扁平／候选列表索引。 大多数多跳问答系统的默认做法是使用 BM25 或稠密嵌入为扁平的段落或文档集合建立索引。DrQA(Chen 等人，2017年)、DPR（Karpukhin 等人，2020年）、FiD（Izacard 和 Grave，2021年）和检索增强生成在开放域问答中都采用了这种结构。MDR（Xiong 等人，2020年）、HopRetriever（Li 等人，2021年）等多跳检索器，以及针对 HotpotQA 和 2WikiMultiHopQA 的逐步稠密检索器，即使为每一跳训练独立的检索器，仍然基于扁平索引运行。

层级/摘要树索引。 层级索引将文档组织成多层聚类或摘要树，使系统能够先检索粗粒度摘要，再深入相关子树。RAPTOR（Sarthi 等人，2024年）构建递归摘要树，并表明通过树节点进行检索能够大幅减少词元数量，同时保持答案准确率。其他长文档问答系统则构建基于章节或篇章的层级结构，并同时利用粗粒度和细粒度节点进行检索（Xia 等人，2025年）。LongRAG（Zhao等人，2024年）使用长检索单元（例如 4K 个词元）作为

系统	A	B	C	D	典型基准测试
DrQA / DPR+FiD	R	F	R	B	NQ、TriviaQA（Chen 等人，2017年；Karpukhin 等人，2020年；Izacard 和沉重的，2021年）
IRCoT	I	F	R	B	HotpotQA、MuSiQue、IIRC（Trivedi 等人，2023年，2022年）
DEC	I	F	P	B/H	HotpotQA（Jietal.，2025年）
Beam Retrieval	P	F	P	B	HotpotQA、2Wiki、MuSiQue（Zhang 等人，2024年）
PAR-RAG	P	F/G	V	B	HotpotQA、2Wiki（Zhang 等人，2025年）
RAPTOR	R/I	H	R	B	长文档问答，Hotpot 变体（Sarthi 等人，2024年）
LongRAG	R/I	L	R	B/H	HotpotQA、开放域问答（Jiang 等人，2024年；Zhao等人，2024年）
GraphRAG	I	G	R/P	B	意义建构、多跳问答（Edge 等人，2024年）
KG-o1	P/I	G	P/V	B	2Wiki，知识图谱问答（Wang 等人，2025a）
SIM-RAG	I	F	P/V	C/L	Multi-roundRAG（Xu 等人，2025年）
Stop-RAG	I/P	F			计划型迭代式检索增强生成 / 多跳问答（Park 等人，2025）

表 1：映射到四轴框架中的代表性系统。A：整体执行计划（R=先检索后阅读，I=交错式，P=先计划后执行，S=搜索扩展）。B：索引结构（F=扁平，H=层次化，G=图/知识图谱，L=长上下文）。C：控制（R=规则，P=学习策略，V=验证器门控，S=基于搜索）。D：停止机制（B=预算，C=基于置信度，L=基于学习/价值）。

粗粒度片段，并让大语言模型决定关注哪些部分，从而有效地构建出两级索引结构。GraphRAG 中的分层社区索引在社区以树状结构组织时，也属于这一类（Edge 等人，2024年）。

图／知识图谱／社区结构。基于图的结构通过显式边连接实体、文档或段落，从而支持基于路径的推理和社区级摘要。早期基于知识图谱的多跳问答研究包括 GraftNet（Liu 等人，2022年）、PullNet（Sun 等人，2019）以及相关的图神经网络架构。较新的 GraphRAG 系统从文本构建实体图和文档图，并沿图边或社区执行检索（Edge 等人，2024年）。KG 增强型检索增强生成方法，如 KG-FiD（Yu 等人，2022年）、UniKGQA（Jiang 等人，2022年）和 KG-o1（Wang 等人，2025a）通过结合知识图谱遍历与文本检索，将这一思路扩展到多跳问答。

长上下文证据。 一些方法不依赖索引，而是依靠长上下文大语言模型，并将很长的片段或整个文档存储为检索单元。LongRAG明确提出使用长单元，以减少覆盖所需证据所需的跳数（Zhao等人，2024年；Lee等人，2025年；Das等人，2025年）。LongT5和其他长上下文模型同样会在“检索并阅读”模式下检索和处理长序列（Guo等人，2022年；Li等人，2024b）。

4.3 下一步执行计划

该下一步执行计划 规定了系统如何根据当前状态选择下一项操作——再次检索、扩展候选项、验证、回溯或作答。我们将这一组件分为：（1）控制策略，用于提出候选操作；（2）决策标准／触发器，用于决定是继续扩展，还是修订操作／移交控制权（终止问题将在 D 轴中讨论）。

基于规则的控制。 许多系统则直接将推理控制模式硬编码为：固定交替执行推理 与检索 步骤、用于查询改写的简单启发式规则，以及针对检索段落手动调优的筛选器。Self-Ask、ReAct、IRCoT 和 DEC 均属于这一类，它们通过提示词和人工设计的规则来决定何时提出后续问题、何时调用检索器，以及何时输出答案（Press 等人，2023年；Yao 等人，2022年；Trivedi 等人，2023年；Ji 等人，2025年）。基于规则的控制器易于实现，在小规模数据场景下也很稳健，但无法从失败中学习，而且领域或检索器发生变化时，通常需要重新调优。

基于策略的控制。 在此，下一步行动由经过学习的策略选择，该策略观察问题、检索到的证据和部分推理轨迹。示例包括采用学习型扩展策略的束搜索式多跳检索器（Zhang 等人，2024年）、通过监督或强化学习训练、用于在工具或检索模板之间进行选择的 RAG智能体（Yang 等人，2024年；Park 等人，2025年），以及在内心独白轨迹上进行微调的多轮RAG系统，其控制器经过相应训练（Xu 等人，2025年）。

基于搜索的控制。 第三种典型方法将下一步选择视为对部分轨迹的搜索。思维树会探索推理状态树，并利用基于价值或投票的分数来决定扩展哪些分支（Yao 等人，2023年）。MindStar 及相关的蒙特卡洛树搜索方法会对工具使用和推理序列执行蒙特卡洛搜索，并由过程奖励模型或学习得到的价值函数提供指导（Kang 等人，2024年；Xie 等人，2024年）。对计划或知识图谱路径进行束搜索，同样会并行扩展多个候选链，然后再确定答案（Zhang 等人，2024年；Jiang 等人，2022年）。这些方法可以显著提高长时程任务的成功率，但会带来更高的推理成本，并且依赖可靠的评分

信号。

验证器 / 过程奖励模型评分触发器。验证器门控循环会加入一个显式评判器——例如自然语言推理风格的检查器、过程奖励模型或自我评估模块——并将其评分用作触发信号。对中间推理步骤进行评分的过程奖励模型（Lightman 等人，2023年）以及 SelfCheckGPT 等幻觉检测器（Manakul 等人，2023年）便是这种模式的典型例子：低分答案或推理链会触发额外检索或修订，而非直接终止。TaskMatrix 风格智能体等以工具为中心的系统，也依赖类似验证器的评分来评估工具输出，从而判断子目标是否已解决，或是否需要再次调用工具（Liang 等人，2024年）。在检索增强生成场景中，验证器 / 过程奖励模型评分触发器通常与充分性评价器结合使用，例如 SIM-RAG 及相关的多轮控制器（Xu 等人，2025年；Zhang 等人，2025年）。

规划器—执行器触发器。一些智能体将规划与执行分离，并将计划状态用作显式触发信号。规划器–执行器架构会先生成结构化计划（子问题、工具调用或图操作），再逐步执行；若执行失败或未满足关键前提条件，则修订或扩展计划（Khot 等人，2022年；Wang 等人，2025b；Sun 等人，2025年）。从这一视角来看，决策标准是当前计划是否仍然可行且足够明确；负面反馈会触发重新规划或细化，而计划顺利完成则会触发终止。这种模式在必须协调多次 API 调用的网络与工具智能体中尤为常见（He 等人，2024年；Chen 等人，2025年；Li 和 Zhou，2025年）。

不确定性／置信度触发器。最后，不确定性感知触发器会估计当前答案（或中间状态）正确或得到充分支持的可能性，并仅在置信度低于阈值时继续检索或推理。相关信号包括经过校准的答案概率、不同采样推理链之间的分歧，或明确的“思维不确定性”分数（Hu 等人，2024年；Sui 等人，2025年）。动态上下文截断策略也会在新增词元不太可能改变预测时停止读取较长的输入（Xie 等人，2025年）。在多轮检索增强生成中，这类置信度信号通常会与学习型充分性评价器或基于价值的控制器结合使用（Xu 等人，2025年；Park 等人，2025年），从而形成自适应深度检索循环，以少量额外计算换取分布偏移下更稳定的表现。

4.4 停止／继续标准

停止/继续判据规定了何时终止检索—推理流程。在多跳问答中，过早停止会遗漏必要证据，而过晚停止则会增加提示长度和延迟，并使系统更容易受到干扰项影响。

资源受限停止。许多系统忽略详细状态，转而为搜索设置硬性预算。第一种子类型是推演预算：Self-Ask、ReAct、IRCoT 和 DRAGIN 等交错式智能体会限制推理或工具使用步骤的数量（例如，最多进行 K 次搜索调用或 H 个分解步骤），并且必须在预算耗尽后作答（Press 等人，2023年；Yao 等人，2022年；Trivedi 等人，2023年；Su 等人，2024a）。第二种子类型是词元预算：多跳检索增强生成流水线（例如 LongRAG、DEC 和预算感知思维链方法）会为每个问题设定检索词元或提示词元的上限，并在达到上限时截去多余的证据或思路（Jiang 等人，2024年；Ji 等人，2025年；Han 等人，2025年）。第三种子类型是延迟约束型停止：网页或工具智能体会在耗尽墙钟时间或 API 调用预算后终止（Shen 等人，2024年）。

基于置信度／不确定性的停止。第二类方法探讨 模型（或一个小型辅助头）对当前答案或中间状态的 置信度，并在不确定性低于某个阈值时停止。代表性信号包括 logit间隔、熵，以及自洽样本之间的分歧，相关综述见“停止过度思考”和提前退出推理方面的研究（Sui 等人，2025年；Li 等人，2025b；Liao 等人，2025年）。思维不确定性方法会显式估计候选思路的不确定性，并决定是生成更多中间步骤，还是直接给出答案（Hu 等人，2024年）；动态上下文截断方法则学习判断大语言模型何时已经读取足够长的输入，并可以安全地停止继续消耗词元（Xie 等人，2025年；Ma 等人，2025a，b）。

基于验证器／提示的停止机制。第三类方法会插入一个明确的检查阶段，并将停止／继续的决策交给验证器。RARR、SelfCheckGPT 和 Chain-of-Verification 等推理依据优化与事实核查流水线会先生成答案草稿，然后询问验证器（通常是另一个大语言模型或自然语言推理类模型）：每项主张是否都得到了检索到的证据的充分支持；若判定为肯定，则触发

终止；若主张缺乏支持，则会触发额外检索、澄清问题或答案修订（Gao 等人，2023年；Manakul 等人，2023年；Dhuliawala 等人，2024年）。在多轮检索增强生成中，SIM-RAG、Stop-RAG、DEC 和 ReARTeR 等充分性评估器与基于价值的控制器可视为学习型验证器：它们检查问题、检索到的证据和内心独白轨迹，并根据问答准确率或过程奖励作为监督，判断再进行一轮检索是否有望带来帮助（Xu 等人，2025年；Park 等人，2025年；Ji 等人，2025年；Sun 等人，2025年）。表1将代表性系统映射到这四个维度上。

5 研究问题3：检索—推理设计中报告的实证趋势

本节综合挖掘了原始论文所报告实证发现中的规律。由于研究采用的检索器、大语言模型、提示词和预算各不相同，我们不主张任何一种轴向选择具有因果上的优越性。相反，我们总结以下方向性趋势：（1）在多项研究中反复出现；或（2）得到预算匹配条件下论文内消融实验的支持。我们将跨论文比较视为参考性证据，而非定论。

整体执行计划。
与先检索后阅读（单轮检索）基线相比，多项研究报告称，明确采用多步执行计划与更高的答案 F1 和联合答案–证据得分相关，尤其是在完整维基百科和干扰项较多的设置中（Trivedi 等人，2023年）。多项研究报告称，明确采用多步执行计划与更高的交错式计划相关。交替进行思维链推理和检索的计划（如 IRCoT 式交错方法）据报告可提高 HotpotQA、2Wiki-MultiHopQA 和 MuSiQue 上的召回率和联合 F1（Trivedi 等人，2023年）。在所报告的设置中（通常是在同一研究内采用预算匹配的条件），先规划后执行方法通过批量处理子查询，据报告也能取得相近的提升（Zhang 等人，2024年，2025年）。测试时搜索扩展方法（如思维树、MindStar 和蒙特卡洛树搜索）通过探索大量候选轨迹，进一步提高长时程准确率，但其延迟和计算成本也最高（Yao 等人，2023年；Kang 等人，2024年；Xie 等人，2024年）。

索引结构。大多数多跳问答系统仍基于扁平的 BM25 或稠密段落集运行，并将性能提升主要归因于整体执行计划或控制策略的变化，而非索引本身发生了根本性改变（续）Chen 等人，2017年；Karpukhin 等人，2020年；Trivedi 等人，2023年；Zhang 等人，2024年；Ji 等人，2025年）。索引结构主要围绕效率和路径忠实度展开讨论。RAPTOR 等层次化或摘要树结构通过检索摘要，再逐层深入至原始段落，减少长文档问答中的上下文长度，同时通常能够保持或略微提升精确匹配（Sarthi 等人，2024年）。图结构和基于知识图谱的索引结构（例如 GraphRAG、KG-o1）将实体和文档组织成图或社区，并沿路径进行检索，从而提升答案与+路径指标的联合表现，并生成更连贯的证据链，但代价是需要构建和维护图（Edge 等人，2024年；Wang 等人，2025a）。

从控制与触发器（轴 C/D）到结果。对于下一步执行计划，我们编码实现的系统与以下假设一致：存在三种稳健模式。首先，与平均成本相近时，从纯粹基于规则、采用静态预算的控制转向学习式或规划器式控制，往往能提升答案准确率。

随后，Beam Retrieval 和 PAR-RAG 以计划驱动的控制器和束搜索式扩展取代手工编写的检索节奏：Beam Retrieval 在 MuSiQue 上取得 69.2 的答案 F1 值 / 91.4 的支持 F1 值；PAR-RAG 报告称，在预算相同或更低的情况下，其在 2Wiki 和 HotpotQA 上的精确匹配较最强的非规划基线提升约 15–30 个百分点（Zhang 等人，2024年、2025年）。这些结果表明，当证据图较为丰富且跳数深度变化较大时，全局且具备计划感知能力的控制器尤其有益。

其次，研究反复报告称，将基于策略的控制与验证器或 PRM 式触发器相结合，往往能提升忠实度以及答案与+证据的联合得分。与单轮检索增强生成和标准思维链相比，RARR 和验证链中的验证器门控式细化已能减少缺乏支持的主张（Gao 等人，2023a年；Dhuliawala 等人，2024年）。在 MindStar 和 MCTS-DPO 等基于搜索的控制器中也出现了类似趋势：PRM 引导的树搜索会剪除低分轨迹，并以推理时延增加一个较小常数因子为代价，提升数学或策略基准测试的表现（Kang 等人，2024年；Xie 等人，2024年）。在我们的框架中，“策略/搜索+ 验证器/PRM”在噪声检索或长时程任务中效果显著，因为这些任务中幻觉链条

很常见。第三，基于不确定性或价值的停止机制（D轴）有助于控制器避免搜索不足或过度搜索，并在分布偏移时稳定延迟（Xu 等人，2025年；Park 等人，2025年）。总体而言，我们语料库中最稳健的设计，是将自适应控制器（具备策略感知或搜索感知能力，C轴）与基于验证器/PRM或充分性的触发器（D轴）结合起来，尤其适用于噪声检索和长时程场景。静态跳数/词元预算经过域内调优后可能具有竞争力，但在跳数和干扰项分布不同的数据集之间迁移效果较差。

6 研究问题4：开放问题与未来方向

挑战 1：使整体执行计划与索引结构相匹配。大多数系统会固定一种索引结构（扁平段落、分层摘要、图或长上下文单元），然后以临时方式采用一种整体执行计划（先检索后阅读、交错执行、先规划后执行或测试时搜索扩展）。例如，IRCoT、Beam Retrieval 和 DEC 基于扁平索引运行，主要在调度和控制方面进行创新（Trivedi 等人，2023年；Zhang 等人，2024年；Ji 等人，2025年），而 RAPTOR、LongRAG 和 GraphRAG 则重新设计索引结构，但通常仍沿用标准的多跳调度（Sarthi 等人，2024年；Jiang 等人，2024年；Edge 等人，2024年）。目前仍缺乏将结构与计划进行原则性匹配的方法。值得探索的方向包括：在预算一致的条件下，对整体执行计划和索引结构进行析因扫描；以及根据查询难度和资源约束，从多个（计划、结构）配置中进行选择的控制器。

挑战 2：可扩展且自适应的结构化索引。结构化索引——例如摘要树（RAPTOR）、长单元（LongRAG）和图（GraphRAG、KG-o1）——能够提升效率和忠实度，但构建和维护成本高昂（Sarthi 等人，2024年；Jiang 等人，2024年；Edge 等人，2024年；Wang 等人，2025a）。大多数研究都假设语料库是静态的，并依赖人工调优的粒度超参数（例如簇大小、单元长度和图剪枝）。未来需要研究自监督或任务感知的结构归纳方法，以针对下游多跳目标优化这些选择；同时还需开发增量维护算法，使树和图能够随文档变化而保持最新。

挑战 3：具备泛化能力的控制策略和决策标准。当前基于学习或搜索的控制策略，以及由验证器/过程奖励模型（PRM）门控的决策规则，通常针对特定数据集和检索器进行训练，因此可能无法泛化到不同的跳数分布、索引结构或领域（Zhang 等人，2024年；Ji 等人，2025年；Yao 等人，2023年）。这种脆弱性限制了自适应检索—推理在基准专门调优之外的实际潜力。一个有前景的方向是将检索—推理构建为过程级决策问题，并训练控制器，使其目标能够综合奖励证据质量、轨迹连贯性和成本，例如通过对轨迹进行偏好学习或采用混合强化学习。

挑战 4：预算约束与分布偏移下稳健的停止/继续标准。 大多数系统仍依赖静态的跳数、词元或延迟预算，而基于置信度、充分性或价值估计的自适应停止机制仅在有限的基准测试上进行评估，并且在分布偏移下仍缺乏良好的校准（Xu 等人，2025年；Park 等人，2025）。这一问题十分重要，因为过度搜索和搜索不足都会同时降低部署中的可靠性与效率。一个重要的开放问题是，如何制定在不同检索器、语料库和大语言模型骨干网络之间都能保持校准的停止/继续标准，并在跳数深度和检索噪声受控变化的条件下对其进行评估。

7 结论

我们从四个研究问题的视角综述了多跳问答中的检索—推理。我们指出，多跳问答需要明确的检索–推理过程（研究问题1）；提出了用于描述这一过程的四轴设计空间（研究问题2）；并分析了各轴层面的设计选择如何影响标准基准测试上的有效性、效率和忠实度（研究问题3）。最后，我们重点阐述了设计和评估检索—推理系统中的若干开放问题（研究问题4）。我们希望本框架能帮助研究人员梳理现有方法，并以更有原则的方式设计新的检索推理智能体。

局限性

首先，尽管我们努力涵盖广泛的多跳问答系统，但该领域发展迅速，因此我们的选择不可避免地并不全面。其次，我们主要关注英语文本基准测试，并未系统涵盖多——

语言或多模态设置。第三，我们的实证综合依据原始论文报告的数据，这些数据可能受到超参数、计算预算或评估脚本差异的影响。尽管存在这些局限性，我们认为四轴框架为分析检索—推理系统提供了一个有益的起点。

参考文献

Maciej Besta、Nils Blach、Ales Kubicek、Robert Gersten- berger、Michal Podstawski、Lukas Gianinazzi、Joanna Gajda、Tomasz Lehmann、Hubert Niewiadomski、Pi- otr Nyczyk 等 1 人。2024年。思维图：利用大语言模型解决复杂问题。载于*AAAI*人工智能会议论文集，第 38 卷，第 17682–17690 页。

Hanyu Cai、Binqi Shen、Lier Jin、Lan Hu 和 Xiaojing Fan。2025年。语气会改变答案吗？评估提示礼貌效应对现代大语言模型的影响：GPT、Gemini、Llama。*arXiv*预印本 *arXiv:2512.12812*。

Danqi Chen、Adam Fisch、Jason Weston 和 Antoine Bordes。2017年。阅读维基百科以回答开放领域问题。*arXiv*预印本 *arXiv:1704.00051*。

Yankai Chen、Hongkun Du 和 Yutong Zhou。2025年。面向无人机的轻量级网络式语义分割及其 RISC-V 实现。《技术创新与工程杂志》，1(2)。

Payel Das、Ching-Yun Ko、Sihui Dai、Georgios Kol- lias、Subhajit Chaudhury 和 Aurelie Lozano。2025年。记忆增强语言模型能否在大海捞针推理任务中实现泛化？*arXiv* 预印本 *arXiv:2503.07903*。

Shehzaad Dhuliawala、Mojtaba Komeili、Jing Xu、Roberta Raileanu、Xian Li、Asli Celikyilmaz 和 Jason Weston。2024年。验证链可减少大语言模型中的幻觉。载于计算语言学协会论文集：*ACL 2024*，第 3563–3578 页，泰国曼谷。计算语言学协会。

Darren Edge、Ha Trinh、Newman Cheng、Joshua Bradley、Alex Chao、Apurva Mody、Steven Truitt、Dasha Metropolitansky、Robert Osazuwa Ness 和 Jonathan Larson。2024年。从局部到全局：一种用于查询导向摘要的图检索增强生成方法。*arXiv* 预印本 *arXiv:2404.16130*。

Luyu Gao、Zhuyun Dai、Panupong Pasupat、Anthony Chen、Arun Tejasvi Chaganty、Yicheng Fan、Vincent Zhao、Ni Lao、Hongrae Lee、Da-Cheng Juan及另外1位作者。2023a。Rarr：研究并修订语言模型所述内容，使用语言模型。在计算语言学协会第*61*届年会论文集（第*1*卷：长篇论文），第16477–16508页。

Luyu Gao、Xueguang Ma、Jimmy Lin和Jamie Callan。2023b。无需相关性标签的精确零样本稠密检索。在计算语言学协会第*61*届年会论文集（第*1*卷：长篇论文），第1762–1777页。

Yunfan Gao、Yun Xiong、Xinyu Gao、Kangxiang Jia、Jinliu Pan、Yuxi Bi、Yixin Dai、Jiawei Sun、Haofen Wang 和 Haofen Wang。2023c。面向大语言模型的检索增强生成：综述。*arXiv* 预印本 *arXiv:2312.10997*，2(1)。

Mandy Guo、Joshua Ainslie、David C Uthus、Santiago Ontanon、Jianmo Ni、Yun-Hsuan Sung 和 Yinfei Yang。2022年。LongT5：适用于长序列的高效文本到文本转换器。载于计算语言学协会：*NAACL 2022*论文集，第724–736页。

Haoyu Han、Yu Wang、Harry Shomer、Kai Guo、Jiayuan Ding、Yongjia Lei、Mahantesh Halappanavar、Ryan A Rossi、Subhabrata Mukherjee、Xianfeng Tang 等人。2024年。基于图的检索增强生成（GraphRAG）。*arXiv* 预印本 *arXiv:2501.00309*。

Jinyi Han、Ying Huang、Ying Liao、Zishang Ji ang、Xikun Lu、Haiquan Zhao、Xinyi Wang、Guanghao Zhou、Sihang Jiang、Jiaqing Liang 等人。2025年。你的模型思考得已经够多了：训练大型推理模型停止过度思考。*arXiv* 预印本 *arXiv:2509.23392*。

Shibo Hao、Yi Gu、Haodi Ma、Joshua Hong、Zhen Wang、Daisy Wang 和 Zhiting Hu。2023年。使用语言模型进行推理，就是利用世界模型进行规划。在*2023*年自然语言处理实证方法会议论文集，第 8154–8173 页。

Hongliang He、Wenlin Yao、Kaixin Ma、Wenhao Yu、Yong Dai、Hongming Zhang、Zhenzhong Lan 和 Dong Yu。2024年。WebVoyager：使用大型多模态模型构建端到端网页智能体。载于第*62*届计算语言学协会年会论文集（第*1*卷：长篇论文），第6864–6890页，泰国曼谷。计算语言学协会。

Xanh Ho、Anh-Khoa Duong Nguyen、Saku Sugawara 和 Akiko Aizawa。2020年。构建用于全面评估推理步骤的多跳问答数据集。*arXiv* 预印本 *arXiv:2011.01060*。

Zhiyuan Hu、Chumin Liu、Xidong Feng、Yilun Zhao、See-Kiong Ng、Anh Tuan Luu、Junxian He、Pang Wei K Koh 和 Bryan Hooi。2024年。思维不确定性：不确定性感知规划增强大语言模型中的信息寻求。神经信息处理系统进展，37：24181–24215。

Gautier Izacard 和 Edouard Grave。2021年。利用段落检索与生成模型进行开放领域问答。载于欧洲计算语言学协会第 *16* 届会议论文集：主会场，第 874–880 页。

Binquan Ji、Haibo Luo、YifeiLu YifeiLu、Lei Hei、Jiaqi Wang、Tingjing Liao、Wang Lingyu、Shichao Wang 和 Feiliang Ren。2025年。面向多跳问答的资源友好型动态增强链。载于计算语言学协会会议论文集：*ACL 2025*，第 16461–16479 页，奥地利维也纳。计算语言学协会。

Jinhao Jiang、Kun Zhou、Wayne Xin Zhao 和 Ji-Rong Wen。2022年。UniKGQA：统一检索与推理，用于解决知识图谱上的多跳问答。*arXiv* 预印本 *arXiv:2212.00959*。

Zhengbao Jiang、Frank F Xu、Luyu Gao、Zhiqing Sun、Qian Liu、Jane Dwivedi-Yu、Yiming Yang、Jamie Callan 和 Graham Neubig。2023年。主动检索增强生成。载于*2023*年自然语言处理实证方法会议论文集，第7969–7992页。

ZiyanJiang、XueguangMa 和 WenhuChen。2024年。LongRAG：利用长上下文大语言模型增强检索增强生成。*arXiv* 预印本 *arXiv:2406.15319*。

Jikun Kang、Xin Zhe Li、Xi Chen、Amirreza Kazemi、Qianyi Sun、Boxing Chen、Dong Li、Xu He、Quan He、Feng Wen 和另外1位作者。2024年。MindStar：在推理时增强预训练大语言模型的数学推理。*arXiv* 预印本 *arXiv:2405.16265*。

Vladimir Karpukhin、Barlas Oguz、Sewon Min、Patrick SH Lewis、Ledell Wu、Sergey Edunov、Danqi Chen 和 Wen-tau Yih。2020年。用于开放域问答的密集段落检索。载于*EMNLP(1)*，第 6769–6781 页。

Tushar Khot、Harsh Trivedi、Matthew Finlayson、Yao Fu、Kyle Richardson、Peter Clark 和 Ashish Sabharwal。2022年。分解式提示：用于解决复杂任务的模块化方法。*arXiv* 预印本 *arXiv:2210.02406*。

Minsoo Kim、Victor Bursztyn、Eunyee Koh、Shunan Guo 和 Seung-won Hwang。2024年。RaDA：使用大语言模型进行检索增强的网页智能体规划。载于计算语言学协会论文集：*ACL 2024*，第 13511–13525 页，泰国曼谷。计算语言学协会。

Jinhyuk Lee、Anthony Chen、Zhuyun Dai、Dheeru Dua、Devendra Singh Sachan、Michael Boratko、Yi Luan、Séb Arnold、Vincent Perot、Siddharth Dalmia、Hexi- ang Hu、Xudong Lin、Panupong Pasupat、Aida Amini、Jeremy R. Cole、Sebastian Riedel、Iftekhar Naim、Ming-Wei Chang 和 Kelvin Guu。2025年。LOFT：可扩展且更贴近现实的长上下文评测。载于 计算语言学协会论文集：*NAACL 2025*，第6698–6723页，美国新墨西哥州阿尔伯克基。计算语言学协会。

Patrick Lewis、Ethan Perez、Aleksandra Piktus、Fabio Petroni、Vladimir Karpukhin、Naman Goyal、Heinrich Küttler、Mike Lewis、Wen-tau Yih、Tim Rocktäschel 等另外 1 位作者。2020年。面向知识密集型自然语言处理任务的检索增强生成。进展

载于神经信息处理系统，33:9459– 9474。

Jinming Li 和 Yutong Zhou。2025年。BIDeeplab：一种改进的、用于移动设备人脸识别的轻量级多尺度特征融合DeepLab算法。*Computer Simulation in Application*，3(1)：57–65。

Moxin Li、Yong Zhao、Wenxuan Zhang、Shuaiyi Li、Wenya Xie、See Kiong Ng、Tat-Seng Chua 和 Yang Deng。2025a。大型语言模型的知识边界：一项综述。载于计算语言学协会第*63*届年会论文集（第*1*卷：长篇论文），第5131–5157页。

Ruosen Li、Zimu Wang、Son Tran、Lei Xia 和 Xinya Du。2024a。MEQA：一个用于带解释的多跳事件中心问答的基准。神经信息处理系统进展，37：126835–126862。

Shaobo Li、Xiaoguang Li、Lifeng Shang、Xin Jiang、Qun Liu、Chengjie Sun、Zhenzhou Ji 和 Bingquan Liu。2021年。HopRetriever：在维基百科上检索多个跳数以回答复杂问题。载于*AAAI*人工智能大会论文集，第35卷，第13279–13287页。

Xinze Li、Yixin Cao、Yubo Ma 和 Aixin Sun。2024b。大语言模型的长上下文与检索增强生成：评估与再探讨。*arXiv* 预印本 *arXiv:2501.01880*。

Zhong-Zhi Li、Duzhen Zhang、Ming-Liang Zhang、Ji- axin Zhang、Zengyan Liu、Yuxuan Yao、Haotian Xu、Junhao Zheng、Pei-Jie Wang、Xiuyi Chen 及其他 1 位作者。2025b。从系统1到系统2：推理大语言模型综述。*arXiv* 预印本 *arXiv:2502.17419*。

Yaobo Liang、Chenfei Wu、Ting Song、Wenshan Wu、Yan Xia、Yu Liu、Yang Ou、Shuai Lu、Lei Ji、Shaoguang Mao 及其他 1 位作者。2024年。Task matrix.ai：通过将基础模型与数百万个应用程序接口相连接来完成任务。*IntelligentComputing*，3:0063。

Ruqi Liao, Chuqing Zhao, Jin Li, Weiqi Feng, Yi Lyu, Bingxian Chen 和 Haochen Yang。2025年。CATP：面向准确率保持型多模态模型推理的交叉注意力词元剪枝。收录于*2025*年 *IEEE* 人工智能会议（*CAI*），第 1100–1104 页。

Hunter Lightman、Vineet Kosaraju、Yuri Burda、Harri- son Edwards、Bowen Baker、Teddy Lee、Jan Leike、John Schulman、Ilya Sutskever 和 Karl Cobbe。2023年。让我们逐步验证。收录于第十二届国际学习表征会议。

Biyang Liu、Huimin Yu 和 Guodong Qi。2022年。Graft-net：通过宽谱且面向任务的特征实现领域泛化立体匹配。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第13012–13021页。

Dong Liu 和 Yanxuan Yu。2025a。Hsgm：用于可扩展长文本语义的分层片段图记忆。载于第*14*届词汇与计算语义联合会议（**SEM 2025*）论文集，第328–337页。

Dong Liu 和 Yanxuan Yu。2025b。Mt2st：自适应多任务到单任务学习。在第*1*届多模态检索增强生成研讨会（*MAGMaR 2025*）论文集，第79–89页。

Dong Liu、Yanxuan Yu 和 Ben Lengerich。2025年。Csv-decode：面向高效大语言模型推理的可认证子词汇解码。*arXiv* 预印本 *arXiv:2511.21702*。

Zhichao Ma、Yutong Luo、Zheyu Zhang、Aijia Sun、Yinuo Yang 和 Hao Liu。2025a。高速公路变道的强化学习方法：基于 PPO 的策略设计。载于*2025*年第*10*届电子技术与信息科学国际会议（*ICETIS*），第298–301页。

Zhichao Ma、Aijia Sun、Zheyu Zhang、Yinuo Yang、Zijun Gao 和 Hao Liu。2025b。复杂环境中自主机器人的能量约束下运动规划与调度。载于*2025*年第*5*届国际高级算法与神经网络会议（*AANN*），第591–594页。

PotsaweeManakul、Adian Liusie 和 Mark Gales。2023年。Selfcheckgpt：面向生成式大语言模型的零资源黑盒幻觉检测。载于*2023*年自然语言处理实证方法会议论文集，第 9004–9017 页。

Vaibhav Mavi、Anubhav Jangra、Adam Jatowt 和其他 1 位作者。2024年。多跳问答。信息检索基础与趋势®，17(5)：457–586。

Tianyue Niu、Ting Liu、Yiming Taclis Luo、Patrick Cheong-Iao Pang、Shuaishuai Huang 和 Ao Xiang。2025年。解读学生认知能力：教育数据挖掘中可解释人工智能算法的比较研究。《科学报告》，15(1)：26862。

Matthew J Page、Joanne E McKenzie、Patrick M Bossuyt、IsabelleBoutron、Tammy CHoffmann、Cynthia D Mulrow、Larissa Shamseer、Jennifer M Tetzlaff、Elie A Akl、Sue E Brennan 和其他 1 位作者。2021年。PRISMA 2020 声明：系统综述报告指南的更新。《英国医学杂志》，372。

Jaewan Park、Solbee Cho 和 Jay-Yoon Lee。2025年。Stop-rag：面向迭代式检索增强生成的基于价值的检索控制。*arXiv* 预印本 *arXiv:2510.14337*。

Ofir Press、Muru Zhang、Sewon Min、Ludwig Schmidt、Noah Smith 和 Mike Lewis。2023年。测量并缩小语言模型中的组合性差距。收录于计算语言学协会论文集：*EMNLP 2023*，第5687–5711页，新加坡。计算语言学协会。

Peng Qi、Xiaowen Lin、Leo Mehr、Zijian Wang 和 Christopher D. Manning。2019。通过迭代查询生成回答复杂开放域问题。收录于*2019*年自然语言处理实证方法会议暨第九届国际联合自然语言处理会议（*EMNLP-IJCNLP*）论文集，第2590–2602页，中国香港。计算语言学协会。Parth Sarthi、Salman Abdullah、Aditi Tuli、Shubh Khanna、Anna Goldie 和 Christopher D Manning。2024年。RAPTOR：用于树组织检索的递归式抽象处理。收录于第十二届国际学习表征会议。

Yongliang Shen、Kaitao Song、Xu Tan、Wenqi Zhang、Kan Ren、Siyu Yuan、Weiming Lu、Dongsheng Li 和 Yueting Zhuang。2024年。Taskbench：用于任务自动化的大语言模型基准。神经信息处理系统进展，37：4540–4574。

Zhengliang Shi、Shuo Zhang、Weiwei Sun、Shen Gao、Pengjie Ren、Zhumin Chen 和 Zhaochun Ren。2024年。面向多跳问答的检索增强生成中的先生成后 grounding。载于计算语言学协会第*62*届年会论文集（第*1*卷：长篇论文），第7339–7353页。

Weihang Su、Yichen Tang、Qingyao Ai、Zhijing Wu 和 Yiqun Liu。2024年a。DragIn：基于大语言模型信息需求的动态检索增强生成。*arXiv*预印本 *arXiv:2403.10081*。

Weihang Su、Yichen Tang、Qingyao Ai、Zhijing Wu 和 Yiqun Liu。2024年b。DRAGIN：基于大语言模型实时信息需求的动态检索增强生成。收录于计算语言学协会第*62*届年会论文集（第*1*卷：长篇论文），第12991–13013页，泰国曼谷。计算语言学协会。

Yang Sui、Yu-Neng Chuang、Guanchu Wang、Jiamu Zhang、Tianyi Zhang、Jiayi Yuan、Hongyi Liu、Andrew Wen、Shaochen Zhong、Na Zou 等。2025年。停止过度思考：大语言模型高效推理综述。*arXiv*预印本 *arXiv:2503.16419*。

Haitian Sun、Tania Bedrax-Weiss 和 William Cohen。2019。Pullnet：通过在知识库和文本中进行迭代检索实现开放域问答。载于 *2019*年自然语言处理实证方法会议暨第九届自然语言处理国际联合会议（*EMNLP-IJCNLP*）论文集，第 2380–2390 页。

Zhongxiang Sun、Qipeng Wang、Weijie Yu、Xiaoxue Zang、Kai Zheng、Jun Xu、Xiao Zhang、Yang Song 和 Han Li。2025年。Rearter：采用可信过程奖励的检索增强推理。载于

第*48*届 *ACM* 信息检索研究与发展国际会议论文集，第1251–1261页。

Harsh Trivedi、Niranjan Balasubramanian、TusharKhot 和 Ashish Sabharwal。2022年。MuSiQue：通过单跳问题组合实现多跳问题。计算语言学协会汇刊，10：539–554。

Harsh Trivedi、Niranjan Balasubramanian、Tushar Khot 和 Ashish Sabharwal。2023年。在计算语言学协会第*61*届年会论文集（第*1*卷：长篇论文）中，交错结合检索与思维链推理，以解决知识密集型多步问题，第10014–10037页。

Lei Wang、Wanyu Xu、Yihuai Lan、Zhiqiang Hu、Yunshi Lan、Roy Ka-Wei Lee 和 Ee-Peng Lim。2023年。计划与求解提示：通过大语言模型改进零样本思维链推理。在计算语言学协会第*61*届年会论文集（第*1*卷：长篇论文）中，第2609–2634页，多伦多，加拿大。计算语言学协会。

Nan Wang、Yongqi Fan、ZongYu Wang、Xuezhi Cao、Xinyan He、Haiyun Jiang、Tong Ruan、JingpingLiu 及另外 1 位作者。2025a。Kg-o1：通过知识图谱集成增强大语言模型中的多跳问答。*arXiv* 预印本 *arXiv:2508.15790*。

Xinyu Wang、Yanzheng Xiang、Lin Gui 和 Yulan He。2025b。Pecan：利用注意力引导的分层加权图实现大语言模型引导的动态进度控制，以处理长文档问答。载于计算语言学协会会议论文集：*ACL 2025*，第 13317–13335 页。

Tomer Wolfson、Mor Geva、Ankit Gupta、Matt Gard-ner、Yoav Goldberg、Daniel Deutch 和 Jonathan Berant。2020 年。Breakitdown：一个问题理解基准。计算语言学协会汇刊，8：183–198。

Yu Xia, Subhojyoti Mukherjee, Zhouhang Xie, Junda Wu, Xintong Li, Ryan Aponte, Hanjia Lyu, Joe Barrow, Hongjie Chen, Franck Dernoncourt 和另外 1 位作者。2025年。从选择到生成：基于大语言模型的主动学习综述。载于计算语言学协会第*63*届年会论文集（第*1*卷：长篇论文），第 14552–14569 页。

Roy Xie、Junlin Wang、Paul Rosu、Chunyuan Deng、Bolun Sun、Zihao Lin 和 BhuwanDhingra。2025年。知晓何时停止：面向大语言模型的动态上下文截断。*arXiv* 预印本 *arXiv:2502.01025*。

Yuxi Xie、Anirudh Goyal、Wenyue Zheng、Min-Yen Kan、Timothy P Lillicrap、Kenji Kawaguchi 和 Michael Shieh。2024年。蒙特卡洛树搜索通过迭代偏好学习提升推理能力。*arXiv* 预印本 *arXiv:2405.00451*。

Wenhan Xiong、Xiang Lorraine Li、Srini Iyer、Jingfei Du、Patrick Lewis、William Yang Wang、Yashar Mehdad、Wen-tau Yih、Sebastian Riedel、Douwe Kiela 等另外 1 位作者。2020年。利用多跳稠密检索回答复杂的开放域问题。*arXiv* 预印本 *arXiv:2009.12756*。

Ran Xu、Hui Liu、Sreyashi Nag、Zhenwei Dai、Yaochen Xie、Xianfeng Tang、Chen Luo、Yang Li、Joyce C. Ho、Carl Yang 和 Qi He。2025年。SimRAG：面向专业领域适配大语言模型的自改进检索增强生成。载于*2025*年美洲计算语言学协会分会会议论文集：人类语言技术（第*1*卷：长篇论文），第 11534–11550 页，美国新墨西哥州阿尔伯克基。计算语言学协会。

Diji Yang、Jinmeng Rao、Kezhen Chen、Xiaoyuan Guo、Yawen Zhang、Jie Yang 和 Yi Zhang。2024年。Im-rag：通过学习内心独白实现多轮检索增强生成。载于第*47*届国际 *ACM SIGIR* 信息检索研究与发展会议论文集，第730–740页。

Zhilin Yang、Peng Qi、Saizheng Zhang、Yoshua Bengio、William Cohen、Ruslan Salakhutdinov 和 Christopher D Manning。2018年。HotpotQA：用于多样化、可解释多跳问答的数据集。载于*2018*年自然语言处理实证方法会议论文集，第2369–2380页。

Shunyu Yao、Dian Yu、Jeffrey Zhao、Izhak Shafran、Tom Griffiths、Yuan Cao 和 Karthik Narasimhan。2023年。思维树：利用大语言模型进行审慎的问题求解。神经信息处理系统进展，36:11809–11822。

Shunyu Yao、Jeffrey Zhao、Dian Yu、Nan Du、Izhak Shafran、Karthik R Narasimhan 和 Yuan Cao。2022年。ReAct：在语言模型中协同推理与行动。载于第十一届国际学习表征会议。

Donghan Yu、Chenguang Zhu、Yuwei Fang、Wenhao Yu、Shuohang Wang、Yichong Xu、Xiang Ren、Yim- ing Yang 和 Michael Zeng。2022年。KG-FiD：在解码器内融合中注入知识图谱，以实现开放域问答。载于计算语言学协会第*60*届年会论文集（第*1*卷：长篇论文），第4961–4974页。

Jiahao Zhang、Haiyang Zhang、Dongmei Zhang、Liu Yong 和 Shen Huang。2024年。面向多跳问答的端到端束搜索检索。载于*Proceed- ings of the 2024 Conference of the North American Chapter of the Association for Computational Lin- guistics:* 人类语言技术（第*1*卷：长篇论文），第1718–1731页。

Ningning Zhang、Chi Zhang、Zhizhong Tan、Xingx-ing Yang、Weiping Deng 和 WenyongWang。2025年。

面向多跳问题答案的可信的规划驱动 RAG 方法。*arXiv*预印本 *arXiv:2504.16787*。

Qingfei Zhao、Ruobing Wang、Yukuo Cen、Daren Zha、Shicheng Tan、Yuxiao Dong 和 Jie Tang。2024年。LongRAG：一种面向长上下文问答的双视角检索增强生成范式。*arXiv* 预印本 *arXiv:2410.18050*。

Denny Zhou、Nathanael Schärli、Le Hou、Jason Wei、Nathan Scales、Xuezhi Wang、Dale Schuurmans、Claire Cui、Olivier Bousquet、Quoc Le 和另外 1 位作者。2022年。由简到繁提示使大语言模型能够进行复杂推理。*arXiv* 预印本 *arXiv:2205.10625*。

Andrew Zhu、Alyssa Hwang、Liam Dugan 和 Chris Callison-Burch。2024年。FanOutQA：面向大语言模型的多跳、多文档问答。*arXiv* 预印本 *arXiv:2402.14116*。

PRISMA 式文献收集

为反复识别与多跳问答（QA）相关的现有综述和概述，我们遵循了适用于自然语言处理领域的 PRISMA（系统综述和荟萃分析优先报告条目）式流程（Page 等人，2021年）。

范围和时间窗口。 我们的目标是收集讨论多跳问答、多步或组合式问答，或广泛使用的多跳基准测试（例如 HotpotQA、2WikiMultiHopQA、MuSiQue）的综述类论文。我们将搜索范围限制在2022年及之后发表的文献，以聚焦大语言模型时代的研究和近期的检索增强系统。

数据来源。 我们检索了多个学术搜索引擎和数字图书馆，包括 Google Scholar（使用标题限定搜索和关键词查询）、ACL Anthology、arXiv，以及可通过 ACM 和 IEEE 门户获取的会议论文集。同一候选论文若出现在多个来源中，我们将其视为单条记录。

搜索查询。 对于每个来源，我们使用了以下关键词组合（此处以 Google Scholar 风格的语法为例），并始终将时间筛选条件设为自*2022*年起：

- 标题限定搜索: (综述 OR 综述 OR 分类体系 OR "系统综述" OR 教程) ("多跳问答" OR 多跳 OR "多跳")
- ("多跳问答" OR 多跳 OR "多跳" OR "多步推理" OR "复杂问答" OR "组合式问答") (综述 OR 综述 OR 分类体系 OR "系统综述" OR 教程) (自然语言处理 OR " 自然语言处理")
- (“HotpotQA” OR “2WikiMultiHopQA” OR “MuSiQue” OR “IIRC” OR “ComplexWebQuestions” OR “QAngaroo” OR “MedHopQA”) (综述 OR 评述 OR 分类体系OR教程OR概述)

选择这些查询是为了：(1) 找出标题中明确标注为综述的论文；(2) 检索涵盖自然语言处理领域多跳问答或复杂问答的更广泛综述和教程；(3) 查找围绕标准多跳基准测试展开的概述性论文。

筛选与资格评估。 我们首先合并了所有查询的结果，并根据标题和发表场所删除完全重复的条目。随后，我们筛查标题和摘要，排除并非综述、评述、分类体系、教程或概述的论文（例如，单项方法论文、纯理论研究或非自然语言处理领域的应用）。在第二轮筛查中，我们阅读了剩余论文的引言和方法部分，以确认多跳问答或复杂问答是其主要研究重点，而非次要话题。我们仅保留英文论文，且这些论文需对多跳问答、多步／组合式问答或上述基准测试进行充分讨论，以便为我们的研究定位提供参考。

数据提取与本文中的使用。对于每篇纳入的论文，我们记录了元数据，例如发表年份、发表场所、目标任务（多跳问答、推理大语言模型、检索增强生成、基准测试）以及作者采用的主要组织维度（例如模型架构、提示策略、训练方案）。这些标注用于我们的背景与相关工作部分，以便（1）将我们的四要素设计框架置于现有分类体系中进行比较；（2）确保我们对多跳问答和推理综述的讨论全面涵盖2022—2025年的文献。

除这一 PRISMA 风格的综述汇编外，还通过针对性搜索纳入了单篇方法论文（例如，具体的检索–推理系统或搜索算法）

这些论文通过定向搜索、正向和反向引文追踪以及领域专业知识加以补充。

B 论文编码与主题分布

在第 3 节的分析中，我们从五个主题维度对每篇检索到的多跳问答／检索增强生成论文进行了编码：

• **T1：“多跳”的定义与边界**。 明确讨论多跳推理的判定标准、所需支持事实或跳数，或如何区分真正的多跳问题与伪影的论文。

•**T2：分解策略**。 提出明确中间表示的论文，例如思维链、问题分解、子目标规划、程序／图／模板或蒙特卡洛树搜索的论文。

• **T3：证据充分性与忠实性**。着重于确保答案获得充分支持，并使推理链与检索到的证据保持一致的论文，通常采用支持事实指标、验证模块或大语言模型评审诊断。

•**T4：检索—推理耦合**。 主要贡献在于检索与推理如何交错进行或受到控制的论文，包括交错式思维链—检索循环（Trivedi 等人，2023年）、基于束搜索的多跳检索器（Zhang 等人，2024年）、分层或图感知检索控制器（Sarthi 等人，2024年；Edge 等人，2024年），以及感知充分性的多轮检索增强生成（Xu 等人，2025年；Ji 等人，2025年）。

• **T5：评估与诊断**。提出多跳推理新型诊断方法、过程层面指标或鲁棒性测试的论文。

在所有已编码论文中，主题 T4（检索–推理耦合）所占数量最多（252 篇论文中的 139 篇，占 55%），其次是 T2（分解策略）和 T5（评估与诊断）。这一分布促使我们决定在正文中以检索–推理过程作为主要分析对象。

主题	High	Med	Low	总计
T1：定义／边界	6	37	209	252
T2：分解策略	32	33	187	252
T3：充分性／忠实性	17	51	184	252
T4：检索–推理耦合	139	14	99	252
T5：评估／诊断	29	38	185	252

表2：各主题维度中被编码为高／中／低匹配度的论文数量。就高匹配度论文而言，检索–推理耦合（T4）是最大的类别，252 篇中有 139 篇（约 55%）的方法主要关注检索与推理如何耦合。

B.1 编码方法

我们通过关键词搜索（Google Scholar、ACL Anthology 和 arXiv），并结合正向／反向引文追踪，构建了论文语料库。去重后，我们筛选标题和摘要，保留有关多跳问答、检索增强问答和智能体检索–推理的研究。对于每篇纳入的论文，我们提取元数据（年份、任务／基准、监督方式和报告指标），并按照由四个组成部分构成的模式对方法进行编码：整体执行计划、索引结构、控制策略以及停止/继续标准。我们还将每篇论文归入一个或多个主题维度（T1–T5；附录 B.1）。为提高一致性，我们对部分论文进行了双人编码，并通过讨论解决分歧；随后将最终编码应用于全部论文。

从 252 篇论文缩减至 104 篇子集。 完整语料库包含 252 篇多跳问答／检索增强生成论文，用于第 3 节的广泛主题分析。对于各轴分布和方法图谱，我们重点分析其中的 104 篇论文子集，这些论文提供了足够的方法细节，能够据此确定四个组成部分（整体执行计划、索引结构、控制策略以及停止/继续标准）。因此，该子集是从更大语料库中筛选出的、细节充分且与框架兼容的部分，而非单独收集的论文。

C 我们设计轴的数量

C.1 总体执行计划

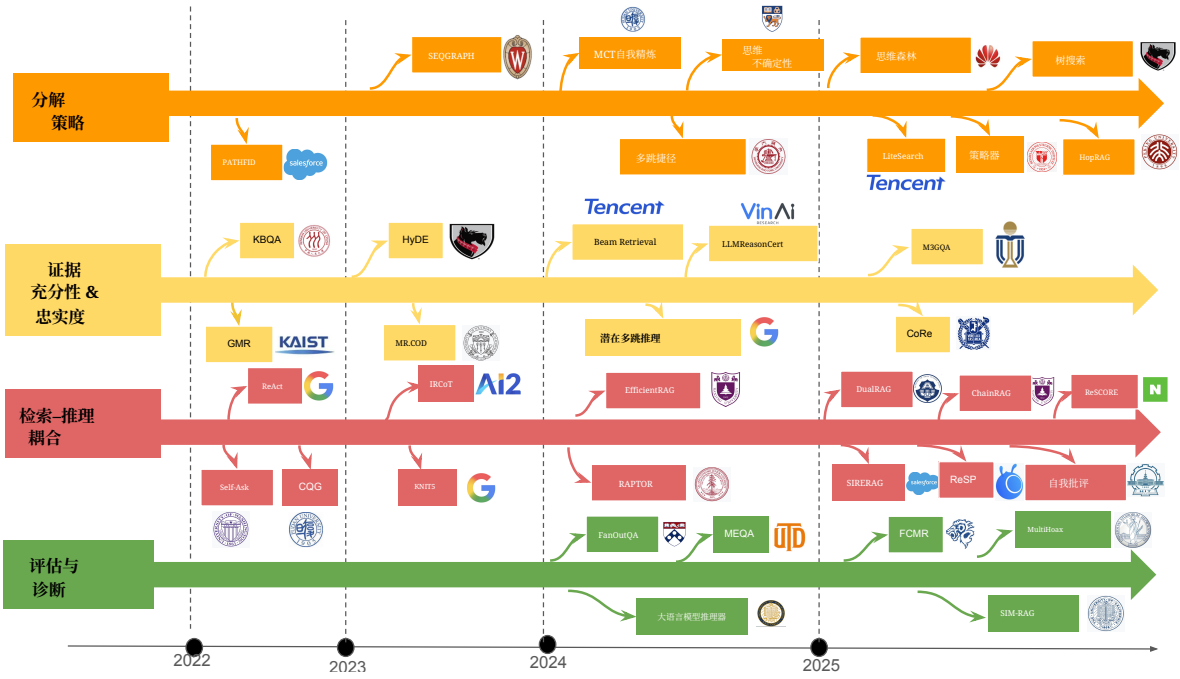

图 4：主要主题及其从 2023年至2025年的发展情况。

数据集	Type	论文数量	主题4论文中的典型作用
HotpotQA	维基百科两跳阅读理解（支持事实）	29	核心多跳基准，用于检索–推理循环以及支持事实忠实度。
2WikiMultiHopQA	混合式知识图谱+维基百科多跳问答	15	测试跨文档与知识图谱–文本推理；通常与 HotpotQA 和 MuSiQue 配合使用。
WebQuestionsSP	基于 Freebase 的知识库问答（1–2 跳）	14	标准知识图谱问答基线；许多主题4系统评估图或知识图谱控制器。
MuSiQue	组合式多跳问答（维基百科）	11	构建了更难的多跳基准由单跳问题组合而成；强调连贯推理。
ComplexWebQuestions(CWQ)	网页+知识库复杂问答（搜索+SPARQL）	11	开放网页环境，包含复杂问题；用于测试分解和网络搜索控制。
MRAMG	多模态检索（图像+文本）	6	将检索—推理从纯文本问答扩展到多模态检索增强生成以及智能体式系统。
StrategyQA	具有隐式多跳推理的布尔问答	5	用于评估控制策略的基准隐式常识多跳推理与分解。
MetaQA	电影领域多跳知识图谱问答（1–3 跳）	4	测试知识图谱遍历策略和束搜索／基于搜索的控制器在大型知识图谱上。
CronQuestions	时序知识图谱问答（时间感知）	4	评估必须处理知识图谱上推理路径的时间约束。
HybridQA	表格+文本混合多跳问答	4	基准测试：结构化表格上的检索—推理以及关联的维基百科段落。

表3：主题4（检索—推理耦合）样本中最常用的数据集。“论文数量”统计了104篇编码论文中提及各数据集的论文数量。

数据集	Type	论文数量
TriviaQA	开放域问答（维基百科 + 网页）	3
Qasper	自然语言处理论文中的长文档问答 自然语言处理论文	3
NarrativeQA	生成式叙事问答（故事）	2
GSM8K	小学数学推理	2
Natural Questions	开放域问答来自搜索日志	2
FB15k-237	知识图谱问答 / 链接预测	2
ALFWorld	基于文本的交互环境	2
EntailmentBank	多步文本推理	2
MQuAKE-CF-3k / MQuAKE-T	蕴含证明 反事实 / 迁移多跳问答	2
FEVER	维基百科事实验证	2
CommonsenseQA	常识选择题问答	2
OpenBookQA	开放书本科学 QA	2

表4：在主题4样本中出现2–3次的数据集。

调度模式（轴A）	论文数量	占比
仅先检索后阅读检索	49	47.1%
仅交错式	24	23.1%
交错式 + 先规划后执行	8	7.7%
先检索后阅读 + 先规划后执行	7	6.7%
交错式 + 测试时搜索扩展	7	6.7%
仅测试时搜索扩展	5	4.8%
先规划后执行 + 测试时搜索扩展	1	1.0%
仅先规划后执行	1	1.0%
其他 / 未指定	2	1.9%
总计	104	100%

表 5：我们的语料库中全部 104 种编码方法的调度模式（轴A）分布（主题 T4：检索—推理耦合）。

C.2 索引结构

索引结构（轴B）	论文数量	占比
图／知识图谱／社区	46	44.2%
扁平／候选列表	29	27.9%
层次化／摘要树	3	2.9%
长上下文	3	2.9%
网页／实时检索	6	5.8%
多模态／其他	5	4.8%
其他 / 未指定	12	11.5%
总计	104	100%

表6：全部104种方法中索引（轴B）的分布。在检索–推理耦合论文中，图和知识图谱式索引最为常见，其次是扁平段落列表。

C.3 下一步执行计划

控制模式（C轴）	论文数量	占比
基于策略	32	30.8%
基于规则	31	29.8%
基于策略 +基于规则	14	13.5%
基于规则 +基于验证器/过程奖励模型	7	6.7%
基于策略 +基于搜索	5	4.8%
基于搜索 + 基于验证器/过程奖励模型	4	3.8%
基于搜索	2	1.9%
基于验证器/过程奖励模型	2	1.9%
未指定	2	1.9%
基于策略 +基于验证器/过程奖励模型	1	1.0%
基于策略 +不确定性门控	1	1.0%
基于规则 +基于搜索	1	1.0%
基于规则 +不确定性门控	1	1.0%
基于策略 +基于规则 + 基于验证器/过程奖励模型	1	1.0%
总计	104	100%

表 7：所有 104 种方法中下一步执行/控制模式的分布（C轴）。

C.4 数据集覆盖范围

数据集	论文数量	占 104 的比例
HotpotQA	35	33.7%
2WikiMultiHopQA	14	13.5%
MuSiQue	12	11.5%
ComplexWebQuestions	12	11.5%
StrategyQA	6	5.8%
IIRC	1	1.0%
FanOutQA	1	1.0%
方法总数	104	100%

表8：我们编码语料库中主要多跳问答基准的覆盖情况。计数基于数据集_摘要 表中 5。

C.5 停止/继续标准

停止／继续模式（D轴）	论文数量	占比
基于预算 + 启发式完成	42	40.4%
基于预算	40	38.5%
基于预算 + 验证器阈值	10	9.6%
基于预算 + 记忆／进度	5	4.8%
启发式完成	5	4.8%
未指定	2	1.9%
总计	104	100%

表9：全部104种方法中停止/继续判据的分布（轴D）。

D 四个组件论文数量

表 10 总结了我们的框架中的四个组件在主题4样本所涵盖的104篇多跳问答/检索增强生成论文中的具体实现方式。沿轴A（整体执行计划）来看，超过半数的系统（56/104，53.8%）仍采用冻结池式的“先检索后阅读”模式；39/104（37.5%）明确交错检索与推理，18/104（17.3%）实施先计划后执行策略，另有13/104（12.5%）在这些模式基础上扩大测试时搜索的规模（例如思维树或蒙特卡洛树搜索）。

轴B（索引结构）显示，图或知识图谱式索引如今略为常见（47/104，45.2%），多于扁平文本索引（33/104，31.7%）；而分层摘要树和长上下文检索单元仍很少见（各3/104）。另有20篇论文（19.2%）不依赖固定的静态索引，而是采用网络搜索、多模态检索或纯参数化知识。

对于C轴（下一步执行计划），超过半数的系统采用学习型全局控制策略（53/104，51.0%），数量相近的系统仍使用基于规则的动作选择（56/104，53.8%），例如固定推理节奏或手工设计的启发式规则；仅有9/104篇论文（8.7%）依赖完全基于搜索的控制器（束搜索、蒙特卡洛树搜索或思维树）。

最后，D轴（停止/继续判据）主要由资源受限停止主导：104个系统中有97个（93.3%）施加了明确的跳数、词元或延迟预算，其中47个（45.2%）将这些预算与简单的“路径完成即结束”启发式规则或束耗尽规则结合起来。基于验证器或过程奖励模型的停止机制仍不太常见，但正在增加（104个系统中有15个，占14.4%）；只有2个系统（1.9%）明确采用基于不确定性或置信度的停止机制。这些分布支持了我们的主张：近期大多数进展都集中在更丰富的控制与规划之上，而索引结构和预算约束停止方案相对保守。

轴／选项	论文数量 104篇中的占比	
（A）整体执行计划		
先检索后阅读	56	53.8%
交错检索 + 推理	39	37.5%
先规划后执行	18	17.3%
测试时搜索扩展 (思维树/蒙特卡洛树搜索/束搜索)	13	12.5%
（B）索引结构		
扁平文本／候选列表	33	31.7%
图／知识图谱／社区索引	47	45.2%
层次树／摘要树	3	2.9%
长上下文单元	3	2.9%
其他（网页／多模态／参数）	20	19.2%
（C）下一步执行计划：控制策略		
基于策略的控制 53		51.0%
基于规则的控制	56	53.8%
基于搜索的控制（束搜索/ 蒙特卡洛树搜索/思维树）	9	8.7%
（D）停止/继续判据		
基于预算的停止机制（跳数 / 词元 / 延迟）	97	93.3%
启发式 / 路径完成即结束规则	47	45.2%
基于进度的规则（记忆／前沿）	5	4.8%
验证器／基于过程奖励模型的停止机制	15	14.4%
基于不确定性／置信度的停止机制	2	1.9%

表10：我们主题4样本中104篇多跳问答/检索增强生成论文的检索—推理设计选择分布，与图1中的四个组件相对应。各模块内计数采用多标签统计：单个系统可能在同一轴上采用多种选择，因此百分比之和不等于100%。

数据集	系统	Year	在此数据集上的报告性能
2WikiMultiHopQA	可信的规划驱动 RAG 方法用于多跳问答（PAR RAG）	2025	2WikiMultiHopQA：精确匹配 0.71，准确率 0.78。
2WikiMultiHopQA	KG-o1：增强多跳问答中的大语言模型与知识图谱	2025	2WikiMultiHopQA：精确匹配 62.40 / F1 值 74.45 / 精确率 72.56 / 召回率 79.45。
HotpotQA	资源友好的动态增强链，用于多跳问答（DEC）	2025	HotpotQA — CoverEM 47.19 / F1 值 50.96 / 准确率† 49.60;. . . HotpotQA — CoverEM 58.52 / F1 值 62.11 / 准确率† 60.32.
HotpotQA	KG-o1：增强多跳问答：利用知识边缘图	2025	HotpotQA：精确匹配 51.00 / F1 值 67.11 / 精确率 69.66 / 召回率 68.18。
MuSiQue	端到端束搜索检索，用于多跳-多跳问题解答	2024	MuSiQue-Ans（测试集）：答案 F1 = 69.2, 支持度 F1 = 91.4（与人工 Sp = 93.9 相当）。
IIRC	将检索与思维链交替结合推理，以应对知识密集型多步问题（IR-CoT)	2023	Flan-T5-XXL：IRCoT 与 OneR → +2.5 IIRC 上的 F1 值，且在分布外数据及较小规模模型上仍保持优势。
StrategyQA	基于大语言模型的搜索助手，采用整体引导的蒙特卡洛树搜索，用于复杂的信息寻求	2025	StrategyQA：精确匹配/F1 = 77.67 / 77.67（GPT-4o-mini，表现最佳的设置）。
FanOutQA	基于大语言模型的搜索助手，采用整体引导的蒙特卡洛树搜索，处理复杂的信息寻求	2025	FanOutQA：准确率 58.38；R-1/2/L = 55.02 / 35.45 / 49.40（优于其他智能体基线）。