移动智能体人工智能的知识驱动推理：概念、方法与发展方向

Guangyuan Liu、Changy uan Zhao、Yinqiu Liu、Dusit Niyato，*IEEE* 会士，以及 Biplab Sikdar，*IEEE* 会士

arXiv:2603.05831v1 [cs.DC] 6 Mar 2026

摘要—**移动智能体人工智能正将自主能力扩展到边缘机器人和无人机（UAV）等资源受限平台。在这些平台上，严格的尺寸、重量、功耗和成本（SWAP-C）约束以及间歇性无线连接，限制了端侧计算和云端访问。现有方法大多致力于优化每轮通信效率，但移动智能体必须在一系列任务中持续保持能力。我们提出一种知识驱动推理框架，从过往执行中提取可复用的决策结构，通过带宽受限的链路进行同步，并将其注入端侧推理，以降低时延、能耗和错误累积。受DIKW启发的分类体系区分了原始观测、单次任务轨迹和持久跨任务知识，并将知识分为检索式、结构化、程序化和参数化表示；每种表示在推理加速与失败风险之间各有不同的权衡。一个关键发现是，知识暴露的效果并非单调变化：知识过少会迫使系统进行代价高昂的试错重规划，而知识过多则会引入相互冲突的线索和错误。无人机案例研究验证了该框架：通过间歇性回传链路同步一个紧凑知识包，使30亿参数机载模型实现了完美的任务可靠性，且推理成本低于无知识端侧推理和以云端为中心的重新规划。**

索引词—**移动智能体人工智能、知识驱动推理、DIKW、程序性知识**

一、引言

智能体人工智能正从云中心化部署转向移动平台，例如智能手机、边缘机器人和无人机 [1]。与可依赖充足计算资源并稳定访问大型外部记忆的云端智能体不同，移动智能体必须在严格的尺寸、重量、能耗和成本（SWAP-C）约束以及间歇性无线连接条件下，完成长时程的感知—推理—行动循环 [2]。因此，移动智能体不仅必须决定推断什么，还必须决定在本地存储什么、适时检索什么，以及在带宽和能耗限制下传输什么。

近期进展提升了移动智能的通信效率，尤其是语义通信：每次交互都会传输当前观测的压缩表示（例如词元、潜变量或图像块），以减少每轮带宽，同时保留与任务相关的内容。然而，每轮压缩主要针对单次交互进行优化，而移动自主性本质上是序贯的：智能体会随时间执行一系列任务，并且必须在目标、环境和用户偏好不断变化时保持胜任能力。因此，系统瓶颈从单次传输转向资源约束不断累积情况下的持续胜任能力。移动智能体不应为每个新任务反复消耗带宽和能耗来重建任务局部上下文，而应识别能够惠及许多未来任务的可复用结构，并以能够将成本分摊到一段时间内的方式传递这些结构。

我们将这一范式称为知识通信，它支持从数据驱动的信号交换向知识驱动的能力演化的转变。在此上下文中，知识广泛涵盖可迁移的结构，例如从累积的任务执行中提炼出的映射、先验、模板或策略，并且它可以同时以隐式（参数化）和显式（可外化）形式存在 [3]。

为使这一视角可操作化，我们采用受数据-信息-知识-智慧（DIKW）启发的分类体系，并针对智能体系统对其进行重新诠释 [4]。数据对应于原始观测。信息对应于执行过程中产生的、以任务为根基的单次任务轨迹与结果，例如交互轨迹、中间产物、工具轨迹和任务结果。上下文是由信息构建的最小工作信息包，用于在单次任务内为推理提供条件。相比之下，知识是从累积的执行信息中提炼出的可复用且持久的资产，旨在支持跨任务迁移，从而使其获取与通信开销能够随时间摊销。这种时间尺度分离进一步意味着可分离的知识生命周期，包括知识抽取（或提升）、知识通信与同步，以及在推理过程中通过激活实现的知识利用。

一个核心挑战在于，知识影响的不仅是准确性，还会影响推理的轨迹级成本。例如，当无人机空中基站在运行时中断后必须重新规划服务顺序时，每增加一个推理步骤都会消耗词元和设备端内存，并可能触发经由受限回传链路传输的信令。更一般地说，在移动智能体流水线中，推理沿着多步轨迹展开，其成本会随着词元生成与计算、记忆与历史记录保留，以及不确定性下的分支与回溯而累积。知识可以通过多种表示形式重塑这一轨迹，包括用于实例复用的检索知识、用于约束感知组合的结构化知识、用于

G. Liu 任职于新加坡南洋理工大学计算与数据科学学院、南洋理工大学能源研究院和跨学科研究生项目（电子邮箱：liug0022@e.ntu.edu.sg）。Y. Liu、C. Zhao 和 D. Niyato 任职于新加坡南洋理工大学计算与数据科学学院（电子邮箱：yin-qiu001@e.ntu.edu.sg、zhao0441@e.ntu.edu.sg、dniyato@ntu.edu.sg）。B. Sikdar 任职于新加坡国立大学电气与计算机工程系（电子邮箱：bsikdar@nus.edu.sg）。

确定性工作流执行的程序性知识，以及用于低时延模式补全的参数化知识。重要的是，知识暴露并非单调变化：有效知识不足会导致搜索密集型推理，而知识过多或关联度较低，则可能带来干扰、分支和幻觉风险。因此，系统的关键问题并非最大化存储知识量，而是在严格的通信和能量预算下控制相关性感知激活。在无线原生部署中，这种激活控制还与链路可用性和信道质量紧密相关，因为知识必须先通过间歇且容量受限的回传链路完成同步，才能在机载设备上使用。

本文的贡献总结如下：

- 我们为移动智能体系统引入一种受DIKW启发的操作性分类法，并依据数据、信息、上下文与知识在SWAP-C和无线约束下的作用及生命周期，阐明它们之间的区别。
- 我们通过将推理建模为承担成本的轨迹，并将知识划分为检索、结构化、过程性和参数化表示，从而形式化知识驱动推理机制，并突出它们对轨迹长度、分支和失效模式的不同影响。
- 我们揭示知识激活中的非单调权衡，并通过间歇性回传链路与运行时中断条件下的无人机空中基站案例研究加以验证；其中，紧凑且可缓存的知识包使机载模型能够以更低的推理成本实现完美的任务可靠性。

II. 知识驱动推理的基础：概念与表示

在本节中，我们给出移动智能体人工智能在SWAP-C与无线约束下进行知识驱动推理的操作性基础。我们首先将推理定义为智能体流水线中的多步轨迹，阐明何为中间状态，以及为何系统成本会沿多步推理与工具使用不断累积。随后，我们通过受DIKW启发的时间尺度分离来区分数据、信息与知识，并将知识划分为四种表示（检索、结构化、过程性与参数化），同时说明其主要系统成本与风险。图1以一个典型的无线移动事件（例如切换）总结了这一基础。

A. 基本概念：智能体流水线中的推理与知识

在智能体人工智能中，推理是指通过一系列中间状态将任务输入转化为决策或输出的过程。具体而言，给定任务规范（例如查询、目标以及可选约束）和可用资源（例如工具和记忆），智能体通过迭代地形成假设、选择操作并更新其内部状态来产生输出。根据智能体的设计，这些中间状态可以表现为自然语言审议、结构化计划、可执行程序、工具调用参数、检索到的证据或验证信号 [5], [6]。

当前的智能体人工智能系统通过若干反复出现的范式实现这一过程，例如生成显式中间步骤的链式审议 [5],在内部推理与外部动作之间交替的工具增强推理、扩展多条候选路径的面向搜索的推理，以及将审议转化为可执行代码的程序辅助推理 [6]。尽管这些范式形式各异，但它们共享一种共同结构，即推理以多步轨迹展开，系统开销沿该轨迹累积。

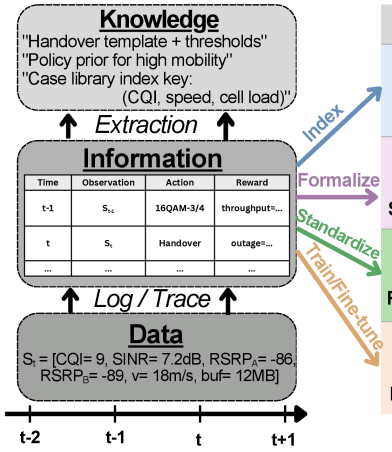
为了以与轨迹视角兼容的方式定义知识，我们采用受DIKW启发的层次结构，并针对智能体推理系统对其重新诠释 [4]。在重复的无线通信任务（如链路自适应和切换）中，数据是原始观测，信息由执行过程中产生的单次任务轨迹与结果构成，而知识是从累积的执行信息中提炼出的可复用且持久的资产，用以支持跨任务迁移，从而使其获取与通信开销能够随时间摊销。这一层次结构进一步表明，知识驱动推理中存在可分离的知识生命周期，包括知识抽取（或提升）、知识通信与同步，以及通过相关性感知激活将受预算约束的知识子集注入正在进行的推理轨迹的知识利用。

B. 知识表示与操作能力

有了这些定义，接下来我们按表示形式对知识进行分类。人们常常将知识视为一种可以增加或减少的单一数量来讨论，但这种抽象不足以支撑智能体推理，因为知识的差异不仅体现在内容上，也体现在形式上。表示形式决定了如何访问知识、知识如何与推理相互作用，以及它会带来哪些系统成本。因此，以下讨论围绕四种表示形式展开，每种形式都着重体现一种独特的操作能力。

1）检索知识：检索知识以实例级记录的形式存储，例如文本块、已编入索引的文档、工具轨迹或已求解样例，并通过语义检索或向量检索机制进行访问 [7]。它在推理中的核心作用是直接复用，而无需重新组合任务。在无线通信中，这对应于检索无线与移动性上下文相似的历史情景，例如信道质量指示（CQI）、信干噪比（SINR）、参考信号接收功率（RSRP）和参考信号接收质量（RSRQ）相近的情景，以及与之相关的动作（如调制与编码方案（MCS）选择、波束切换和切换触发）及其观测结果。当某项任务与已存储实例高度相似时，智能体可以检索该记录并将其作为参考，甚至直接作为近乎完整的解决方案，从而省去大部分中间推理轨迹。

这种表示方式在移动场景中尤其具有吸引力，因为它能够通过



Type	Representation	Reasoning Role	Cost and Risk
Retrieval	Instance-level memory store (indexed cases, documents, tool traces)	Retrieve-and-adapt shortcut, reuse without full recomposition, reduces exploration and backtracking when well matched	• Retrieval latency and overhead • Memory sync overhead • Mis-retrieval anchoring
Structured	Explicit symbolic structures (graphs, schemas, operator trees, factor graphs)	Prunes infeasible branches, stabilizes decomposition order, enforces compositional validity and safety constraints	• Constraint checking overhead • Over-pruning under missing variables
Procedural	Executable workflows (recipes, macros, programs, control-flow templates)	Trajectory compression via deterministic execution, predictable step budget, verification-friendly intermediate checkpoints	• Tool orchestration latency • Procedure version drift • Wrong procedure selection
Parametric	Implicit model parameters (networks, predictors, unfolded networks)	Low-latency pattern completion, fast default decision path when within training support	• Inference energy and update cost • Low auditability • Overconfidence under shift

图 1. 面向移动智能体的知识驱动推理的运行基础。受 DIKW 启发的时间尺度层级以一个典型的无线移动事件为基础：每轮观测构成数据，限定于事件范围内的执行轨迹构成信息，而从信息中提炼出的可复用结构则构成知识。图中展示了四种知识表示及其运行作用、系统成本/风险和提取途径：索引（检索）、形式化（结构化）、标准化（程序化）和参数更新（参数化）。

近乎直接的复用来减少计算量和时延。然而，从系统角度看，这种效率伴随着两个固有的弱点，且影响重大。首先，检索知识的覆盖范围有限，因为其有效性取决于记忆或数据库中是否存在足够相似的实例。其次，它对精确度十分敏感，因为不匹配的检索结果可能会使智能体依赖不恰当的解决方案（例如，在干扰发生变化时采用过于激进的 MCS），进而产生置信度很高但不正确的输出。

2）结构化知识：结构化知识以明确的符号形式表示，例如运算符树、方程、图或约束网络 [8]。它编码的不是针对特定任务实例、可直接执行的解决方案，而是某一任务类别的结构形式，即各组成部分如何组合以及它们之间必须满足哪些关系。在无线通信中，这种结构可以具体化为明确的约束和关系模型，例如用于切换触发的可行性规则、将链路质量与可接受速率选择关联起来的可靠性约束，或链路之间的干扰耦合图。常见格式包括方程的解析树、程序的抽象语法树，以及用于约束推理的因子图。

当任务具有较强的组合结构时，例如多步工具程序、符号变换或约束下规划，这种表示尤其重要。从移动系统的角度来看，结构化知识颇具吸引力，因为它紧凑、易于解释，并且能够自然地在不同任务间重复使用。然而，它通常不会指定完整的执行路径。相反，它充当一种脚手架，通过尽早剪除无效选项，并借助程序化激活或受控搜索，为后续的轨迹压缩做好准备，从而稳定探索过程。

3）程序性知识：程序性知识以可执行程序的形式表示，例如分步操作指南、控制流模板或可复用的工具宏。它并不规定什么是有效的，而是规定接下来该做什么。因此，其核心推理作用是通过确定性执行压缩轨迹。在无线通信中，程序性知识对应于标准化决策工作流，例如先进行测量聚合，再针对波束切换或切换触发进行候选项评估、决策和验证。激活适当的程序后，智能体便可用一系列简短的规定操作取代开放式探索，从而减少分支、回溯和中间状态累积。

这种表示对移动智能体人工智能尤为有效，因为它能带来可预测且节省预算的推理。由于执行路径简短而稳定，系统可以在 SWAP-C约束下更好地分配词元生成、记忆访问和工具调用的预算，同时也能简化验证，因为各个中间阶段都有明确定义。此外，程序性资源能够自然地在不同任务间重复使用，尤其是当任务共享常见子程序时，例如感知到行动工具链、规划基元或标准化的变换模式。

然而，程序性知识会使瓶颈从推导转移到选择。如果智能体激活了错误的程序，或在不兼容的假设下激活了正确的程序（例如，在高移动性场景下应用低移动性切换程序），由此产生的执行可能会迅速失败，或者更严重地，生成看似合理却不正确的输出。因此，有效利用程序性知识需要进行相关性感知激活和轻量级有效性检查，在确定性执行 [9]之前确认其适用性。

4）参数化知识：参数化知识以分布式表征的形式隐式编码在模型参数中，这些表征通过预训练和微调获得。模型并不显式呈现中间结构，而是将统计规律内化，从而能够直接将输入映射为输出。在无线通信中，这通常体现为学习策略或预测器，将无线电和移动性观测映射为动作或性能

估计，从而无需显式检索即可快速决策。具体示例包括：经过训练的信道预测网络，利用历史测量数据预测未来信道状态；以及移动性感知调度策略，根据用户轨迹模式学习分配资源 [5]。因此，其核心推理作用是在透明度有限的情况下进行低时延推理。当输入处于已学习模式的支持范围内时，参数化知识可以将可能很长的推理轨迹压缩为简短的推理过程，从而将时延降至最低。

从移动系统的视角来看，这种表示颇具吸引力，因为它在推理时无需外部检索或显式的结构化接口，从而避免了额外的通信与记忆访问开销。然而，在分布偏移、任务表述含糊或安全关键条件下，其有效性会下降。在无线系统中，此类偏移可能源于时变干扰、移动性、拓扑变化以及不断演变的业务负载，这使得纯粹的参数化捷径在缺乏外部依据时变得脆弱。由于中间推理状态未必显式，参数化推理对验证、审计和回滚的支持有限，从而在紧张的资源预算下使可靠性管理更为复杂。因此，尽管参数化知识在稳定工况下提供了高效基线，但当鲁棒性与可控性成为首要系统目标时，通常仍需辅以检索机制、结构化约束或程序化检查 [5]。

第三. 知识塑造推理轨迹：调制机制与成本权衡

在阐明不同形式的知识如何表示以及它们提供何种推理能力之后，我们现在考察知识如何在实际智能体系统中重塑推理轨迹。一个洞见是，知识并非仅仅缩短推理链。相反，它改变了智能体形成假设、消解不确定性并确定动作的方式。因此，知识会影响轨迹长度、分支行为以及主导失效模式，而这些正是SWAP-C约束下系统成本的主要来源。这些影响总结于图2，其中对比了缺乏有效知识时以搜索为主的轨迹，以及若干由知识调制的捷径。

A. 知识如何调节推理轨迹

检索知识通过实例复用重塑轨迹。当存在高度匹配的记录时，智能体可以用简短的“检索并适配”流程取代多步推导，从而大幅压缩中间推理过程。在无线通信中，这意味着检索具有相似无线与移动性上下文的历史情景，并将相关决策作为当前情景的强先验加以复用。这一机制是检索增强推理系统的基础。例如， [7]的作者提出了Self-RAG，该方法让模型决定何时需要检索，并在注入检索到的证据之前对其进行评判，从而降低推理深度和 token 消耗。此外， [10] 研究了模型何时应依赖内部信号而非外部检索，并表明当模型内部置信度较低时，自适应检索能够同时提升效率和准确率；例如，智能体遇到训练期间覆盖不足的陌生信道或移动性条件时便是如此。与此同时，检索也会引入一种对相关性敏感的失效模式。当实例匹配度较低时，智能体可能会受不恰当模板的影响，进而产生高置信度错误。 [11] 的作者报告称，不匹配证据可能导致此类失败，并提出基于链的验证，以过滤不可靠输入，避免其影响后续步骤。总体而言，这些结果表明，检索能够实现最显著的轨迹压缩，但其可靠性取决于相关性感知选择和轻量级验证。

结构化知识通过围绕明确的方程级或程序级结构重新组织推理，重塑推理轨迹。通过将算子角色和可容许变换外显化，它固定了分解顺序并稳定了中间承诺，从而在选择规则或工具之前减少不确定性。在无线系统中，这类结构可具体表现为显式约束图或因子图式表述，用于编码调度和资源分配中的干扰耦合、可行性约束及服务需求。这一机制体现在将符号约束注入生成过程的结构感知推理系统中。例如， [12]探索了符号限制下的受约束解码，以稳定收敛； [9]则采用可编程抽象语法树约束，以减少无效轨迹。重要的是，结构化知识并不会完全消除探索，而是通过稳定中间承诺、减少早期发散，改变探索发生的位置。

程序性知识通过提供可执行工作流来重塑推理轨迹。选定合适的程序后，智能体便可从开放式探索转向一系列简短的确定性步骤，从而获得可预测的成本特征并提高可控性。在无线移动性控制中，程序性知识通常对应于标准化决策工作流，例如先进行测量滤波，再进行候选项评估和规则检查（如迟滞和触发时间），然后再决定是否执行波束切换或切换操作 [13]。这一范式是程序辅助推理和工具引导推理的基础，其中可复用程序取代了冗长的自然语言推理。例如， [6]证明，生成可执行程序能够降低推理深度并支持基于执行的验证，同时在预算紧张时稳定推理时延。

参数化知识通过隐式模式补全重塑轨迹。当输入落在模型的学习支持集内时，智能体可以以最少的显式中间结构生成输出，从而在无需外部记忆访问的情况下实现低时延。在无线通信领域，深度展开通过将迭代算法求解器映射为可训练网络，为参数化知识提供了一条典型路径，例如展开投影梯度式MIMO检测器（如DetNet），使决策

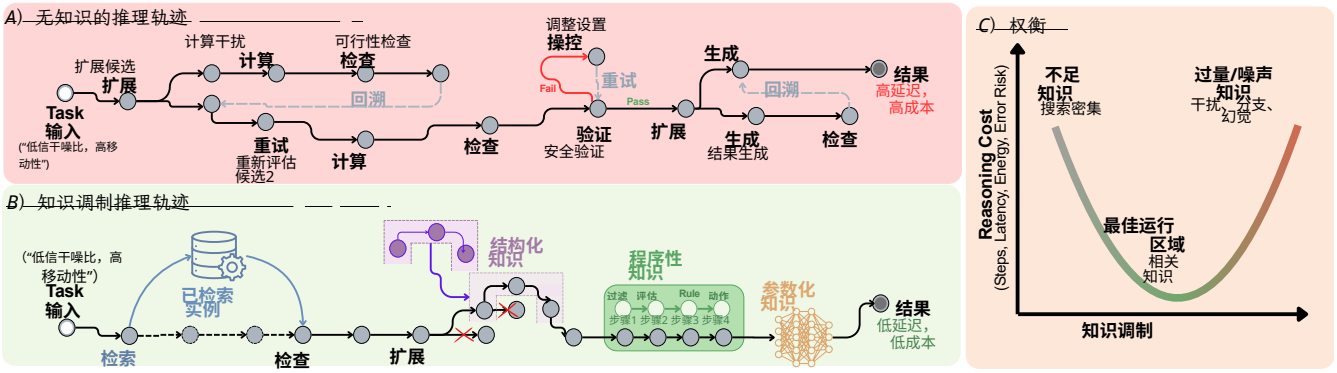


图2. 在一个具有代表性的无线移动事件中（例如低信干噪比和高移动性下的切换），知识调制推理轨迹及其非单调权衡。面板A展示了缺乏有效知识时搜索主导的轨迹：智能体扩展候选项，反复进行计算和可行性检查，并通过重试和安全验证进行回溯，导致高时延和高成本。面板B展示了知识调制后的轨迹：检索知识通过匹配实例实现“检索并适配”的捷径；结构化知识通过施加约束剪除不可行分支；程序性知识将推理压缩为具有可预测步骤预算的确定性工作流；参数化知识则在情境熟悉时提供快速模式补全。这些共同带来低时延和低成本。面板C概述了知识调制与推理成本（步骤、时延、能耗和错误风险）之间的非单调权衡：有效知识过少会导致搜索密集型推理，而知识过多或带有噪声则会引发干扰、分支和幻觉；通过激活相关知识，可达到最佳工作区域。

能够通过固定深度前向传递生成 [14]。与此同时，由于可能不存在用于验证的中间锚点，这一路径的可控性有限。近期研究考察了参数化知识如何与外部记忆交互。例如，[15] 研究了将内部表征与检索相结合的半参数化设计，以平衡时延与可靠性。

这四种机制通过一个具有代表性的无线移动事件（例如低信干噪比和高移动性下的切换）共同呈现在图2（B面板）中：检索知识支持检索并适应的捷径，结构化知识剪除不可行分支，程序性知识将推理压缩为具有可预测步骤预算的确定性工作流，而参数化知识则在所处状态较为熟悉时提供快速的模式补全。

在实践中，这些机制并非相互排斥，往往会在同一个推理过程中共存。一种常见的系统控制模式是先快速提出参数化假设。若能找到高度匹配的实例，便会调用检索知识。如果检索知识不可用或不足，则使用结构化知识将任务构成和允许的操作外化，从而稳定后续决策。最后，一旦确定了合适的工作流，就会激活程序性知识，使剩余的推理能够通过一段简短且确定的序列来执行。并非每项任务都会触发所有阶段。简单输入通常通过参数化捷径解决，而新颖或复杂的情形则会逐步激活检索、结构和程序。这种协同堆栈视角对移动场景尤为重要，因为激活顺序会直接影响 SWAP-C约束下的效率、可靠性和可控性。然而，这种分层激活也意味着，每增加一个知识源，既会带来收益，也会产生开销。激活目标明确时，堆栈会迅速收敛；激活过度或不匹配时，则可能引入协调成本和竞争性锚点，从而降低而非提升性能。这种张力引出了我们接下来要分析的非单调权衡。

B. 推理作为一种具有成本的轨迹，以及知识激活的非单调权衡

在移动智能体流水线中，推理是一种具有成本的轨迹，通过一系列中间步骤 [6]将任务输入转换为有效输出。在无线赋能智能体中，这些中间状态还可能进一步触发测量访问、信令交换或边云同步，因此在严格的预算约束下会直接转化为时延和能耗开销。

从系统角度来看，推理成本主要来自三个方面。

1）词元生成和计算成本，该成本随中间步骤数量的增加而增长，并直接影响时延和能耗。2）内存与历史记录保留成本，该成本取决于后续步骤必须保留多少中间状态。移动智能体依赖有限的输入窗口和受限的设备端内存。随着轨迹增长，智能体要么截断历史记录，从而损害可靠性；要么反复检索并重新注入历史记录，从而增加通信开销。3）分支与回溯成本，该成本会在不确定性下产生。在缺乏强先验的情况下，智能体可能探索多条候选路径，在失败后回溯，并执行重复检查，因此词元和工具调用都耗费在导航上，而非最终输出上。

因此，移动平台上的推理优化自然可以表述为轨迹塑形。知识是实现这种塑形的主要工具，因为它能够向推理循环中注入先验、结构和捷径。然而，其影响并非单调，因为激活不仅会改变轨迹长度，还会改变分支行为，如图2（C面板）所示。

当有效知识不足时，智能体便处于搜索主导状态。推理依赖探索、扩展和回溯，导致轨迹较长、分支因子较高 [6]。因此，时延和能耗会增加，累积推理错误也更容易发生。

随着有效知识激活程度的提高，推理成本会迅速下降。相关检索可以缩短轨迹中的部分路径 [7], , 结构化约束可以剪除不可行的分支 [12], , 程序性知识则可以将探索转化为更紧凑且更具确定性的执行路径 [6]。在这一阶段，知识会同时缩短轨迹并减少分支，从而提升效率与稳定性。

然而，知识数量进一步增加可能会降低性能。弱对齐检索可能诱发锚定效应和高置信度幻觉 [10], [11], , 而过多或相互冲突的约束可能增加协调开销并触发额外的分支 [9]。智能体可能在相互竞争的策略之间摇摆，或采用不恰当的模板，从而同时增加推理成本和错误风险。

这些观察表明，系统面临的主要挑战并非知识存储，而是知识激活控制。系统必须在通信和能量预算严格受限的情况下，调节开放哪些知识、何时激活这些知识，以及如何验证其适用性。因此，面向移动智能体人工智能推理的知识工程，其目标并不是最大化知识量，而是优化相关性感知激活。

第四. 案例研究：低空无线网络中的知识驱动人工智能推理

我们考察低空无人机空中基站任务中的知识驱动推理。无人机必须反复决定下一步为哪个地面用户簇提供服务、应在哪里悬停以提供服务，以及在回传链路间歇中断时是否请求跨回合知识。在这一场景中，我们必须确保服务能够达到目标速率，同时回传链路决定无人机何时可以与Home地面站同步知识。此外，运行时中断会使预先计算的决策失效，因此，可靠运行取决于端侧宏观推理，而非单次云端规划。

A. 通过确定性执行进行宏观推理

图3(a)展示了一个8×8区域，其中每个单元格为500 m×500 m。Home位于单元格(4,4)，四个地面用户簇布置在靠近四角的位置。无人机以100 m的固定高度飞行。接入链路采用10 MHz带宽和30 dBm发射功率。仅当所选悬停航点处的接入速率超过8 Mbps时，才认为该用户簇已得到服务。回传链路采用20 MHz带宽和33 dBm发射功率。仅当回传速率超过5 Mbps时，知识传输才可行；该速率是在单个决策间隔内交付知识包所需的最低速率。这种传输是间歇性的，因为

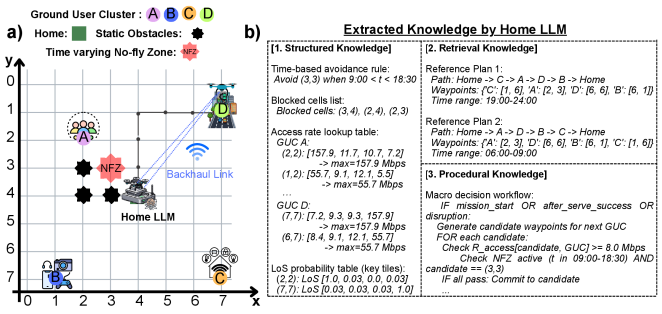


图3. 低空无人机空中基站案例研究与 Home 提升的知识。(a) 一个 8×8 的服务区域，其中包含 Home、四个地面用户簇、静态障碍物和一个随时间变化的禁飞区网格。接入链路决定服务可行性，而间歇性回传链路决定无人机何时可以与 Home 同步知识。(b) 基于 Home 的历史任务日志提升而来的精简知识包，其中包括结构化知识、检索知识和程序性知识。

回传链路质量随无人机与基地的几何关系以及视距概率而变化。

环境中有三个静态障碍网格，以及一个位于 (3,3) 的时变禁飞区（NFZ）网格。机载设备已知禁飞区位置，但机载数据未提供其激活时间表。该时间表只能从历史任务日志中获取，因此会被提取并作为知识存储在 Home。

我们还对仅在执行期间才会显现的运行时中断进行建模：用户簇可能在无人机抵达时漂移一个网格，或者由于临时散射或遮挡，规划航点处测得的信噪比可能低于预测值，从而使当前服务顺序和悬停航点失效，并触发重新规划。

本案例研究中的推理属于宏观层面重新规划，而非低层运动。在每个决策点，机载模型都会输出一项宏观决策，其中包括下一个要服务的用户簇、悬停航点，以及在回传链路允许时是否请求知识。随后，确定性运动模块会执行该决策，引导无人机移动至选定航点，同时避开障碍网格。这种设计使推理轨迹易于理解，并将推理与路径执行分离。

我们通过统计大语言模型轨迹中的推理步骤和推理令牌来衡量推理成本，其中，推理步骤是一个标记的宏观推理步骤，而不是无人机移动步骤。违例表示约束违例，包括在其有效时间窗口内进入禁飞区、发生障碍物碰撞，以及服务尝试失败（即选定悬停航点处的接入速率未达到 8 Mbps 阈值）。

B. Home 中的知识提升与实验设置

Home 从累积的任务日志中离线提炼可复用的结构，并将其存储为紧凑的知识包。在我们的实现中，我们使用 Gemini 2.0 Flash 在 Home 充当知识提取器和摘要器 [1]。这种异构模型分配反映了现实的部署情况：Home

[1]https://blog.google/innovation-and-ai/models-and-research/google-deepmind/google-gemini-ai-update-december-2024/.

地面站拥有充足的计算资源，可运行能力强大的抽取模型；而机载智能体受限于轻量级模型，以满足无人机的 SWAP-C 预算。在此任务中，我们将 DIKW 表示具体化如下。数据是机载可用的逐决策信号，包括无人机位置、剩余簇、剩余步骤预算、本地时钟，以及上一步动作的结果标签，例如已移动、受阻或已服务。信息是回合轨迹，包括已访问的网格、选定的悬停航点、测得的接入和回传链路结果、中断指示以及终止原因。知识是从累积的信息中提炼并持久保存、可在不同回合中重复使用的结构。

如图3(b)所示，知识包包含三种显式形式。结构化知识将可行性和约束外化，包括禁飞区激活规则、障碍物列表、回传可行性图，以及用于航点与用户簇配对的接入可行性查询表。检索知识存储少量成功的参考任务计划，以服务顺序和航点序列表示，并记录实现的吞吐量。程序性知识提供简短的服务与验证工作流，包括接入链路检查、知识请求的回传链路检查、合法性检查，以及提交或拒绝。这些形式旨在重塑推理轨迹，使机载智能体减少对试错重规划的依赖，更多地采用查询、复用和确定性验证。

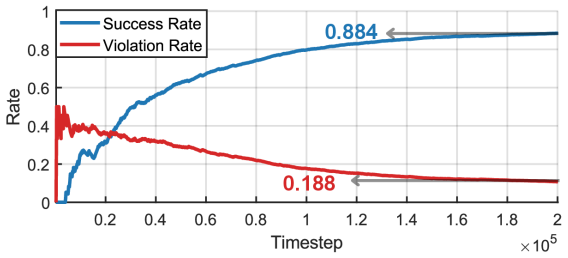
对于机载宏观推理，我们使用 Qwen2.5-3B[2]。在回传链路支持传输时，Qwen 有 k 会请求并缓存 Home知识包；而 Qwen 无 k 则在不接收该知识包的情况下运行相同的宏观推理循环。为隔离参数化知识，我们还在机载端运行 Gemini 2.0 Flash，且不注入显式知识包，记为 Gemini no k。该基线体现了更强的隐式先验，这些先验能够提升决策质量，但与通过 SWAP-C约束促进的显式知识相比，其可外化程度较低，也较难进行预算管理。

为提供具有竞争力的云端基线，我们纳入了一种云端重新规划方法，称为 Home重规划，该方法在 Home 上使用 Gemini 2.0 Flash 实现。发生中断时，无人机通过上行链路发送精简的观测摘要，Home 重新规划，随后通过下行链路发送更新后的宏观计划。如果中断发生时回传链路处于中断状态，闭环就会断开，无人机必须继续执行过期计划，直到到达可行的检查点。我们还纳入了一个非大语言模型基线 DRL PPO，它是一种步级策略，输出微观移动动作，且不提供推理轨迹。与在宏观决策层级运行的大语言模型方法不同，PPO 在逐瓦片移动层级运行，因此其动作粒度更细，推理成本也无法直接比较；纳入 PPO 的目的是评估任务完成情况和约束满足情况，而非推理效率。PPO 在评估前经过 200,000 个环境步的训练。

每种方法均在 50 个推理回合中进行评估，并报告平均值。回合开始时间随机化，因此禁飞区边界并非固定；各方法所受干扰均遵循相同的采样过程。

[2]https://huggingface.co/Qwen/Qwen2.5-3B.

1)



2)

Method	Success Rate	Violation Rate	Reasoning Steps	Reasoning Tokens
DRL_PPO	0.884	0.188	-	-
home_replan	0.64	0.36	60	52.2
Qwen_no_k	0.6	0.4	49	83.2
Qwen_with_k	1	0	38	50.0
Gemini_no_k	0.8	0.04	196	242.4

图 4. 间歇性回传链路和运行时中断下的主要结果。面板 1）展示了 PPO 在 200,000 个环境步内的成功率和违规率训练曲线。面板 2）汇总了各方法在每种方法 50 个推理回合上的平均评估结果。违规情况包括违反禁飞区规定、发生障碍物碰撞，以及未达到 8 Mbps 接入速率目标的服务尝试失败。推理步骤和推理令牌根据大语言模型的宏观推理轨迹进行测量，其中的推理步骤并非无人机移动步数。

C. 结果与分析：可靠性、推理成本与无线权衡

图 4 显示，DRL PPO 在训练后达到 0.884 的成功率，但其违规率仍为 0.188。这表明，步骤级控制能够学会完成任务，但在潜在的时变约束和无线通信可行性方面仍然脆弱。

缺少显式知识时，Qwen 无 k 的可靠性较低，违规率较高。其宏观推理轨迹也更长、更冗长，这与必须即时推断可行性结构时采用试错重规划的情况一致。启用 Home 提升的知识后，Qwen 有 k 实现了零违规的完美可靠性，同时减少了推理步骤数量和每个推理步骤消耗的词元数。这一模式符合预期机制：结构化查询和程序化验证能及早排除不可行选项，而检索引用则减少了以搜索为主的分支。

Gemini 无 k 基线阐明了参数化知识的作用。与 Qwen 无 k 相比，它提高了可靠性，这表明更强的隐式先验可以在一定程度上弥补显式结构的缺失。然而，它的推理成本要高得多，推理步骤更多，且每个推理步骤消耗的词元也多得多。这一差距表明，参数化知识有助于提升性能，但当智能体必须在不确定性下反复重新决策时，它不如显式提升的知识那样易于控制，成本效率也较低。

云端重新规划基线凸显了一项与计算正交的无线局限。尽管 Home 可以借助更强大的模型进行重新规划，但重新规划依赖于中断发生时及时可用的上行链路和下行链路。回传链路中断

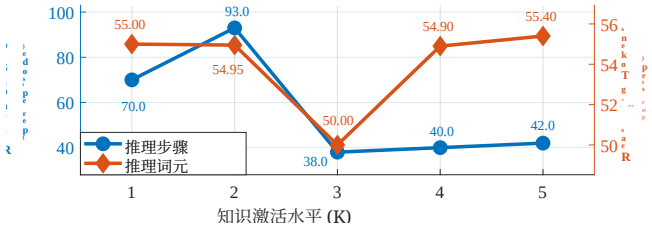


图5. 当调整 Qwen 有 k 的知识暴露水平 K 时，非单调推理权衡。适度的知识暴露可使推理步骤数和每个推理步骤的词元数达到最少；知识过少或过多都会因搜索密集型重新规划和冗余导致的协调而增加推理成本。

恰恰会在需要修复时打破闭环，这解释了为什么在间歇性连接条件下，云端重新规划可能不如借助缓存知识进行设备端推理可靠。

最后，如图5所示，我们改变 Qwen 有 k 的知识暴露水平 K 。每提高一个级别，注入机载提示的知识包都会增加一些内容： $K=1$ 仅包含核心结构化约束（禁飞区规则、障碍物列表）， $K=2$增加回传可行性图， $K=3$ 进一步加入程序性服务与验证工作流以及接入可行性查找， $K=4$ 加入检索参考计划， $K=5$则包含完整知识包以及额外条目和扩展注释。结果证实，推理成本存在非单调权衡。在中等暴露水平（$K=3$）下，智能体的推理步骤最少，且每步词元数最低，因为知识包提供了充足的可行性结构和程序性指导，支持高效决策。在较低的 K 下，智能体缺少关键约束和参考计划，因此只能回到搜索密集型重新规划，产生更长、更冗长的推理轨迹。在较高的 K 下，重复或关联较弱的条目会引入冗余和协调开销，导致推理轨迹出现更多分支且消耗更多词元，因为智能体需要协调相互竞争的线索。除了推理成本，这组测试还揭示了无线通信权衡。提高 K 会增加知识同步的一次性回传负载：从 $K=1$ 时的125个词元升至 $K=5$ 时的1,234个词元。尽管如此，即使是最大的知识包，也仍小于每次中断事件中 Home重规划所需的1,870词元闭环交换。此外，知识包可以缓存；当回传链路不可用时，其同步成本可分摊到所有后续重新规划过程中。相比之下，每当发生新的中断，云端重新规划都必须重复承担上行链路和下行链路的通信成本。

V. 结论

本文提出了面向移动智能体人工智能的知识驱动推理。通过受DIKW启发的分类体系，我们将知识视为检索式、结构化、程序化和参数化表示，并指出每种表示都会从分支、验证和成本可预算性等方面，以不同方式重塑推理轨迹。一项重要发现是，知识暴露并非单调变化：适度且匹配良好的激活能够最大限度地降低推理成本，而知识暴露不足或过度都会降低效率和可靠性。无人机空中基站案例研究证实，通过间歇性回传链路同步的显式、可缓存知识，在可靠性和效率方面均优于无知识端侧推理和云中心式重新规划。这些结果表明，相关性感知激活控制和通信感知知识封装应成为无线原生智能体系统的首要设计原则。未来研究方向包括自适应激活门控，以联合优化知识选择和无线资源分配；以及基于增量任务经验进行在线知识提升。

参考文献

[1] Z. Zhang, Q. Dai, X. Bo, C. Ma, R. Li, X. Chen, J. Zhu, Z. Dong, and J.-R. Wen, “基于大语言模型的智能体记忆机制综述,” *ACM* 信息系统汇刊, 第43卷，第6期，2025年9月。

[ ] 10 1145 3748302 在线。访问地址: [2] https://doi.org/。 / Y. Hu, S. Liu, Y. Yue, G. Zhang, B. Liu, F. Zhu, J. Lin, H. Guo, S. Dou, Z. Xi 等，“人工智能智能体时代的记忆”，*arXiv*预印本 *arXiv:2512.13564*，2025年。[3] M. Farahani and R. Johansson, “解读检索增强语言模型中参数化记忆与非参数化记忆的相互作用”，载于*2024*年自然语言处理实证方法会议论文集，Y. Al-Onaizan、M. Bansal 和 Y.-N. Chen 编。美国佛罗里达州迈阿密：计算语言学协会，2024年11月，第16 966–16 977页。 [在线]。访问地址: https://aclanthology.org/2024.emnlp-main.943/ [4] J. Rowley，“智慧层级：DIKW层级的表征”，《信息科学杂志》，第33卷，第2期，第163–180页，2007年。[5] H. Luo, Y. Yan, Y. Bian, W. Feng, R. Zhang, Y. Liu, J. Wang, G. Sun, D. Niyato, H. Yu 等，“无线通信与网络中的人工智能推理：综述与展望”，*arXiv*预印本 *arXiv:2509.09193*，2025年。[6] S. Gao, J. Dwivedi-Yu, P. Yu, X. E. Tan, R. Pasunuru, O. Golovneva, K. Sinha, A. Celikyilmaz, A. Bosselut, and T. Wang，“通过抽象链推理高效使用工具”，载于第*31*届国际计算语言学会议，2025年，第2727–2743页。[7] A. Asai, Z. Wu, Y. Wang, A. Sil, and H. Hajishirzi，“Self-RAG：通过自我反思学习检索、生成与批判”，载于国际学习表征会议，2024年。[8] H. Han, Y. Xie, H. Liu, X. Tang, S. Nag, W. Headden, Y. Li, C. Luo, S. Ji, Q. He 等，“基于图的推理：构建隐式知识结构以增强大语言模型推理能力”，*CoRR*，2025年。[9] S. Nagy, T. Zhou, N. Polikarpova, and L. D’Antoni，“ChopChop：用于语义约束语言模型输出的可编程框架”，《*ACM*编程语言论文集》，第10卷，POPL，第1905–1932页，2026年。[10] A. Mallen, A. Asai, V. Zhong, R. Das, D. Khashabi, and H. Hajishirzi，“何时不应信任语言模型：探究参数化记忆与非参数化记忆的有效性”，载于计算语言学协会第*61*届年会论文集（第一卷：长篇论文），2023年，第9802–9822页。[11] S. Dhuliawala, M. Komeili, J. Xu, R. Raileanu, X. Li, A. Celikyilmaz, and J. Weston，“验证链可减少大语言模型中的幻觉”，载于计算语言学协会会议论文集：*ACL 2024*，2024年，第3563–3578页。[12] D. Banerjee, T. Suresh, S. Ugare, S. Misailovic, and G. Singh，“Crane：受约束的大语言模型生成推理”，*arXiv*预印本 *arXiv:2502.09061*，2025年。[13] D. Paulius and Y. Sun，“服务机器人学中的知识表示综述”，《机器人与自主系统》，第118卷，第13–30页，2019年。[14] S. Deka, K. Deka, N. T. Nguyen, S. Sharma, V. Bhatia, and N. Rajatheva，“面向下一代无线通信系统的深度展开技术综述”，*arXiv*预印本 *arXiv:2502.05952*，2025年。[15] J. Xie, K. Zhang, J. Chen, R. Lou, and Y. Su，“善变的变色龙还是固执的树懒：揭示大语言模型在知识冲突中的行为”，载于第十二届国际学习表征会议，2023年。