

## 大模型赋能计算机生成兵力决策行为建模综述

凌艳香, 陈丽\*, 陈江明, 支烽耀, 陆正之, 黄小霞  
(国防科技大学 试验训练基地, 陕西 西安 710106)

**摘要:** 计算机生成兵力 (computer generated force, CGF) 是军事仿真系统的核心, 传统建模方法存在知识表征僵化、高质量样本稀缺、决策复杂性建模不足、行为进化能力欠缺等瓶颈, 大模型为破解上述难题提供了新范式。本文从数据知识增强、决策智能生成、能力迭代进化三个维度系统阐明大模型赋能路径, 围绕感知、决策、行动、角色、记忆五个关键模块, 详细阐述基于大模型的 CGF 决策行为建模框架, 梳理各模块技术实现路线与代表性研究成果, 归纳关键技术特点与应用现状, 并从决策实时性、决策质量、决策逼真性、决策评估体系、决策风险控制五个方向提出未来研究重点, 可为智能 CGF 研究与军事仿真智能化升级提供系统性参考。

**关键词:** 计算机生成兵力; 大模型; 决策行为建模; 军事仿真; 智能体; 具身智能

**中图分类号:** E919 **文献标志码:** A **文章编号:** 1001-2486(2026)03-252-17

## Large language model empowered decision-making behavior modeling for computer generated force: a survey

LING Yanxiang, CHEN Li\*, CHEN Jiangming, ZHI Fengyao, LU Zhengzhi, HUANG Xiaoxia  
(Test Center, National University of Defense Technology, Xi'an 710106, China)

**Abstract:** CGF (computer generated force) is the core component of military simulation systems. Traditional modeling methods suffer from bottlenecks including rigid knowledge representation, scarcity of high-quality samples, insufficient modeling of decision complexity, and lack of behavioral evolution capability. Large language models provide a new paradigm to address these issues. This paper systematically clarified the enabling paths of large models from three dimensions: data and knowledge enhancement, decision intelligence generation, and capability iterative evolution. Focusing on five key modules—perception, decision-making, action, role, and memory—this paper elaborated on the large-language-model-based CGF decision-making behavior modeling framework, sorted out the technical implementation routes and representative research achievements of each module, and summarized key technical characteristics and application status. Furthermore, future research directions were proposed from five aspects: decision real-time performance, decision quality, decision fidelity, decision evaluation system, and decision risk control. The findings can provide a systematic reference for intelligent CGF research and the intelligent upgrading of military simulation.

**Keywords:** computer generated force; large language model; decision-making behavior modeling; military simulation; intelligent agent; embodied AI

计算机生成兵力 (computer generated force, CGF) 是由计算机生成和控制的自主或半自主智能体, 可用于扮演敌、我、友各方作战单元<sup>[1-4]</sup>。作为现代军事仿真系统的核心组件, CGF 技术已广泛集成于 OneSAF、VBS、AFSIM、XSim、“墨子”等国内外仿真平台<sup>[5-6]</sup>, 为构建规模化、高对抗的

动态仿真环境提供了重要支撑。CGF 的核心价值在于其自主决策能力的真实性与合理性, 能够模拟真实作战单元在战场环境中的认知与行动过程, 而实现这一能力的关键在于决策行为建模。

所谓决策行为, 是指智能体在特定环境中, 基于自身认知基础、预设目标及场景约束, 做出判断

收稿日期: 2026-01-19

基金项目: 国家自然科学基金资助项目 (62402507); 陕西省创新人才推进计划青年科技新星资助项目 (2024ZC-KJXX-072); 国防科技大学自主创新基金资助项目 (24-ZZCX-JDZ-50)

第一作者: 凌艳香 (1988—), 女, 新疆乌鲁木齐人, 副研究员, 博士, 硕士生导师, E-mail: lingyanxiang@nudt.edu.cn

\*通信作者: 陈丽 (1992—), 女, 宁夏银川人, 助理研究员, 博士, E-mail: chenli18@nudt.edu.com

引用格式: 凌艳香, 陈丽, 陈江明, 等. 大模型赋能计算机生成兵力决策行为建模综述 [J]. 国防科技大学学报, 2026, 48(3): 252-268.

Citation: LING Y X, CHEN L, CHEN J M, et al. Large language model empowered decision-making behavior modeling for computer generated force: a survey [J]. Journal of National University of Defense Technology, 2026, 48(3): 252-268.

并采取对应行动的一系列行为过程,涵盖不同智能体、不同场景下的各类决策活动。决策行为建模是采用形式化描述、算法实现等技术手段,对智能体决策行为的认知规律、逻辑流程、影响因素及输出机制进行量化表征,构建可复现、可计算的决策模型的过程。其核心目的是通过模型模拟智能体决策过程,实现对决策行为结果的预测、分析与优化。

本文的研究对象为 CGF 决策行为建模技术。CGF 作为特殊的决策主体,其决策行为既遵循普通人类“感知—判断—决策—行动”的基本认知链路,又因军事任务而呈现鲜明的独特性。如表 1 所示,CGF 决策行为建模与普通人类决策行为建模的差异可归纳为建模目标、约束规则、可控性与可解释性、对抗性四方面。以普通游戏中的非玩家角色 (non-player character, NPC) 为例,游戏 NPC 以娱乐交互、剧情推进为目标,行为逻辑允许模糊化、程序化甚至夸张化的决策输出;但 CGF 以实战化推演、战术训练、指挥训练等为目标,其行为必须高度符合真实作战逻辑,严格遵循现行作战条令和战术规程,不能逾越军事法理和战术能力边界,同时也应具备可解释的逻辑链路。

表 1 CGF 与普通人类的决策行为建模差异

Tab. 1 Differences of decision-making behavior modeling between CGF and ordinary people

方面	CGF 决策行为建模	普通人类决策行为建模
建模目标	实现作战行为的可信模拟,服务于作战推演、虚拟训练等军事应用	贴近自然人的决策习惯、情绪波动及非理性特征,服务于行为预测、心理分析、社会模拟、人机交互等应用
约束规则	必须嵌入作战条令、战术规范、装备性能、协同规则、军事法理等刚性约束,严格限制模型输出的行为边界,不允许出现不符合军事逻辑、超出作战能力的决策行为	允许融入情绪、偏好、随机决策等非理性因素,无严格的刚性规则约束,模型输出可具有一定的模糊性与随机性
可控性与可解释性	决策链路应清晰可追溯,行为结果应在可控范围内,以满足军事训练评估、战术复盘与推演验证等现实需求	可接受模型输出的概率性、模糊性,决策链路无须可解释性
对抗性	面向强对抗、非合作的战场环境	多面向静态或弱对抗性环境

如表 2 所示,当前主流 CGF 决策行为建模方法大致可分为四类:①基于知识规则,如有限状态机<sup>[3,7-8]</sup>、行为树<sup>[9-12]</sup>等,依赖专家预设的“条件—动作”逻辑框架,虽然结构清晰、可解释性强,但固化的知识难以快速适配多场景泛化需求,而且随着战场模拟要素的增多,规则组合也可能陷入“爆炸”困境。②基于认知架构,如 Soar、ACT-R 等<sup>[1,13]</sup>,通过符号系统模拟人类思维过程,逻辑性强、结果可追溯、可靠性高,但建模复杂且基于特定场景构建的符号化规则具有强场景依赖性,泛化性差。③基于运筹规划,如分层任务网络<sup>[9]</sup>,将作战决策视为完全理性的寻优过程,对决策目标、约束条件与优化准则进行数学建模,再通过运筹算法求解最优策略。此类方法逻辑严谨且结果精确,但存在完全信息、静态策略等强假设条件,与真实战场的高度不确定性存在脱节。④基于机器学习,如强化学习、模仿学习等<sup>[10,14-15]</sup>,以数据驱动为核心,通过模型训练、环境交互试错的方式,学习特定场景下的策略。此类方法是当前 CGF 决策行为建模研究的热点,虽然无须依赖专家先验知识且具备自主学习优化能力,但模型性能受限于训练数据分布,跨场景、突发态势等情况下可能面临性能大幅下降问题。

表 2 主流 CGF 决策行为建模方法

Tab. 2 Main methods of CGF decision-making behavior modeling

方法类型	优点	缺点	适用场景
基于知识规则	结构清晰、可解释性强	泛化能力弱、规则组合爆炸	相对固定、要素简单的场景
基于认知架构	逻辑严谨、结果可追溯、可靠性高	建模复杂、场景依赖性强、泛化能力弱	强调逻辑推理严谨性的场景
基于运筹规划	逻辑严谨、结果精确	存在完全信息、静态策略等假设	环境可控、决策目标明确且约束清晰的场景
基于机器学习	不依赖先验知识,自主学习优化	跨场景泛化能力弱	数据丰富、需要自适应决策的场景

综上,现有研究显著推动了 CGF 技术的进步,但在实际场景中还存在诸多瓶颈,大模型 (large language model, LLM) 技术<sup>[16]</sup>为突破 CGF 决策行为建模困境提供了新的可能。LLM 依托海量数据预训练与超大规模参数,不仅具备少样本推理与自然语言生成等能力,更隐式编码了大

量世界知识,在战场态势感知等军事领域展现巨大潜力<sup>[9-10,17-22]</sup>。与此同时,随着提示学习<sup>[23-24]</sup>、工具学习<sup>[25-28]</sup>、智能体<sup>[29-33]</sup>等技术的快速发展,LLM 正从文本处理模型进化为具备“感知—决策—执行—反思”闭环能力的智能决策中枢。对于 CGF 决策行为建模任务,LLM 可通过少量示例数据快速掌握复杂作战决策任务,缓解知识规则与训练数据稀缺难题,其内嵌的人类社会常识有助于弥补传统决策模型在心理建模与群体互动等方面的缺失,而其作为可编排智能体的灵活性,也为构建可演化、可交互、可扩展的智能 CGF 奠定了技术基础。

尽管 LLM 在 CGF 决策行为建模方面具备广阔前景,但当前相关研究多呈分散化、碎片化,尚未形成系统性综述成果,缺乏对现有研究成果的梳理归纳和关键技术的凝练。对此,本文首先立足军事仿真发展趋势,指出 CGF 决策行为建模面临的新挑战;然后,从感知、决策、行动、角色、记忆五个关键模块,详细阐述基于 LLM 的 CGF 决策行为建模框架、技术实现路径,并系统梳理相关研究进展;最后,分析 LLM 落地 CGF 领域可能面临的核心问题,并探讨潜在研究方向。本文旨在为该领域的研究人员提供系统性研究视图与技术引导,为新一代智能 CGF 的研发奠定基础。

## 1 CGF 决策行为建模面临的挑战

随着数字化、智能化技术的不断发展,军事仿真真正从物理层面的场景还原,逐步迈向以认知对抗为核心的博弈推演新阶段。这要求 CGF 不能仅仅停留于行为结果的重现,而应成为具有决策逻辑、经验直觉甚至个性化决策风格的“认知体”,要能够脱离固定脚本的束缚,在对抗中学习、适应并演化。然而,距离实现上述目标,CGF 决策行为建模研究仍面临诸多新挑战,主要表现在决策知识表征、决策复杂性建模、决策行为进化以及高质量决策样本构建四方面。

### 1.1 决策知识表征

知识是决策活动的基础,如何将真实兵力决策过程中涉及的显性知识和隐性知识有效嵌入 CGF,已经成为制约其认知水平提升的关键问题。具体而言:

一方面是显性知识的精准编码。显性知识包括作战条令、战术规则、作战预案等明确表述为语言的知识内容,编码旨在保留语义的前提下将语言描述转化为计算机可识别、可处理的形式。在实际编码过程中,作战人员了解战术逻辑与战场

语境,但对计算机建模方法较为陌生;建模人员擅长计算机技术,但缺乏军事实操经验和背景知识。这种偏差容易导致决策知识在编码转化中出现描述失真、语义偏差、关键信息缺失等问题,进而影响 CGF 与真实兵力的决策一致性。此外,传统机械的“if-then-else”规则编码模式容易造成作战策略固化,丧失决策过程应有的灵活性与适应性。

另一方面是隐性知识的提取表征。隐性知识是作战人员在长期实践中形成的直觉判断、经验诀窍等难以用语言精准表达的知识,具有模糊性、情境依赖性等典型特征,是真实作战中指挥决策艺术的核心体现。然而,这类知识还未形成成熟有效的提取表征路径,导致难以转化为 CGF 可解析、可执行的认知决策逻辑。例如,资深飞行员在近距离格斗中能根据敌机的飞行姿态、速度变化,直觉判断其规避意图并提前调整机动轨迹,这种隐性的“格斗直觉”很难通过显性规则编码嵌入 CGF。

### 1.2 决策复杂性建模

真实战场上,“应该怎么做”和“实际怎么做”之间往往存在差异,兵力决策过程并非单纯的理性逻辑推理,而是受知识储备、风险偏好、心理状态等多重主客观因素约束的复杂认知过程。因此,如何刻画这种理性与非理性交织的复杂决策过程,是提高 CGF 决策行为真实性的主要内容。这里的难点主要体现为以下两个维度:

一是角色驱动的差异化决策建模。真实场景中,指挥层级、兵种属性、作战经验、性格特质等因素都会造就不同的兵力角色,其相同决策目标下的决策风格存在显著差异。现有 CGF 决策模型多基于统一规则体系或奖励优化准则构建,决策结果存在明显同质化倾向,导致对抗中的预置性强。

二是决策过程中的人因变量嵌入。当前 CGF 决策模型多以最优化决策为核心,依赖运筹规划、强化学习等纯理性寻优算法构建决策逻辑。但真实作战中指挥员在连续遭受敌方打击、伤亡较大情况下,可能因心理压力产生“盲目反击”或“过度收缩防线”等非理性决策。因此,忽略知识局限、心理压力等关键人因变量对决策过程的影响,会导致 CGF 决策行为虽逻辑自洽,但与真实兵力决策脱节,难以匹配实际作战场景需求。

### 1.3 决策行为进化

进化能力是 CGF 适配动态对抗的核心需求。真实兵力并非被动执行预设战术的“自动机”,而

是能够基于实战经验持续学习的认知主体。这种“实践—反思—迭代”类人化的认知进化能力,是实现智能 CGF 不可或缺的一环。当前 CGF 决策模型普遍呈现“冻结式智能”特征,一旦部署就处于固定状态,缺乏对自身决策行为结果的反思提炼。具体表现为:一是无记忆性,多数 CGF 仅依赖当前态势快照进行即时决策,没有类似于人类的长期记忆库,无法存储环境交互历史,也无法调用历史经验来应对当前决策任务。这种从零开始的决策,极大限制了决策效率。二是无反思机制,无法对决策失败的情况进行归因分析,也无法对决策成功的经验进行提炼总结,导致 CGF 决策模型停留于初始部署状态。

#### 1.4 高质量决策样本构建

在数据驱动、智能建模主导的时代,决策样本数据是支撑 CGF 决策模型训练、调优、验证的基础,其质量和数量直接决定了模型的性能和泛化能力。高质量军事决策样本的稀缺性,已成为 CGF 智能化发展的关键瓶颈。

一是实战决策数据获取。实战决策数据具有天然的不可重复性、高风险性,其获取过程受限于作战场景、安全保密等多重因素,因此往往只能收集到部分兵力的行动轨迹、交战结果等片段信息。这导致现有决策数据多以碎片化、局部化的信息片段形式存在,难以构建完整的“感知—决策—行动—反馈”闭环行为数据链,无法完整复现真实兵力作战决策的全流程逻辑。

二是高质量标注样本生成。同一决策行为在不同战场态势、不同作战任务背景下,其合理性可能存在显著差异,“何为合理的决策”缺乏统一通用的量化标注依据。这导致决策数据样本标注过程周期长、人工成本高。同时,标注人员的个体经验差异、认知偏好等主观因素,也容易产生标签偏差,这进一步降低了标注样本质量。

三是动态测试样本构建。现有 CGF 决策模型的测试样本多为“固定敌方兵力部署、固定作战场景”的静态脚本,而真实战场中,敌方兵力会动态调整部署,战场环境会随天气、地形变化而改变,现有测试样本无法模拟这类动态场景,这导致 CGF 决策模型在复杂动态场景下的适应性、鲁棒性缺乏测试支撑。

## 2 LLM 赋能路径

LLM 具备开放域知识迁移、自然语言交互以及自主推理规划能力,有望提供新的 CGF 决策行为建模范式。表 3 列出了代表性 LLM 决策

技术在 CGF 中的应用分析<sup>[34-45]</sup>。总体来说,LLM 对 CGF 决策行为建模的赋能路径可以归结为数据知识增强、决策智能生成、能力迭代进化三个维度。

表 3 代表性 LLM 决策技术在 CGF 中的应用分析

Tab.3 Analysis of representative LLM based decision-making techniques in CGF

技术	CGF 应用分析
检索增强生成 <sup>[34-40]</sup>	解决 CGF 知识更新与幻觉问题。构建外部决策知识库,通过知识检索、上下文融合增强 CGF 决策结果质量;无须重新训练,只需更新外部知识库
反思 <sup>[41-43]</sup>	让 CGF 自主优化决策。针对失败的决策结果,分析失败原因,生成结构化反思内容,并将其纳入下一轮决策输入,避免重复错误;无须额外训练数据,自动进行反思,对人工规则依赖少
思维链 <sup>[44-45]</sup>	让 CGF 学会分步推理。将真实兵力的隐性思维逻辑转化为显性推理路径描述,以提示文本方式驱动 LLM 完成 CGF 决策;无须额外训练数据,推理过程透明且可解释,多场景适配性强

1) 数据知识增强。一方面,LLM 经海量文本预训练后具备强大的少样本/零样本推理能力,仅需少量作战决策示例或自然语言指令,即可快速掌握战术决策逻辑,适配不同军兵种、不同作战场景的 CGF 建模需求,这能极大缓解作战决策样本稀缺难题。另一方面,思维链(chain of thought, CoT)技术<sup>[44-45]</sup>通过引导 LLM 生成分步推理的逻辑链路,可将复杂作战决策任务拆解为一系列子步骤,模拟人类指挥员的决策思考过程,为决策知识表征提供一种高效途径。此外,检索增强生成<sup>[34-40]</sup>(retrieval augmented generation, RAG)能让 LLM 与外部知识库联动,实施通用推理与知识检索协同下的决策,这能够加强 CGF 与军事决策知识的深度耦合。

2) 决策智能生成。ReAct<sup>[43]</sup>作为 CoT 进阶版本,通过“推理—行动—反馈”闭环,使 LLM 不仅能生成决策,还能模拟执行效果并据此迭代优化后续决策,这能够还原真实作战中临机调整的动态决策过程,提升 CGF 在复杂环境下的适应性。角色扮演大模型技术<sup>[46-49]</sup>,通过提示引导 LLM “扮演”某一具有特定身份、背景、性格、目标或知识结构的角色,并在此角色约束下生成符合该角色认知逻辑、行为习惯和语言风格的输出。同样,

通过提示可将真实兵力的军兵种属性、指挥层级、作战经验、性格特征等注入 CGF,使其生成符合兵力个体特质的决策,这为解决 CGF 长期存在的决策同质化问题提供了可能。

3)能力迭代进化。作为 LLM 复盘反思技术, Reflexion<sup>[41]</sup>让智能体在完成任务后,对自身行为进行事后反思,识别失败原因或优化空间,并在后续决策中调整策略。这能使 CGF 对决策行为、执行结果、对抗胜负进行复盘分析,具备迭代优化、自主进化能力。人类反馈强化学习<sup>[50]</sup> (reinforcement learning with human feedback, RLHF)等 LLM 轻量级增量微调技术,无须重新训练整个模型,通过少量数据来更新局部参数,进而实现 CGF 决策能力的持续升级。此外,CGF 历次对抗的态势数据、决策日志、反思结果等信息,能以向量数据库方式形成可长期存储、快速调用的经验知识库,通过 RAG 技术让 CGF 具备记忆能力,无须从零开始决策。

### 3 基于 LLM 的 CGF 决策行为建模框架

本节提出基于 LLM 的 CGF 决策行为建模框架,旨在突破传统技术路径的局限,为智能化 CGF 决策行为模型构建提供思路借鉴。该框架主要由感知、决策、行动三个核心功能模块构成,并通过角色与记忆两个支撑模块,分别实现决策行为的对齐与进化。表 4 列出了当前部分相关工作。具体而言:①感知模块,对仿真态势数据进行语义解析与情境理解;②决策模块,基于感知结果和记忆内容,生成符合角色约束与任务目标的高层行为策略;③行动模块,将高层行为策略转化为军事仿真系统可识别的底层执行指令,通过指令接口传送至仿真引擎,驱动 CGF 完成相应动作,实现闭环决策交互;④角色模块,定义 CGF 所扮演兵力角色的基础属性、任务目标、行为准则等信息,为其他模块提供稳定的上下文语义约束;⑤记忆模块,作为外挂式知识库,可通过经验回溯、错误归因与策略提炼等方式,为 CGF 提供类人化的反思与进化能力。

各模块之间存在的数据交互逻辑如图 1 所示,其中:①角色模块通过定义 CGF 感知范围、感知概率、目标识别优先级、指挥层级、兵种特性、作战经验、装备性能参数、作战协同要求等特征,对感知、决策、行动三个模块提供顶层约束条件;②感知模块要在角色模块提供的感知约束以及记忆模块提供的历史经验之下,向决策模块输出态势感知结果;③决策模块接收感知模块的实时态

表 4 相关工作分类

Tab.4 Classification of related work

模块	子方向	代表性工作
感知	基于视觉模型的态势图像转换	文献[51-54]
	基于 LLM 的结构化数据重构	文献[55-57]
	多模态 LLM 端对端感知	文献[58]
决策	军事知识融合	文献[59-62]
	推理与规划	文献[54,63-67]
	决策反馈优化	文献[68-71]
	协同决策	文献[72-75]
行动	语义到代码转化	文献[52]
	场景约束	文献[76-77]
	行动规划优化	文献[54,56]
角色	角色档案构建	文献[46-49]
	角色特征注入	文献[46-47,78-79]
记忆	记忆建模	文献[34-35,80]
	记忆检索	文献[36-39]
	记忆演化	文献[41-42,81-83]

势结果、角色模块的决策约束以及记忆模块的经验知识和历史态势,生成决策方案并下发至行动模块,同时将本次决策全流程逻辑写入记忆模块,完成经验积累;④行动模块在决策方案基础上结合角色模块的行动约束,将高层决策细化为可执行的具体行动指令序列,通过指令接口驱动 CGF 仿真实体执行,同时将行动执行结果、战场反馈写入记忆模块,为后续迭代优化提供依据;⑤记忆模块全程存储态势各模块处理结果,形成完整的“感知—决策—行动—反馈”闭环数据链,同时为 CGF 感知、决策和行动过程提供历史经验。

图 2 展示了一个基于 LLM 的指挥员 CGF 决策模型示例,模拟了指挥员个体的决策过程,基于战场态势感知结果生成作战计划和行动指令。采用 LLM 单智能体架构,按照分析式和直觉式两种决策模式,其中:①直觉式决策模式,适用时间紧迫条件下的决策任务。指挥员根据战场态势,凭借自身的直觉认识和经验知识直接决策生成可行的解决方案,代表了一种基于直觉经验的快速反应思维活动。这里的直觉经验可具象化基于历史案例形成的决策预案,具体实现可以采用检索匹配算法或者 RAG 框架。②分析式决策模式,适用时间较为充裕条件下

的决策任务,综合考虑各方面因素,运用作战决策模型计算分析找到最优方案,代表了一种基于计算分析的谨慎思维活动。具体实现可以在

图 1 基本框架基础上进行深化扩展,例如人机协作模块能够提供人类指挥员与 CGF 的交互协作接口。

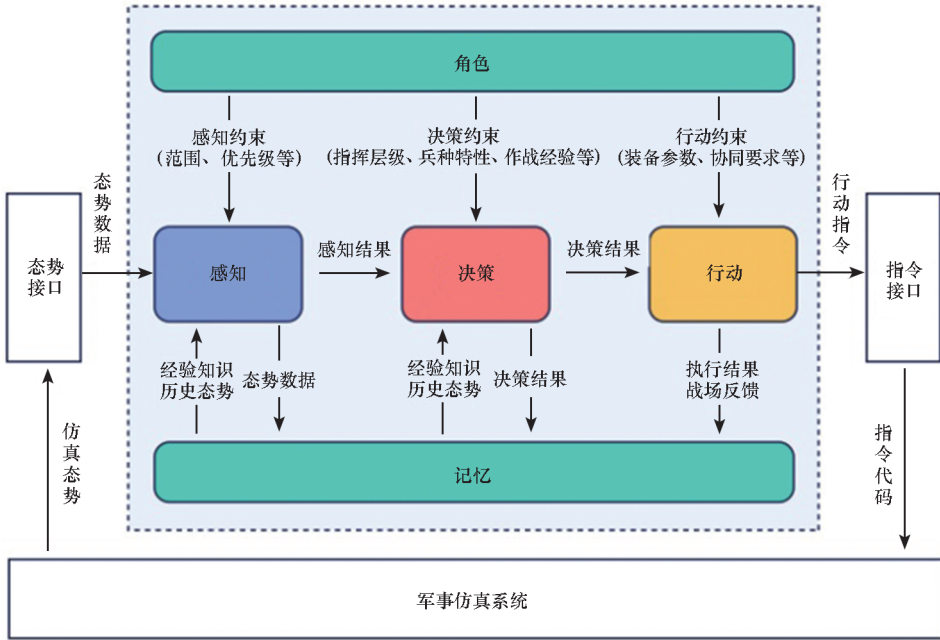


图 1 基于 LLM 的 CGF 决策行为建模框架

Fig. 1 Framework of LLM based CGF decision-making behavior modeling

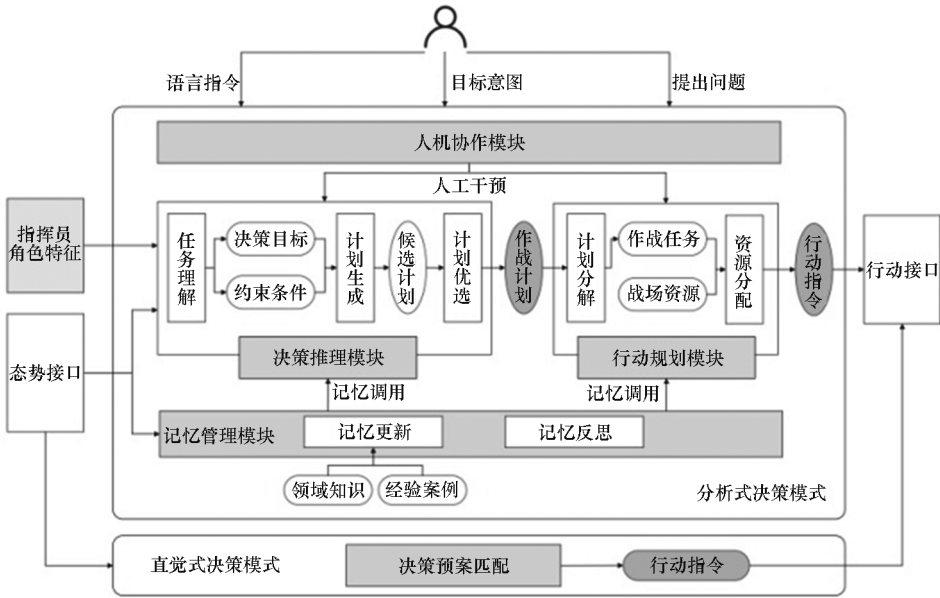


图 2 指挥员 CGF 决策模型框架示例

Fig. 2 An example framework for a commander-oriented CGF decision-making model

图 3 展示了一个基于 LLM 的指挥机构 CGF 群体决策模型示例。与指挥员个体决策相比,指挥机构的群体决策存在显著差异:决策主体是多个席位或部门,来自不同的业务领域,具备多样化知识经验,信息处理上需整合多渠道信息,决策结果由多个席位或部门协作完成。采用 LLM 多智能体协作框架,以多个智能体模拟指挥群体,通过

信息交互、协作机制的设计,实现协商博弈过程下的决策结果生成。①智能体角色分工:指挥机构中包括情报分析、态势评估、推理决策、资源调度、作战评估等智能体。每个智能体角色分工不同,相互之间存在信息交互,可通过提示输入方式,由 LLM 来模拟不同智能体的业务职能。②信息公告机制:统一的信息共享平台,作为多智能体之间

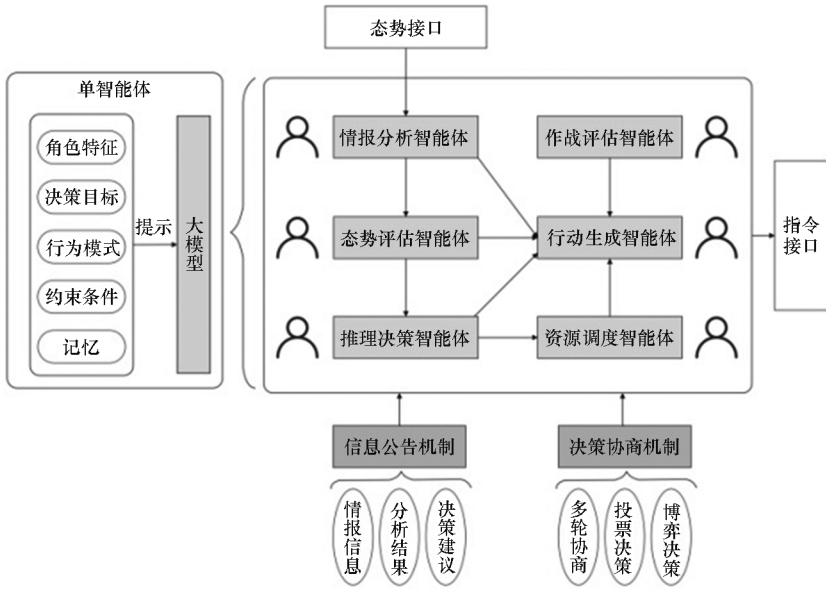


图 3 指挥机构 CGF 群体决策模型框架示例

Fig. 3 An example framework for a commander organization-oriented CGF decision-making model

信息交互的基础。所有智能体将收集到的情报信息、分析结果、决策建议等上传至该平台,同时也可以从平台上获取其他智能体的信息,确保信息的及时共享和流通。③决策协商机制:包括多轮协商、投票决策、博弈决策等交互协作机制。

### 3.1 感知

在当前军事仿真系统架构下,CGF 的感知输入可分为结构化和非结构化两类态势数据。如表 5 所示:①结构化态势数据,由仿真系统直接生成,通常以键值对、JSON、XML、数据库记录或消息协议(如 DIS、HLA)形式传输,语义明确,易于直接映射为自然语言描述。典型来源包括仿真引擎生成的离散化事件流,以及 CGF 所搭载或关联的虚拟传感器模型(如雷达、红外、电子侦察设备)输出的探测结果。②非结构化态势数据,主要指虚拟战场全局或局部视图、仿真回放视频等,信息丰富但语义隐含,需特定模型进行解析处理。

感知模块的任务是将原始态势数据转化为结构清晰、语义连贯且符合自然语言范式的标准化内容,为后续 LLM 决策提供可解释的上下文输入。其中,对结构化态势数据,重点在于语义增强与语言化重构;对非结构化态势数据,关键在于多模态理解与高层语义提取。如图 4 所示,根据数据类型与处理方式的不同,现有方法可归纳为三类典型范式:基于专用视觉模型的态势图像转换、基于 LLM 的结构化态势数据重构以及基于多模态 LLM 的端对端态势感知。

表 5 典型 CGF 态势数据

Tab. 5 Typical situation data of CGF

类型	来源	典型内容
结构化态势数据	仿真事件流数据	①平台状态变化,如机动指令、速度/航向变更、损伤等级更新;②武器与装备操作,如雷达开机/关机、导弹发射、通信建立;③指挥控制信息,如上级下达的任务指令;④交战结果,如命中判定、毁伤评估
	虚拟传感器探测结果	①目标的身份、位置、速度、姿态等运动学参数;②目标属性,如平台类型、威胁等级、电磁特征;③探测置信度、遮蔽/干扰状态等环境影响信息
非结构化态势数据	态势图像/视频	①电子地图叠加的敌我兵力分布;②虚拟场景渲染图,如第一/第三人称视角图像;③仿真回放视频

1) 基于专用视觉模型的态势图像转换:通过专用视觉模型对态势图像进行特征提取,再利用文本生成模型将其转化为自然语言形式的态势摘要。例如,Zhang 等<sup>[51]</sup>为具身智能体感知模块设计了一种专用转换器,其利用二维图像构建三维点云,进而提取关键物体状态等高层语义信息,最终映射为结构化的场景文字描述;Liang 等<sup>[52]</sup>结合基于对比语言-图像预训练(contrastive language-image pre-training, CLIP)模型的开放词

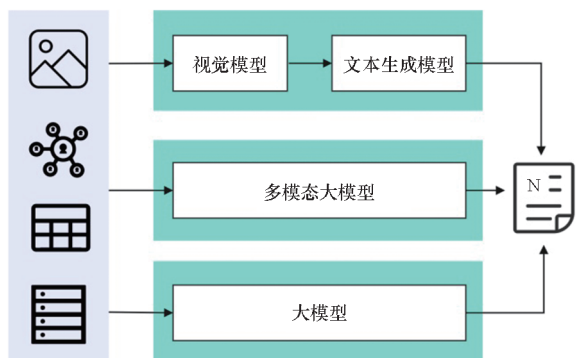


图4 感知模块

Fig.4 Inception module

汇目标检测器 ViLD,可从输入图像中生成带类别与空间关系标签的对象列表;Yoo 等<sup>[53]</sup>则采用引导式语言-图像预训练(bootstrapped language-image pre-training, BLIP)模型,将输入图像编码为向量,并通过与历史记忆库中的图文对进行跨模态检索,增强对态势图像的理解;Wu 等<sup>[54]</sup>利用 YOLOv5 目标检测器识别图像中的可见实体,并生成“存在性+属性”形式的简洁语言描述。

2) 基于 LLM 的结构化态势数据重构:通过 LLM 和解析规则提示文本,将结构化、符号化态势数据转化为自然语言描述,并进行语义增强。文献[55-57]均使用 LLM,直接解析仿真环境反馈的包含对象、位置、状态的结构化数据列表,并转换为自然语言描述,实现对场景的完整语义表达。

3) 基于多模态 LLM 的端对端态势感知:直接采用多模态 LLM 对异构态势数据进行端对端理解。例如,Chen 等<sup>[58]</sup>将原始态势图像输入多模态 LLM,模型内部完成感知、理解、推理一体化处理,直接输出高层决策结果或态势判断,显著简化了感知过程。

### 3.2 决策

决策模块是 CGF 实现智能行为的核心环节,其主要任务是在感知模块提供的态势理解基础上,结合 CGF 的角色身份、任务目标、作战规则与历史经验等上下文信息,生成合理可行的行动策略。LLM 驱动的决策模块在 CGF 建模中具有显著优势:一是表达灵活性强,可处理模糊、非结构化指令,并生成符合人类语言习惯的解释;二是泛化能力好,在少量示例或规则引导下即可适应新任务、新场景,有助于缓解军事决策数据稀缺问题;三是内嵌常识知识与认知能力,为建模指挥员直觉、部队文化、敌我心理等难以形式化的复杂因素奠定基础,进而提升决

策行为真实性。

CGF 作为真实兵力的代理,其决策模块同样应具备可靠性和有效性。然而,LLM 有可能因幻觉问题<sup>[84]</sup>生成违反交战规则或军事常识的决策,需要通过知识约束、输出验证等方式加以限制;同时,军事决策本身具有高度的复杂性、动态性与对抗性,仅依赖 LLM 的通用推理能力难以适配多场景下的决策性能需求。以下重点从军事知识融合、推理与规划、决策反馈优化以及协同决策四个方面,系统探讨基于 LLM 的 CGF 决策模块构建与优化。

#### 3.2.1 军事知识融合

LLM 通用知识与 CGF 决策需求之间存在领域差距,军事知识融合的目的是将军事规则、作战条令等决策知识转化为 CGF 决策模块可直接利用的信息资源。具体方法可归纳为两种,即静态化知识嵌入和动态化知识检索。

1) 静态化知识嵌入,即将决策知识固化到 CGF 决策流程。例如,Wu 等<sup>[62]</sup>针对医疗决策场景,基于医学知识图谱和临床指南,设计了知识驱动的 LLM 协作框架,通过动态协商机制整合领域知识。Zhu 等<sup>[59]</sup>直接将人工撰写的文本化游戏攻略注入 LLM 提示词,同时记录成功的动作序列,并将其总结为文本记忆以优化后续规划,有助于解决开放域游戏决策中的知识匮乏问题。

2) 动态化知识检索,即通过动态检索外部知识库、实时数据或高质量决策任务示例,为决策提供知识支持。这类方法主要通过 RAG 框架<sup>[80]</sup>实现。例如,文献[60]利用 RAG 设计了跨领域的检索模块,使 LLM 能够针对决策需求,从法律、医疗等领域精准获取专业知识支撑;文献[61]则从离线数据集中检索高回报的状态作为目标状态,并利用生成模型以这些目标为条件生成子轨迹以进行规划,显著提升策略泛化能力。

#### 3.2.2 推理与规划

推理和规划是 CGF 从简单响应向类人决策跃升的关键支撑。一方面,CGF 决策是由态势判断、威胁评估、目标优先级排序、资源调配、风险权衡等环节构成的严密逻辑链条,若仅依赖 LLM 的端到端生成能力,其结果易受提示文本措辞或内部随机性等因素影响,导致推理跳跃甚至违反基本军事原则。另一方面,作战任务通常具有高度复合性,比如“夺取要点并建立防御”这一任务可能包含机动、突击、警戒等多个子阶段,时间跨度可能覆盖数分钟至数小时,LLM 很难一次性生成有效的完整行动策略,也容易因局部最优陷入全

局失效。因此,必须采用一系列强化手段来提升 LLM 决策的可信性、可控性和复杂任务应对能力。

1) 结构化推理框架:通过构建结构化的推理框架,提升 CGF 推理过程的可控性和置信度。例如,Zhang 等<sup>[63]</sup>将 LLM 推理过程转化为一个推理图,量化每个决策步骤的置信度,并通过剪枝低置信路径筛选出最优决策链路;Wang 等<sup>[64]</sup>进一步提出增强训练框架 LTA-Thinker,构建了基于可学习先验的隐式思维生成架构,并引入基于分布的定向优化范式,增强了复杂推理性能;Hou 等<sup>[85]</sup>提出 ThinkPrune 方法,利用强化学习修剪冗余推理步骤,聚焦核心推理节点,从而提升了决策效率和准确性。

2) 分层规划:将复杂决策目标分层拆解,解决长程规划问题。例如,Wu 等<sup>[54]</sup>提出的“规划—排除—跟踪”框架,通过分步规划、排除无效路径以及跟踪目标完成进度,显著提高了规划精度;Zhou 等<sup>[86]</sup>则将决策空间表征为树形结构,利用蒙特卡罗树实现自上到下的搜索遍历,同时结合基于语言模型的价值函数与自我反思机制,实现高效探索与增强决策;Ye 等<sup>[87]</sup>将多任务决策分为任务层和操作层,在各层中通过事后反思调整规划路径,优化了多目标规划的效果。此外,文献[65-67]将规划领域定义语言(planning domain definition language, PDDL)与 LLM 融合,通过生成 PDDL 中的领域文件和问题文件,使用 Fast Downwards 等求解器实现精准的任务规划。

### 3.2.3 决策反馈优化

反馈优化旨在提升 CGF 决策适应性与鲁棒性。初始决策后,环境是动态变化的,因此需要通过环境反馈、专家评估等多源信号构建闭环优化体系,实现决策策略的迭代。

1) 强化学习驱动决策优化:这类方法核心是将强化学习思想与 LLM 结合。例如,Xi 等<sup>[68]</sup>提出 AgentGym-RL 框架,通过多轮强化学习训练 LLM 智能体,在长期决策任务中利用环境反馈动态调整策略,显著提升了策略的累积奖励。

2) 人类反馈融合:结合人类偏好或任务优先级来优化决策,确保 CGF 决策逻辑与真实兵力对齐。Jiao 等<sup>[69]</sup>提出以思维为中心的偏好优化方法,通过收集人类对 LLM 内部思维过程合理性的偏好反馈,优化决策逻辑;Yuan 等<sup>[88]</sup>在多目标路径规划任务中,融入了用户在时间、成本等方面的偏好,驱使 LLM 生成个性化决策方案。

3) 决策经验复用:通过召回历史决策经验来

优化决策。文献[70-71]中,LLM 在决策过程中首先召回历史相似场景的决策结果,然后筛选出最优决策作为当前决策参考,不仅减小了重复计算和错误决策的概率,还通过积累和复用经验,逐步提升了 CGF 的整体智能水平。

### 3.2.4 协同决策

通过设计多智能体架构,明确不同角色分工与通信逻辑,实现多主体协同决策。文献[30]系统梳理了单智能体集中式和多智能体协商分布式两种典型的协同决策架构,前者适用于任务简单、信息共享要求低的场景,而后者则更适合于复杂、动态且需要频繁信息交换的任务环境,为 CGF 协同决策架构设计提供了框架借鉴。Becker 等<sup>[72]</sup>采用模块化设计,支持多智能体的任务分配、信息共享与协同决策,提高了决策系统整体的灵活性和扩展性。Aratchige 等<sup>[73]</sup>提出“分工—通信—协作”三层架构,旨在提升多智能体决策的协作效率。Pokharel 等<sup>[74]</sup>针对多智能体之间的决策冲突,设计了“提案—讨论—投票—修正”四阶段协议,确保最终决策结果的一致性。Hong 等<sup>[75]</sup>构建了 HLSMAC 星际争霸多智能体挑战平台,通过训练 LLM 智能体在高层战略决策中进行分工协作,优化团队整体的决策效能。

## 3.3 行动

行动模块是 CGF 决策输出向仿真系统执行的“最后一公里”,其核心任务是将决策模块输出的行动策略转化为符合仿真引擎接口规范的指令序列。行动模块的作用不仅在于实现“决策—执行”的闭环,更在于确保所生成动作的技术可行性、军事合理性和物理可信性,具体来说就是要符合仿真系统接口协议,契合 CGF 角色特征和任务目标,满足装备能力、地形环境、动力学等约束。

LLM 的自然语言理解与结构化生成能力,能够直接解析高层语义策略,并依据上下文自动生成细粒度行动指令,尤其在少样本条件下的泛化性、可解释性方面具有较大优势。但 LLM 原生的自由文本生成容易产生格式错误或逻辑漏洞,还需要多层次约束与优化才能确保从决策到执行的可靠转化。当前相关工作主要围绕三方面开展:一是语义到代码转化,解决行动与仿真实接口适配性问题;二是场景约束,解决行动的物理可信性和军事合理性问题;三是行动规划优化,解决行动执行的质效问题。

1) 语义到代码转化:仿真系统自身具有严格的语法规则和确定性逻辑,比如条件判断、循环执行等,这就决定了 CGF 仿真实体可执行的动作指

令必须严格适配仿真系统接口。Liang 等<sup>[52]</sup>提出“代码作为策略”框架,可将高层语义策略转化为 Python 代码指令,通过代码化输出提高了行动指令的精确性,同时增强了对复杂任务逻辑的支持能力。此外,研究人员还通过形式化模板约束、后处理验证机制等方式,进一步确保行动指令可被仿真引擎正确执行。

2) 场景约束:通过融入 CGF 相关的装备性能参数、战场地理信息、物理规则、交战规则等场景约束信息,提升行动指令的物理可信性和军事合理性。例如,Ahn 等<sup>[76]</sup>提出“按能力执行,而非按指令执行”的动作生成理念,通过特定数据集上的预训练,为模型设立能力边界,使 LLM 能够生成既可行又符合场景的自然语言动作指令,而非拘泥于对语义指令的字面遵循。Xiang 等<sup>[77]</sup>将 LLM 与世界模型<sup>[89]</sup>融合,其中世界模型负责模拟行动指令执行的物理效果,通过模拟的环境反馈验证行动指令的可信度。

3) 行动规划优化:面对相对复杂的任务目标,可分步拆解行动序列,通过搜索算法、迭代优化等机制,筛选最优行动路径,提高行动执行效果和效率。文献[56]将 LLM 作为零样本规划模型,将自然语言策略拆解为分步骤行动序列,并提出一种基于现有示范案例的动作生成流程,通过语义翻译将规划策略精准映射至智能体可执行的指令。文献[54]提出“规划—排除—跟踪”三阶段方法,首先将高层策略拆解为候选行动序列,接着移除不符合物理规则、仿真接口的无效指令,最后通过实时跟踪行动执行进度与目标匹配情况,动态调整后续指令。该方法提高了行动执行的灵活性,还确保了整体任务目标的有效达成。

### 3.4 角色

角色模块的任务是为每个 CGF 赋予明确的个性化特征,从而确保感知、决策与行动各模块的角色一致性。具体来说,这里的个性化特征包括三个维度:身份属性,如 CGF 所属的军种、职级、装备配置等;认知特征,如决策风格、风险偏好、经验水平等;行为边界,如任务权限、能力边界等。

CGF 的角色模块主要通过 LLM 与角色提示词实现。LLM 天然支持自然语言形式的角色描述,便于军事专家定义角色,其少样本泛化能力还能够支持同一底座模型快速适配不同角色。具体来说,相关工作主要包括两部分:一是角色档案构建,二是角色特征注入。

1) 角色档案构建:文献[46]使用心理量表构建角色的人格数据,并通过蒸馏已有角色扮演智

能体的方式生成角色对应的对话数据,两部分数据分别从静态人格描述和动态行为两个维度来刻画角色特征。文献[47]提出经验重构方法,利用 LLM 基于角色描述进一步生成角色相关的场景行为数据,包括事件描述、地点、时间、背景、互动等要素,能够扩展角色数据的丰富性。此外,研究人员还尝试构建通用角色数据集,为个性化决策模型构建提供训练测试的数据支撑,例如文献[48]依据 MBTI 人格类型生成虚拟角色描述档案,Wang 等<sup>[49]</sup>则构建了 100 个特定角色的专属档案。

2) 角色特征注入,主要包括两类方法:一是利用角色档案数据微调 LLM<sup>[46-47]</sup>;二是通过提示学习,将角色档案数据注入 LLM 的上下文输入。相比微调,提示学习所需数据量较小,无须大规模训练算力,同时还具备在多种角色间切换的灵活性。例如,Zheng 等<sup>[78]</sup>设计了特定于对话者和受众两种角色的提示词,探究不同角色提示的效果;Min 等<sup>[79]</sup>研究发现随机替换示例中的标签几乎不会降低模型性能,标签空间、输入文本的分布特征和序列的整体格式才是影响最终任务表现的关键驱动因素,因此提出通过合理设计演示样例来引导 LLM,使其生成符合角色设定的行动,有效降低了角色特征嵌入过程对数据体量的要求。

### 3.5 记忆

记忆模块是 CGF 模型实现经验复用、行为一致性与行为进化的关键组件,其核心任务是在仿真过程中,对 CGF 的感知信息、决策过程、行动结果等交互历史进行存储和运用,从而赋予 CGF 记住过去、理解现在、规划未来的类人记忆能力。记忆模块的本质是 CGF 与仿真环境交互的动态信息系统,需重点解决三方面问题:一是记忆建模,将先验知识、交互历史等数据转化为具有语义关联的知识表示,以支撑高阶推理;二是记忆检索,以决策目标驱动进行记忆内容检索,避免无差别调用引入噪声;三是记忆演化,实现记忆内容迭代更新。

#### 3.5.1 记忆建模

记忆建模是对 CGF 交互历史、先验知识等数据进行系统化存储、结构化组织与语义化表征的过程,目的是将零散的经验转化为可被后续决策调用的有效知识。

文献[80]提出包含短期记忆、长期记忆与元记忆的三层架构,明确记忆系统应具备“感知—存储—提炼—调用”的闭环能力,为 CGF 记忆建

模提供了理论框架。Hu 等<sup>[34]</sup>借鉴卡尔曼“直觉 + 理性”双系统认知理论,构建双超图结构:前者为基于主题对齐的知识图谱,刻画实体概念间的语义关系;后者为事件 - 行为关联网,记录“情境 - 决策 - 结果”三元组,通过图神经网络实现跨记忆融合与类比推理,支持“类似态势下曾采取何种行动并导致何结果”的记忆复用。Li 等<sup>[35]</sup>进一步提出 StructRAG 框架,模拟人类工作记忆机制,在推理阶段对检索到的非结构化数据进行动态结构化重构表征,再输入 LLM 进行推理,显著提升知识密集型任务中的决策质量。

### 3.5.2 记忆检索

记忆检索是依据当前决策任务需求,从记忆库中精准定位、提取并调用相关内容的过程,直接决定了记忆资源的复用效率与决策支持的有效性。现有研究主要围绕三方面开展:检索结果与决策目标的相关性、多源记忆关联检索、检索机制的多任务适配性。

具体来说, Lee 等<sup>[36]</sup>提出“先规划、后检索”范式,即 LLM 首先生成高层检索计划,再依据各子目标定向检索匹配的记忆内容,有效减少了无关信息干扰,增强了检索结果与决策目标的相关性。Wang 等<sup>[37]</sup>针对复杂决策场景提出了 RichRAG 框架,能从记忆库中联合挖掘多源关联信息,生成内容丰富的上下文。Zhang 等<sup>[39]</sup>则提出任务自适应 RAG 机制,在运行时根据当前任务动态构建轻量级知识图谱作为检索索引,实现“任务感知”的记忆检索。此外, Tang 等<sup>[38]</sup>探索了端到端的 LLM 记忆检索方法,提出了自检索框架,通过自监督学习内化检索语料库,将检索过程转化为序列化段落生成,并执行相关性评估以实现重排序,使单个 LLM 在无须外部工具的情况下自主完成记忆定位与提取,极大简化了记忆系统架构。

### 3.5.3 记忆演化

记忆演化是指记忆系统随仿真进程与任务推进,通过经验提炼、反思优化、结构重组等方式,实现记忆内容迭代更新的动态过程。

文献[41]借鉴人类自我反思机制,使智能体在任务失败、任务低效或面临复杂任务情况下自动触发反思,按照轨迹回溯、问题归因、总结提炼、记忆整合的流程,依托 LLM 完成高阶知识生成。文献[42]使智能体在面对新任务时先执行“思考”,对决策任务进行分解和推理;然后开展“精炼”,检索相关记忆,滤除噪声信息并重组记忆结构;最后实施“行动”,输出决策结果并执行相应

动作。通过三步循环迭代,主动优化记忆质量。

针对长时间交互下记忆冗余问题, Suzgun 等<sup>[81]</sup>设计的动态“备忘录”机制,能够提供筛选关键过程知识、删除冗余信息等自适应记忆精炼方法; Xi 等<sup>[82]</sup>优先保留关键互动信息,舍弃或进一步压缩不重要的记忆,实现对记忆存储的动态管理。文献[90]提出了语义无损的记忆压缩方法,通过熵感知过滤机制剔除低语义价值的交互内容,并依据语义相似度、时间邻近度将关联记忆进行合并。文献[91]提出主动式自主记忆压缩管理方法,按照“关键信息留存 + 冗余信息剔除 + 中等信息摘要”的分层压缩模式,实现了记忆规模的动态精简。Wang 等<sup>[83]</sup>针对具身智能体,提出一种新型长短期记忆系统,其中长期记忆以三维场景图的形式全面捕捉环境表征,短期记忆则动态记录物体位置与状态的变化,同时还设计了遗忘曲线和剪枝策略,自动删除短期记忆中的低价值内容。

## 4 潜在研究方向

LLM 驱动的 CGF 决策行为建模,是大模型技术与军事仿真领域的新兴交叉方向,尽管展现出了巨大应用潜力,但目前仍处于探索起步阶段,尚未形成成熟的技术框架与标准范式。由于 LLM 训练算力要求高、推理时延长、存在幻觉等固有特性,加之 CGF 决策行为模拟任务的复杂性和高难度,当前研究仍面临诸多亟待解决的关键挑战,还需大量研究人员的系统性研究。以下将主要从 CGF 决策行为建模的实时性、质量、逼真性、评估体系以及风险控制五方面论述潜在研究方向,旨在为相关研究人员提供思路借鉴。

### 4.1 决策实时性

LLM 的性能表现与其参数规模高度相关,超大规模参数虽能支撑复杂决策能力,但也导致推理链路冗长,对算力资源提出了高要求。这一特性直接造成其决策响应时延通常处于秒级甚至分钟级,而 CGF 决策尤其是单兵、单装、单平台等战术战斗级决策对响应时间的需求普遍聚焦于秒级、毫秒级,二者存在较大差异。同时,如何在有限算力环境下平衡决策效果和决策实时性,也是一大困难。

未来研究可从以下两方面寻求突破:①轻量化模型定制适配。通过模型蒸馏、量化压缩、结构化剪枝等技术手段,在保留核心军事决策能力的前提下,构建面向 CGF 特定场景的轻量化 LLM,降低推理阶段的算力消耗与时延成本。②分层协

同推理架构设计。采用即时响应、深度规划两种差异化决策模式,将高频简单决策任务交由轻量模型或规则引擎快速处理,将低频复杂任务通过 LLM 完成,实现实时性与决策效果的动态平衡。

#### 4.2 决策质量

军事决策是多目标、多约束、高动态的复杂智能活动,而 LLM 存在“幻觉”、通用知识向军事领域知识的适配偏差等固有缺陷。这导致当前 LLM 驱动的 CGF 决策结果在精准性、可靠性、鲁棒性等方面仍存在显著不足。可能的决策质量优化方向:

1) 面向 LLM 的 CGF 决策知识生成,即结合 LLM 的输入特征与理解偏好,研究新型决策知识表征与生成技术。在表征形式上,将分散的作战条令、战术经验、装备操作规范等转化为 LLM 易理解的自然语言描述、结构化提示模板、场景化案例脚本等形式。在知识来源上,整合真实战场演练数据、仿真推演日志、战术教材文本、专家访谈记录等多源数据,通过数据清洗、语义对齐、场景化重构等处理,生成高质量知识。在生成机制上,构建“专家引导-机器生成-人工校验”的协同生成框架,充分结合 LLM 与人类专家的优势,提高决策知识生成的效率。

2) 大小模型协同 CGF 决策架构,即融合运用 LLM 与传统 CGF 决策行为建模方法,实现优势互补。一方面,可研究“LLM+轻量化专用模型”的分层决策增强方法:LLM 负责复杂态势理解、任务分解等高层决策任务,知识规则推理、运筹优化以及强化学习等专用模型则负责低时延、输出确定性高的底层决策任务,各自优势互补,提升决策的可靠性与鲁棒性。另一方面,可考虑设计决策结果多维交叉验证机制,通过冲突消解、加权投票等机制,筛选最优决策方案,提升决策结果的可靠性。

#### 4.3 决策逼真性

逼真性并非仅指行为表面的“像”,更应深入至决策机制层面,体现真实作战环境中人类决策者的认知逻辑、心理动态与交互规律。因此,逼真性的内涵有两重:一是再现真实场景下决策活动的复杂性,既有理性分析,也掺杂情绪波动、风险偏好、认知偏差、疲劳状态等非理性因素。同时,个体决策嵌套于群体互动结构之中,受到上下级指令、同级协同、跨单位协作关系乃至文化背景等多重社会性变量的影响。这种“理性+非理性”“个体+群体”多重因素叠加,构成了动态、非线

性的军事决策生态。二是兵力角色决策行为的模拟逼真度,即在给定角色设定下 CGF 能否生成符合该角色认知模式与行为逻辑的决策输出。针对这些挑战,未来可围绕三方面开展深化研究:

1) 嵌入人因变量的 CGF 决策行为建模。结合军事心理学、认知行为科学的研究成果,构建涵盖风险偏好、心理状态、疲劳累积效应、作战经验等级等核心要素的人因模型,通过提示学习、增量微调等方式将其嵌入 LLM 决策链路,使 CGF 决策呈现符合角色定位的非理性倾向与个体差异特征。

2) 群体交互环境下的协同与博弈决策建模。真实战场中的决策极少孤立发生,而是嵌入复杂的群体交互网络之中,可结合多智能体强化学习、博弈论等理论,构建群体决策架构,复现兵力内部因任务优先级、资源或指挥体系等因素引发的协商、妥协甚至对抗过程。此外,可考虑模拟不同 CGF 角色因信息不对称而产生的误判、试探、欺骗等行为。

3) CGF 角色对齐方法。从静态属性、动态属性以及行为范式等维度,系统性构建典型角色档案库,在此基础上,研究 CGF 角色对齐方法,例如:角色感知的提示学习,以上下文提示方式引导 CGF 模型生成角色适配的决策路径;角色一致性约束机制,通过基于角色规则的过滤、重排序等方式,剔除不符合角色身份的输出;高效角色微调,利用少量高质量角色档案数据对 LLM 进行高效微调,实现角色专属的行为风格固化。

#### 4.4 决策评估体系

当前 LLM 赋能 CGF 决策行为建模领域缺乏完善的评估体系,包括评估指标、测试数据集、测试场景科目以及支持动态对抗验证的仿真环境。这不仅制约了不同方法之间的横向比较与性能归因,更使得模型在复杂、高动态战场环境中的鲁棒性、适应性与合理性难以被有效验证。

未来需尽快构建一套 CGF 决策模型评估体系,包括:①多维可量化的评估指标,应涵盖功能性(CGF 是否有效完成指定决策任务)、逼真性(CGF 决策行为与给定角色的相似度)、鲁棒性(CGF 决策模型在态势突变等扰动下的稳定性)、智能性(CGF 在态势理解、任务规划、策略生成等方面的能力)等维度。②标准测试数据集,例如基于真实历史战例、专家推演数据等,提炼态势演变、决策记录、结果反馈等信息,形成示范样本;基于规则或强化学习等智能体,生成大量、可控变量的对抗轨迹,对人工数据进行增强。③系统化测

试科目,例如设计基础决策能力科目,测试 CGF 决策模型的基本认知与逻辑能力;通过火力分配、路径规划等中阶科目,测试 CGF 战术理解与规划决策能力;构建红蓝对抗、突发事件处置等高阶科目,测试复杂条件下的 CGF 决策表现。④动态仿真测试环境,基于已有军事仿真系统,设计通用灵活的嵌入接口,允许接入不同架构的 CGF 决策行为模型,支持多粒度数据观测和结果分析等。

#### 4.5 决策风险控制

LLM 自身存在固有安全风险与可信挑战<sup>[92]</sup>,而军事场景的数据高敏感性与决策强约束性,进一步放大了此类问题的影响。LLM 驱动 CGF 决策建模的核心决策风险性主要体现为三方面:一是军事敏感信息泄露。CGF 决策建模需融合作战条令、装备参数、演练数据等敏感信息,在数据处理、模型训练、知识检索等环节,若防护措施不足,易造成信息泄露。二是隐式推理不可解释。LLM 的核心是神经网络,黑箱式的隐式推理使 CGF 决策模型无法有效解释“为何生成该决策”,尤其当 CGF 出现决策偏差或错误时,难以快速定位问题根源并进行针对性修正。三是非确定性输出易引发违规。LLM 生成特性使其在相同态势与约束下仍可能输出差异化结果,部分结果易突破作战条令、装备性能等硬性约束,导致 CGF 决策与客观实际脱节。

聚焦 CGF 决策行为建模的场景需求,围绕决策风险控制,未来可开展三方面研究:①全流程敏感信息防护体系,比如在面向 LLM 的提示词设计、记忆检索等环节,对 CGF 行为建模涉及的军事数据进行分类脱敏,设计访问权限,在决策输出阶段增加敏感信息过滤等。②CGF 决策推理链路建模,立足作战条令、战术原则等军事实际,抽象凝练 CGF 标准化决策流程,设计显式的军事决策推理链路,进而使 LLM 可解释、可追溯。③CGF 决策输出合规校验,比如在 CGF 决策与行动模块间增设合规校验节点,开展军事规则与战术合理性双重校验,剔除违规决策,优化非确定性输出,保障决策合规有效。

## 5 总结

针对大模型赋能 CGF 决策行为建模这一新兴交叉领域,开展了系统性综述研究。结合军事仿真发展趋势,深入剖析了该领域在知识表征、高质量样本获取、决策复杂性建模及行为进化四方面的核心挑战,厘清了传统建模方法的局限与技

术突破的迫切性。在此基础上,构建了基于 LLM 的 CGF 决策行为建模框架,从感知、决策、行动、角色、记忆五个模块,系统梳理了各模块的技术实现路径与代表性研究成果。最后,从决策实时性、质量、逼真性、评估体系及风险控制五个维度,明确了该领域亟待突破的关键问题与潜在研究方向。

大模型为构建可演化、可交互、高逼真的新一代智能 CGF 筑牢了技术根基。然而,当前研究仍处于探索阶段,大模型固有特性与军事决策特殊性的适配问题尚未完全解决,仍需持续攻关。未来,人工智能、军事心理学、运筹学等跨学科深度协作,是推动该领域走向成熟的关键。本综述旨在为相关研究提供系统性参考,期望更多学者聚焦这一重要方向,通过技术创新突破现有瓶颈,助力军事仿真系统向更高智能水平演进。

## 参考文献 (References)

- [1] 许凯,曾俊杰,杨伟龙,等.面向计算机生成兵力的认知行为模型架构研究综述[J].系统仿真学报,2021,33(2):239-248.  
XU K, ZENG J J, YANG W L, et al. Overview of CGF-oriented cognitive architecture [J]. Journal of System Simulation, 2021, 33(2): 239-248. (in Chinese)
- [2] 张琪,曾俊杰,许凯,等.基于机器学习的计算机生成兵力行为建模研究综述[J].系统仿真学报,2021,33(2):280-287.  
ZHANG Q, ZENG J J, XU K, et al. Behavior modeling for computer generated forces based on machine learning [J]. Journal of System Simulation, 2021, 33(2): 280-287. (in Chinese)
- [3] 王钦钊,张心路,郭傲兵. Game AI 在计算机生成兵力中的应用研究[J].计算机应用研究,2020,37(增刊1):17-18,5.  
WANG Q Z, ZHANG X L, GUO A B. Research on application of Game AI in computer generated forces [J]. Application Research of Computers, 2020, 37 (Suppl 1): 17-18, 5. (in Chinese)
- [4] 高昂,段莉,张国辉,等.计算机生成兵力行为建模发展现状[J].计算机工程与应用,2019,55(19):43-51.  
GAO A, DUAN L, ZHANG G H, et al. Development status of computer generated force behavior modeling [J]. Computer Engineering and Applications, 2019, 55(19): 43-51. (in Chinese)
- [5] 赵震,赵晓林,高茂林,等.军事仿真实验技术综述[J/OL].现代防御技术,2025:1-13.(2025-07-17)[2025-12-12].<https://link.cnki.net/urlid/11.3019.tj.20250716.1716.004>.  
ZHAO Z, ZHAO X L, GAO M L, et al. Overview of the techniques of military simulation experiments [J/OL]. Modern Defence Technology, 2025: 1-13. (2025-07-17) [2025-12-12]. <https://link.cnki.net/urlid/11.3019.tj.20250716.1716.004>. (in Chinese)
- [6] 刘波韬,李定主,王学文,等.基于 XSim 仿真平台的榴弹炮建模与弹道仿真[J].火力与指挥控制,2022,

- 47(3): 173-179.
- LIU B T, LI D Z, WANG X W, et al. Howitzer modeling and trajectory simulation based on XSimStudio simulation platform[J]. Fire Control & Command Control, 2022, 47(3): 173-179. (in Chinese)
- [7] 严宗睿, 冯伟强, 潘宜宏. 航母编队反潜仿真军事规则研究[J]. 现代防御技术, 2017, 45(3): 186-192.
- YAN Z R, FENG W Q, PAN X H. Military rules for ASW simulation of aircraft carrier formation[J]. Modern Defence Technology, 2017, 45(3): 186-192. (in Chinese)
- [8] 孙鹏, 谭玉玺, 汤磊. 基于有限状态机的作战实体模型行为规则可视化建模[J]. 指挥控制与仿真, 2015, 37(2): 27-30.
- SUN P, TAN Y X, TANG L. Visual modeling of combat entities behavior model rules based on finite state machine[J]. Command Control & Simulation, 2015, 37(2): 27-30. (in Chinese)
- [9] 罗鑫, 张雷, 韩斌, 等. 分队级决策行为建模研究现状与展望[J/OL]. 计算机工程与应用, 2025: 1-27. (2025-12-30)[2026-01-01]. <https://link.cnki.net/urlid/11.2127.TP.20251229.1751.011>.
- LUO X, ZHANG L, HAN B, et al. Research status and prospects of squad-level decision-making behavior modeling[J/OL]. Computer Engineering and Applications, 2025: 1-27. (2025-12-30)[2026-01-01]. <https://link.cnki.net/urlid/11.2127.TP.20251229.1751.011>. (in Chinese)
- [10] 刘大勇, 董志明, 郭齐胜, 等. LLM 赋能的战术兵棋决策 Agent 构建方法[J/OL]. 系统仿真学报, 2025: 1-16. (2025-07-24)[2026-01-01]. <https://link.cnki.net/urlid/11.3092.v.20250724.1354.004>.
- LIU D Y, DONG Z M, GUO Q S, et al. Construction approach of LLM-empowered tactical wargame decision-making agents[J/OL]. Journal of System Simulation, 2025: 1-16. (2025-07-24)[2026-01-01]. <https://link.cnki.net/urlid/11.3092.v.20250724.1354.004>. (in Chinese)
- [11] 刘奇东, 汪敏, 范红旗, 等. 基于自循环行为树框架的蓝军空中兵力行为建模[J]. 指挥控制与仿真, 2025, 47(3): 145-154.
- LIU Q D, WANG M, FAN H Q, et al. Enemy air force behavior modeling based on self-looping behavior tree framework[J]. Command Control & Simulation, 2025, 47(3): 145-154. (in Chinese)
- [12] 曹朋飞, 邸彦强, 孟宪国, 等. 基于BDD的作战实体行为一体化建模方法研究[J]. 计算机仿真, 2024, 41(7): 6-13, 49.
- CAO P F, DI Y Q, MENG X G, et al. Research on integrated modeling method of combat entity behavior based on BDD[J]. Computer Simulation, 2024, 41(7): 6-13, 49. (in Chinese)
- [13] 吴涛, 孙向军, 赵斯强. Soar 在水面舰艇 CGF 防空决策行为模型构建中的应用[J]. 指挥控制与仿真, 2013, 35(2): 108-112.
- WU T, SUN X J, ZHAO S Q. The application of Soar in surface warship CGF defense decision-making behavior modeling[J]. Command Control & Simulation, 2013, 35(2): 108-112. (in Chinese)
- [14] 王宇琨, 王泽, 董力维, 等. 基于分层的智能建模方法的多机空战行为建模[J]. 系统仿真学报, 2023, 35(10): 2249-2261.
- WANG Y K, WANG Z, DONG L W, et al. Research on multi-aircraft air combat behavior modeling based on hierarchical intelligent modeling methods[J]. Journal of System Simulation, 2023, 35(10): 2249-2261. (in Chinese)
- [15] 王成飞, 董亚卓, 苏千叶, 等. 海战仿真中的智能对抗行为建模方法研究[J]. 指挥控制与仿真, 2022, 44(1): 79-85.
- WANG C F, DONG Y Z, SU Q Y, et al. Research on modeling method of intelligent confrontation behavior in naval battle simulation[J]. Command Control & Simulation, 2022, 44(1): 79-85. (in Chinese)
- [16] ZHAO W X, ZHOU K, LI J Y, et al. A survey of large language models[EB/OL]. (2023-03-31)[2026-01-01]. <https://arxiv.org/abs/2303.18223>.
- [17] 刘大勇, 董志明, 郭齐胜, 等. 大模型赋能战术对抗仿真实验体系架构及技术路径研究[J/OL]. 计算机科学, 2025: 1-15. (2025-09-02)[2026-01-01]. <https://link.cnki.net/urlid/50.1075.tp.20250901.1830.004>.
- LIU D Y, DONG Z M, GUO Q S, et al. Research on the architecture and technology pathways for em-powering tactical adversarial simulation experiments with LLMs[J/OL]. Computer Science, 2025: 1-15. (2025-09-02)[2026-01-01]. <https://link.cnki.net/urlid/50.1075.tp.20250901.1830.004>. (in Chinese)
- [18] 成志宇, 陈星霖, 王菁, 等. 一种基于知识图谱的检索增强生成情报问答技术[J]. 计算机科学, 2025, 52(1): 87-93.
- CHENG Z Y, CHEN X L, WANG J, et al. Retrieval-augmented generative intelligence question answering technology based on knowledge graph[J]. Computer Science, 2025, 52(1): 87-93. (in Chinese)
- [19] 叶云, 邓宁. 基于预演理论与反事实反思的大模型态势预测和决策方法[J]. 兵工自动化, 2025, 44(12): 103-107.
- YE Y, DENG N. A large language model-based posture prediction and decision-making method integrating pre-simulation theory and counterfactual reflection[J]. Ordnance Industry Automation, 2025, 44(12): 103-107. (in Chinese)
- [20] 杨光飞, 孙畅, 刘振东, 等. 面向网络空间认知战的大语言模型: 技术与挑战[J]. 指挥与控制学报, 2024, 10(6): 643-652.
- YANG G F, SUN C, LIU Z D, et al. Large language models for cognitive warfare in cyberspace: technologies and challenges[J]. Journal of Command and Control, 2024, 10(6): 643-652. (in Chinese)
- [21] 吴婷婷, 於志文, 徐健. 水下群体智能[J/OL]. 智能系统学报, 2026: 1-21. (2026-01-04)[2026-01-05]. <https://link.cnki.net/urlid/23.1538.TP.20260104.0939.004>.
- WU T T, YU Z W, XU J. Underwater swarm intelligence[J/OL]. CAAI Transactions on Intelligent Systems, 2026: 1-21. (2026-01-04)[2026-01-05]. <https://link.cnki.net/urlid/23.1538.TP.20260104.0939.004>. (in Chinese)
- [22] 庞宁, 肖卫东, 赵翔, 等. 基于混合智能体的作战行动方案生成方法[J]. 指挥与控制学报, 2025, 11(2): 181-190.
- PANG N, XIAO W D, ZHAO X, et al. Course of action

- based on mixture-of-agents [J]. *Journal of Command and Control*, 2025, 11(2): 181–190. (in Chinese)
- [23] LOU R Z, ZHANG K, YIN W P. Is prompt all you need? No. A comprehensive and broader view of instruction learning[EB/OL]. (2023–03–21) [2026–01–05]. <https://arxiv.org/abs/2303.10475v2>.
- [24] QIAO S F, OU Y X, ZHANG N Y, et al. Reasoning with language model prompting: a survey [EB/OL]. (2023–09–18) [2026–01–05]. <https://arxiv.org/abs/2212.09597>.
- [25] QIN Y J, HU S D, LIN Y K, et al. Tool learning with foundation models [EB/OL]. (2024–08–06) [2026–01–05]. <https://arxiv.org/abs/2304.08354>.
- [26] SCHICK T, DWIVEDI-YU J, DESSÌ R, et al. Toolformer; language models can teach themselves to use tools[EB/OL]. (2023–02–09) [2026–01–05]. <https://arxiv.org/abs/2302.04761>.
- [27] XU C W, XU Y C, WANG S H, et al. Small models are valuable plug-ins for large language models[EB/OL]. (2023–05–15) [2026–01–05]. <https://arxiv.org/abs/2305.08848>.
- [28] QIAO S F, GUI H H, LV C F, et al. Making language models better tool learners with execution feedback[EB/OL]. (2024–03–14) [2026–01–05]. <https://arxiv.org/abs/2305.13068>.
- [29] GAO C, LAN X C, LI N, et al. Large language models empowered agent-based modeling and simulation: a survey and perspectives[EB/OL]. (2023–12–19) [2026–01–05]. <https://arxiv.org/abs/2312.11970>.
- [30] GUO T C, CHEN X Y, WANG Y Q, et al. Large language model based multi-agents: a survey of progress and challenges[EB/OL]. (2024–04–19) [2026–01–05]. <https://arxiv.org/abs/2402.01680>.
- [31] HUANG X, LIU W W, CHEN X L, et al. Understanding the planning of LLM agents: a survey [EB/OL]. (2024–02–05) [2026–01–05]. <https://arxiv.org/abs/2402.02716>.
- [32] SUMERS T R, YAO S Y, NARASIMHAN K, et al. Cognitive architectures for language agents[EB/OL]. (2024–03–15) [2026–01–05]. <https://arxiv.org/abs/2309.02427>.
- [33] PARK J S, O'BRIEN J C, CAI C J, et al. Generative agents: interactive simulacra of human behavior [EB/OL]. (2023–08–06) [2026–01–05]. <https://arxiv.org/abs/2304.03442>.
- [34] HU H, FENG Y F, LI R X, et al. Cog-RAG: cognitive-inspired dual-hypergraph with theme alignment retrieval-augmented generation[EB/OL]. (2025–12–16) [2026–01–05]. <https://arxiv.org/abs/2511.13201>.
- [35] LI Z Q, CHEN X A, YU H Y, et al. StructRAG: boosting knowledge intensive reasoning of LLMs via inference-time hybrid information structurization [EB/OL]. (2024–10–25) [2026–01–05]. <https://arxiv.org/abs/2410.08815>.
- [36] LEE M, AN S, KIM M S. PlanRAG: a plan-then-retrieval augmented generation for generative large language models as decision makers[EB/OL]. (2024–06–18) [2026–01–05]. <https://arxiv.org/abs/2406.12430>.
- [37] WANG S T, YU X, WANG M, et al. RichRAG: crafting rich responses for multi-faceted queries in retrieval-augmented generation[EB/OL]. (2024–10–01) [2026–01–06]. <https://arxiv.org/abs/2406.12566>.
- [38] TANG Q Y, CHEN J W, LI Z Q, et al. Self-retrieval: end-to-end information retrieval with one large language model[EB/OL]. (2024–11–04) [2026–01–06]. <https://arxiv.org/abs/2403.00801>.
- [39] ZHANG J, TANG B, SHAO W Z, et al. TAdaRAG: task adaptive retrieval-augmented generation via on-the-fly knowledge graph construction [EB/OL]. (2025–11–16) [2026–01–06]. <https://arxiv.org/abs/2511.12520>.
- [40] GAO Y F, XIONG Y, GAO X Y, et al. Retrieval-augmented generation for large language models: a survey [EB/OL]. (2024–03–27) [2026–01–06]. <https://arxiv.org/abs/2312.10997>.
- [41] SHINN N, CASSANO F, BERMAN E, et al. Reflexion: language agents with verbal reinforcement learning[EB/OL]. (2023–10–10) [2026–01–06]. <https://arxiv.org/abs/2303.11366>.
- [42] WEI T X, SACHDEVA N, COLEMAN B, et al. Evomemory: benchmarking LLM agent test-time learning with self-evolving memory[EB/OL]. (2025–11–25) [2026–01–06]. <https://arxiv.org/abs/2511.20857>.
- [43] YAO S Y, ZHAO J, YU D, et al. ReAct: synergizing reasoning and acting in language models[C]//Proceedings of the International Conference on Learning Representations (ICLR), 2023.
- [44] WEI J, WANG X Z, SCHUURMANS D, et al. Chain-of-thought prompting elicits reasoning in large language models[C]//Proceedings of the 36th International Conference on Neural Information Processing Systems, 2022; 24824–24837.
- [45] CHU Z, CHEN J C, CHEN Q L, et al. A survey of chain of thought reasoning: advances, frontiers and future[EB/OL]. (2023–09–27) [2026–01–06]. <https://arxiv.org/pdf/2309.15402v2>.
- [46] RAN Y T, WANG X T, XU R, et al. Capturing minds, not just words: enhancing role-playing language models with personality-indicative data [C]//Proceedings of the Findings of the Association for Computational Linguistics: EMNLP 2024, 2024; 14566–14576.
- [47] SHAO Y F, LI L Y, DAI J Q, et al. Character-LLM: a trainable agent for role-playing[C]//Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, 2023; 13153–13187.
- [48] SHEN C L, XIE G F, ZHANG X, et al. On the decision-making abilities in role-playing using large language models[EB/OL]. (2024–02–29) [2026–01–06]. <https://arxiv.org/abs/2402.18807>.
- [49] WANG N, PENG Z Y, QUE H R, et al. RoleLLM: benchmarking, eliciting, and enhancing role-playing abilities of large language models[C]//Proceedings of the Findings of the Association for Computational Linguistics: ACL 2024, 2024; 14743–14777.
- [50] OUYANG L, WU J, JIANG X, et al. Training language models to follow instructions with human feedback [C]//Proceedings of the 36th International Conference on Neural Information Processing System, 2022; 27730–27744.
- [51] ZHANG H X, DU W H, SHAN J M, et al. Building cooperative embodied agents modularly with large language models[C]//Proceedings of the International Conference on Learning Representations (ICLR), 2024.
- [52] LIANG J, HUANG W L, XIA F, et al. Code as policies;

- language model programs for embodied control [ C ]// Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), 2023; 9493–9500.
- [53] JANG J, PARK W J, WOO H, et al. Exploratory retrieval-augmented planning for continual embodied instruction following [ C ]// Proceedings of the Advances in Neural Information Processing Systems 37, 2024; 67034–67060.
- [54] WU Y, MIN S Y, BISK Y, et al. Plan, eliminate, and track; language models are good teachers for embodied agents[EB/OL]. (2023–05–07) [2026–01–06]. <https://arxiv.org/abs/2305.02412>.
- [55] KELLEY R. Behavior trees enable structured programming of language model agents[EB/OL]. (2024–04–11) [2026–01–06]. <https://arxiv.org/abs/2404.07439>.
- [56] HUANG W L, ABBEEL P, PATHAK D, et al. Language models as zero-shot planners; extracting actionable knowledge for embodied agents[EB/OL]. (2022–03–08) [2026–01–06]. <https://arxiv.org/abs/2201.07207>.
- [57] SONG C H, WU J M, WASHINGTON C, et al. LLM-planner: few-shot grounded planning for embodied agents with large language models[EB/OL]. (2023–03–30) [2026–01–06]. <https://arxiv.org/abs/2212.04088>.
- [58] CHEN L T, WANG L, DONG H, et al. Introspective tips: large language model for in-context decision making [EB/OL]. (2023–05–19) [2026–01–06]. <https://arxiv.org/abs/2305.11598>.
- [59] ZHU X Z, CHEN Y T, TIAN H, et al. Ghost in the Minecraft: generally capable agents for open-world environments via large language models with text-based knowledge and memory[EB/OL]. (2023–06–01) [2026–01–06]. <https://arxiv.org/abs/2305.17144>.
- [60] MISRAHI A, CHIRKOVA N, LOUIS M, et al. Adapting large language models for multi-domain retrieval-augmented-generation[EB/OL]. (2025–04–03) [2026–01–06]. <https://arxiv.org/abs/2504.02411>.
- [61] GUO L, SHAN Y X, ZHU Z B, et al. RAD: retrieval high-quality demonstrations to enhance decision-making[EB/OL]. (2025–07–21) [2026–01–06]. <https://arxiv.org/abs/2507.15356>.
- [62] WU X, HUANG T Z, DENG L J, et al. A knowledge-driven adaptive collaboration of LLMs for enhancing medical decision-making[EB/OL]. (2025–09–18) [2026–01–06]. <https://arxiv.org/abs/2509.14998>.
- [63] ZHANG C Q, SHU C, SHAREGHI E, et al. All roads lead to Rome: graph-based confidence estimation for large language model reasoning[EB/OL]. (2025–09–16) [2026–01–06]. <https://arxiv.org/abs/2509.12908>.
- [64] WANG J Q, JI B Q, LUO H B, et al. LTA-Thinker: latent thought-augmented training framework for large language models on complex reasoning[EB/OL]. (2025–12–16) [2026–01–06]. <https://arxiv.org/abs/2509.12875>.
- [65] ZHENG H S, MISHRA S, ZHANG H, et al. NATURAL PLAN: benchmarking LLMs on natural language planning[EB/OL]. (2024–06–06) [2026–01–06]. <https://arxiv.org/abs/2406.04520>.
- [66] VALMEEKAM K, MARQUEZ M, OLMO A, et al. PlanBench: an extensible benchmark for evaluating large language models on planning and reasoning about change[EB/OL]. (2023–11–26) [2026–01–06]. <https://arxiv.org/abs/2206.10498>.
- [67] BENYAMIN Y, MORDOCH A, SHPERBERG S S, et al. Toward PDDL planning copilot[EB/OL]. (2025–09–16) [2026–01–06]. <https://arxiv.org/abs/2509.12987>.
- [68] XI Z H, HUANG J X, LIAO C Y, et al. AgentGym-RL: training LLM agents for long-horizon decision making through multi-turn reinforcement learning[EB/OL]. (2025–09–10) [2026–01–06]. <https://arxiv.org/abs/2509.08755>.
- [69] JIAO K C, FANG Z R, LIU J H, et al. TCPO: thought-centric preference optimization for effective embodied decision-making[EB/OL]. (2025–09–10) [2026–01–06]. <https://arxiv.org/abs/2509.08500>.
- [70] DUAN J H, WANG S Q, DIFFENDERFER J, et al. ReTA: recursively thinking ahead to improve the strategic reasoning of large language models [ C ]// Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics; Human Language Technologies (Volume 1: Long Papers), 2024; 2232–2246.
- [71] CUI S J, HE A Y, XU S, et al. Self-guided function calling in large language models via stepwise experience recall[EB/OL]. (2025–08–21) [2026–01–06]. <https://arxiv.org/abs/2508.15214>.
- [72] BECKER J, KAESBERG L B, BAUER N, et al. MALLM: multi-agent large language models framework [EB/OL]. (2025–12–15) [2026–01–06]. <https://arxiv.org/abs/2509.11656>.
- [73] ARATCHIGE R M, ILMINI W M K S. LLMs working in harmony: a survey on the technological aspects of building effective LLM-based multi agent systems[EB/OL]. (2025–03–13) [2026–01–06]. <https://arxiv.org/abs/2504.01963>.
- [74] POKHAREL A, DANTU R, ZAMAN S, et al. Achieving unanimous consensus through multi-agent deliberation [EB/OL]. (2025–04–02) [2026–01–06]. <https://arxiv.org/abs/2504.02128>.
- [75] HONG X X, WANG Y G, JIN D X, et al. HLSMAC: a new StarCraft multi-agent challenge for high-level strategic decision-making[EB/OL]. (2025–09–16) [2026–01–06]. <https://arxiv.org/abs/2509.12927>.
- [76] AHN M, BROHAN A, BROWN N, et al. Do as I can, not as I say: grounding language in robotic affordances[EB/OL]. (2022–08–16) [2026–01–06]. <https://arxiv.org/abs/2204.01691>.
- [77] XIANG J N, TAO T H, GU Y, et al. Language models meet world models; embodied experiences enhance language models[EB/OL]. (2023–10–28) [2026–01–06]. <https://arxiv.org/abs/2305.10626>.
- [78] ZHENG M Q, PEI J X, LOGESWARAN L, et al. When "a helpful assistant" is not really helpful: personas in system prompts do not improve performances of large language models[EB/OL]. (2024–10–09) [2026–01–06]. <http://arxiv.org/abs/2311.10054>.
- [79] MIN S, LYU X X, HOLTZMAN A, et al. Rethinking the role of demonstrations: what makes in-context learning work [EB/OL]. (2022–10–20) [2026–01–06]. <https://arxiv.org/abs/2202.12837>.
- [80] ZHANG Z Y, BO X H, MA C, et al. A survey on the memory mechanism of large language model based agents[EB/OL]. (2024–04–21) [2026–01–06]. <https://arxiv.org/abs/2404.13501>.
- [81] SUZGUN M, YUKSEKGONUL M, BIANCHI F, et al.

- Dynamic cheatsheet: test-time learning with adaptive memory[EB/OL]. (2025-04-10)[2026-01-06]. <https://arxiv.org/abs/2504.07952>.
- [82] XI R, WANG X H. Livia: an emotion-aware AR companion powered by modular AI agents and progressive memory compression[EB/OL]. (2025-08-12)[2026-01-06]. <https://arxiv.org/abs/2509.05298>.
- [83] WANG Z X, YU B, ZHAO J Z, et al. KARMA: augmenting embodied AI agents with long-and-short term memory systems[EB/OL]. (2025-03-21)[2026-01-06]. <https://arxiv.org/abs/2409.14908>.
- [84] HUANG L, YU W J, MA W T, et al. A survey on hallucination in large language models: principles, taxonomy, challenges, and open questions[J]. *ACM Transactions on Information Systems*, 2025, 43(2): 1-55.
- [85] HOU B R, ZHANG Y, JI J B, et al. ThinkPrune: pruning long chain-of-thought of LLMs via reinforcement learning[EB/OL]. (2025-04-02)[2026-01-06]. <https://arxiv.org/abs/2504.01296>.
- [86] ZHOU A, YAN K, SHLAPENTOKH-ROTHMAN M, et al. Language agent tree search unifies reasoning acting and planning in language models[EB/OL]. (2025-06-06)[2026-01-06]. <https://arxiv.org/abs/2310.04406>.
- [87] YE S C, YU C, KE K Q, et al. H2R: hierarchical hindsight reflection for multi-task LLM agents[EB/OL]. (2025-09-16)[2026-01-06]. <https://arxiv.org/abs/2509.12810>.
- [88] YUAN L Q, HAN D J, BRINTON C G, et al. LLMAP: LLM-assisted multi-objective route planning with user preferences[EB/OL]. (2025-09-14)[2026-01-06]. <https://arxiv.org/abs/2509.12273>.
- [89] DING J T, ZHANG Y K, SHANG Y, et al. Understanding world or predicting future? A comprehensive survey of world models[J]. *ACM Computing Surveys*, 2025, 58(3): 1-38.
- [90] LIU J Q, SU Y F, XIA P, et al. SimpleMem: efficient lifelong memory for LLM agents[EB/OL]. (2026-01-05)[2026-01-06]. <https://arxiv.org/abs/2601.02553>.
- [91] VERMA N. Active context compression: autonomous memory management in LLM agents[EB/OL]. (2026-01-12)[2026-01-13]. <https://arxiv.org/abs/2601.07190>.
- [92] MA X J, GAO Y F, WANG Y X, et al. Safety at scale: a comprehensive survey of large model and agent safety[EB/OL]. (2025-02-02)[2026-01-13]. <https://arxiv.org/abs/2502.05206>.