

基于大语言模型的集群协同决策研究

邹通¹, 丁学良^{2,3}, 戴瀚苏², 李超勇¹

(1.浙江大学电气工程学院, 杭州 310027; 2.中国航空工业集团公司成都飞机设计研究所, 成都 610091;

3.空天飞行器技术航空科技重点实验室, 成都 610091)

摘要: 针对大规模动态作业环境中多域协同决策的复杂性与实时性挑战, 提出了基于大语言模型的集群协同决策框架 (Courses of Action-Large Language Models, COA-LLM), 通过提示词工程、环境态势文本化与多级指令解析, 构建了"感知-决策-执行"闭环响应系统。突破了传统目标分配方法仅能处理离散任务的限制, 实现了离散任务分配与连续空间决策的统一。最终, 在仿真系统中构建了集群协同场景, 使用了 GPT 与 DeepSeek 等通用大模型, 并与其他基于优化算法和机器学习的算法进行了对比实验, 实验结果证明了 COA-LLM 框架的可行性和有效性, 框架的输入灵活性与输出可解释性, 为智能指控系统的发展提供了新范式。

关键词: 集群协同; 任务分配; 大语言模型; 多域协同; 知识嵌入; DeepSeek

中图分类号: V279; E91 文献标志码: A 文章编号:

DOI:

Research on Swarm Operations Decision-Making based on Large Language Models

Zou Tong¹, Ding Xueliang^{2,3}, Dai Hansu², Li Chaoyong¹

(1. College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China;

2. Avic Chengdu Aircraft Design & Research Institute, Chengdu 610091, China;

3. Aviation Key Laboratory of Science and Technology on Aerospace Vehicle, Chengdu 610091, China)

Abstract: To address the complexity and real-time challenges of multi-domain collaborative decision-making in large-scale dynamic environment, a swarm cooperation decision-making framework based on large language models, named Courses of Action-Large Language Models (COA-LLM), is proposed. By leveraging prompt engineering, textual conversion of environment situations, and multi-level instruction parsing, a "perception-decision-execution" closed-loop response system is constructed. This framework overcomes the limitations of traditional target allocation methods, which are restricted to handling discrete tasks, and achieves the unification of discrete task allocation and continuous spatial decision-making. Finally, a swarm cooperation scenario is built in a simulation system, where general-purpose large models such as GPT and DeepSeek are employed. Comparative experiments are conducted against algorithms based on optimization and machine learning. The experimental results demonstrate the feasibility and effectiveness of the COA-LLM framework, highlighting its input flexibility and output interpretability, which provides a new paradigm for the development of intelligent command and control systems.

Keywords: swarm cooperation; task allocation; LLM; Multi-Domain Collaboration; knowledge embedding; DeepSeek

收稿日期: xxxx-xx-xx 修回日期: xxxx-xx-xx

基金项目: 国家自然科学基金(U2441205)

第一作者: 邹通, 男, 硕士研究生, 研究方向为集群决策。E-mail: 22210160@zju.edu.cn

通讯作者: 李超勇, 男, 研究员, 博士生导师。E-mail: Chaoyong@zju.edu.cn

0 引言

随着任务装备的发展,复杂环境的多域协同特性对指挥与协调系统(Command and Control, C2)提出了更高要求,亟需实现动态多变环境下的资源快速整合与优化。作为决策核心的行动方案生成(Courses of Action, COA)生成,要求系统兼具实时态势感知与动态规划能力^[1]。

传统方法主要依赖基于规则的优化模型、动态规划算法、博弈理论和强化学习等技术。例如,Chen 等^[2]建立了考虑效能最优性和任务收益的资源分配模型;Zhong 等^[3]通过将动态规划和多子群蚁群算法相结合,研究了时变约束下的任务分配问题;Duan 等^[4]将捕食-被捕食混合粒子群算法(PP-PSO)应用于混合策略纳什解的求解,用以解决多智能体动态任务分配问题;Orafa^[5]等则采用 Q 学习算法处理大规模决策问题。然而,这些方法仍存在显著不足:传统优化算法(如匈牙利算法)仅适用于离散目标分配,难以在连续空间生成控制指令;基于规则的风险评估依赖人工定义优先级规则,面对动态环境时缺乏适应性;而强化学习方法由于需要对异构智能体分别训练模型,在大规模协同场景中难以快速部署。

大型语言模型(Large Language Models, LLM)是自然语言处理领域的一项变革性技术^[6-7]。其卓越的语义理解和生成能力为智能决策开辟了新途径。该技术已在灾难响应^[8-9]和机器人技术^[10]等领域展现出强大潜力。在家庭服务场景中,Ahn^[11]和 Mees^[12]等设计的 SayCan 框架通过 LLM 与可供性函数的结合,可以根据用户指令完成清洁、整理等任务。在游戏《我的世界》中,Wang 等^[13]设计了基于 GPT-4 的 Voyager 智能体,利用自动课程、技能库和迭代提示机制等,实现了在游戏世界中的自主探索。

军事领域也在加速应用 LLM 技术:Palantir 公司发布的人工智能平台 AIP(artificial intelligence platform)^[14],能够从数据分析、态势理解、任务规划等方面为决策提供全流程支撑,被认为是未来军事指控大模型的一种发展雏形。美国 Scale AI 公司推出“多诺万(Donovan)”军事决策辅助系统^[15],通过多源数据融合为指挥官提供优化的作战方案和战略建议。国内源亨科技发布“天机大模型”^[7],通过对基座大模型的指令微调 and 反馈训练,融合军事术语常识、数字战场图谱、战略条令法规等国防领域多维度数据,构建了面向军事场景的专用大模型。Lamparth 等^[16]在模拟真实战争场景下,通过多角度验证表明 LLM 生成的作战方案与人类决策具有高度一致性。Vinicius 等^[17]提出了一种基于生成式预训练 LLM 的 COA 制定方法,并在星际

争霸的小规模对抗场景中进行了验证。

基于此,本文提出了一种基于大语言模型的集群协同决策框架(COA-LLM),该方法利用 LLM 的强大语义理解与生成能力,通过多层级整合解决传统方法的问题。本文的主要贡献包括:

1) 提出了一种基于大语言模型的集群智能决策框架,为行动方案生成问题提供了新的研究范式;

2) 以 LLM 为核心,将目标分配、风险评估、路径规划等传统分散模块整合为语义驱动的统一框架,通过自然语言指令解决了动态环境中多模块协同的实时性与适应性问题;

3) 通过构造多种类本文指令,实现了离散任务分配与连续空间决策的统一;

4) 将传统控制问题转化为文本生成问题,在保证可解释性的同时,利用 LLM 的泛化能力覆盖复杂场景。

1 智能体协同系统建模

本文研究红方(主动方)与蓝方(响应方)的动态交互过程。研究场景以飞行器为主要单元,本章重点对飞行器类单元的运动学特性、动力学模型及控制约束进行建模与分析。

1.1 动力学建模

设定双方智能体数量分别为 N^a 和 N^b ,场景空间约束为 $W \times L \times H$ 三维区域。

交互过程采用离散时间模型,时间步长 $t_k = kT$, $k=0, \dots, K$ 且 T 为预定的时间间隔,仿真采用三维质点模型,单位 i 的运动模型可描述如下:

$$\begin{cases} x_i(k+1) = x_i(k) + Tv_i(k)\cos\phi_i(k)\cos\gamma_i(k) \\ y_i(k+1) = y_i(k) + Tv_i(k)\sin\phi_i(k)\cos\gamma_i(k) \\ z_i(k+1) = z_i(k) + Tv_i(k)\sin\gamma_i(k) \\ v_i(k+1) = v_i(k) + V_i(k) \\ \phi_i(k+1) = \phi_i(k) + \psi_i(k) \\ \gamma_i(k+1) = \gamma_i(k) + \Upsilon_i(k) \end{cases} \quad (1)$$

式中, $[x_i(k), y_i(k), z_i(k)]^T$ 表示位置, $V_i(k)$ 、 $\psi_i(k)$ 和 $\Upsilon_i(k)$ 代表对应变量在 T 时间间隔内的变化量。 $v_i(k)$ 、 $\phi_i(k)$ 和 $\gamma_i(k)$ 分别代表飞行器的速度、航向角和俯仰角。

设 v_i^{\max} 、 ϕ_i^{\max} 与 γ_i^{\max} 分别为单位 i 的速度、航向角及俯仰角上限,则需满足:

$$\begin{cases} v_i(k) \leq v_i^{\max} \\ \phi_i(k) \leq \phi_i^{\max} \\ \gamma_i(k) \leq \gamma_i^{\max} \end{cases} \quad (2)$$

底层控制指令表示为

$\mathbf{u}_i^x(k)=[v_i^{\text{exp}}(k), \phi_i^{\text{exp}}(k), \gamma_i^{\text{exp}}(k)]$, 分别代表期望的速度大小、俯仰角以及偏航角, 本场景中限定各单位速度大小不变, $v_i^{\text{exp}}(k)$ 为恒值。

决策中枢以目标点 $[x_t, y_t, z_t]$ 的形式发出指令, 为满足控制器接口需求, 通过以下公式进行转换:

$$\begin{cases} \phi_i^{\text{exp}}(k) = \arctan 2(y_t - y_i(k), x_t - x_i(k)) \\ \gamma_i^{\text{exp}}(k) = \arctan\left(\frac{z_t - z_i(k)}{\sqrt{(x_t - x_i(k))^2 + (y_t - y_i(k))^2}}\right) \end{cases} \quad (3)$$

定义单位 i 的状态为:

$$\lambda_i^x(k)=[x_i^x(k), y_i^x(k), z_i^x(k), p_i^x(k), w_i^x(k)]^T \quad (4)$$

式中, $p_i^x(k)$ 和 $w_i^x(k)$ 代表单位剩余的耐久度和剩余攻击次数, $X \in \{R, B\}$ 分别代表红方和蓝方阵营。

1.2 交互规则建模

设 $d_{ij}^{XY}(k)$ 与 $\theta_{ij}^{XY}(k)$ 分别表示 k 时刻红方单位 X_i 与蓝方单位 Y_j 间的距离和视线角, $t_i^{\text{last}}(k)$ 为 X_i 上一次攻击时间, $w_i^x(k)$ 为 X_i 的剩余攻击次数。当 X_i 满足以下条件时可对 Y_j 发动攻击:

$$\begin{cases} d_{ij}^{XY}(k) \leq D_i^x \\ \theta_{ij}^{XY}(k) \leq \theta_i^x \\ k - t_i^{\text{last}}(k) \geq T_i^{\text{fire}} \\ w_i^x(k) \geq 1 \end{cases} \quad (5)$$

式中, D_i^x 为最大攻击距离, θ_i^x 为最大限制视线角, T_i^{fire} 为最小攻击间隔时间。

若 X_i 成功攻击 Y_j , 则 Y_j 耐久度的更新方式表示如下:

$$\begin{cases} p_j^y(k+1) = p_j^y(k) - \Delta p_j^y(k+1) \\ \Delta p_j^y(k+1) = \sum_{i=1}^{N^x} P_{ij}^{xy}(k) A_i^x(k) \end{cases} \quad (6)$$

式中, $P_{ij}^{xy}(k) \in \{0, 1\}$, $P_{ij}^{xy}(k) = 1$ 表示单位 i 对单位 j 发动了攻击, $A_i^x(k)$ 代表单位 i 的攻击力。当 $p_j^y(k) < 0$ 时, 单位 j 失效并清离环境。

2 基于大语言模型的智能决策

本文基于 LLM 的强大文本处理能力, 将高层次的、模糊的指挥决策问题转化为具体的、可执行的任务规划方案, 提出了基于大语言模型的集群协同决策框架 (COA-LLM)。本框架的核心目标是通过 LLM 的语义理解能力, 解决传统方法中动态任务分解、多模块协同与异构智能体协同决策三大问题。其设计原则包括:

1) 通过系统提示词 (System Message) 重构 LLM, 使其具备理解环境信息并生成行动方案的能力;

力;

2) 以自然语言为接口, 通过构造多样化的指令类型, 仅需输入观测信息即可获取控制指令, 使单个模型统一了目标分配、风险评估、路径规划等任务;

3) 通过领域知识库与物理规则限制 LLM 的自由度, 确保生成指令的可行性;



图 1 仿真系统框架

Fig. 1 Simulation System Framework

如图 1 所示, 仿真系统的迭代遵循 OODA 循环^[18] (Observe 观察, Orient 判断, Decide 决策, Act 行动), 环境中观测到的信息输入 COA-LLM, 由 COA-LLM 统一进行判断与决策环节, 最后模型输出的指令作用于环境中的各单位, 循环迭代至仿真终止。

COA-LLM 的应用主要由 4 部分组成。1) 提示词工程: 创建能够有效引导 LLM 处理和生成任务规划的提示。2) 环境态势文本化: 将环境态势 (各单位位置分布等信息) 转化为结构化的文本形式, 形成 LLM 能够理解的指令输入。3) 指令解析: 读取 LLM 的回复词, 解析其中的有效指令。4) 指令执行: 各单位执行指令进行仿真。

2.1 提示词工程

提示词工程 (Prompt Engineering) 是指通过塑造合适的提示, 引导 LLM 生成符合领域规则输出的关键技术^[19]。在多智能体协同决策场景中, 其核心挑战在于将抽象的策略原则 (如"资源优化配置"、"系统效能维持") 转化为 LLM 可解析的显式约束。本文主要通过系统提示词 (System Message) 设定模型的角色、行为模式和任务目标, 使其能够理解并完成复杂环境下的决策任务, 具备理解动态信息并生成行动方案的能力。设计的系统提示词由角色定义、单元特性描述、文本规范和任务提示四

部分组成。

2.1.1 角色塑造

如图 2 所示，角色塑造部分通过自然语言定义模型的虚拟身份与任务目标，使其模拟协同决策中枢的逻辑推理过程。在动态决策的应用场景中，主要确立模型作为“智能决策辅助系统”的功能属性，以结构化文本的形式引导 LLM 的决策思路。

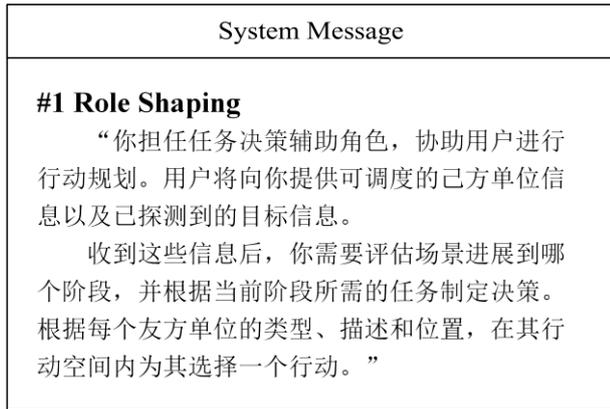


图 2 系统提示词的角色塑造部分
Fig. 2 Role Shaping section of System Message

2.1.2 智能体信息

如图 3 所示，智能体信息部分以规范的 JSON 格式将各类协同单元的参数信息嵌入 LLM。智能体信息包括单元类型名称、功能描述和动作空间 (Act Space) 三部分，动作空间定义了该类型单元允许执行的操作指令集。

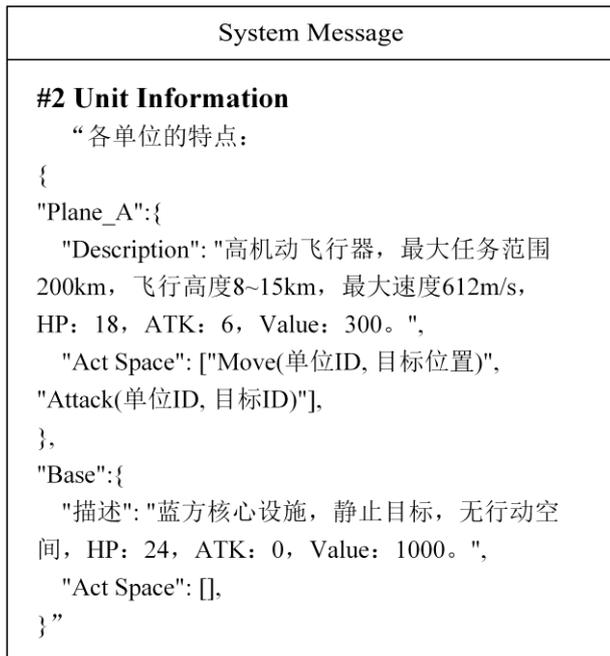


图 3 系统提示词的智能体信息部分
Fig. 3 Unit Information section of System Message

图 3 系统提示词的单位信息部分
Fig. 3 Unit Information section of System Message

通过文本的方式，可以自定义各单位可执行的指令类型，解决了传统目标分配问题只能处理离散目标（如“单位 R_i 追踪单位 B_i ”），无法生成包含连续坐标点的指令（如“移动到坐标 $[x_t, y_t, z_t]$ ”）的问题，实现了离散任务分配与连续空间决策的统一。

2.1.3 文本规范

文本规范部分将各类型指令的执行逻辑嵌入 LLM，通过结构化模板与语法约束，给予 LLM 回复实例，确保其生成的指令具备可解析性与可执行性；并要求 LLM 为其决策提供推理依据，确保决策过程的逻辑可验证，具体实现如图 4 所示。

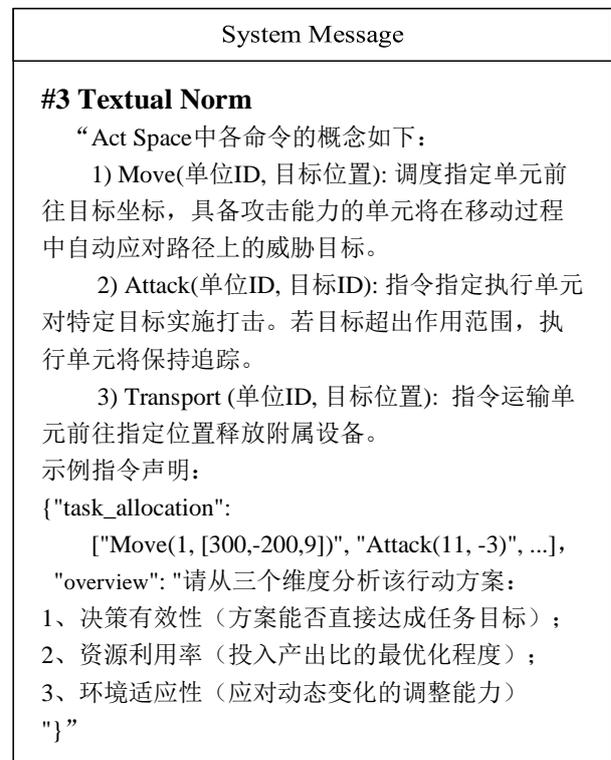


图 4 系统提示词的文本规范部分
Fig. 4 Textual Norm section of System Message

2.1.4 提示信息

如图 5 所示，提示信息部分通过嵌入高阶战术原则与效能评估规则，弥补 LLM 多智能体决策领域知识的不足，赋予 LLM 动态适应环境态势的深度推理能力。

除此之外，文本交互形式提供了用户灵活的可操作性，用户可制定“优先处理特定目标”、“无人设备协同工作”等文本信息，甚至随环境变化动态

调整提示内容, 引导模型输出符合用户预期的决策指令。克服了基于规则的专家系统的静态规则依赖 (需要预定义所有可能场景), 能够灵活适应动态变化的任务场景。

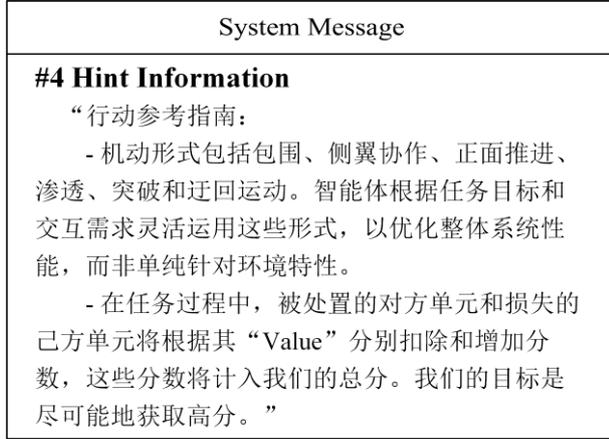


图 5 系统提示词的提示信息部分
Fig. 5 Hint Information section of System Message

2.2 环境态势文本化

环境态势文本化模块是连接仿真环境与 LLM 的关键数据接口, 其核心任务是将多维度的环境信息转化为结构化、机器可解析的文本输入, 确保模型能够准确感知环境变化并生成合理决策。

实现流程包含以下关键步骤: 1) 筛选任务单元: 包括可进行控制的我方设备和在感知范围内的对方单位。2) 数据标准化: 将仿真环境中的多模态数据 (位置、类型、状态) 按 JSON 格式统一编码。最终得到的文本用户信息 (User Message) 如图 6 所示, 用户信息输入被 System Message 重构过的 LLM。

2.3 返回词解析

LLM 的回复信息 (Assistant Message) 如图 7 所示, 返回词解析模块是将自由文本指令转化为可

执行命令的核心环节。由于 LLM 本质上是概率生成模型, 其输出可能包含格式错误、语义矛盾或违背任务约束的指令。

本模块通过多级验证机制与异常恢复策略, 实现高鲁棒性的指令解析与错误隔离。本模块的核心在于错误处理机制, 针对 LLM 可能生成的五类典型错误 (表 1), 设计分级恢复策略。

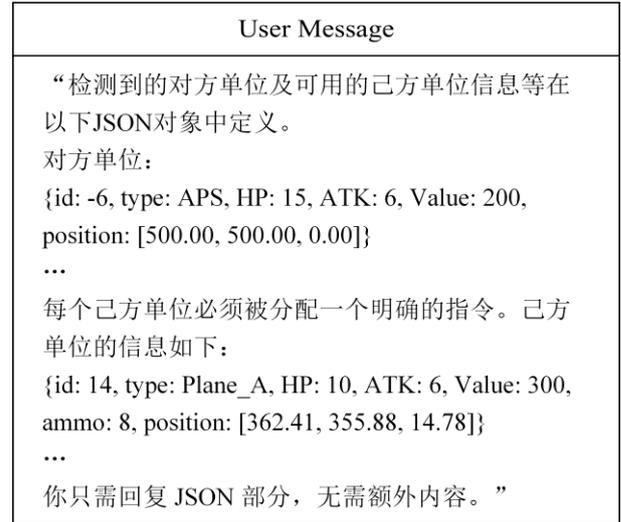


图 6 用户信息
Fig. 6 User Message

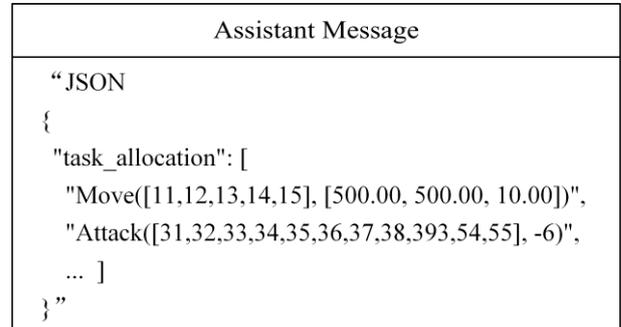


图 7 回复信息
Fig. 7 Assistant Message

表 1 错误处理机制

Table 1 Error handling mechanism

错误类型	检测方法	恢复策略
回复格式错误	JSON 解析异常	丢弃当前回复, 重新请求 LLM 生成
单条指令格式错误	正则匹配失败	忽略该指令并记录日志
非法单位 ID	与存活单位列表比对缺失	忽略该指令并记录日志
动作越界	指令类型不在单位动作空间中	忽略该指令并记录日志
参数越界	坐标/目标超出地图范围或高度限制	自动修正至最近合法值

2.4 指令执行

指令执行是 COA-LLM 框架的核心模块, 负责将 LLM 生成的高层次语义指令 (如 Move、Attack)

转化为底层可执行动作, 并确保其符合动力学约束与任务规则。LLM 作为决策中枢, 在任务过程中以固定时间间隔 ($T_{dec} = nT_{sim}, n \in \mathbf{N}^+$) 介入环境, 仅

实现从环境信息（图 6）到控制指令（图 7）的映射。生成指令后，由指令执行模块在仿真步长（ T_{sim} ）内持续驱动单位行为，直至达到终止条件，指令执行主要通过规则类编码实现。

1) 移动指令转换

将 LLM 生成的高层次移动指令（如“Move”）通过公式(3)转化为期望航向角（ ϕ_i^{exp} ）与俯仰角（ γ_i^{exp} ），并结合动力学模型（公式(1)）动态调整控制参数，进而实现飞行器在三维空间内的简单路径规划。

2) 规则指令验证

主要包括两类核心指令处理，在每个 T_{sim} 间隔进行规则判断：

(1) “Transport”指令：判断运载单位是否达到目标点，满足条件即释放小型飞行器；

(2) “Attack”指令：攻击类单位根据公式(5)判断是否满足对目标的触发条件，符合条件则执行相应操作。

3) 指令保持

由于决策周期（ T_{dec} ）与仿真步长（ T_{sim} ）存在数量级差异，为避免因决策延迟导致的动作震荡问题，本研究通过有限状态机^[21]方法管理单位行为状态。通过建立完备的状态转移机制，确保控制指令能够持续稳定执行直至任务终止。具体而言，本研究为单元设计了层次化的任务状态集合，每个状态均设有严格的行为约束规则：

(1) “Move”：单位向指令给定的目标点飞行。执行类单位处于该状态时，在满足操作约束的条件下自动响应沿途目标。单位位置与目标点的欧式距离 $d \leq \varepsilon$ 时移动终止，切换至待命（“Wait”）状态。

(2) “Attack”：执行类设备特有状态，单位仅针对指令指定的目标进行操作，持续向目标坐标点移动并执行任务，目标消失或超出探测范围时执行终止，切换至待命（“Wait”）状态。

(3) “Transport”：运载类单位特有状态，单位向指令给定的目标点飞行，抵达目标点或遭遇威胁时，在当前坐标释放附属单位，运输终止，退出任务区域。

(4) “Wait”：等待指令状态，单位在当前位置保持待机，自动响应临近目标。

3 仿真及结果分析

为了验证 LLM 在多智能体协同场景中的决策效能，本研究设计了动态交互仿真环境。基于 Matlab 平台构建了高精度仿真系统，通过 API 接口^[22]集成通用 LLM 的决策能力。

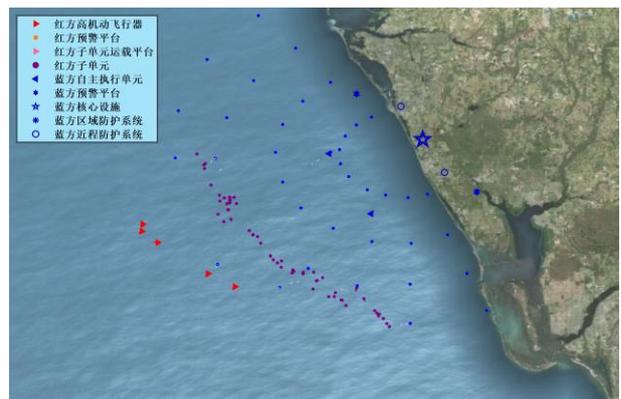
红方采用 LLM 与其他基准对比方法作为指挥

中枢，决策间隔 30s；蓝方部署基于规则的专家系统^[23]，各单位依托预设战术逻辑实现分布式自主决策。

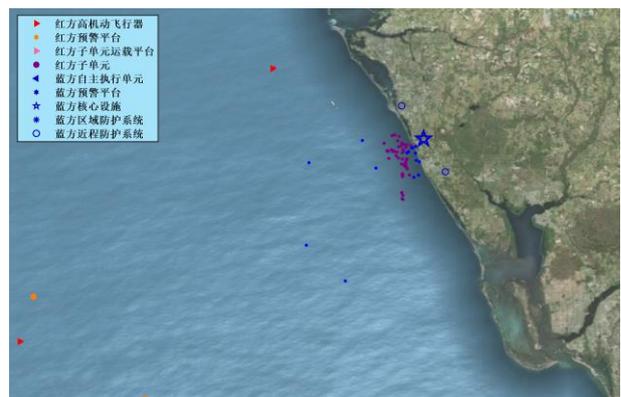
3.1 任务想定



(a) Proximity stage



(b) Dynamic interaction



(c) Final stage

图 8 场景流程节点

Fig. 8 Scenario process node

交互场景设定在 900km×900 km×25km 的三维

空间内, 主动方初始部署于坐标(0, 0, 0)区域; 响应方的核心设施位于(725, 600, 0)。实验过程设计如下(场景流程节点见图 9):

红方通过监测系统获取蓝方的大致位置后, 迅速组织由高机动飞行器、预警平台及可部署子单元系统(搭载多组小型执行单元)构成的联合任务群, 向目标区域推进; 蓝方随即启动全域监测, 机动单元快速部署并构建防御阵列; 红方释放子单元群, 双方随即进入动态交互状态; 最终, 红方突破响应方防御体系, 达成目标任务。

3.2 场景参数设置

双方的实验单元配置如表 2 及表 3 所示。

表 2 红方主要单元与配置

序号	单位类型	数量
1	高机动飞行器	5
2	预警平台	2
3	侦察型子单元	120
4	任务执行子单元	200
5	子单元运载平台	8

表 3 蓝方主要单元与配置

序号	单位类型	数量
1	核心设施	1
2	区域防护系统	2
3	近程防护系统	4
4	高机动飞行器	2
5	自主执行单元	40

其中, 子单元运载平台作为场景支持单元, 不直接参与任务执行, 每个平台搭载 15 架侦察型子单元和 25 架任务执行子单元, 用于快速投送子单元至任务区域。此外, 同组单元共享感知范围, 感知范围内的对方单元将被自动识别为任务目标。

3.3 评估指标

当满足如下任意条件时仿真终止:

- 1) 蓝方核心设施被摧毁;
- 2) 红方全部子单元损失;
- 3) 仿真时间超过预设上限 5500 秒。

仿真结束后, 根据任务完成过程中双方单元的损耗情况评估测试组决策效能, 各类型单元依据其功能重要性被赋予预设权重值(Value)。

效能评估指标包括:

- 1) 任务收益: 处置对方单元累计获得的正面评估值;

- 2) 损耗代价: 己方单元损失累计的负面评估值;

- 3) 总得分: 任务收益与损耗代价的代数和。

3.4 仿真结果分析

3.4.1 LLM 决策样例分析

本研究通过设计包含解释性模板的提示词(见图 4), 要求 LLM 为其决策提供推理依据。为系统评估 LLM(以 o1-mini 模型为例)生成决策指令的合理性, 本研究选取一次完整仿真过程中的两个典型阶段, 结合可视化呈现与模型自解释(参见附录 A、B "overview"部分)文本, 从两方面开展分析: (1) 决策本身的合理性; (2) 模型推理过程与最终指令的逻辑一致性。

- 1) 初始推进阶段(场景时间 745 秒)

LLM 回复见附录 A, COA 样本可视化结构如图 9 所示, 模型生成的部署方案展现出系统的战术考量: 通过分散子单元投放点位, 实现任务区域的最优覆盖; 预警平台在保持前出侦察态势的同时兼顾安全性与监测效率(优化监测范围的重叠); 而高机动飞行器则集中部署于子单元投放区域后方, 形成局部效能优势。

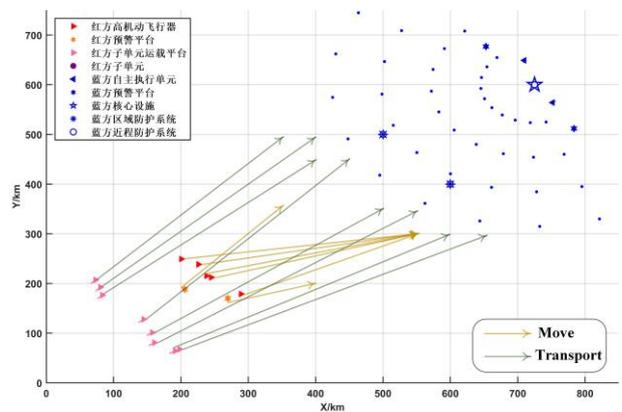


图 9 逼近阶段的决策样例

Fig. 9 Decision sample of the approaching stage

- 2) 动态交互阶段(场景时间 1525 秒)

LLM 回复见附录 B, COA 样本可视化结构如图 10 所示, 此时子单元已全部投放完毕, 高价值单元在子单元掩护下展开行动。高机动飞行器集中应对对方高机动性目标, 充分发挥机动优势控制关键节点; 同时子单元执行动态占位任务, 既向核心区域集中填补防御空隙, 又通过多层次的战术配置, 为高价值单元行动构建立体保护屏障。

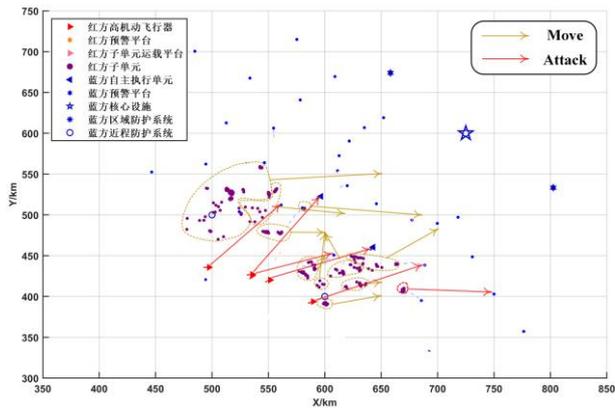


图 10 动态交互阶段的决策样例
Fig. 10 Decision samples of dynamic interaction stage

经对比分析，模型在两个阶段的决策指令均符合基本行动逻辑，且其自解释文本与生成的指令呈现高度一致性，证实了模型决策过程的逻辑连贯性。

3.4.2 基准方法对比

为验证 LLM 决策的战术有效性，本实验选取两类典型的多智能体决策算法作为基准：基于博弈论的改进 Kuhn-Munkres (KM) 算法[24]与平均场多智能体强化学习 (Mean Field MARL) [25]。其中，KM 算法通过矩阵建模双方交互，利用二分图匹配求解最优任务分配；平均场 MARL 则通过局部观测 (自身状态与邻域平均行为) 实现分布式决策。同时，横向对比多种通用大语言模型 (如 DeepSeek-R1、o1-mini 等) 的零样本决策能力。所有方法遵循 30 秒固定决策间隔约束，在统一实验环境下进行测试，每个基准共进行了 50 次评估，本节讨论的所有结果均代表这 50 次评估中每个指标的平均值和标准偏差。

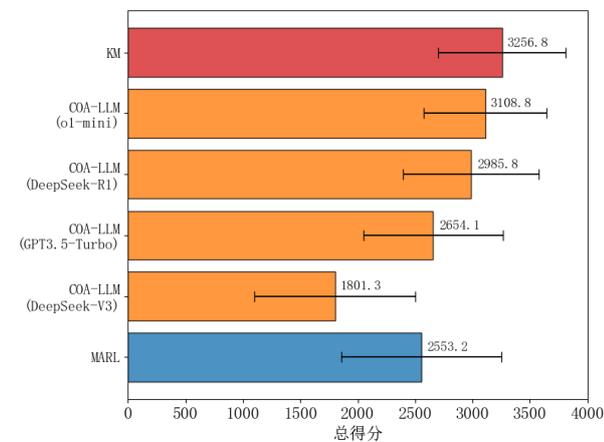


图 11 各基线的总得分对比
Fig. 11 Total score comparison of baselines

图 11 展示了各基线方法在该场景下的总得分均值及标准差。从整体表现来看，COA-LLM 系列模型 (含 o1-mini、DeepSeek 和 GPT3.5-Turbo) 在多智能体协同场景中的决策效能显著优于强化学习算法 (MARL)。这一差异主要源于 MARL 算法的固有局限性：其需要针对异构单位类型分别训练独立模型，在面对大规模单位协同时存在收敛性不足的问题。值得注意的是，传统 KM 算法在本场景中表现最优，这得益于其基于增广路径搜索的最大权匹配算法具有严格的数学最优性保证，能够为静态任务分配问题提供全局最优解。本实验将决策建模为任务分配问题 (即每个己方单位分配一个对方单位)，从而直接应用 KM 算法。然而，该建模方式存在一定局限性：预警平台、子单元运载平台等非直接执行单元无法纳入任务分配框架，因此采用预设巡航坐标点的固定策略进行控制。这种方式牺牲了部分决策灵活性，也降低了场景的异构复杂度。值得注意的是，COA-LLM 框架下的 o1-mini 模型展现出与 KM 算法相近的决策效能，这一结果充分验证了 COA-LLM 在复杂多智能体协同决策任务中的有效性和实用性。

在 COA-LLM 框架下，本研究共采用了 4 种大模型方法。图 12 展示了 4 种通用大模型在 50 次评估中决策时间 (生成 COA 耗时) 的分布情况。其中，DeepSeek-R1 模型的决策效能较优，但由于其独特的“深度思考”机制，其决策时间显著高于其他模型，同时受服务器波动影响，决策时间稳定性较差；DeepSeek-V3 的决策效能相对欠佳，常因忽略邻近目标而生成直接攻击核心设施的指令，导致己方单元遭受较大损耗；GPT3.5-Turbo 的决策速率最快，作为参数量较小的模型，在处理大规模文本输入时，推理时间显著低于大参数量模型[26]。而 o1-mini 凭借其快速响应能力和优异的数学推理能力，在整体表现中最为突出。

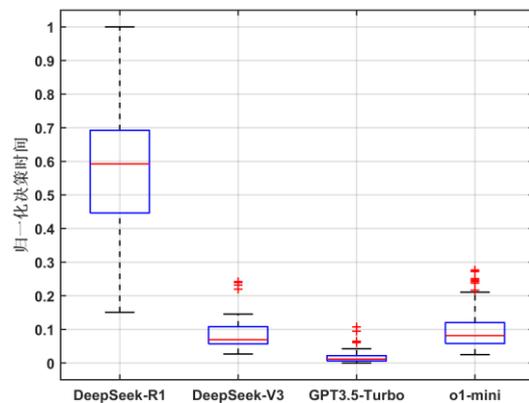


图 12 各模型决策时间分布
Fig. 12 Decision Time distribution of each Model

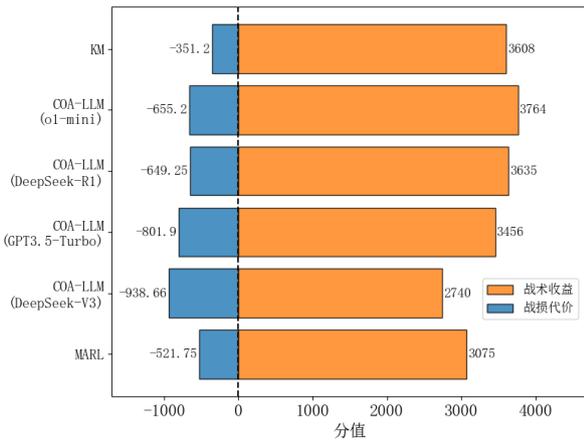


图 13 各基线方法得分分布

Fig. 13 Score distribution of baseline methods

图 13 展示了各基线方法在该场景中的得分分布情况, 其中橙色区域代表处置对方单元获得的任务收益, 蓝色区域则代表己方单元损耗带来的负面代价。结果显示, 智能决策框架下的模型在处置对方单元方面可以媲美甚至超越 KM 算法, 但与 KM 和 MARL 方法相比, 该框架下的模型倾向于造成更多己方单元损耗, 这可能与其在任务执行过程中的动态调整能力有限有关。

此外, 从执行效率的角度来看, COA-LLM 的文本交互方式赋予了其显著的灵活性和即时适应能力。相比之下, MARL 方法需要经过 2000 次模拟仿真训练才能达到最佳性能, 而 COA-LLM 在快速部署方面具有明显优势。更重要的是, COA-LLM 展现出卓越的适应性, 能够快速应对双方单元配置的变化以及场景转换, 而无需进行重新训练。

3.4.3 LLM“幻觉”分析

LLM 在处理复杂任务时可能出现细节错误, 如无法按照提示词要求生成符合模板的指令(见表 1)。本研究针对五类典型生成错误, 统计了不同 LLM 在 300 次回复和 5000 条指令中的错误概率, 结果如表 5 所示。

表 4 各模型错误率统计

Table 5 Error rate statistics of each model

模型名称	回复格式	指令格式	非法 ID	动作越界	参数越界
DS-R1	0%	0%	0.48%	0.24%	0.78%
DS-V3	0%	0.34%	0.18%	0.26%	1.66%
GPT3.5	5.7%	0.26%	5.32%	7.08%	12.60%
o1-mini	0%	0.74%	1.32%	0.52%	0.26%

回复格式和指令格式错误率能够反映模型的结构化输出能力, 而其余三类错误则体现模型对参

数的理解和记忆能力。从纵向对比来看, 各模型在“参数越界”上的错误率最高, 说明参数范围的理解是普遍存在的难点。相比之下, LLM 在结构化输出问题上的表现优于逻辑约束问题。

横向对比各模型的表现, GPT3.5 的错误率显著高于其他模型, 但其在图 11 中的得分仍处于较好水平。这表明, 尽管 GPT3.5 在指令生成时存在较高的错误率, 但其剩余指令仍能保持相对稳定的决策能力。

3.4.3 决策频率敏感性分析

图 14 展示了 COA-LLM 框架(以 GPT3.5-Turbo 为对象)在不同决策间隔时间下的平均得分及其标准差(通过误差条表示), 直观反映了决策间隔对系统性能的影响。实验结果表明, 平均得分随着决策间隔的增加呈现单调下降趋势, 这表明更频繁的决策(即较短的决策间隔)能够更快地响应环境变化, 从而获得更高的得分, 这与常识一致。

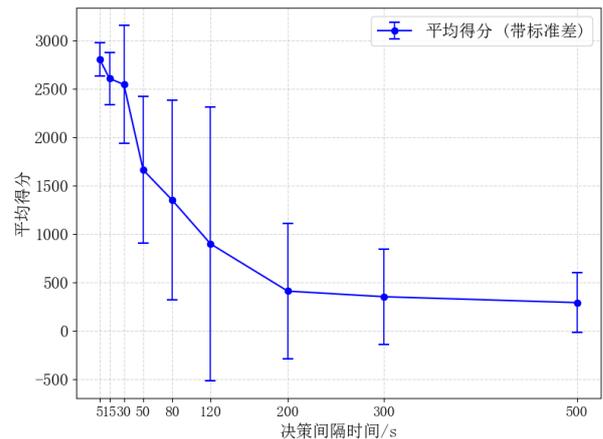


图 14 不同决策间隔的效能对比

Fig. 14 Performance comparison across different decision intervals

此外, 标准差随决策间隔的变化呈现出先增大后减小的趋势。在短决策间隔下系统表现出较强的适应性和稳定性; 在中等决策间隔下系统对环境变化的敏感性增加, 表现波动较大; 而在长决策间隔下, 系统的适应性降低, 但表现趋于稳定, 此时, 系统的行为可能趋于简单化(如无法动态适应环境变化), 导致得分波动减少, 方差逐渐减小。

4 结论

本文针对集群协同决策需求, 提出了一种基于大语言模型的集群协同决策框架 COA-LLM, 将目标分配、威胁评估、路径规划等传统分散模块整合为语义驱动的统一框架。通过设计适配多智能体决策的提示工程, COA-LLM 能够生成切实可行的行

动方案。本文通过一系列模拟实验对框架进行了验证,结果表明,与传统的优化算法和强化学习方法相比,COA-LLM 能够有效地整合环境态势信息,在复杂动态环境中展现较优的适应性和决策效率。未来研究将重点优化 COA-LLM 的推理能力,计划结合 Agent 技术和检索增强生成技术,进一步提升模型对任务规则和任务经验的理解能力,从而推动集群协同任务决策技术的创新发展。

参考文献:

- [1] Virtanen K, Karelahti J, Raivio T. Modeling air combat by a moving horizon influence diagram game[J]. *Journal of guidance, control, and dynamics*, 2006, 29(5): 1080-1091.
- [2] Chen Xia, Hu Yongxin. Study on the task assignment for multi-UCAV in air-combat[C]//2012 Fifth International Symposium on Computational Intelligence and Design. IEEE, 2012, 1: 448-451.
- [3] Zhong Li, Luo Qi, Wen Dong, et al. A task assignment algorithm for multiple aerial vehicles to attack targets with dynamic values[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2012, 14(1): 236-248.
- [4] Duan Haibin, Li Pei, Yu Yaxiang. A predator-prey particle swarm optimization approach to multiple UCAV air combat modeled by dynamic game theory[J]. *IEEE/CAA Journal of Automatica Sinica*, 2015, 2(1): 11-18.
- [5] Orafa S, Yazdanpanah M J, Lucas C, et al. Development of reinforcement learning methods in control and decision making in the large scale dynamic game environments[C]//2006 IEEE Conference on Computer Aided Control System Design. IEEE, 2006: 850-855.
- [6] Devlin J, Chang M W, Lee K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers). 2019: 4171-4186.
- [7] 高志强, 沈佳楠, 姬纬通, 等. 大模型技术的军事应用综述[J]. *南京航空航天大学学报*, 2024,56(5): 801-814.
Gao Zhiqiang, Shen Jianan, Ji Weitong, et al. Review of Military Applications of Foundation Model Technology[J]. *Journal of Nanjing University of Aeronautics & Astronautics*, 2024, 56(5): 801-814.
- [8] Ningsih A K, Hadiana A I. Disaster tweets classification in disaster response using bidirectional encoder representations from transformer (BERT)[C]//IOP Conference Series: Materials Science and Engineering. IOP Publishing, 2021, 1115(1): 012032.
- [9] Zhou Bing, Lei Zou, Ali Mostafavi, et al. VictimFinder: Harvesting rescue requests in disaster response from social media with BERT[J]. *Computers, Environment and Urban Systems*, 2022, 95: 101824.
- [10] Huang Wenlong, Xia Fei, Xiao Ted, et al. Inner monologue: Embodied reasoning through planning with language models[J]. *arXiv preprint arXiv: 2207.05608*, 2022.
- [11] Ahn M, Brohan A, Brown N, et al. Do as i can, not as i say: Grounding language in robotic affordances[J]. *arXiv preprint arXiv: 2204.01691*, 2022.
- [12] Mees O, Borja-Diaz J, Burgard W. Grounding language with visual affordances over unstructured data[C]//2023 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2023: 11576-11582.
- [13] Wang Guanzhi, Xie Yuqi, Jiang Yunfan, et al. Voyager: An open-ended embodied agent with large language models[J]. *arXiv preprint arXiv: 2305.16291*, 2023.
- [14] PALANTIR INC. Deploying full spectrum AI in days with AIP bootcamps[EB/OL].(2023-10-12)[2024-06-17]. <https://www.palantir.com/platforms/aip/>.
- [15] 王亚坤, 陈浩, 葛悦涛, 等. 2023 年人工智能领域科技发展综述[J]. *战术导弹技术*, 2024, (01): 20-32+67.
Wang Yashen, Chen Hao, Ge Yuetao, et al. Overview of artificial intelligence science and technology development in 2023[J]. *Tactical Missile Technology*, 2024, (01): 20-32+67.
- [16] Lamparth M, Corso A, Ganz J, et al. Human vs. machine: Language models and wargames[J]. *arXiv e-prints*, 2024: arXiv: 2403.03407.
- [17] Goecks V G, Waytowich N. Coa-gpt: Generative pre-trained transformers for accelerated course of action development in military operations[C]//2024 International Conference on Military Communication and Information Systems (ICMCIS). IEEE, 2024: 01-10.
- [18] Johnson J. Automating the OODA loop in the age of intelligent machines: reaffirming the role of humans in command-and-control decision-making in the digital age[J]. *Defence Studies*, 2023, 23(1): 43-67.
- [19] Ye Qinyuan, Maxamed Axmed, Reid Pryzant, et al. Prompt engineering a prompt engineer[J]. *arXiv preprint arXiv:2311.05661*, 2023.
- [20] 王友朋. 面向位置服务的无人机飞行路线规划方法[D]. 南京: 南京航空航天大学, 2021.
Wang Youpeng. Flight Path Planning for UAVs to

Provide Location Service[D]. NanJing: Nanjing University of Aeronautics and Astronautics, 2021.

[21] 涂芬芬, 张霖, 张志英, 等. 基于运动控制卡的多轴机床控制系统[J]. 南京航空航天大学学报, 2012, 44(S1): 146-150.

[22] DeepSeek. DeepSeek API 文档 [EB/OL]. <https://api-docs.deepseek.com/zh-cn>, 2025-03-07.

[23] 王炫, 王维嘉, 宋科璞, 等. 基于进化式专家系统树的无人机空战决策技术[J]. 兵工自动化, 2019, 38(01): 42-47.
Wang Xuan, Wang Weijia, Song Kepu, et al. UAV Air Combat Decision Based on Evolutionary Expert System Tree[J]. Ordnance Industry Automation, 2019, 38(01): 42-47.

[24] Zhang Tao, Li Chengchao, Ma Dongying, et al. An optimal task management and control scheme for military operations with dynamic game strategy[J]. Aerospace Science and Technology, 2021, 115: 106815.

[25] Yang Yaodong, Luo Rui, Li Minne, et al. Mean field multi-agent reinforcement learning[C]//International conference on machine learning. PMLR, 2018: 5571-5580.

[26] Kaplan J, McCandlish S, Henighan T, et al. Scaling laws for neural language models[J]. arXiv preprint arXiv: 2001.08361, 2020.

附录 A: 抵近阶段 LLM 回复样本

Scene time = 745.00 seconds:

LLM Response:

```json

{

```
"task_allocation": [
 "Transport(1, [350, 500, 20])",
 "Transport(2, [400, 500, 20])",
 "Transport(3, [400, 450, 20])",
 "Transport(4, [450, 450, 20])",
 "Transport(5, [500, 350, 20])",
 "Transport(6, [550, 350, 20])",
 "Transport(7, [600, 300, 20])",
 "Transport(8, [650, 300, 20])",
 "Move(9, [350, 350, 10])",
 "Move(10, [400, 200, 10])",
 "Move([11, 12, 13, 14], [550, 300, 12])"
```

],

"overview": "

1. **Feasibility**: The plan directly achieves the operational objective of positioning our forces advantageously. By moving the Carriers to various coordinates, we are setting up a logistical network that can support future combat operations.

2. **Resource Efficiency**: The benefits and risks are balanced in this COA. increase overall surveillance coverage. The E2D aircraft, valuable for their reconnaissance capabilities, are placed in positions that leverage their detection range without exposing them to unnecessary risks. The F35Cs, our primary combat assets, are grouped to maintain a strong defensive posture while minimizing their exposure to potential threats.

3. **Environment Adaptability**: The COA is designed to adapt to the dynamic nature of the environment. By positioning our Carriers at various locations, we create a flexible logistical network that can be adjusted based on future enemy movements or operational needs. The E2D aircraft's positioning allows for continuous monitoring of the battlefield, enabling us to respond to any emerging threats. The grouped F35Cs can quickly disperse or engage as needed, providing a high degree of adaptability to changing battlefield conditions."

}  
```

附录 B: 交互阶段 LLM 回复样本

Scene time = 1525.00 seconds:

LLM Response:

```json

{

```
"task_allocation": [
 "Attack(11, -4)",
 "Attack(12, -5)",
 "Attack(13, -21)",
 "Attack(14, -22)",
 "Attack(15, -23)",
 "Move([22, 28], [680, 500, 7.8])",
 "Move([31, 32, 33, 34, 35, 36, 37, 40, 42, 43, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55], [620, 500, 6.9])",
 "Move([56, 57, 58, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70], [700, 480, 7.7])",
 "Move([72, 74, 76, 80, 82, 83, 85, 88, 89, 90, 91, 94], [650, 400, 5.5])",
 "Move([100, 101, 103, 104, 106, 108, 109, 110, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164], [700, 480, 7.7])",
 "Attack([111, 112, 113, 114, 116, 117, 118, 119, 120, 122, 123, 124, 126, 127, 128, 129, 130, 131, 132, 134, 135], -31)",
 "Move([137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150], [600, 450, 7.7])",
 "Move([165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 181, 183, 188, 190],
```

[650, 420, 6.8])",  
"Move([201, 202, 203, 204, 206, 207, 208, 209,  
210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220,  
221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231,  
234, 235, 238, 239, 240, 241], [600, 480, 7.8])",  
"Move([243, 244, 245, 246, 247, 248, 249, 250,  
251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261,  
262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272,  
273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283,  
284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294,  
295, 296, 297, 298, 300, 301, 302, 303, 304, 305, 306,  
307, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320,  
321, 322, 324, 326, 328, 329, 330, 331, 332, 333, 334,  
335], [650, 550, 6.9])"

]  
"overview": "

1. **Feasibility**: The plan directly addresses the operational objective by prioritizing the elimination of enemy F35Cs and MQ9s, which are critical threats due to their mobility and firepower. By deploying F35Cs to attack these high-value targets, we aim to reduce the enemy's offensive capabilities significantly. Simultaneously, UAVs are positioned to engage additional MQ9s and provide reconnaissance, ensuring a comprehensive approach to weakening the enemy's overall strength. The strategy is designed to systematically dismantle the enemy's mobile units,

paving the way for a subsequent attack on their base.

2. **Resource Efficiency**: The plan directly addresses the operational objective by prioritizing the elimination of enemy F35Cs and MQ9s, which are critical threats due to their mobility and firepower. UAVs are positioned to engage additional MQ9s and provide reconnaissance, ensuring a comprehensive approach to weakening the enemy's overall strength. This strategy is feasible as it leverages the long-range capabilities of our F35Cs and the expendable nature of UAVs to achieve our goals.

3. **Environment Adaptability**: The COA is designed to adapt to the dynamic nature of the battlefield. By positioning UAVs at various altitudes and locations, we create a flexible screen that can shift to counter emerging threats or reinforce areas as needed. The F35Cs maintain mobility, allowing them to adjust their targets based on real-time battlefield developments. This adaptability ensures that our forces can respond effectively to enemy movements or unexpected changes in the combat environment, maintaining a strong offensive posture while protecting our assets."

}  
``