通过策略参数化提示影响大语言模型多智能体对话

Hongbo Bo 布里斯托大学 布里斯托，英国
hongbo.bo@bristol.ac.uk

Jingyu Hu 布里斯托大学 布里斯托，英国
jingyu.hu@bristol.ac.uk

Weiru Liu 布里斯托大学 布里斯托，英国
weiru.liu@bristol.ac.uk

arXiv:2603.09890v1 [cs.AI] 10 Mar 2026

摘要

大语言模型（LLMs）已成为多智能体系统的新范式。然而，现有关于基于大语言模型的多智能体行为的研究依赖临时性提示，缺乏原则性的策略视角。与强化学习不同，我们探究提示即动作能否被参数化，从而构建一种由状态-动作对序列组成的轻量策略，以在无需训练的情况下影响对话行为。我们的框架将提示视为由大语言模型执行的动作，并根据智能体的当前状态，通过五个组件动态构建提示。为检验参数化控制的有效性，我们基于五项指标评估对话流程：响应性、反驳、证据使用、非重复和立场转变。我们使用不同的大语言模型驱动的智能体，在两个与公众相关的讨论场景中开展实验，结果表明提示参数化能够影响对话动态。该结果表明，策略参数化提示提供了一种简单而有效的机制来影响对话过程，这将有助于多智能体系统在社会模拟方向上的研究。

关键词

大语言模型、多智能体系统、社会模拟

1 引言

传统的多智能体模拟通常依赖显式建模或强化学习来训练策略，使智能体能够学习如何回应和互动。相比之下，大语言模型（LLM）已成为多智能体系统的一种新范式，智能体无需额外训练即可天然具备自然语言生成和知识检索能力。基于这一潜力，近期研究通过为大语言模型分配角色、任务和指令，广泛将基于大语言模型的多智能体系统用于社会模拟任务。

这些方法中的大语言模型智能体通常通过临时提示彼此通信，重点关注智能体能否契合人类偏好，以及能否表现出协作 [12, 28,32], 谈判 [1, 5, 19] 等行为，并提供合理计划 [6, 11]。然而，这些方法缺乏一种有原则的框架，无法将通信策略视为策略，以系统地控制智能体行为。由于没有充分关注如何有意塑造和优化智能体对话，因此难以预测智能体行为、优化沟通模式，并将洞见迁移到不同任务中。

为解决这一局限，我们希望提出一种有原则的方法，用于概念化并落实智能体通信策略，使我们能够正式定义、比较和优化多智能体对话中的不同策略。具体而言，本研究关注如何利用策略参数化提示来影响多智能体讨论对话中的对话行为。

本研究将输入提示本身视为一种通过轻量策略参数化生成的动作。具体而言，我们将提示分解为五个组件：任务和角色描述（T）、对话历史记忆（M）、外部知识库（D）、规则模板（R）和权重（W）。通过对这些组件进行自适应参数化，使其对大语言模型智能体产生不同程度的影响，我们能够直接设定不同因素对对话话语风格的重要程度，从而在无需额外训练的情况下调节其对话行为。

在这一策略框架下，我们构建了具有不同立场和知识库的多角色智能体，并让它们围绕与公众相关的问题开展多轮对话。为进一步提升适应性，我们设计了一种自适应权重调度器，根据时间趋势和行为反馈，在对话过程中自动调整对视角/材料/证据的依赖程度。为量化不同控制策略的影响，我们提出了一组评估指标，包括响应性、反驳、非重复性、证据使用和立场转变。这些指标使我们能够系统地比较不同提示控制模式下的对话差异，并观察群体立场随时间的演变。

我们设计的框架旨在回答以下研究问题。问题1：提示控制能否作为一种轻量的策略参数化形式，用于调节基于大语言模型的多智能体系统的对话行为？问题2：不同的提示控制策略（例如规则模板和权重调度）是否会导致显著的行为差异，例如响应性、证据使用和立场演变方面的变化？为回答这些问题，我们设计了两个讨论场景，分别涉及一般性的公共土地资源利用和教育资源分配；每个场景都包含三个由不同大语言模型驱动、角色设定各异的智能体。在每个场景中，智能体围绕任务、记忆和证据组件动态构建的提示，进行多轮对话。我们通过启用或禁用规则模板、调整权重参数，系统地改变控制策略，并依据提出的五项评估指标对生成的对话进行评估。结果表明，提示参数化能够有效影响对话，不同策略会在反驳、证据使用和立场转变方面形成不同模式。

[0]收录于 AAMAS 2026

总体而言，本研究的核心在于将提示控制视为一种轻量策略，以实现对大语言模型多智能体对话的结构化调节，从而探索一条不同于传统基于训练的方法的社会模拟路径。

2 相关工作

大语言模型智能体与社会模拟。 近期研究探索了将大语言模型用作社会模拟中的智能体，旨在通过模拟社会行为来探究真实的社会现象。 [23]利用具备记忆能力的大语言模型构建生成式智能体，并发现智能体会涌现出信息传播、关系形成以及协调参与活动等社会行为。类似研究 [7, 20, 21] 也使用大语言模型构建沙盒环境，以研究社会行为。 [24]通过设计提示链，构建了一个由 1,000 个角色画像组成的模拟社区，用以生成发帖、回复以及反社会动作等行为。 S^3 系统 [8] 利用大语言模型智能体模拟社交网络中用户的情绪、态度和交互行为，并赋予系统中的智能体感知信息环境的能力。CAMEL [17]是一种多智能体角色扮演框架，旨在利用启发式提示来初始化角色、任务和对话格式，从而使基于大语言模型的聊天智能体能够协作完成复杂任务。 [3]将多智能体大语言模型用于模拟智能体群体，结果表明，这类群体能够通过去中心化交互形成社会惯例和集体偏见。 [25] 将社会模拟扩展至大规模，涵盖超过 10k 个智能体和 5million 次交互，并分析了可信的个体行为与涌现的社会行为。其他研究 [12, 19] 验证了多智能体系统在协作和辩论任务中的有效性。然而，这些系统中的提示通常是临时设计的，缺乏将其作为策略进行系统化处理的原则。

大语言模型智能体形式化。许多研究尝试将基于大语言模型的智能体抽象为决策理论框架，灵感来自强化学习（RL）或信念-愿望-意图（BDI）模型。例如，大语言模型被视为将状态映射为自然语言动作的策略，也可作为世界模型，与蒙特卡洛树搜索结合进行规划搜索 [11]。BDIPrompting [14] 将 BDI 模型融入提示设计，以提升大语言模型的主动行动规划能力和透明度。Reflexion[26] 是一种通过语言反馈强化大语言模型智能体的框架，它将策略参数化为记忆，使智能体能够通过自我反思进行试错学习。 [6] 研究大语言模型能否成为博弈论中的理性参与者：通过提供偏好和规则等输入来构建信念与愿望，进而规划最优动作。 [30] 将大语言模型与强化学习（RL）相结合，使智能体能够学习狼人杀游戏中的策略性语言通信和决策。这些研究关注的问题是：大语言模型能否理性地行动，或与人类决策保持一致；但它们没有探讨如何在多智能体场景中控制大语言模型。此外，在这些研究中，大语言模型用于生成决策，而非执行动作。

3 提示控制作为轻量策略参数化

本节介绍我们提出的方法如何通过策略参数化提示影响大语言模型多智能体对话。图1展示了我们方法的总体框架。

3.1 多智能体形式化

我们将多智能体讨论形式化为一种可控的状态—动作过程，其中策略通过提示构造直接指定。具体而言，大语言模型多智能体对话被形式化为一种提示参数化过程：我们定义 N 个智能体参与 K 轮讨论。对于每一轮对话$k \in \{1,\ 2..... K\}$，每个智能体 A_i都有相应的状态 $s_{(k)i}$，该状态由智能体记忆、证据和任务设置组成。随后，策略 π_i将状态映射为构造的提示，作为智能体的动作 $a_i^{(k)}$。我们通过规则模板 R和权重向量W的调整，利用提示实现对智能体行为的参数化控制。

3.1.1 智能体。 每个智能体都表示为四元组 $\mathrm{A}_i = \{Q,\ T_i,\ D_i,\ LLM_i\}$，其中 Q是用户提供的所有智能体之间的全局讨论查询，T_i 包含任务描述和智能体 A_i的角色设定，D_i 是智能体 A_i 完整的角色特定知识数据，LLM_i 是动作执行器，用于生成由智能体 A_i 开展的对话。由于本研究专门考察基于大语言模型的智能体，我们将大语言模型视为智能体能力的一部分，智能体可以调用它来执行动作（类似于调用其他函数等）。

3.1.2 状态与检索。 在第 k 轮，对话记忆记为 $M_{(k)}$，由此前所有 k 轮的话语组成。智能体从完整知识库 D_i 中通过嵌入相似度搜索检索出一个子集 $\hat{}D_i^{(k)}$：

$$\hat{D}_i^{(k)} = \begin{cases} \text{Top-n}_{c \in C(D_i)} \cos(\phi(c), \phi(Q)), & k = 1, \\ \text{Top-n}_{c \in C(D_i)} \cos(\phi(c), \phi(M^{(k)})), & k > 1, \end{cases}$$

其中，$\mathrm{C}(D_i)$ 是通过分段获得的文本块集合，$\phi(\cdot)$ 是检索增强生成（RAG） [16]中使用的嵌入函数。

智能体 A_i 的状态 由 $s_i^{(k)} = \{T_i, Q, \hat{}M^{(k)}, \hat{}D^{(k)i}\}$ 组成，其中，$\hat{}M^{(k)} \subseteq M^{(k)}$ 表示为当前轮次提取的对话记忆子集，其中包含近期对话；而在初始化时，$\hat{}M^{(1)}$ 为空。本研究采用共享消息池 设置 [10],，所有智能体均可访问共同的对话历史；因此，记忆 $M^{(k)}$ 是共享的，而非个体独有的。我们的方法还可扩展至其他通信配置，例如分层式、去中心化或集中式配置。

3.1.3 提示即动作。我们采用提示即动作的视角：策略 π_i将状态 $s_i^{(k)}$映射为动作 $a_i^{(k)}$，其中动作是构造出的提示：$a_i^{(k)} = \pi_i(s_i^{(k)})$。此处，策略 π_i基于 $< R \oplus W_{(k)\,i} >$生成动作 $a_i^{(k)}$，对应状态 s_i；其中，R是可选的规则模板，而 $W_{(k)\,i} = w_{(k)\,iT}, w_{(k)\,iM}, w_{(k)\,iD}$是权重，用于控制角色设定、记忆或检索到的知识的强调程度。算子 $\oplus$ 表示将规则和权重耦合为单一提示规范。具体而言，$\oplus$ 会生成 (a) 附加在提示中、紧邻受影响对象的分层微指令。

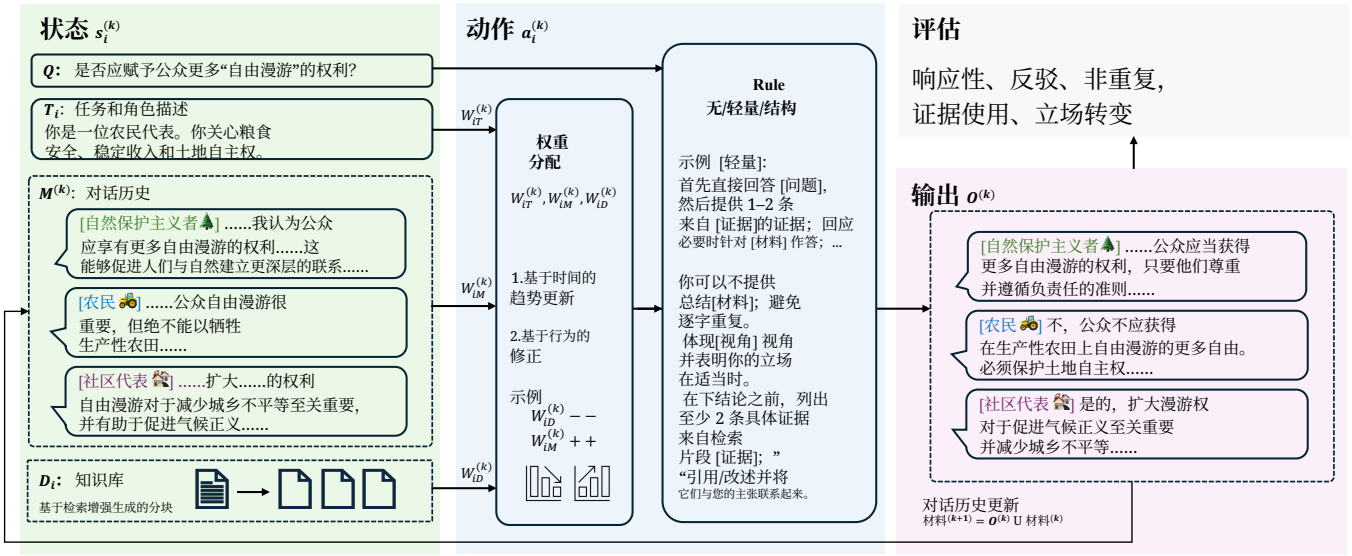


图1：总体框架展示了从智能体的状态表示到动作生成、基于大语言模型的动作执行以及评估的过程。

当权重较低/较高时，[视角]/[材料]/[证据] 区块（见第 3.2.2 节），以及（b）启用时插入 [规则] 区块的可选规则骨架（见第 3.2.1 节）。通过这种方式，策略参数化提示提供了一种轻量机制，用于引导基于大语言模型的智能体的对话过程。本研究中的策略将生成一系列状态-动作对，这些对可视为强化学习策略学习中的一条轨迹。需要指出的是，在我们的框架中，轨迹长度由所需的对话轮次数量决定。

3.1.4 动作执行与输出。 动作 $a_i^{(k)}$（提示）由大语言模型执行，产生具体的自然语言输出$o_i^{(k)}$ = 大语言模型$_i(a_i^{(k)})$。在每一轮 k 中，所有智能体遵循相同的动作生成来生成各自的动作，以产生各自的输出 $o_i^{(k)}$。这些输出随后被聚合并追加到 $M^{(k)}$，以形成全局对话记忆 $M(k+1)$，供下一轮使用，进而影响后续检索$M(k+1)=\{\cup_{i=1}^{N} o(k)i\} \cup M(k)$。

3.2 策略参数化

3.2.1 规则模板 *R.* 除四个信息源 Q、T、M 和 D 外，我们引入一个可选组件 R（规则模板），旨在以可控方式引导智能体的交互行为。使用 R 可使智能体指定输出格式以及信息的使用方式，从而在不改变底层模型参数的情况下塑造其对话行为。我们设计了三种规则模板，分别代表递增程度的结构约束：

- 无：不给出明确的结构指令；智能体直接基于拼接的信息块 $[T]$, $[M]$, $[D]$, $[Q]$ 生成回应，无需任何额外的顺序或格式控制。
- 轻量：仅通过规定基本的回应顺序和长度限制来提供最少的结构。“首先直接回答 [问题], , 然后从 [证据]中提供 1–2 条证据；如有必要，回应 [材料] ； 回应最多限制为 N 句话。”
- 结构：通过将讨论分解为特定类别的要点（支持、反对、冲突、合作），强制采用详细的推理结构。“首先按顺序从 [材料]中提取四类要点（每类不超过 3 个）：1）支持目标的论据，2）威胁目标的论据，3）未解决的冲突点，4）潜在的合作机会；然后基于这些要点生成不超过 3 句的回应，并优先引用 [证据]。”

这使我们能够采用轻量策略，在对话推进过程中动态生成提示（或动作），从而引导智能体之间逐渐形成的讨论模式。

3.2.2 权重设计。 为了更好地完善策略对智能体行为的影响，我们为每个智能体 A$_i$定义一个权重集 $W_i = \{w(k)iT, w(k)iM, w_{iD}^{(k)}\}$。每个权重 $w \in W_i$取值于 $[0,2]$，并可动态调整 A$_i$在对话过程中对相应组件（T、M 和 D）的依赖程度。随后，每个权重 w都会映射到一个三级信号系统：若 $w \in [0,0.85]$，则为低；若 $w \in [0.85,1.25]$，则为中；若 $w \in [1.25,2]$，则为高。每个等级都对应三个组件的特定行为指令。下文概述了各组件在各等级下的简化指令，完整指令见附录 B。

- 视角：– 低：不明示角色设定；聚焦论据；– 中：体现指定角色的视角，并在相关时表达立场；– 高：先明确说明角色/立场，再进行论证；
- 材料：– 低：可不作总结直接回复；避免重复；

– 中：考虑先前讨论以保持上下文连贯性；– 高：以简要回顾开头并解决待决要点；

• 证据：– 低：若检索片段并非必要，可不引用；– 中：适当时使用检索片段支持关键主张；– 高：在下结论前提供具体证据项；

3.2.3 自适应权重。 为了更好地体现对话不断演变的特性，我们引入基于时间的趋势更新和基于行为的校正，以更新提示权重。

基于时间的趋势更新。 我们假设智能体在早期轮次应更多依赖 D 来确立其立场，而在后期轮次更多依赖对话历史 M 来参与进行中的辩论。因此，我们将权重定义为：

$$w_{iM}^{(k)} = \min\{w_{iM}^{(0)} + 0.1k,\ 2.0\},$$
$$w_{iD}^{(k)} = \max\{w_{iD}^{(0)} - 0.1k,\ 0.5\},$$
$$w_{iT}^{(k)} \equiv w_{iT}^{(0)}.$$

基于行为的校正。 观察智能体先前的回应后，我们应用由 α 参数化的更新算子 $\min(\cdot)$（也可以定义其他更新算子）。具体而言：

• 如果智能体在上一轮未使用证据 $k-1$：

$$w_{iD}^{(k)} \leftarrow \min\{w_{iD}^{(k)} + \alpha,\ 2.0\}$$

• 如果智能体未回应材料：

$$w_{iM}^{(k)} \leftarrow \min\{w_{iM}^{(k)} + \alpha,\ 2.0\}$$

3.3 评估

我们通过五项提出的评估指标，评估策略参数化的有效性。这些指标针对智能体在状态-动作对 (s, a) 上执行的输出 o 进行评估。我们使用大语言模型作为评判模型，并以文本嵌入模型作为嵌入后端进行评估。每项指标 $m(o_i^{(k)})$ 均衡量一个不同的行为维度：

• **响应性（回应）** $m\text{resp}_{o_{(k)i}}$：如果 $o_{(k)i}$ 回应了 $M_{(k)}$ 中最近的话语，则返回 1，否则返回 0。向评判模型提供前一个和当前的话语，要求其做出二元决策。

• **反驳** $m\text{rebut}_{o_{(k)i}}$：如果评判模型判定 $o_{(k)i}$ 明确反对 $M_{(k)}$ 中最近的话语，则返回 1，否则返回 0。该指标直接衡量“是否发生了反驳”。

• **非重复（Non-rep.）** $m\text{nrep}_{o_{(k)i}}$：衡量 $o_{(k)i}$ 相对于智能体自身先前动作 $o_{(k-1)i}$ 的新颖性，定义为当前话语与智能体先前话语之间相似度的 1 减去该相似度。相似度定义为字符串重叠 *over*()和嵌入余弦相似度 *cos*()中的最大值，并对重复的句首施加额外惩罚。定义为

$$m_{\text{nrep}} o_i^{(k)} = 1 - \max\{over(o_i^{(k)}, o_i^{(k-1)}), cos(o_i^{(k)}, o_i^{(k-1)})\}$$

• **证据使用（证据）** $m\text{evid}_{o_{(k)i}}$：如果检索到的知识组件中的关键词组 D_i 出现在 $o_{(k)i}$ 中，则返回 1，否则返回 0。

• **立场转变** m立场$_{o_{(k)i}}$：计算 $o_{(k)i}$ 的嵌入与角色描述 T_i 的嵌入之间的余弦相似度。追踪各轮次的 m 立场，可以揭示智能体是否始终与其初始立场一致，或逐渐偏离初始立场。

对给定智能体的每项指标，均取其在所有对话轮次中的平均值；总体性能则通过计算所有智能体的均值获得。

4 实验

本节介绍所提出方法的实验设置与评估。实现代码见补充材料。

4.1 实验设置

为通过实验评估所提出的、用于规范智能体行为的策略参数化方法的有效性，我们设计了两个场景：土地资源利用 （土地）和教育资源分配 （教育）。针对每个场景，我们设置了三个具有不同角色画像的智能体，并为其配置任务（T）和知识库（D）。知识库收集自公开可用的材料，包括政府政策文件、博客和相关网站，并使用 ChatGPT-5[1] [22] 进一步进行总结和补充，以确保内容一致、清晰。在处理这些材料的同时，ChatGPT-5 还会识别关键的利益相关方视角，并将其整理为相应的智能体角色和任务描述（T）。有关这些知识库的更多信息，请参阅附录 A。每个场景中的三个智能体分别由三种不同的大语言模型驱动：通义千问3-8B[2] [31], Llama3-8B[3] [4], 和 Mistral-7B[4] [15], ，如表2所示。

表2：两个场景中的智能体及大语言模型分配情况。

Land		教育	
智能体	**LLM**	**智能体**	**LLM**
农民	Qwen3	乡村教师	Qwen3
自然保护主义者	Llama3	城市家长	Llama3
社区代表	Mistral	政策制定者	Mistral

每个智能体都配有各自特定角色的外部知识库，并围绕争议性话题查询 Q（例如，“是否应将农田转化为森林？”）进行10轮对话。对于检索到的知识（$\hat{D}_i^{(k)}$），我们采用检索增强生成（RAG）设置：在每一轮中，智能体检索排名前3的

[1]https://openai.com/index/introducing-gpt-5/ [2]Qwen3-8B https://qwen.ai/research [3]Meta-Llama-3-8B https://www.llama.com/models/llama-3/ [4]Mistral 7B https://mistral.ai/news/announcing-mistral-7b

表1：在 5 个查询、10 轮对话中，不同规则模板下策略的整体表现。数值为 5 次实验的均值 ± 标准差。由于反驳和证据为二元值（0/1），均值即为 1 的比例，因此标准差相对于均值可能显得较大。

		Land					教育				
查询	Rule	回应	反驳	非重复	证据	立场	回应	反驳	非重复	证据	立场
Q1	None	$0.88_{\pm 0.33}$	$0.19_{\pm 0.40}$	$0.47_{\pm 0.42}$	$0.10_{\pm 0.30}$	$0.52_{\pm 0.08}$	$0.88_{\pm 0.33}$	$0.25_{\pm 0.43}$	$0.45_{\pm 0.41}$	$0.13_{\pm 0.34}$	$0.52_{\pm 0.09}$
	轻量	$0.90_{\pm 0.30}$	$0.21_{\pm 0.41}$	$0.33_{\pm 0.35}$	$0.32_{\pm 0.47}$	$0.51_{\pm 0.09}$	$0.87_{\pm 0.34}$	$0.27_{\pm 0.44}$	$0.41_{\pm 0.36}$	$0.25_{\pm 0.44}$	$0.52_{\pm 0.11}$
	结构	$0.74_{\pm 0.44}$	$0.27_{\pm 0.45}$	$0.58_{\pm 0.39}$	$0.31_{\pm 0.46}$	$0.49_{\pm 0.09}$	$0.89_{\pm 0.32}$	$0.09_{\pm 0.29}$	$0.54_{\pm 0.37}$	$0.17_{\pm 0.37}$	$0.52_{\pm 0.11}$
Q2	None	$0.87_{\pm 0.34}$	$0.37_{\pm 0.48}$	$0.47_{\pm 0.44}$	$0.11_{\pm 0.31}$	$0.49_{\pm 0.08}$	$0.85_{\pm 0.35}$	$0.15_{\pm 0.36}$	$0.41_{\pm 0.41}$	$0.37_{\pm 0.48}$	$0.48_{\pm 0.07}$
	轻量	$0.89_{\pm 0.31}$	$0.31_{\pm 0.47}$	$0.32_{\pm 0.34}$	$0.27_{\pm 0.44}$	$0.45_{\pm 0.06}$	$0.89_{\pm 0.32}$	$0.06_{\pm 0.24}$	$0.35_{\pm 0.35}$	$0.41_{\pm 0.49}$	$0.45_{\pm 0.08}$
	结构	$0.79_{\pm 0.41}$	$0.17_{\pm 0.38}$	$0.64_{\pm 0.34}$	$0.05_{\pm 0.21}$	$0.46_{\pm 0.07}$	$0.90_{\pm 0.30}$	$0.06_{\pm 0.24}$	$0.68_{\pm 0.30}$	$0.21_{\pm 0.41}$	$0.52_{\pm 0.07}$
Q3	None	$0.90_{\pm 0.30}$	$0.00_{\pm 0.00}$	$0.42_{\pm 0.42}$	$0.09_{\pm 0.28}$	$0.47_{\pm 0.07}$	$0.88_{\pm 0.33}$	$0.17_{\pm 0.38}$	$0.47_{\pm 0.39}$	$0.10_{\pm 0.30}$	$0.49_{\pm 0.06}$
	轻量	$0.90_{\pm 0.30}$	$0.08_{\pm 0.27}$	$0.38_{\pm 0.38}$	$0.19_{\pm 0.39}$	$0.45_{\pm 0.05}$	$0.89_{\pm 0.31}$	$0.15_{\pm 0.36}$	$0.36_{\pm 0.35}$	$0.25_{\pm 0.43}$	$0.49_{\pm 0.05}$
	结构	$0.86_{\pm 0.35}$	$0.04_{\pm 0.20}$	$0.70_{\pm 0.32}$	$0.15_{\pm 0.35}$	$0.45_{\pm 0.07}$	$0.89_{\pm 0.32}$	$0.06_{\pm 0.24}$	$0.61_{\pm 0.32}$	$0.24_{\pm 0.43}$	$0.51_{\pm 0.05}$
Q4	None	$0.79_{\pm 0.41}$	$0.43_{\pm 0.50}$	$0.34_{\pm 0.38}$	$0.10_{\pm 0.30}$	$0.49_{\pm 0.06}$	$0.85_{\pm 0.35}$	$0.13_{\pm 0.34}$	$0.53_{\pm 0.35}$	$0.10_{\pm 0.30}$	$0.47_{\pm 0.11}$
	轻量	$0.68_{\pm 0.47}$	$0.51_{\pm 0.50}$	$0.29_{\pm 0.35}$	$0.50_{\pm 0.50}$	$0.50_{\pm 0.08}$	$0.88_{\pm 0.33}$	$0.17_{\pm 0.37}$	$0.39_{\pm 0.33}$	$0.35_{\pm 0.48}$	$0.45_{\pm 0.12}$
	结构	$0.76_{\pm 0.43}$	$0.37_{\pm 0.48}$	$0.54_{\pm 0.36}$	$0.28_{\pm 0.45}$	$0.49_{\pm 0.07}$	$0.88_{\pm 0.33}$	$0.03_{\pm 0.18}$	$0.65_{\pm 0.28}$	$0.17_{\pm 0.38}$	$0.49_{\pm 0.11}$
Q5	None	$0.83_{\pm 0.37}$	$0.40_{\pm 0.49}$	$0.46_{\pm 0.41}$	$0.08_{\pm 0.27}$	$0.44_{\pm 0.07}$	$0.78_{\pm 0.42}$	$0.07_{\pm 0.25}$	$0.48_{\pm 0.40}$	$0.09_{\pm 0.28}$	$0.44_{\pm 0.07}$
	轻量	$0.87_{\pm 0.34}$	$0.43_{\pm 0.50}$	$0.32_{\pm 0.35}$	$0.13_{\pm 0.33}$	$0.44_{\pm 0.04}$	$0.85_{\pm 0.36}$	$0.00_{\pm 0.00}$	$0.39_{\pm 0.36}$	$0.27_{\pm 0.44}$	$0.42_{\pm 0.05}$
	结构	$0.87_{\pm 0.34}$	$0.23_{\pm 0.42}$	$0.68_{\pm 0.35}$	$0.20_{\pm 0.40}$	$0.47_{\pm 0.07}$	$0.89_{\pm 0.31}$	$0.09_{\pm 0.28}$	$0.61_{\pm 0.34}$	$0.22_{\pm 0.42}$	$0.50_{\pm 0.07}$
总体	None	$0.85_{\pm 0.35}$	$0.28_{\pm 0.45}$	$0.43_{\pm 0.42}$	$0.10_{\pm 0.29}$	$0.48_{\pm 0.07}$	$0.85_{\pm 0.36}$	$0.15_{\pm 0.36}$	$0.47_{\pm 0.39}$	$0.16_{\pm 0.37}$	$0.48_{\pm 0.09}$
	轻量	$0.85_{\pm 0.36}$	$0.31_{\pm 0.46}$	$0.33_{\pm 0.36}$	$0.28_{\pm 0.45}$	$0.47_{\pm 0.07}$	$0.88_{\pm 0.33}$	$0.13_{\pm 0.34}$	$0.38_{\pm 0.35}$	$0.30_{\pm 0.46}$	$0.47_{\pm 0.09}$
	结构	$0.80_{\pm 0.40}$	$0.22_{\pm 0.41}$	$0.62_{\pm 0.36}$	$0.20_{\pm 0.40}$	$0.47_{\pm 0.08}$	$0.89_{\pm 0.31}$	$0.07_{\pm 0.25}$	$0.62_{\pm 0.33}$	$0.20_{\pm 0.40}$	$0.51_{\pm 0.09}$

从知识库中检索与当前主题及近期对话记忆相关的段落（遵循第3.1.2小节），并将其附加到提示中。提示根据第3节介绍的框架动态构建。我们通过改变规则模板 R、同时将权重参数固定为 $w_T = 1.0, w_M = 1.0, w_D = 1.0$来比较不同的提示控制策略。在评估中，我们使用Llama3作为评判模型，沿评估指标对每个对话轮次进行分类，并采用嵌入模型all-MiniLM-L6-v2[5] [27] 计算评估指标的语义相似度。各项指标的数值表示所有智能体和所有轮次的平均得分。对于每种情景，我们针对与该情景对应的五个公开主题查询，对每个查询进行五次独立运行。以下是每种情景的5个主题查询，它们是根据涵盖多个方面的材料内容人工拟定的：

- 土地：
  – 问题1：‘是否应该给予公众更多自由漫游的权利？’– 问题 2：‘ 是否应该限制农民在敏感生态区域扩大农地？’– 问题3：‘农民在应对气候变化方面应承担哪些责任？’– 问题4：‘未来50年英国的土地规划应如何制定？’– 问题5：‘是否应该在私人农地上修建公共绿道？’

- 教育：
  – 问题1：‘政府应如何在设施匮乏的农村学校和竞争激烈的城市学校之间分配有限的教育经费？’– 问题2：‘考虑到农村学生面临的不利处境和城市学生承受的巨大压力，是否应减少对标准化考试的依赖？’– 问题3：‘公立大学是否应调整学费政策，在保持对城市家庭公平的同时，改善低收入农村学生的入学机会？’– 问题4：‘政府是否应先投资建设农村学校的数字基础设施并引入AI导师，再在城市学校推广？’– 问题5：‘教育政策是否应优先为学生支持服务提供经费（例如，为农村地区建设寄宿学校，为城市学校提供心理健康服务）？’

[5]https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2

4.2 主研究

为探讨有关策略参数化有效性的研究问题，主研究比较了策略中采用不同规则 和权重 时的性能。

4.2.1 规则的有效性。根据表1中的结果，我们可以观察到，对于同一主题查询使用不同规则会在各项指标上产生差异。这表明，智能体的行为可以通过策略生成过程进行参数化（RQ1）。其次，对于RQ2，不同的规则策略会带来不同的行为权衡。在非重复性方面，结构的性能最佳，这意味着结构清晰的规则和强约束

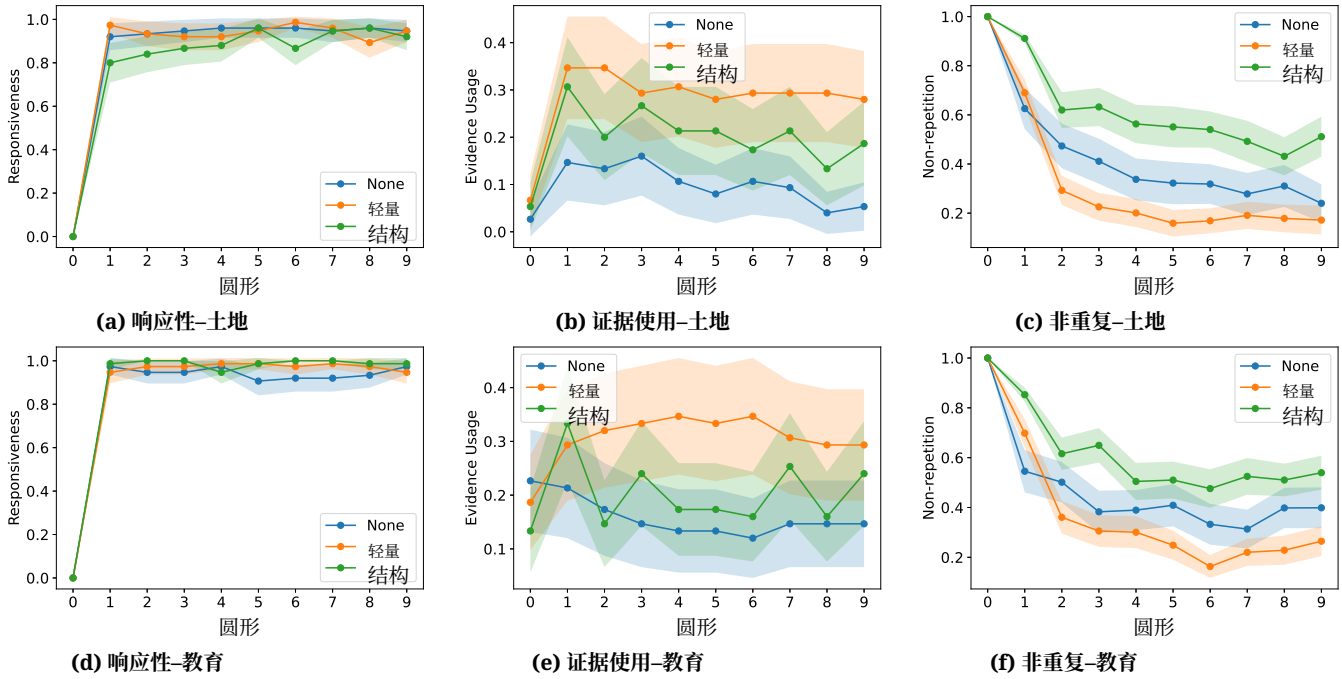


图2：不同规则模板下三项对话指标的轮次变化。每条曲线显示均值（含95%置信区间）。

提示可以减少重复。对于证据使用，轻量能显著提升性能，这表明轻量规则会鼓励使用外部知识，但过度结构化可能会抑制证据使用。在反驳方面，轻量和结构表现出更高的总体反驳率，体现出更具互动性和论辩性的交流。对于立场一致性，所有条件下的得分都较为相近（总体为0.47–0.51），这意味着规则模板主要影响互动风格，而非核心问题立场。最后，这些结果展示了提示作为一种可解释的社会模拟影响因素的潜力。

图2展示了响应性、证据使用和非重复性在对话轮次中的变化。在所有三种规则设置下，响应性都相对稳定。然而，对于证据使用，波动程度会随着结构约束水平的提高而增加：无条件下的轨迹最平稳，而结构 条件下的变化最为显著。对于非重复性，三种条件下的数值都先下降，随后趋于稳定，其中结构 条件下始终保持最高水平，而轻量条件下最低。

4.2.2 权重敏感性。 我们进一步研究不同权重配置是否会影响性能：在土地 场景下，逐一改变每个权重，同时保持另外两个权重不变。我们在正文中报告总体结果，完整的主题级指标见附录C。

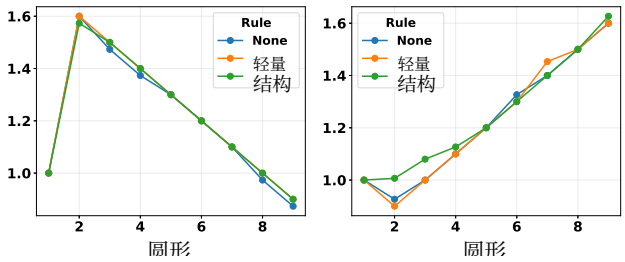
从表3中可以观察到：1. 无论权重如何变化，响应性始终高于0.8，波动极小。2. 当W_T增至1.5时，反驳率显著上升，这表明当角色设定（T）得到更多强调时，智能体更有可能参与冲突并进行反驳。在无立场条件下也会出现同样的情况：较高的W_T也会使智能体更加“忠于角色”，其立场也更明确、更稳定。3. 关于证据使用和W_D，我们观察到一种交叉效应：当$W_D = 0.5$时，轻量 条件下的证据使用率更高；而当$W_D = 1.5$时，无条件的表现相对更好。这表明，即使权重较弱，规则也能促进证据整合；而在缺少规则时，则需要更强的权重来推动证据使用。

4.2.3 自适应权重的有效性。 为探究不同权重对智能体行为的影响效果，我们采用自适应权重，将初始参数设为$w_T = 1.0$, $w_M = 1.0$ $w_D = 1.0$ $\alpha = 0.2$, , and，并在土地 场景下按照与主研究相同的设置开展实验。我们在图3中给出了w_M和w_D变化的示例，完整的实验结果见表4。

如表4所示，总体均值与未使用自适应控制时获得的结果相近，表明自适应权重并未显著改变平均性能。然而，在考察逐轮轨迹（图4中的结果）时，我们观察到，就证据使用而言，曲线动态与主研究中的不同：在最后一轮，分数低于主研究，原因是在自适应设置下， w_D随时间下降。而在初始轮次中，当权重w_D设置过低时，分数会上升。这一模式在None条件中尤为明显。这些发现表明，自适应权重能够根据其配置有效调控对话过程。

表3：土地 场景中不同权重和规则模板下策略的整体表现。详细实验结果见附录C。

权重	Rule	回应	反驳	非重复	证据	立场
$w_T = 1.0$	None	$0.85_{\pm 0.35}$	$0.28_{\pm 0.45}$	$0.43_{\pm 0.42}$	$0.10_{\pm 0.29}$	$0.48_{\pm 0.07}$
$w_M = 1.0$	轻量	$0.85_{\pm 0.36}$	$0.31_{\pm 0.46}$	$0.33_{\pm 0.36}$	$0.28_{\pm 0.45}$	$0.47_{\pm 0.07}$
$w_D = 1.0$	结构	$0.80_{\pm 0.40}$	$0.22_{\pm 0.41}$	$0.62_{\pm 0.36}$	$0.20_{\pm 0.40}$	$0.47_{\pm 0.08}$
$w_T = 1.0$	None	$0.86_{\pm 0.34}$	$0.28_{\pm 0.45}$	$0.30_{\pm 0.35}$	$0.30_{\pm 0.46}$	$0.47_{\pm 0.08}$
$w_M = 1.0$	轻量	$0.84_{\pm 0.37}$	$0.33_{\pm 0.47}$	$0.31_{\pm 0.34}$	$0.28_{\pm 0.45}$	$0.44_{\pm 0.10}$
$w_D = 1.5$	结构	$0.82_{\pm 0.38}$	$0.17_{\pm 0.37}$	$0.49_{\pm 0.40}$	$0.17_{\pm 0.37}$	$0.48_{\pm 0.07}$
$w_T = 1.0$	None	$0.86_{\pm 0.35}$	$0.28_{\pm 0.45}$	$0.47_{\pm 0.40}$	$0.12_{\pm 0.32}$	$0.51_{\pm 0.08}$
$w_M = 1.5$	轻量	$0.86_{\pm 0.34}$	$0.34_{\pm 0.47}$	$0.36_{\pm 0.36}$	$0.26_{\pm 0.44}$	$0.46_{\pm 0.08}$
$w_D = 1.0$	结构	$0.82_{\pm 0.38}$	$0.20_{\pm 0.40}$	$0.49_{\pm 0.40}$	$0.20_{\pm 0.40}$	$0.49_{\pm 0.07}$
$w_T = 1.5$	None	$0.81_{\pm 0.39}$	$0.45_{\pm 0.50}$	$0.26_{\pm 0.37}$	$0.05_{\pm 0.23}$	$0.55_{\pm 0.07}$
$w_M = 1.0$	轻量	$0.83_{\pm 0.37}$	$0.45_{\pm 0.50}$	$0.30_{\pm 0.34}$	$0.28_{\pm 0.45}$	$0.52_{\pm 0.08}$
$w_D = 1.0$	结构	$0.74_{\pm 0.44}$	$0.47_{\pm 0.50}$	$0.36_{\pm 0.39}$	$0.16_{\pm 0.37}$	$0.54_{\pm 0.07}$
$w_T = 1.0$	None	$0.87_{\pm 0.34}$	$0.31_{\pm 0.46}$	$0.31_{\pm 0.39}$	$0.09_{\pm 0.29}$	$0.48_{\pm 0.07}$
$w_M = 0.5$	轻量	$0.85_{\pm 0.35}$	$0.36_{\pm 0.48}$	$0.34_{\pm 0.35}$	$0.32_{\pm 0.47}$	$0.46_{\pm 0.07}$
$w_D = 1.0$	结构	$0.79_{\pm 0.41}$	$0.23_{\pm 0.42}$	$0.52_{\pm 0.39}$	$0.23_{\pm 0.42}$	$0.49_{\pm 0.07}$
$w_T = 1.0$	None	$0.85_{\pm 0.35}$	$0.32_{\pm 0.47}$	$0.41_{\pm 0.41}$	$0.10_{\pm 0.29}$	$0.50_{\pm 0.08}$
$w_M = 1.0$	轻量	$0.86_{\pm 0.35}$	$0.31_{\pm 0.46}$	$0.34_{\pm 0.35}$	$0.39_{\pm 0.49}$	$0.47_{\pm 0.08}$
$w_D = 0.5$	结构	$0.82_{\pm 0.39}$	$0.16_{\pm 0.36}$	$0.57_{\pm 0.38}$	$0.19_{\pm 0.39}$	$0.49_{\pm 0.06}$
$w_T = 0.5$	None	$0.87_{\pm 0.34}$	$0.32_{\pm 0.47}$	$0.30_{\pm 0.39}$	$0.08_{\pm 0.27}$	$0.50_{\pm 0.07}$
$w_M = 1.0$	轻量	$0.88_{\pm 0.32}$	$0.37_{\pm 0.48}$	$0.30_{\pm 0.36}$	$0.36_{\pm 0.48}$	$0.47_{\pm 0.07}$
$w_D = 1.0$	结构	$0.81_{\pm 0.39}$	$0.27_{\pm 0.44}$	$0.49_{\pm 0.40}$	$0.27_{\pm 0.44}$	$0.51_{\pm 0.08}$



(a) W_D **(b)** W_M

图3：自适应权重变化示例。Q1主题下，农民代理的 W_D 和 W_M 在10轮次中的变化。

表4：在土地场景中采用自适应权重设置的策略的总体表现，取所有智能体的平均值。详细结果见附录G。

Rule	回应	反驳	非重复	证据	立场
None	$0.86_{\pm 0.35}$	$0.26_{\pm 0.44}$	$0.45_{\pm 0.39}$	$0.18_{\pm 0.38}$	$0.48_{\pm 0.07}$
轻量	$0.85_{\pm 0.36}$	$0.39_{\pm 0.49}$	$0.37_{\pm 0.35}$	$0.29_{\pm 0.46}$	$0.45_{\pm 0.07}$
结构	$0.82_{\pm 0.38}$	$0.13_{\pm 0.34}$	$0.52_{\pm 0.40}$	$0.16_{\pm 0.37}$	$0.48_{\pm 0.07}$

4.2.4 权重区间与 α. 我们对自适应权重范围的选择直接源于对权重尺度的定义：我们将1.0视为中档值，因此权重自然落在$[0, 2]$之内。人们也可以等价地将尺度重新缩放到其他数值，但这只需调整档位阈值（目前

表 5：不同 α 值的影响。

α	Rule	回应	Reb.	非重复	证据	立场
0.05	None	0.83 ± 0.38	0.24 ± 0.43	0.54 ± 0.38	0.15 ± 0.35	0.49 ± 0.08
	轻量	0.83 ± 0.37	0.25 ± 0.43	0.42 ± 0.36	0.21 ± 0.41	0.46 ± 0.08
	结构	0.84 ± 0.37	0.10 ± 0.30	0.62 ± 0.36	0.18 ± 0.39	0.49 ± 0.06
0.1	None	0.83 ± 0.37	0.20 ± 0.40	0.54 ± 0.37	0.17 ± 0.38	0.49 ± 0.08
	轻量	0.84 ± 0.37	0.24 ± 0.42	0.41 ± 0.35	0.28 ± 0.45	0.46 ± 0.08
	结构	0.83 ± 0.37	0.08 ± 0.28	0.58 ± 0.37	0.20 ± 0.40	0.49 ± 0.06
0.4	None	0.84 ± 0.37	0.24 ± 0.43	0.47 ± 0.38	0.21 ± 0.41	0.48 ± 0.07
	轻量	0.84 ± 0.37	0.25 ± 0.43	0.41 ± 0.36	0.28 ± 0.45	0.46 ± 0.08
	结构	0.83 ± 0.38	0.08 ± 0.27	0.56 ± 0.37	0.20 ± 0.40	0.49 ± 0.06

0.85 和 1.25）。阈值与 α 共同发挥作用。因此，我们只需确定一个参数并调整另一个参数。我们选择固定阈值，并在表 5 中测试了多个 α 取值：所有设置下的总体指标都保持稳定，而 α 主要影响权重轨迹。不同的 α 取值决定权重跨越层级的频率以及权重变化的平滑程度。这表明，α 主要控制适应的平滑程度，并根据其配置调整对话动态。

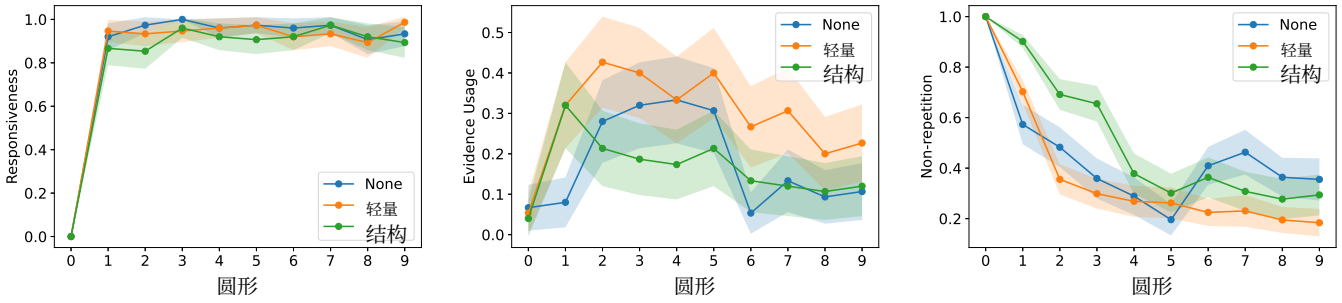
4.3 骨干大语言模型变化

我们针对土地 场景，使用不同的骨干大语言模型配置开展了两项额外实验。首先，我们测试了同质化配置 *S1*，其中所有智能体共享同一个骨干大语言模型（Qwen3）。表 6 的结果显示，与我们的主要异质化配置相比，该配置通常会导致响应性、反驳和非重复性较低。这表明，模型多样性有助于开展更丰富、互动性更强的讨论，而使用统一的骨干大语言模型则会导致对话行为缺乏动态性。

其次，我们测试了一种替代性异质化配置，以检验不同的异构骨干大语言模型选择是否能够保持所观察到的趋势。我们将Llama3、Mistral、Qwen3重新分配给农民、自然保护主义者、社区代表，并将该配置记为*S2*。同样，通过将Mistral、Qwen3、Llama3重新分配给相同的角色，我们得到配置*S3*。连同主研究，这三种设置（主研究、*S2*、*S3*）确保每个智能体都与不同的骨干大语言模型配对。表 6 的结果表明，尽管结果与主研究相比略有差异，但总体结论保持一致。此外，异构骨干网络配置继续优于同质的全Qwen3设置，证实骨干网络大语言模型的多样性会带来更稳健、更有效的对话行为。

在附录 F 中，我们提供了关于以下方面的额外实验结果：（1）不同大语言模型骨干网络对智能体层面行为的影响；（2）在包含更多智能体的更大规模多智能体系统中，智能体层面的性能。结果表明，尽管在不同骨干网络设置下，智能体层面的性能会出现轻微变化，但总体行为仍主要受到规则的影响。

为避免自我评估偏差，评判模型仅用于评估，不属于所提出的框架。我们进一步指出，五项评估指标中有三项基于嵌入，因此不依赖评判模型。不过，我们仍使用三个独立评判模型对系统进行了评估



(a) 响应性 **(b) 证据使用** **(c) 非重复**

图4：三个对话指标随自适应权重变化的轮次变化。

表 6：同质（*S*1）与异构（*S*2、 *S*3）骨干大语言模型变体之间的策略整体表现对比。结果以均值±标准差的形式报告。详细性能见附录E。

	Rule	回应	反驳	非重复	证据	立场
S1	None	$0.67_{\pm 0.47}$	$0.28_{\pm 0.45}$	$0.25_{\pm 0.39}$	$0.14_{\pm 0.35}$	$0.48_{\pm 0.07}$
	轻量	$0.59_{\pm 0.49}$	$0.26_{\pm 0.44}$	$0.26_{\pm 0.37}$	$0.16_{\pm 0.36}$	$0.44_{\pm 0.08}$
	结构	$0.74_{\pm 0.44}$	$0.12_{\pm 0.33}$	$0.38_{\pm 0.41}$	$0.10_{\pm 0.29}$	$0.47_{\pm 0.07}$
S2	None	$0.71_{\pm 0.45}$	$0.26_{\pm 0.44}$	$0.39_{\pm 0.40}$	$0.06_{\pm 0.23}$	$0.48_{\pm 0.09}$
	轻量	$0.72_{\pm 0.45}$	$0.30_{\pm 0.46}$	$0.33_{\pm 0.35}$	$0.42_{\pm 0.49}$	$0.46_{\pm 0.09}$
	结构	$0.70_{\pm 0.46}$	$0.16_{\pm 0.37}$	$0.61_{\pm 0.36}$	$0.22_{\pm 0.42}$	$0.50_{\pm 0.09}$
S3	None	$0.79_{\pm 0.41}$	$0.23_{\pm 0.42}$	$0.46_{\pm 0.40}$	$0.12_{\pm 0.32}$	$0.47_{\pm 0.08}$
	轻量	$0.85_{\pm 0.36}$	$0.24_{\pm 0.43}$	$0.35_{\pm 0.36}$	$0.24_{\pm 0.43}$	$0.46_{\pm 0.07}$
	结构	$0.84_{\pm 0.36}$	$0.08_{\pm 0.27}$	$0.67_{\pm 0.34}$	$0.24_{\pm 0.43}$	$0.50_{\pm 0.06}$

在每组运行5次的受控10轮对话设置下。不同评判者之间的结果高度一致，表明我们的结论不受评判模型选择的影响，具有稳健性。

4.4 消融研究

在前面的实验中考察了规则 R 和权重参数 W 的影响后，我们现在通过在*Land*场景中进行一系列消融研究，探究其余三个组件：任务 T、记忆 M和外部知识库 D。图5报告了消融研究的结果，其中我们选择性地移除了 T、 M和 D。结果表明，每个组件都以独特的方式发挥作用。 T显著提高了反驳频率和立场一致性，验证了清晰的人物与任务描述会促使其与其他智能体产生冲突，并保持稳定的立场。 D 提高了证据使用率，表明证据检索能够促进有依据的讨论。 M 会导致更高的重复率，因为当 M 加入后，智能体会利用此前对话轮次中的信息生成回应。组件组合体现了互补性： $D+T$ 带来了最均衡的性能，同时改善了反驳、证据使用和立场。然而，同时启用这三个维度，在各项指标上的提升仍然有限（例如响应性，这是由于使用了 M），这表明 T、 M和 D的效果可能会部分相互抵消。总体而言，这些结果凸显了不同组件各自可解释的作用。我们还在 EDU 场景中进行了消融研究，其结论与 Land 设置下的结论一致，部分实验结果见附录H。

5 讨论

本研究所提出的、用于影响基于大语言模型的多智能体系统的政策生成过程参数化框架，其重要意义在于重新定义了语言模型在社会模拟中的角色。语言模型不再只是文本生成器，而是具有可调参数的社会行动者。通过我们的框架，每个智能体的对话响应过程（由提示即动作产生）都可以映射为一组含义明确的变量。未来的工作可以通过微调 [2] 和推理时干预 [13] 等技术扩展我们的框架，以定制智能体策略参数。该框架使响应能够由一组定义清晰的认知与社会策略加以控制，从而使社会模拟更加多样和灵活，并为可控、可测量且具有理论依据的社会实验提供机制。

6 结论

本研究提出了一种轻量的策略参数化框架，将提示即动作视为影响基于大语言模型的多智能体对话的手段。通过自适应地组合不同组成部分来构建提示，我们表明，无需额外训练即可系统性地影响对话行为。在两个场景中开展的实验表明，不同的组成部分设置会产生不同的效果。总体而言，我们的研究结果凸显了策略参数化提示作为一种有效且可解释的机制，能够引导大语言模型驱动的多智能体对话系统，并为可控社会模拟提供了一个有前景的方向。

7 致谢

谨此致谢经济与社会研究理事会（ESRC）的资助。项目编号：ES/W002639/1。Hongbo Bo 由 ESRC社会数字未来中心（ES/W002639/1）资助，Jingyu Hu 由 EPSRC-DTP（EP/W524414/1/2894964）资助。Weiru Liu 部分由 ESRC社会数字未来中心（ES/W002639/1）资助。感谢 Keri Facer 提供文档列表并参与有益的讨论。

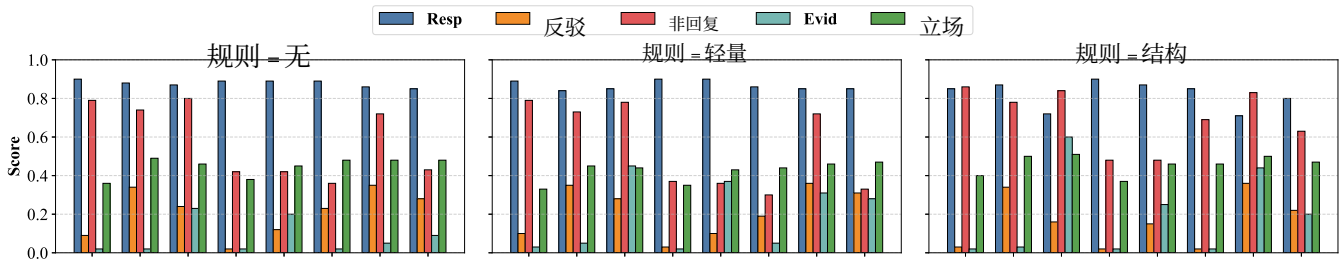


图5：视角、材料和证据组件的消融研究。

参考文献

[1] Sahar Abdelnabi、Amr Gomaa、Sarath Sivaprasad、Lea Schönherr 和 Mario Fritz。2023。LLM-Deliberation：通过交互式多智能体谈判博弈评估大语言模型。（2023）。[2] Ahmed Agiza、Mohamed Mostagir 和 Sherief Reda。2024。Politune：分析数据选择和微调对大语言模型中经济和政治偏见的影响。载于*AAAI/ACM*人工智能、伦理与社会会议论文集，第7卷。2–12。[3] Ariel Flint Ashery、Luca Maria Aiello 和 Andrea Baronchelli。2025。大语言模型群体中涌现的社会惯例与集体偏见。《科学进展》11，20（2025），9368，eadu。[4] Abhimanyu Dubey、Abhinav Jauhri、Abhinav Pandey、Abhishek Kadian、Ahmad Al-Dahle、Aiesha Letman、Akhil Mathur、Alan Schelten、Amy Yang、Angela Fan 等。2024。Llama 3 模型家族。*arXiv* 电子预印本（2024），arXiv–2407。[5] Andrew Estornell 和 Yang Liu。2024。多大语言模型辩论：框架、原则与干预措施。神经信息处理系统进展37（2024），28938–28964。[6] Caoyun Fan、Jindou Chen、Yaohui Jin 和 Hao He。2024。大语言模型能否在博弈论中充当理性参与者？系统性分析。载于*AAAI*人工智能会议论文集，第38卷。17960–17967。[7] Linxi Fan、Guanzhi Wang、Yunfan Jiang、Ajay Mandlekar、Yuncong Yang、Haoyi Zhu、Andrew Tang、De-An Huang、Yuke Zhu 和 Anima Anandkumar。2022。MineDojo：利用互联网规模知识构建开放式具身智能体。神经信息处理系统进展35（2022），18343–18362。[8] Chen Gao、Xiaochong Lan、Zhihong Lu、Jinzhu Mao、Jinghua Piao、Huandong Wang、Depeng Jin 和 Yong Li。2023。S3：配备大语言模型赋能智能体的社交网络模拟系统。*arXiv* 预印本 *arXiv:2307.14984*（2023）。[9] Sven Gronauer 和 Klaus Diepold。2022。多智能体深度强化学习：综述。《人工智能评论》55，2（2022），895–943。[10] Taicheng Guo、Xiuying Chen、Yaqi Wang、Ruidi Chang、Shichao Pei、Nitesh V Chawla、Olaf Wiest 和 Xiangliang Zhang。2024。基于大语言模型的多智能体：进展与挑战综述。*arXiv* 预印本 *arXiv:2402.01680*（2024）。[11] Shibo Hao、Yi Gu、Haodi Ma、Joshua Jiahua Hong、Zhen Wang、Daisy Zhe Wang 和 Zhiting Hu。2023。基于语言模型的推理就是基于世界模型的规划。*arXiv* 预印本 *arXiv:2305.14992*（2023）。[12] Sirui Hong、Mingchen Zhuge、Jonathan Chen、Xiawu Zheng、Yuheng Cheng、Ceyao Zhang、Jinlin Wang、Zili Wang、Steven Ka Shing Yau、Zijuan Lin 等。2024。MetaGPT：面向多智能体协作框架的元编程。国际学习表征会议（ICLR）。[13] Jingyu Hu、Mengyue Yang、Mengnan Du 和 Weiru Liu。2025。大语言模型中政治观点的细粒度解读。*arXiv* 预印本 *arXiv:2506.04774*（2025）。[14] Minsu Jang、Youngwoo Yoon、Jaewoo Choi、Hyobin Ong 和 Jaehong Kim。2023。基于信念—愿望—意图模型的结构化提示方法，用于主动且可解释的任务规划。载于第*11*届国际人机智能体交互会议论文集。375–377。[15] Albert Jiang、Patrick von Platen、Nazar Habib、Teven Le Scao、Thomas Wolf 和 Mistral AI。2023。Mistral 7B。*arXiv* 预印本 *arXiv:2310.06825*（2023）。https: //arxiv.org/abs/2310.06825[16] Patrick Lewis、Ethan Perez、Aleksandra Piktus、Fabio Petroni、Vladimir Karpukhin、Naman Goyal、Heinrich Küttler、Mike Lewis、Wen-tau Yih、Tim Rocktä schel 等。2020。面向知识密集型自然语言处理任务的检索增强生成。神经信息处理系统进展33（2020），9459–9474。[17] Guohao Li、Hasan Hammoud、Hani Itani、Dmitrii Khizbullin 和 Bernard Ghanem。2023。CAMEL：用于探索“大语言模型社会”的交际智能体。神经信息处理系统进展36 (2023), 51991–52008. [18] Xinyi Li、Sai Wang、Siqi Zeng、Yu Wu 和 Yi Yang。2024年。《基于大语言模型的多智能体系统综述：工作流、基础设施与挑战》。*Vicinagearth* 1, 1 (2024), 9.[19] Tian Liang、Zhiwei He、Wenxiang Jiao、Xing Wang、Yan Wang、Rui Wang、Yujiu Yang、Shuming Shi 和 Zhaopeng Tu。2023年。《通过多智能体辩论鼓励大语言模型中的发散性思维》。*arXiv preprint arXiv:2305.19118* (2023)。[20] Jiaju Lin、Haoran Zhao、Aochi Zhang、Yiting Wu、Huqiuyue Ping 和 Qin Chen。2023年。《Agentsims：用于大语言模型评估的开源沙盒》。*arXiv preprint arXiv:2308.04026* (2023).[21] Ruibo Liu、Ruixin Yang、Chenyan Jia、Ge Zhang、Denny Zhou、Andrew M Dai、Diyi Yang 和 Soroush Vosoughi。2023年。《在模拟人类社会中训练社会对齐语言模型》。*arXiv preprint arXiv:2305.16960* 2 (2023). [22] OpenAI。2025年。ChatGPT-5。https://chat.openai.com/。大语言模型。[23] Joon Sung Park、Joseph O’Brien、Carrie Jun Cai、Meredith Ringel Morris、Percy Liang 和 Michael S Bernstein。2023年。《生成式智能体：人类行为的交互式模拟体》。载于第*36*届*ACM*用户界面软件与技术年度研讨会论文集。1–22.[24] Joon Sung Park、Lindsay Popowski、Carrie Cai、Meredith Ringel Morris、Percy Liang 和 Michael S Bernstein。2022年。《社会模拟体：为社会计算系统创建带有虚拟人群的原型》。载于第*35*届*ACM*用户界面软件与技术年度研讨会论文集。1–18.[25] Jinghua Piao、Yuwei Yan、Jun Zhang、Nian Li、Junbo Yan、Xiaochong Lan、Zhihong Lu、Zhiheng Zheng、Jing Yi Wang、Di Zhou 等。2025年。《Agentsociety：大规模模拟大语言模型驱动的生成式智能体，助力深入理解人类行为与社会》。*arXiv preprint arXiv:2502.08691* (2025).[26] Noah Shinn、Federico Cassano、Ashwin Gopinath、Karthik Narasimhan 和 Shunyu Yao。2023年。《Reflexion：通过语言反馈进行强化学习的语言智能体》。神经信息处理系统进展36 (2023), 8634–8652。[27] Wenhui Wang、Furu Wei、Li Dong、Hangbo Bao、Nan Yang 和 Ming Zhou。2020年。《MiniLM：面向任务无关压缩的预训练变换器深度自注意力蒸馏》。神经信息处理系统进展33 (2020), 5776–5788.[28] Qingyun Wu、Gagan Bansal、Jieyu Zhang、Yiran Wu、Beibin Li、Erkang Zhu、Li Jiang、Xiaoyun Zhang、Shaokun Zhang、Jiale Liu 等。2024年。《AutoGen：通过多智能体对话赋能新一代大语言模型应用》。载于首届语言建模会议。[29] Yiwen Wu、Kevin McAreavey、Hongbo Bo、Weiru Liu 和 Ryan McConville。2025 年。《用于虚假新闻检测的多智能体深度强化学习》。载于*2025*年国际神经网络联合会议（*IJCNN*）。IEEE，1–8。[30] Zelai Xu、Chao Yu、Fei Fang、Yu Wang 和 Yi Wu。2023年。《运用强化学习进行狼人杀游戏策略性游戏的语言智能体》。*arXiv preprint arXiv:2310.18940* (2023).[31] An Yang、Anfeng Li、Baosong Yang、Beichen Zhang、Binyuan Hui、Bo Zheng、Bowen Yu、Chang Gao、Chengen Huang、Chenxu Lv 等。2025年。《通义千问3技术报告》。*arXiv preprint arXiv:2505.09388* (2025).[32] John Yang、Carlos E Jimenez、Alexander Wettig、Kilian Lieret、Shunyu Yao、Karthik Narasimhan 和 Ofir Press。2024年。《SWE-agent：智能体-计算机接口赋能自动化软件工程》。神经信息处理系统进展37 (2024), 50528–50652。[33] Changxi Zhu、Mehdi Dastani 和 Shihan Wang。2024年。《带通信的多智能体深度强化学习综述》。自主智能体与多智能体系统38, 1 (2024), 4。

A 智能体外部知识补充信息

土地情景的数据汇编自与英国土地利用和农业相关的公开材料。来源包括政府出版物，例如《土地利用未来》报告 [6] 、英国环境、食品与农村事务部近期发布的土地利用框架咨询文件[7], , 以及农民、自然保护和公众通行权组织的政策文件与报告（例如全国农民联盟[8], 、绿色联盟[9], 和林地信托[10]）。

教育情境的数据汇编自公开的英文报告和有关教育公平、经费及学习成果的国际数据集。主要来源包括《经合组织教育概览》 *2023*[11], 、世界银行《世界发展指标》：政府教育支出（占 *GDP* 的百分比）[12], 以及世界银行教育公平与入学机会开放知识库[13]。这些来源共同提供了关于全球各教育体系支出水平、教师分布、城乡差异和学生表现的实证证据。

B 权重的增强指令

不同权重等级的详细指令：

- T– 高：明确从 [T] 视角发言。首先表明立场和角色特定的优先事项，然后给出理由进行论证。– 中：体现 [T] 视角，并在适当时表明立场。– 低：可以不明说 [T] 视角；重点放在论据上。

- M：– 高：先用 1–2 句话概述 [M] 中的近期轮次，并直接回应尚未解决的问题。– 中：结合近期讨论 [M] ，避免重复先前内容。– 低：可以不总结 [M]而直接回应；避免逐字重复。

- 证据;– 高：在下结论之前，至少列出 2 条来自检索片段的具体证据 [证据]；引用或改述这些内容，并将其与主张联系起来。– 中：证据（优先）：如有相关检索片段 [证据] ，请用其支持关键主张。),– 低：证据（可选）：如果并非必要，可以不引用检索片段 [证据]。

[6]https://assets.publishing.service.gov.uk/media/5a7c31f740f0b674ed20f740/10-634-land-use-futures-summary.pdf [7] https://consult.defra.gov.uk/land-use-framework/land-use-consultation/ [8] https://www.nfuonline.com/updates-and-information/farming-for-britains-future-our-key-asks-of-government/ [9] https://green-alliance.org.uk/project/the-future-of-uk-land-use/ [10] https://www.woodlandtrust.org.uk/state-of-uk-woods-and-trees/ [11] https://www.oecd.org/education/education-at-a-glance/ [12] https://data.worldbank.org/indicator/SE.XPD.TOTL.GD.ZS [13] https://openknowledge.worldbank.org/

C 权重敏感性

表7、8、9、10、11和12展示了权重敏感性实验的结果详情。

表 7： $w_T = 1.0,\ \ w_M = 1.0,\ \ w_D = 1.5$

查询	Rule	回应	反驳	非重复	证据	立场
Q1	None	$0.90_{\pm 0.30}$	$0.33_{\pm 0.47}$	$0.30_{\pm 0.35}$	$0.37_{\pm 0.48}$	$0.52_{\pm 0.08}$
	轻量	$0.83_{\pm 0.37}$	$0.31_{\pm 0.47}$	$0.23_{\pm 0.33}$	$0.49_{\pm 0.50}$	$0.47_{\pm 0.08}$
	结构	$0.81_{\pm 0.40}$	$0.19_{\pm 0.40}$	$0.54_{\pm 0.40}$	$0.29_{\pm 0.46}$	$0.51_{\pm 0.08}$
Q2	None	$0.90_{\pm 0.30}$	$0.34_{\pm 0.47}$	$0.32_{\pm 0.36}$	$0.39_{\pm 0.49}$	$0.46_{\pm 0.07}$
	轻量	$0.88_{\pm 0.33}$	$0.32_{\pm 0.47}$	$0.35_{\pm 0.33}$	$0.25_{\pm 0.44}$	$0.40_{\pm 0.06}$
	结构	$0.87_{\pm 0.33}$	$0.14_{\pm 0.35}$	$0.50_{\pm 0.40}$	$0.17_{\pm 0.38}$	$0.46_{\pm 0.06}$
Q3	None	$0.90_{\pm 0.30}$	$0.03_{\pm 0.16}$	$0.34_{\pm 0.36}$	$0.33_{\pm 0.47}$	$0.47_{\pm 0.09}$
	轻量	$0.89_{\pm 0.32}$	$0.05_{\pm 0.21}$	$0.40_{\pm 0.34}$	$0.13_{\pm 0.34}$	$0.42_{\pm 0.12}$
	结构	$0.89_{\pm 0.31}$	$0.02_{\pm 0.14}$	$0.52_{\pm 0.38}$	$0.11_{\pm 0.32}$	$0.45_{\pm 0.07}$
Q4	None	$0.73_{\pm 0.45}$	$0.53_{\pm 0.50}$	$0.29_{\pm 0.35}$	$0.15_{\pm 0.35}$	$0.50_{\pm 0.08}$
	轻量	$0.71_{\pm 0.46}$	$0.51_{\pm 0.50}$	$0.28_{\pm 0.34}$	$0.23_{\pm 0.42}$	$0.49_{\pm 0.10}$
	结构	$0.69_{\pm 0.46}$	$0.33_{\pm 0.47}$	$0.40_{\pm 0.40}$	$0.12_{\pm 0.33}$	$0.49_{\pm 0.06}$
Q5	None	$0.89_{\pm 0.32}$	$0.17_{\pm 0.37}$	$0.27_{\pm 0.34}$	$0.24_{\pm 0.43}$	$0.42_{\pm 0.04}$
	轻量	$0.89_{\pm 0.32}$	$0.45_{\pm 0.50}$	$0.29_{\pm 0.34}$	$0.28_{\pm 0.45}$	$0.41_{\pm 0.05}$
	结构	$0.86_{\pm 0.35}$	$0.15_{\pm 0.35}$	$0.48_{\pm 0.39}$	$0.13_{\pm 0.33}$	$0.46_{\pm 0.07}$
总体	None	$0.86_{\pm 0.34}$	$0.28_{\pm 0.45}$	$0.30_{\pm 0.35}$	$0.30_{\pm 0.46}$	$0.47_{\pm 0.08}$
	轻量	$0.84_{\pm 0.37}$	$0.33_{\pm 0.47}$	$0.31_{\pm 0.34}$	$0.28_{\pm 0.45}$	$0.44_{\pm 0.10}$
	结构	$0.82_{\pm 0.38}$	$0.17_{\pm 0.37}$	$0.49_{\pm 0.40}$	$0.17_{\pm 0.37}$	$0.48_{\pm 0.07}$

表8： $w_T = 1.0,\ \ w_M = 1.5,\ \ w_D = 1.0$

查询	Rule	回应	反驳	非重复	证据	立场
Q1	None	$0.86_{\pm 0.35}$	$0.27_{\pm 0.45}$	$0.57_{\pm 0.37}$	$0.27_{\pm 0.44}$	$0.56_{\pm 0.08}$
–	轻量	$0.88_{\pm 0.33}$	$0.44_{\pm 0.50}$	$0.38_{\pm 0.37}$	$0.38_{\pm 0.49}$	$0.49_{\pm 0.07}$
	结构	$0.83_{\pm 0.37}$	$0.22_{\pm 0.42}$	$0.53_{\pm 0.40}$	$0.35_{\pm 0.48}$	$0.51_{\pm 0.08}$
Q2	None	$0.87_{\pm 0.33}$	$0.37_{\pm 0.48}$	$0.46_{\pm 0.41}$	$0.05_{\pm 0.23}$	$0.49_{\pm 0.06}$
	轻量	$0.88_{\pm 0.33}$	$0.28_{\pm 0.45}$	$0.36_{\pm 0.37}$	$0.23_{\pm 0.42}$	$0.43_{\pm 0.06}$
	结构	$0.86_{\pm 0.35}$	$0.25_{\pm 0.43}$	$0.47_{\pm 0.41}$	$0.12_{\pm 0.33}$	$0.49_{\pm 0.06}$
Q3	None	$0.89_{\pm 0.31}$	$0.01_{\pm 0.12}$	$0.52_{\pm 0.38}$	$0.09_{\pm 0.29}$	$0.49_{\pm 0.06}$
	轻量	$0.87_{\pm 0.33}$	$0.09_{\pm 0.29}$	$0.45_{\pm 0.36}$	$0.09_{\pm 0.29}$	$0.43_{\pm 0.06}$
	结构	$0.88_{\pm 0.33}$	$0.01_{\pm 0.08}$	$0.50_{\pm 0.40}$	$0.10_{\pm 0.30}$	$0.47_{\pm 0.06}$
Q4	None	$0.76_{\pm 0.43}$	$0.45_{\pm 0.50}$	$0.30_{\pm 0.38}$	$0.06_{\pm 0.24}$	$0.51_{\pm 0.08}$
	轻量	$0.81_{\pm 0.39}$	$0.56_{\pm 0.50}$	$0.28_{\pm 0.34}$	$0.39_{\pm 0.49}$	$0.50_{\pm 0.09}$
	结构	$0.70_{\pm 0.46}$	$0.37_{\pm 0.48}$	$0.41_{\pm 0.40}$	$0.29_{\pm 0.45}$	$0.52_{\pm 0.06}$
Q5	None	$0.89_{\pm 0.31}$	$0.29_{\pm 0.45}$	$0.48_{\pm 0.40}$	$0.12_{\pm 0.33}$	$0.48_{\pm 0.09}$
	轻量	$0.87_{\pm 0.34}$	$0.33_{\pm 0.47}$	$0.36_{\pm 0.34}$	$0.17_{\pm 0.38}$	$0.42_{\pm 0.07}$
	结构	$0.85_{\pm 0.36}$	$0.13_{\pm 0.33}$	$0.56_{\pm 0.36}$	$0.15_{\pm 0.35}$	$0.47_{\pm 0.06}$
总体	None	$0.86_{\pm 0.35}$	$0.28_{\pm 0.45}$	$0.47_{\pm 0.40}$	$0.12_{\pm 0.32}$	$0.51_{\pm 0.08}$
	轻量	$0.86_{\pm 0.34}$	$0.34_{\pm 0.47}$	$0.36_{\pm 0.36}$	$0.26_{\pm 0.44}$	$0.46_{\pm 0.08}$
	结构	$0.82_{\pm 0.38}$	$0.20_{\pm 0.40}$	$0.49_{\pm 0.40}$	$0.20_{\pm 0.40}$	$0.49_{\pm 0.07}$

表9： $w_T = 1.5,\quad w_M = 1.0,\quad w_D = 1.0$

查询	Rule	回应	反驳	非重复	证据	立场
Q1	None	$0.88_{\pm 0.33}$	$0.33_{\pm 0.47}$	$0.32_{\pm 0.39}$	$0.08_{\pm 0.27}$	$0.61_{\pm 0.05}$
	轻量	$0.83_{\pm 0.38}$	$0.41_{\pm 0.49}$	$0.27_{\pm 0.33}$	$0.34_{\pm 0.47}$	$0.56_{\pm 0.07}$
	结构	$0.90_{\pm 0.30}$	$0.27_{\pm 0.45}$	$0.37_{\pm 0.38}$	$0.25_{\pm 0.44}$	$0.59_{\pm 0.07}$
Q2	None	$0.86_{\pm 0.35}$	$0.30_{\pm 0.46}$	$0.21_{\pm 0.34}$	$0.00_{\pm 0.00}$	$0.54_{\pm 0.06}$
	轻量	$0.89_{\pm 0.31}$	$0.31_{\pm 0.46}$	$0.26_{\pm 0.33}$	$0.25_{\pm 0.44}$	$0.50_{\pm 0.06}$
	结构	$0.79_{\pm 0.41}$	$0.48_{\pm 0.50}$	$0.29_{\pm 0.37}$	$0.08_{\pm 0.27}$	$0.51_{\pm 0.05}$
Q3	None	$0.72_{\pm 0.45}$	$0.41_{\pm 0.49}$	$0.23_{\pm 0.37}$	$0.01_{\pm 0.08}$	$0.55_{\pm 0.07}$
	轻量	$0.84_{\pm 0.37}$	$0.31_{\pm 0.47}$	$0.36_{\pm 0.36}$	$0.29_{\pm 0.45}$	$0.53_{\pm 0.09}$
	结构	$0.63_{\pm 0.48}$	$0.45_{\pm 0.50}$	$0.40_{\pm 0.41}$	$0.14_{\pm 0.35}$	$0.52_{\pm 0.07}$
Q4	None	$0.75_{\pm 0.44}$	$0.60_{\pm 0.49}$	$0.21_{\pm 0.35}$	$0.15_{\pm 0.35}$	$0.55_{\pm 0.07}$
	轻量	$0.75_{\pm 0.44}$	$0.60_{\pm 0.49}$	$0.26_{\pm 0.35}$	$0.31_{\pm 0.46}$	$0.53_{\pm 0.06}$
	结构	$0.61_{\pm 0.49}$	$0.57_{\pm 0.50}$	$0.33_{\pm 0.38}$	$0.15_{\pm 0.36}$	$0.54_{\pm 0.07}$
Q5	None	$0.86_{\pm 0.35}$	$0.61_{\pm 0.49}$	$0.32_{\pm 0.37}$	$0.03_{\pm 0.18}$	$0.51_{\pm 0.07}$
	轻量	$0.85_{\pm 0.35}$	$0.61_{\pm 0.49}$	$0.35_{\pm 0.35}$	$0.19_{\pm 0.40}$	$0.47_{\pm 0.07}$
	结构	$0.79_{\pm 0.41}$	$0.56_{\pm 0.50}$	$0.39_{\pm 0.38}$	$0.17_{\pm 0.38}$	$0.53_{\pm 0.06}$
总体	None	$0.81_{\pm 0.39}$	$0.45_{\pm 0.50}$	$0.26_{\pm 0.37}$	$0.05_{\pm 0.23}$	$0.55_{\pm 0.07}$
	轻量	$0.83_{\pm 0.37}$	$0.45_{\pm 0.50}$	$0.30_{\pm 0.34}$	$0.28_{\pm 0.45}$	$0.52_{\pm 0.08}$
	结构	$0.74_{\pm 0.44}$	$0.47_{\pm 0.50}$	$0.36_{\pm 0.39}$	$0.16_{\pm 0.37}$	$0.54_{\pm 0.07}$

表10： $w_T = 1.0, w_M = 0.5, w_D = 1.0$

查询	Rule	回应	反驳	非重复	证据	立场
Q1	None	$0.88_{\pm 0.33}$	$0.31_{\pm 0.46}$	$0.37_{\pm 0.41}$	$0.23_{\pm 0.42}$	$0.56_{\pm 0.05}$
	轻量	$0.87_{\pm 0.34}$	$0.35_{\pm 0.48}$	$0.40_{\pm 0.37}$	$0.33_{\pm 0.47}$	$0.51_{\pm 0.08}$
	结构	$0.75_{\pm 0.43}$	$0.26_{\pm 0.44}$	$0.53_{\pm 0.40}$	$0.35_{\pm 0.48}$	$0.51_{\pm 0.08}$
Q2	None	$0.89_{\pm 0.31}$	$0.44_{\pm 0.50}$	$0.29_{\pm 0.39}$	$0.08_{\pm 0.27}$	$0.46_{\pm 0.05}$
	轻量	$0.90_{\pm 0.30}$	$0.31_{\pm 0.46}$	$0.35_{\pm 0.34}$	$0.21_{\pm 0.41}$	$0.44_{\pm 0.04}$
	结构	$0.87_{\pm 0.34}$	$0.21_{\pm 0.41}$	$0.52_{\pm 0.39}$	$0.31_{\pm 0.47}$	$0.51_{\pm 0.07}$
Q3	None	$0.90_{\pm 0.30}$	$0.00_{\pm 0.00}$	$0.38_{\pm 0.42}$	$0.05_{\pm 0.23}$	$0.46_{\pm 0.06}$
	轻量	$0.90_{\pm 0.30}$	$0.03_{\pm 0.16}$	$0.36_{\pm 0.37}$	$0.23_{\pm 0.42}$	$0.45_{\pm 0.06}$
	结构	$0.89_{\pm 0.32}$	$0.01_{\pm 0.12}$	$0.60_{\pm 0.36}$	$0.09_{\pm 0.29}$	$0.46_{\pm 0.07}$
Q4	None	$0.78_{\pm 0.42}$	$0.55_{\pm 0.50}$	$0.29_{\pm 0.39}$	$0.07_{\pm 0.26}$	$0.50_{\pm 0.05}$
	轻量	$0.70_{\pm 0.46}$	$0.57_{\pm 0.50}$	$0.25_{\pm 0.35}$	$0.67_{\pm 0.47}$	$0.49_{\pm 0.09}$
	结构	$0.67_{\pm 0.47}$	$0.51_{\pm 0.50}$	$0.39_{\pm 0.40}$	$0.22_{\pm 0.42}$	$0.51_{\pm 0.05}$
Q5	None	$0.89_{\pm 0.31}$	$0.25_{\pm 0.44}$	$0.24_{\pm 0.35}$	$0.03_{\pm 0.16}$	$0.44_{\pm 0.03}$
	轻量	$0.89_{\pm 0.31}$	$0.54_{\pm 0.50}$	$0.32_{\pm 0.34}$	$0.15_{\pm 0.35}$	$0.42_{\pm 0.04}$
	结构	$0.77_{\pm 0.42}$	$0.18_{\pm 0.39}$	$0.54_{\pm 0.40}$	$0.17_{\pm 0.37}$	$0.46_{\pm 0.07}$
总体	None	$0.87_{\pm 0.34}$	$0.31_{\pm 0.46}$	$0.31_{\pm 0.39}$	$0.09_{\pm 0.29}$	$0.48_{\pm 0.07}$
	轻量	$0.85_{\pm 0.35}$	$0.36_{\pm 0.48}$	$0.34_{\pm 0.35}$	$0.32_{\pm 0.47}$	$0.46_{\pm 0.07}$
	结构	$0.79_{\pm 0.41}$	$0.23_{\pm 0.42}$	$0.52_{\pm 0.39}$	$0.23_{\pm 0.42}$	$0.49_{\pm 0.07}$

表11： $w_T = 1.0,\quad w_M = 1.0,\quad w_D = 0.5$

查询	Rule	回应	反驳	非重复	证据	立场
Q1	None	$0.86_{\pm 0.35}$	$0.19_{\pm 0.39}$	$0.42_{\pm 0.41}$	$0.20_{\pm 0.40}$	$0.54_{\pm 0.08}$
	轻量	$0.88_{\pm 0.33}$	$0.31_{\pm 0.47}$	$0.33_{\pm 0.35}$	$0.55_{\pm 0.50}$	$0.51_{\pm 0.10}$
	结构	$0.76_{\pm 0.43}$	$0.31_{\pm 0.46}$	$0.57_{\pm 0.38}$	$0.42_{\pm 0.49}$	$0.52_{\pm 0.07}$
Q2	None	$0.82_{\pm 0.39}$	$0.50_{\pm 0.50}$	$0.45_{\pm 0.42}$	$0.11_{\pm 0.32}$	$0.49_{\pm 0.08}$
	轻量	$0.87_{\pm 0.34}$	$0.33_{\pm 0.47}$	$0.32_{\pm 0.34}$	$0.18_{\pm 0.39}$	$0.44_{\pm 0.05}$
	结构	$0.79_{\pm 0.41}$	$0.12_{\pm 0.33}$	$0.57_{\pm 0.37}$	$0.05_{\pm 0.23}$	$0.48_{\pm 0.06}$
Q3	None	$0.89_{\pm 0.32}$	$0.02_{\pm 0.14}$	$0.44_{\pm 0.42}$	$0.06_{\pm 0.24}$	$0.51_{\pm 0.10}$
	轻量	$0.89_{\pm 0.32}$	$0.01_{\pm 0.12}$	$0.41_{\pm 0.37}$	$0.15_{\pm 0.36}$	$0.48_{\pm 0.09}$
	结构	$0.85_{\pm 0.36}$	$0.01_{\pm 0.12}$	$0.60_{\pm 0.36}$	$0.19_{\pm 0.40}$	$0.47_{\pm 0.06}$
Q4	None	$0.81_{\pm 0.39}$	$0.54_{\pm 0.50}$	$0.28_{\pm 0.36}$	$0.04_{\pm 0.20}$	$0.49_{\pm 0.05}$
	轻量	$0.81_{\pm 0.39}$	$0.45_{\pm 0.50}$	$0.28_{\pm 0.35}$	$0.64_{\pm 0.48}$	$0.50_{\pm 0.07}$
	结构	$0.87_{\pm 0.34}$	$0.22_{\pm 0.42}$	$0.54_{\pm 0.38}$	$0.12_{\pm 0.33}$	$0.49_{\pm 0.05}$
Q5	None	$0.89_{\pm 0.31}$	$0.37_{\pm 0.48}$	$0.44_{\pm 0.39}$	$0.06_{\pm 0.24}$	$0.47_{\pm 0.07}$
	轻量	$0.85_{\pm 0.35}$	$0.42_{\pm 0.49}$	$0.33_{\pm 0.35}$	$0.45_{\pm 0.50}$	$0.44_{\pm 0.06}$
	结构	$0.83_{\pm 0.38}$	$0.12_{\pm 0.33}$	$0.56_{\pm 0.39}$	$0.17_{\pm 0.38}$	$0.49_{\pm 0.07}$
总体	None	$0.85_{\pm 0.35}$	$0.32_{\pm 0.47}$	$0.41_{\pm 0.41}$	$0.10_{\pm 0.29}$	$0.50_{\pm 0.08}$
	轻量	$0.86_{\pm 0.35}$	$0.31_{\pm 0.46}$	$0.34_{\pm 0.35}$	$0.39_{\pm 0.49}$	$0.47_{\pm 0.08}$
	结构	$0.82_{\pm 0.39}$	$0.16_{\pm 0.36}$	$0.57_{\pm 0.38}$	$0.19_{\pm 0.39}$	$0.49_{\pm 0.06}$

表12： $w_T = 0.5,\quad w_M = 1.0,\quad w_D = 1.0$

查询	Rule	回应	反驳	非重复	证据	立场
Q1	None	$0.89_{\pm 0.31}$	$0.31_{\pm 0.46}$	$0.36_{\pm 0.41}$	$0.18_{\pm 0.39}$	$0.56_{\pm 0.07}$
	轻量	$0.90_{\pm 0.30}$	$0.51_{\pm 0.50}$	$0.32_{\pm 0.36}$	$0.27_{\pm 0.44}$	$0.49_{\pm 0.08}$
	结构	$0.83_{\pm 0.37}$	$0.33_{\pm 0.47}$	$0.55_{\pm 0.38}$	$0.34_{\pm 0.47}$	$0.57_{\pm 0.08}$
Q2	None	$0.89_{\pm 0.32}$	$0.38_{\pm 0.49}$	$0.29_{\pm 0.39}$	$0.03_{\pm 0.16}$	$0.48_{\pm 0.06}$
	轻量	$0.90_{\pm 0.30}$	$0.30_{\pm 0.46}$	$0.22_{\pm 0.33}$	$0.46_{\pm 0.50}$	$0.43_{\pm 0.05}$
	结构	$0.85_{\pm 0.36}$	$0.20_{\pm 0.40}$	$0.51_{\pm 0.40}$	$0.12_{\pm 0.33}$	$0.49_{\pm 0.06}$
Q3	None	$0.87_{\pm 0.33}$	$0.01_{\pm 0.08}$	$0.31_{\pm 0.40}$	$0.04_{\pm 0.20}$	$0.50_{\pm 0.07}$
	轻量	$0.90_{\pm 0.30}$	$0.02_{\pm 0.14}$	$0.39_{\pm 0.39}$	$0.32_{\pm 0.47}$	$0.48_{\pm 0.09}$
	结构	$0.87_{\pm 0.33}$	$0.06_{\pm 0.24}$	$0.54_{\pm 0.40}$	$0.25_{\pm 0.44}$	$0.49_{\pm 0.08}$
Q4	None	$0.79_{\pm 0.41}$	$0.60_{\pm 0.49}$	$0.24_{\pm 0.37}$	$0.13_{\pm 0.34}$	$0.51_{\pm 0.06}$
	轻量	$0.83_{\pm 0.37}$	$0.53_{\pm 0.50}$	$0.29_{\pm 0.36}$	$0.58_{\pm 0.49}$	$0.49_{\pm 0.07}$
	结构	$0.71_{\pm 0.45}$	$0.44_{\pm 0.50}$	$0.39_{\pm 0.39}$	$0.50_{\pm 0.50}$	$0.51_{\pm 0.07}$
Q5	None	$0.89_{\pm 0.31}$	$0.32_{\pm 0.47}$	$0.29_{\pm 0.37}$	$0.02_{\pm 0.14}$	$0.42_{\pm 0.03}$
	轻量	$0.88_{\pm 0.33}$	$0.51_{\pm 0.50}$	$0.27_{\pm 0.34}$	$0.19_{\pm 0.40}$	$0.42_{\pm 0.04}$
	结构	$0.80_{\pm 0.40}$	$0.32_{\pm 0.47}$	$0.49_{\pm 0.41}$	$0.13_{\pm 0.34}$	$0.49_{\pm 0.07}$
总体	None	$0.87_{\pm 0.34}$	$0.32_{\pm 0.47}$	$0.30_{\pm 0.39}$	$0.08_{\pm 0.27}$	$0.50_{\pm 0.07}$
	轻量	$0.88_{\pm 0.32}$	$0.37_{\pm 0.48}$	$0.30_{\pm 0.36}$	$0.36_{\pm 0.48}$	$0.47_{\pm 0.07}$
	结构	$0.81_{\pm 0.39}$	$0.27_{\pm 0.44}$	$0.49_{\pm 0.40}$	$0.27_{\pm 0.44}$	$0.51_{\pm 0.08}$

证据 消融研究

表13提供了消融研究的详细信息。

表13：消融研究

	Rule	回应	反驳	非重复	证据	立场
None	None	$0.90_{\pm 0.30}$	$0.09_{\pm 0.29}$	$0.79_{\pm 0.16}$	$0.00_{\pm 0.00}$	$0.36_{\pm 0.09}$
	轻量	$0.89_{\pm 0.31}$	$0.10_{\pm 0.30}$	$0.79_{\pm 0.14}$	$0.03_{\pm 0.16}$	$0.33_{\pm 0.08}$
	结构	$0.85_{\pm 0.35}$	$0.03_{\pm 0.17}$	$0.86_{\pm 0.11}$	$0.01_{\pm 0.08}$	$0.40_{\pm 0.08}$
仅视角	None	$0.88_{\pm 0.33}$	$0.34_{\pm 0.47}$	$0.74_{\pm 0.20}$	$0.00_{\pm 0.05}$	$0.49_{\pm 0.09}$
	轻量	$0.84_{\pm 0.37}$	$0.35_{\pm 0.48}$	$0.73_{\pm 0.23}$	$0.05_{\pm 0.21}$	$0.45_{\pm 0.09}$
	结构	$0.87_{\pm 0.33}$	$0.34_{\pm 0.47}$	$0.78_{\pm 0.18}$	$0.03_{\pm 0.17}$	$0.50_{\pm 0.08}$
仅证据	None	$0.87_{\pm 0.34}$	$0.24_{\pm 0.43}$	$0.80_{\pm 0.15}$	$0.23_{\pm 0.42}$	$0.46_{\pm 0.08}$
	轻量	$0.85_{\pm 0.35}$	$0.28_{\pm 0.45}$	$0.78_{\pm 0.14}$	$0.45_{\pm 0.50}$	$0.44_{\pm 0.08}$
	结构	$0.72_{\pm 0.45}$	$0.16_{\pm 0.36}$	$0.84_{\pm 0.14}$	$0.60_{\pm 0.49}$	$0.51_{\pm 0.06}$
仅材料	None	$0.89_{\pm 0.31}$	$0.00_{\pm 0.06}$	$0.42_{\pm 0.37}$	$0.00_{\pm 0.06}$	$0.38_{\pm 0.08}$
	轻量	$0.90_{\pm 0.30}$	$0.03_{\pm 0.17}$	$0.37_{\pm 0.36}$	$0.01_{\pm 0.12}$	$0.35_{\pm 0.08}$
	结构	$0.90_{\pm 0.30}$	$0.01_{\pm 0.10}$	$0.48_{\pm 0.38}$	$0.01_{\pm 0.08}$	$0.37_{\pm 0.09}$
M+D	None	$0.90_{\pm 0.31}$	$0.12_{\pm 0.33}$	$0.42_{\pm 0.40}$	$0.20_{\pm 0.40}$	$0.45_{\pm 0.09}$
	轻量	$0.90_{\pm 0.30}$	$0.10_{\pm 0.30}$	$0.35_{\pm 0.36}$	$0.37_{\pm 0.48}$	$0.43_{\pm 0.09}$
	结构	$0.87_{\pm 0.34}$	$0.15_{\pm 0.36}$	$0.48_{\pm 0.39}$	$0.25_{\pm 0.43}$	$0.46_{\pm 0.07}$
M+T	None	$0.89_{\pm 0.32}$	$0.23_{\pm 0.42}$	$0.36_{\pm 0.41}$	$0.02_{\pm 0.14}$	$0.48_{\pm 0.10}$
	轻量	$0.86_{\pm 0.35}$	$0.20_{\pm 0.40}$	$0.30_{\pm 0.35}$	$0.05_{\pm 0.22}$	$0.44_{\pm 0.10}$
	结构	$0.85_{\pm 0.36}$	$0.01_{\pm 0.12}$	$0.69_{\pm 0.31}$	$0.01_{\pm 0.09}$	$0.47_{\pm 0.08}$
D+T	None	$0.86_{\pm 0.34}$	$0.35_{\pm 0.48}$	$0.72_{\pm 0.20}$	$0.05_{\pm 0.21}$	$0.48_{\pm 0.07}$
	轻量	$0.85_{\pm 0.36}$	$0.36_{\pm 0.48}$	$0.72_{\pm 0.20}$	$0.30_{\pm 0.46}$	$0.46_{\pm 0.07}$
	结构	$0.71_{\pm 0.46}$	$0.36_{\pm 0.48}$	$0.83_{\pm 0.17}$	$0.44_{\pm 0.50}$	$0.51_{\pm 0.06}$

E骨干大语言模型

表14和表15是骨干大语言模型变体实验的结果详情。

表14：同质（S1）：所有骨干大语言模型均为 Qwen3 时各策略的整体表现。在 *Land* 场景中，报告所有智能体及所有轮次的均值±标准差。

查询	Rule	回应	反驳	非重复	证据	立场
Q1	None	$0.81_{\pm 0.40}$	$0.35_{\pm 0.48}$	$0.18_{\pm 0.35}$	$0.13_{\pm 0.33}$	$0.49_{\pm 0.09}$
	轻量	$0.68_{\pm 0.47}$	$0.28_{\pm 0.45}$	$0.23_{\pm 0.36}$	$0.28_{\pm 0.45}$	$0.44_{\pm 0.08}$
	结构	$0.67_{\pm 0.47}$	$0.04_{\pm 0.20}$	$0.37_{\pm 0.42}$	$0.09_{\pm 0.29}$	$0.47_{\pm 0.07}$
Q2	None	$0.68_{\pm 0.47}$	$0.25_{\pm 0.43}$	$0.21_{\pm 0.37}$	$0.33_{\pm 0.47}$	$0.47_{\pm 0.08}$
	轻量	$0.64_{\pm 0.48}$	$0.30_{\pm 0.46}$	$0.22_{\pm 0.34}$	$0.25_{\pm 0.43}$	$0.42_{\pm 0.06}$
	结构	$0.68_{\pm 0.47}$	$0.00_{\pm 0.00}$	$0.36_{\pm 0.42}$	$0.01_{\pm 0.12}$	$0.46_{\pm 0.07}$
Q3	None	$0.76_{\pm 0.43}$	$0.00_{\pm 0.00}$	$0.47_{\pm 0.42}$	$0.00_{\pm 0.00}$	$0.49_{\pm 0.08}$
	轻量	$0.77_{\pm 0.42}$	$0.01_{\pm 0.12}$	$0.42_{\pm 0.40}$	$0.03_{\pm 0.18}$	$0.46_{\pm 0.06}$
	结构	$0.85_{\pm 0.36}$	$0.03_{\pm 0.18}$	$0.52_{\pm 0.39}$	$0.22_{\pm 0.42}$	$0.46_{\pm 0.06}$
Q4	None	$0.53_{\pm 0.50}$	$0.60_{\pm 0.49}$	$0.16_{\pm 0.34}$	$0.15_{\pm 0.36}$	$0.50_{\pm 0.04}$
	轻量	$0.12_{\pm 0.33}$	$0.56_{\pm 0.50}$	$0.17_{\pm 0.34}$	$0.21_{\pm 0.41}$	$0.51_{\pm 0.07}$
	结构	$0.75_{\pm 0.43}$	$0.39_{\pm 0.49}$	$0.28_{\pm 0.40}$	$0.05_{\pm 0.23}$	$0.50_{\pm 0.04}$
Q5	None	$0.55_{\pm 0.50}$	$0.20_{\pm 0.40}$	$0.25_{\pm 0.38}$	$0.11_{\pm 0.32}$	$0.44_{\pm 0.05}$
	轻量	$0.75_{\pm 0.43}$	$0.14_{\pm 0.35}$	$0.26_{\pm 0.35}$	$0.01_{\pm 0.12}$	$0.39_{\pm 0.05}$
	结构	$0.73_{\pm 0.44}$	$0.14_{\pm 0.35}$	$0.38_{\pm 0.40}$	$0.09_{\pm 0.29}$	$0.45_{\pm 0.07}$
总体	None	$0.67_{\pm 0.47}$	$0.28_{\pm 0.45}$	$0.25_{\pm 0.39}$	$0.14_{\pm 0.35}$	$0.48_{\pm 0.07}$
	轻量	$0.59_{\pm 0.49}$	$0.26_{\pm 0.44}$	$0.26_{\pm 0.37}$	$0.16_{\pm 0.36}$	$0.44_{\pm 0.08}$
	结构	$0.74_{\pm 0.44}$	$0.12_{\pm 0.33}$	$0.38_{\pm 0.41}$	$0.10_{\pm 0.29}$	$0.47_{\pm 0.07}$

表15：异构：不同骨干网络变体下策略的整体表现。报告 *Land* 场景下所有智能体和所有轮次的均值±标准差。

查询	Rule	回应	反驳	非重复	证据	立场
Q1	None	$0.75_{\pm 0.44}$	$0.35_{\pm 0.48}$	$0.37_{\pm 0.40}$	$0.07_{\pm 0.25}$	$0.50_{\pm 0.10}$
	轻量	$0.81_{\pm 0.40}$	$0.31_{\pm 0.46}$	$0.37_{\pm 0.35}$	$0.49_{\pm 0.50}$	$0.44_{\pm 0.10}$
	结构	$0.79_{\pm 0.41}$	$0.12_{\pm 0.33}$	$0.61_{\pm 0.37}$	$0.23_{\pm 0.42}$	$0.55_{\pm 0.07}$
Q2	None	$0.70_{\pm 0.46}$	$0.31_{\pm 0.46}$	$0.36_{\pm 0.39}$	$0.07_{\pm 0.26}$	$0.46_{\pm 0.09}$
	轻量	$0.69_{\pm 0.47}$	$0.28_{\pm 0.45}$	$0.26_{\pm 0.32}$	$0.33_{\pm 0.47}$	$0.46_{\pm 0.08}$
	结构	$0.67_{\pm 0.47}$	$0.11_{\pm 0.31}$	$0.61_{\pm 0.34}$	$0.28_{\pm 0.45}$	$0.48_{\pm 0.09}$
Q3	None	$0.85_{\pm 0.35}$	$0.01_{\pm 0.08}$	$0.46_{\pm 0.40}$	$0.01_{\pm 0.08}$	$0.48_{\pm 0.08}$
	轻量	$0.85_{\pm 0.35}$	$0.00_{\pm 0.00}$	$0.37_{\pm 0.37}$	$0.24_{\pm 0.43}$	$0.49_{\pm 0.07}$
	结构	$0.72_{\pm 0.45}$	$0.05_{\pm 0.21}$	$0.73_{\pm 0.30}$	$0.25_{\pm 0.43}$	$0.47_{\pm 0.09}$
Q4	None	$0.57_{\pm 0.50}$	$0.47_{\pm 0.50}$	$0.40_{\pm 0.39}$	$0.10_{\pm 0.30}$	$0.51_{\pm 0.08}$
	轻量	$0.49_{\pm 0.50}$	$0.47_{\pm 0.50}$	$0.30_{\pm 0.36}$	$0.65_{\pm 0.48}$	$0.51_{\pm 0.07}$
	结构	$0.59_{\pm 0.49}$	$0.34_{\pm 0.47}$	$0.52_{\pm 0.38}$	$0.19_{\pm 0.40}$	$0.52_{\pm 0.09}$
Q5	None	$0.69_{\pm 0.47}$	$0.17_{\pm 0.38}$	$0.35_{\pm 0.39}$	$0.04_{\pm 0.20}$	$0.45_{\pm 0.08}$
	轻量	$0.77_{\pm 0.42}$	$0.43_{\pm 0.50}$	$0.36_{\pm 0.36}$	$0.40_{\pm 0.49}$	$0.40_{\pm 0.08}$
	结构	$0.71_{\pm 0.46}$	$0.19_{\pm 0.40}$	$0.58_{\pm 0.37}$	$0.16_{\pm 0.37}$	$0.47_{\pm 0.08}$
总体	None	$0.71_{\pm 0.45}$	$0.26_{\pm 0.44}$	$0.39_{\pm 0.40}$	$0.06_{\pm 0.23}$	$0.48_{\pm 0.09}$
	轻量	$0.72_{\pm 0.45}$	$0.30_{\pm 0.46}$	$0.33_{\pm 0.35}$	$0.42_{\pm 0.49}$	$0.46_{\pm 0.09}$
	结构	$0.70_{\pm 0.46}$	$0.16_{\pm 0.37}$	$0.61_{\pm 0.36}$	$0.22_{\pm 0.42}$	$0.50_{\pm 0.09}$

F 智能体级性能

为了进一步评估策略参数化框架在智能体数量更多的系统中的有效性，我们新增了两个分别由Llama3和Mistral驱动的智能体（Farmer 2和Farmer 3）。Farmer 1与Farmer 2被设计为具有相似的角色画像，而Farmer 3则被赋予不同的行为角色以引入多样性。三个农民代理共享一个知识库。表16中的结果表明，策略参数化提示仍然有效：每个智能体在不同规则下的各项指标仍表现出差异。表17展示了各智能体在不同骨干大语言模型下各项策略的整体表现。

表16：添加两个农民代理后各智能体策略的性能。

Rule	智能体	回应	反驳	非重复	证据	立场
None	Comm.	0.90 ± 0.00	0.08 ± 0.04	0.84 ± 0.03	0.28 ± 0.14	0.46 ± 0.05
	Conser.	0.87 ± 0.03	0.24 ± 0.17	0.71 ± 0.06	0.02 ± 0.03	0.49 ± 0.06
	Farmer 1（Qwen3）	0.88 ± 0.02	0.50 ± 0.33	0.52 ± 0.15	0.13 ± 0.13	0.59 ± 0.04
	Farmer 2（Llama）	0.87 ± 0.03	0.63 ± 0.21	0.59 ± 0.16	0.10 ± 0.10	0.68 ± 0.06
	Farmer 3（Mistral）	0.88 ± 0.02	0.26 ± 0.17	0.80 ± 0.08	0.33 ± 0.20	0.51 ± 0.02
轻量	Comm.	0.90 ± 0.01	0.09 ± 0.07	0.63 ± 0.03	0.21 ± 0.14	0.40 ± 0.03
	Conser.	0.88 ± 0.02	0.21 ± 0.21	0.71 ± 0.07	0.17 ± 0.04	0.44 ± 0.06
	Farmer 1(Qwen3)	0.86 ± 0.04	0.58 ± 0.28	0.39 ± 0.06	0.58 ± 0.17	0.55 ± 0.05
	Farmer 2(Llama)	0.87 ± 0.05	0.62 ± 0.25	0.49 ± 0.09	0.64 ± 0.17	0.62 ± 0.05
	Farmer 3(Mistral)	0.89 ± 0.02	0.17 ± 0.10	0.65 ± 0.04	0.54 ± 0.08	0.49 ± 0.03
结构	Comm.	0.90 ± 0.00	0.10 ± 0.04	0.83 ± 0.03	0.11 ± 0.05	0.42 ± 0.02
	Conser.	0.78 ± 0.03	0.19 ± 0.07	0.78 ± 0.02	0.26 ± 0.08	0.48 ± 0.01
	Farmer 1(Qwen3)	0.78 ± 0.06	0.23 ± 0.13	0.74 ± 0.04	0.33 ± 0.14	0.55 ± 0.03
	Farmer 2(Llama)	0.80 ± 0.04	0.25 ± 0.16	0.75 ± 0.05	0.40 ± 0.10	0.61 ± 0.03
	Farmer 3(Mistral)	0.89 ± 0.02	0.27 ± 0.17	0.80 ± 0.05	0.34 ± 0.10	0.48 ± 0.02

G自适应权重

表18展示了自适应权重有效性的实验详情。

表19：EDU 场景上的消融研究

查询	Rule	回应	Reb.	非重复	证据	立场
M+D	None	0.89 ± 0.32	0.02 ± 0.15	0.42 ± 0.39	0.06 ± 0.23	0.46 ± 0.08
	轻量	0.89 ± 0.32	0.01 ± 0.12	0.39 ± 0.36	0.14 ± 0.35	0.45 ± 0.09
	结构	0.89 ± 0.31	0.02 ± 0.15	0.46 ± 0.38	0.05 ± 0.22	0.45 ± 0.08
T+D	None	0.85 ± 0.35	0.46 ± 0.50	0.74 ± 0.24	0.13 ± 0.34	0.50 ± 0.08
	轻量	0.85 ± 0.35	0.34 ± 0.47	0.69 ± 0.25	0.55 ± 0.50	0.48 ± 0.09
	结构	0.80 ± 0.40	0.30 ± 0.46	0.80 ± 0.17	0.31 ± 0.46	0.53 ± 0.08
T+M	None	0.86 ± 0.35	0.17 ± 0.38	0.45 ± 0.39	0.09 ± 0.29	0.46 ± 0.10
	轻量	0.89 ± 0.31	0.10 ± 0.30	0.37 ± 0.35	0.08 ± 0.27	0.44 ± 0.09
	结构	0.88 ± 0.33	0.04 ± 0.20	0.65 ± 0.30	0.01 ± 0.11	0.46 ± 0.08

表17：不同骨干大语言模型下各智能体策略的整体表现

Rule	智能体	回应	反驳	非重复	证据	立场
None	Comm.（Mistral）	0.48 ± 0.20	0.02 ± 0.04	0.22 ± 0.10	0.14 ± 0.13	0.41 ± 0.03
	Conser.（Llama）	0.72 ± 0.28	0.23 ± 0.34	0.30 ± 0.19	0.26 ± 0.26	0.48 ± 0.04
	农民（通义千问）	0.80 ± 0.12	0.58 ± 0.35	0.24 ± 0.06	0.04 ± 0.08	0.54 ± 0.03
轻量	Comm.（Mistral）	0.40 ± 0.16	0.04 ± 0.07	0.22 ± 0.05	0.23 ± 0.21	0.39 ± 0.04
	Conser.(Llama)	0.65 ± 0.31	0.21 ± 0.29	0.31 ± 0.13	0.13 ± 0.14	0.44 ± 0.04
	农民(通义千问)	0.73 ± 0.29	0.53 ± 0.38	0.26 ± 0.08	0.11 ± 0.12	0.51 ± 0.05
结构	Comm.(mistral)	0.57 ± 0.12	0.03 ± 0.06	0.37 ± 0.07	0.17 ± 0.24	0.41 ± 0.04
	Conser.(Llama)	0.78 ± 0.07	0.20 ± 0.28	0.43 ± 0.10	0.11 ± 0.09	0.48 ± 0.02
	农民(通义千问)	0.86 ± 0.02	0.13 ± 0.15	0.35 ± 0.09	0.00 ± 0.01	0.51 ± 0.03
None	Comm.(通义千问)	0.52 ± 0.19	0.01 ± 0.02	0.25 ± 0.14	0.03 ± 0.06	0.42 ± 0.03
	Conser.(Mistral)	0.83 ± 0.14	0.19 ± 0.20	0.53 ± 0.06	0.07 ± 0.08	0.45 ± 0.03
	农民(Llama)	0.79 ± 0.17	0.58 ± 0.36	0.38 ± 0.08	0.08 ± 0.11	0.57 ± 0.03
轻量	Comm.(通义千问)	0.59 ± 0.18	0.02 ± 0.05	0.22 ± 0.09	0.31 ± 0.26	0.41 ± 0.04
	Conser.(Mistral)	0.87 ± 0.05	0.20 ± 0.24	0.45 ± 0.05	0.21 ± 0.14	0.42 ± 0.04
	农民(Llama)	0.71 ± 0.30	0.66 ± 0.34	0.33 ± 0.06	0.74 ± 0.06	0.55 ± 0.04
结构	Comm.(通义千问)	0.61 ± 0.06	0.06 ± 0.07	0.45 ± 0.09	0.24 ± 0.12	0.44 ± 0.04
	Conser.(Mistral)	0.86 ± 0.04	0.22 ± 0.20	0.74 ± 0.10	0.01 ± 0.02	0.48 ± 0.03
	农民(Llama)	0.62 ± 0.13	0.21 ± 0.11	0.64 ± 0.06	0.41 ± 0.08	0.58 ± 0.02
None	Comm.(Llama)	0.75 ± 0.12	0.03 ± 0.05	0.38 ± 0.04	0.13 ± 0.08	0.44 ± 0.03
	Conser.(通义千问)	0.76 ± 0.13	0.34 ± 0.35	0.33 ± 0.19	0.10 ± 0.12	0.46 ± 0.06
	农民 (Mistral)	0.85 ± 0.07	0.31 ± 0.28	0.65 ± 0.08	0.12 ± 0.08	0.52 ± 0.03
轻量	Comm.(Llama)	0.89 ± 0.01	0.05 ± 0.09	0.32 ± 0.07	0.25 ± 0.18	0.42 ± 0.02
	Conser.(通义千问)	0.76 ± 0.26	0.37 ± 0.37	0.28 ± 0.12	0.04 ± 0.04	0.44 ± 0.05
	农民 (Mistral)	0.90 ± 0.01	0.30 ± 0.26	0.46 ± 0.13	0.42 ± 0.23	0.51 ± 0.03
结构	Comm.（Llama）	0.88 ± 0.02	0.00 ± 0.01	0.63 ± 0.04	0.35 ± 0.20	0.47 ± 0.02
	Conser.（通义千问）	0.78 ± 0.03	0.12 ± 0.09	0.59 ± 0.05	0.08 ± 0.10	0.49 ± 0.01
	农民（Mistral）	0.87 ± 0.05	0.12 ± 0.08	0.78 ± 0.06	0.29 ± 0.11	0.53 ± 0.01

表18：自适应权重设置下的整体表现。报告所有智能体和所有轮次在 *Land*场景中的均值±标准差。

查询	Rule	回应	反驳	非重复	证据	立场
Q1	None	$0.88_{\pm 0.33}$	$0.21_{\pm 0.41}$	$0.50_{\pm 0.38}$	$0.23_{\pm 0.42}$	$0.53_{\pm 0.07}$
	轻量	$0.89_{\pm 0.32}$	$0.51_{\pm 0.50}$	$0.36_{\pm 0.35}$	$0.47_{\pm 0.50}$	$0.50_{\pm 0.07}$
	结构	$0.83_{\pm 0.38}$	$0.17_{\pm 0.38}$	$0.54_{\pm 0.40}$	$0.15_{\pm 0.36}$	$0.51_{\pm 0.07}$
Q2	None	$0.87_{\pm 0.33}$	$0.38_{\pm 0.49}$	$0.49_{\pm 0.39}$	$0.19_{\pm 0.40}$	$0.48_{\pm 0.07}$
	轻量	$0.89_{\pm 0.31}$	$0.39_{\pm 0.49}$	$0.39_{\pm 0.33}$	$0.17_{\pm 0.38}$	$0.44_{\pm 0.05}$
	结构	$0.81_{\pm 0.40}$	$0.11_{\pm 0.32}$	$0.49_{\pm 0.39}$	$0.17_{\pm 0.38}$	$0.47_{\pm 0.07}$
Q3	None	$0.90_{\pm 0.30}$	$0.00_{\pm 0.00}$	$0.39_{\pm 0.39}$	$0.12_{\pm 0.33}$	$0.45_{\pm 0.06}$
	轻量	$0.90_{\pm 0.30}$	$0.07_{\pm 0.26}$	$0.43_{\pm 0.36}$	$0.27_{\pm 0.44}$	$0.44_{\pm 0.06}$
	结构	$0.89_{\pm 0.31}$	$0.02_{\pm 0.14}$	$0.54_{\pm 0.40}$	$0.10_{\pm 0.30}$	$0.45_{\pm 0.06}$
Q4	None	$0.79_{\pm 0.41}$	$0.41_{\pm 0.49}$	$0.35_{\pm 0.37}$	$0.17_{\pm 0.38}$	$0.49_{\pm 0.05}$
	轻量	$0.70_{\pm 0.46}$	$0.55_{\pm 0.50}$	$0.33_{\pm 0.35}$	$0.42_{\pm 0.49}$	$0.49_{\pm 0.08}$
	结构	$0.79_{\pm 0.41}$	$0.18_{\pm 0.39}$	$0.47_{\pm 0.40}$	$0.24_{\pm 0.43}$	$0.51_{\pm 0.07}$
Q5	None	$0.86_{\pm 0.35}$	$0.31_{\pm 0.47}$	$0.51_{\pm 0.38}$	$0.17_{\pm 0.37}$	$0.46_{\pm 0.06}$
	轻量	$0.87_{\pm 0.34}$	$0.45_{\pm 0.50}$	$0.35_{\pm 0.34}$	$0.13_{\pm 0.34}$	$0.41_{\pm 0.05}$
	结构	$0.79_{\pm 0.41}$	$0.18_{\pm 0.39}$	$0.55_{\pm 0.40}$	$0.15_{\pm 0.35}$	$0.47_{\pm 0.06}$
总体	None	$0.86_{\pm 0.35}$	$0.26_{\pm 0.44}$	$0.45_{\pm 0.39}$	$0.18_{\pm 0.38}$	$0.48_{\pm 0.07}$
	轻量	$0.85_{\pm 0.36}$	$0.39_{\pm 0.49}$	$0.37_{\pm 0.35}$	$0.29_{\pm 0.46}$	$0.45_{\pm 0.07}$
	结构	$0.82_{\pm 0.38}$	$0.13_{\pm 0.34}$	$0.52_{\pm 0.40}$	$0.16_{\pm 0.37}$	$0.48_{\pm 0.07}$

H 在 EDU 上的消融研究

表19展示了 EDU 场景上的消融研究结果。