arXiv:2603.01505v1 [cs.RO] 2 Mar 2026

FATE：面向物理基础机器人课程的闭环可行性感知任务生成与主动修复

Bingchuan Wei [*1] **Bingqi Huang** [*1] **Jingheng Ma** [2] **Zeyu Zhang** [3] **Sen Cui** [4]

摘要

近期，生成式仿真领域的突破利用大语言模型（LLM）生成多样化的机器人任务课程，但这些开环范式往往会产生语言上连贯、物理上却不可行的目标，这源于任务规范缺乏物理基础或目标表述不匹配。为解决这一关键局限，我们提出 FATE（可行性感知任务生成）——一种闭环、自纠正框架，将任务生成重新构想为迭代式验证与改进过程。与将生成和验证拆分为离散阶段的传统方法不同，FATE 将通用具身智能体直接嵌入生成循环，主动确保生成课程的物理接地性。FATE 实现了一套顺序审计流程：首先验证静态场景属性（例如物体可供性、布局兼容性），随后通过仿真中的具身交互验证执行可行性。FATE 的性能关键在于，当检测到不可行任务时，会部署主动修复模块，自主调整场景配置或策略规范，将不可行的方案转化为物理上有效的任务实例。大量实验验证了 FATE 能生成语义多样且具有物理基础的任务课程，同时与最先进的生成式基线方法相比，大幅降低了执行失败率。

[1]清华大学航空航天工程学院，中国北京 [2]清华大学集成电路学院，中国北京 [3]北京通用人工智能研究院（BIGAI）通用人工智能国家重点实验室，中国北京 [4]清华大学自动化系，中国北京。通讯作者：Sen Cui< cuis@mail.tsinghua.edu.cn>。

预印本。*2026*年*3*月*3*日。

1. 引言

具身人工智能的范式正从静态环境转向开放式、程序化生成的世界（Wang 等人，2023a；b）。基于大语言模型的开环生成器能够提出多样化任务提案，但其中许多任务因静态不一致（例如指令与场景不匹配、资产可供性违规或前置条件未满足）而不可行。即使初始设置看似合理，任务仍可能在运行时因机器人可达性限制、接触不稳定性或规划／学习故障而失败。因此，不受约束的开环机制会生成无效数据并浪费计算资源。

基于上述观察，我们认为，任务规范合理性（即任务设定与现实世界常识的一致性）和任务目标的可执行性（即通过机器人交互实现目标的实际可行性）必须充分纳入闭环任务生成流程。这样既能确保生成的任务在语言上连贯，也能使其符合物理接地要求，从而解决开环生成的根本局限。

为克服这一挑战，我们提出将可行性视为一个涵盖静态感知和运行时执行的分层约束满足问题。具体而言，视觉合理性并不意味着可解性；一个连贯的场景仍可能因目标不可达或不易察觉的碰撞而无法完成。因此，我们针对图1所示的两种不同失效模式进行处理：**静态幻觉**（几何／语义违规）和**动态偏差**（运行时运动学或交互故障）。

我们提出**FATE** （可行性感知任务生成），这是一种用于分层可行性对齐的人在回路框架，采用解耦策略：首先将方案投射到静态有效流形上，然后通过执行中反馈优化任务。具体而言，首先由**静态对齐**操作器纠正感知违规（例如，物体相互穿透）。随后，**动态对齐**审查器监控执行过程，并迭代调整求解器参数，以解决运行时故障。由此，任务生成从随机采样转变为闭环课程。

我们利用生成式模拟器和通用型

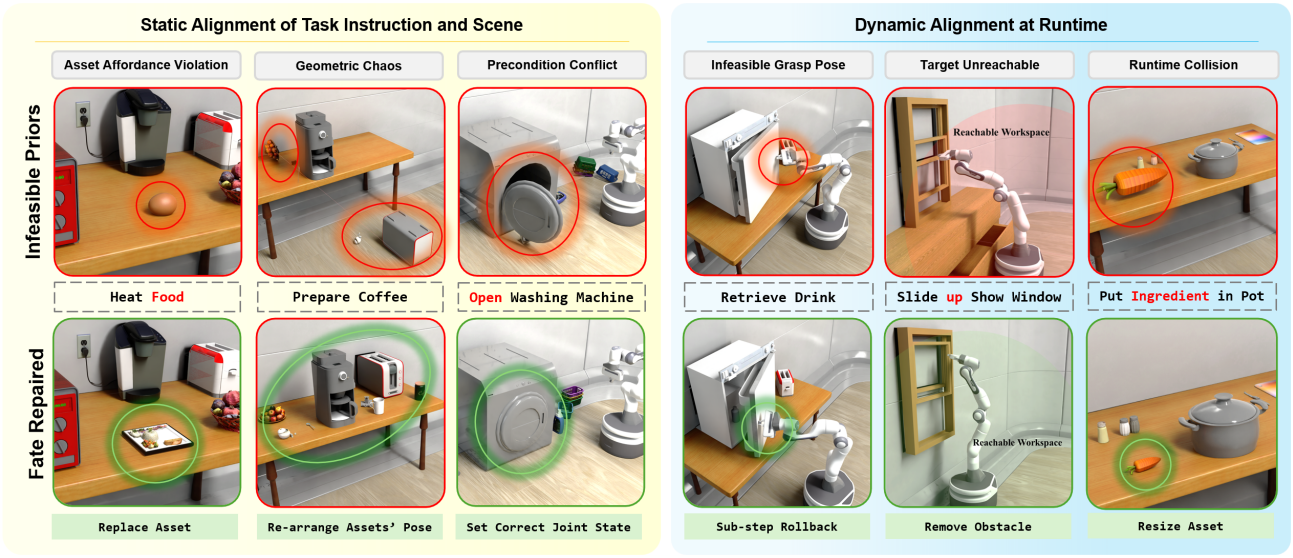


图*1*。 **FATE 从静态和动态两个维度实现任务可行性对齐。（左）静态对齐：** 系统通过解决语义违规（例如替换生鸡蛋）、几何混乱和前置条件冲突，修正初始场景先验。**（右）动态对齐：** FATE 通过迭代回滚机制处理运行时执行失败，例如高扭矩抓取、不可达目标和碰撞。**红色**标示初始不可行之处；**绿色** 标示经 FATE 修复后可执行的环境。

具身基础模型构建 FATE，并在一个旨在对可行性感知生成进行压力测试的新基准上评估该框架。我们从三个维度验证 FATE：**生成多样性、可行性感知**和**精化效能**。实验表明，与专家设计的基准和开环生成器相比，FATE 保持了较强的语言和视觉多样性，同时通过分层审查与修复显著缩小了端到端可行性差距。我们还报告了审计器的诊断准确率，并量化了各阶段通过修复挽救不可行提案的频率。本研究的贡献主要有三点：

1. 我们提出一种新的问题设定——**可行性感知开放式任务生成**，主动考虑由不合理的任务设定和不可行的任务目标共同造成的可行性差距。2. 我们提出一种人在回路框架**FATE**，将双阶段审计与修复相结合，并通过迭代优化过程主动诊断失效模式，以确保实际可行性。3. 实验表明，**FATE**能够有效识别并纠正不可行任务，减少流程各阶段的仿真浪费，同时显著提升下游策略学习的性能。

2. 相关工作

利用大语言模型程序化生成仿真环境的方式已取得显著发展。早期系统如 **GenSim** （Wang 等人，2023a）和 **RoboGen**（Wang 等人，2023b）展示了根据文本生成仿真代码的潜力。近期研究进一步拓展了这一领域：**GenSim2** （University，2025b）通过整合多模态大语言模型（例如 GPT-4V ），提升场景理解与推理能力；**AgentGen** （字节跳动，2025）提出双向进化策略，以生成难度梯度平滑的任务；**ReGen**（Laboratory，2025）提出“逆向设计”范式，从期望的智能体行为推断仿真环境；而 **Factorsim** （University，2024）则采用因子化部分可观测马尔可夫决策过程表示来生成模块化仿真代码。

尽管在结构连贯性和难度扩展方面已取得这些进展，这些方法在很大程度上仍以“前馈”方式运行。它们缺乏一个主动的可行性审计器 来检测生成的约束（例如，目标位置、物体质量）何时违反机器人的物理极限。虽然**DrEureka**（Ma 等人，2024）实现了面向仿真到现实迁移的奖励调优和域随机化自动化，但它假定任务本身是有效的。FATE 通过引入自适应修复模块来解决这一局限，该模块会主动修改任务定义，以确保在训练之前具备物理可执行性。

另一条研究路线着眼于扩增稀缺的人类演示。**MimicGen**（英伟达，2024）及其近期扩展版本 **DemoGen** （University，2025a）和**SkillMimicGen**（Garrett 等，2024）利用数据增强（例如刚体变换、技能分割）

将有限的源演示扩展为大型数据集。**DexMimicGen**（Jiang 等，2025）进一步将这一方法扩展至双手灵巧操作。尽管这些方法效果显著，但它们依赖演示——只能生成由人类演示过的任务的变体。相比之下，FATE无需演示，且完全开放。它旨在发现全新的任务语义（例如“用烤面包机加热马克杯”——人类可能从未演示过的任务），通过从头合成任务、环境和解决策略来实现这一目标，并将多模态大语言模型审计器用作可扩展的人类验证代理。

视频生成的兴起催生了使用合成视频作为训练监督信号的方法。**DreamGen**（英伟达，2025）对视频世界模型进行微调，以生成未来帧用于机器人学习；**Gen2Real**（香港中文大学深圳，2025）和**LuciBot**（阿默斯特，2025）直接从生成视频中提取策略；**IRASim**（Zhu 等人，2025a）则通过扩散 Transformer 实现细粒度交互建模。然而，生成式视频模型容易产生“物理幻觉”——生成视觉上看似合理、但物理上不可能的状态转变（Knightnemo，2025）。FATE 不将这些生成模型视为真值，而是将其视为假设。通过让具身大脑（审查器）参与闭环，FATE 验证生成目标的物理一致性（类似于**SEEA-R1**（卡内基梅隆大学，2025）和**MindJourney**（UMass-Embodied-AGI，2025）中的推理），并在目标违反物理定律时加以修复，确保“想象”的课程在语义上有意义、具有物理基础且可学习。

3. 问题形式化

本节首先介绍必要的符号，最后给出总体问题的正式定义。

3.1. 可行性感知任务

我们认为，机器人任务不仅仅是一条语义指令，而是由语言意图、物理环境和受约束解空间构成的耦合系统。

定义 3.1（可行性感知任务）。 任务实例定义为一个元组 $\tau = (\mathcal{I}, \mathcal{S}, \Omega_\pi)$，其中 $\mathcal{I}$表示语义-时间规范， $\mathcal{S}$ 表示物理场景配置，Ω_π 表示受约束执行策略流形。

指令 $\mathcal{I}$ 作为语义-时间规范。 </style}指令 被形式化为一对 (， $\Psi_\mathcal{I}$)。其中， $\mathcal{L} \in \mathcal{V}^*$表示提供高层意图的自然语言词元序列。关键在于， $\mathcal{L}$ 映射到一个隐式的、通常不可微的语义谓词泛函 $\Psi_\mathcal{I}$： $\Xi \to \{0,1\}$，其中 Ξ 是状态-动作轨迹的空间。当且仅当一条轨迹 $\xi \in \Xi$ 满足以下条件时，$\Psi_\mathcal{I}(\xi) = 1$。这封装了任务的时序逻辑约束，而无需假设特定的分解方式。

场景 $\mathcal{S}$ 作为语义化关节配置。 场景 $\mathcal{S}$ 定义为一组语义化关节实体 $\mathcal{S} = \{(M_j, q_j, \Lambda_j)\}_{j=1}^{N}$。每个实体由其物理流形和惯性属性 M_j、其运动学配置状态 $q_j \in \mathcal{Q}$（涵盖李群 $SE(3)$ 和关节空间 $\mathbb{R}^{d_j}$），以及语义可供性映射Λ_j： $\mathcal{M}_j \to \mathbb{L}_{sem}$组成。映射 Λ_j 将几何特征与语义标签（例如可抓取性、关节类型）关联起来，从而提供必要的物理支撑。

流形 Ω_π 作为受约束策略。 与探索通用参数空间 Θ的标准强化学习不同，生成式仿真对解的搜索施加结构性约束。我们将$\Omega_\pi \subset \Theta$ 定义为由生成器的代码规范所诱导的紧致子流形（例如，特定的奖励结构、原语选择或规划时域）。Ω_π 有效地划定了智能体学习算法的允许搜索空间 。通过将Ω_π 纳入任务定义，我们将所生成的实例不仅表述为一个待求解的问题，而且表述为一个自包含的学习单元。该表述认为，一个有效的开放式任务必须内在地编码其自身的可解性边界。

3.2. 可行策略与任务可行性

具身任务中的有效性并非二元的，而是测度论意义上的。只有当解集在智能体的探索能力下具有不可忽略的体积时，一个任务才具有实际可行性。

定义 3.2（可行策略集）。 对于给定任务 τ，可行策略集 $\Phi_\tau \subset \Omega_\pi$ 是在由 $\mathcal{S}$诱导的场景动力学下以高概率满足语义谓词 $\Psi_\mathcal{I}$ 的策略子集：

$$\Phi_\tau = \left\{ \theta \in \Omega_\pi \mid \mathbb{E}_{\xi \sim (\pi_\theta, \mathcal{S})}[\Psi_\mathcal{I}(\xi)] \geq 1 - \epsilon \right\} \quad (1)$$

其中， ϵ 是执行随机性的容差参数， π_θ 是由 θ参数化的策略。

为量化寻找此类策略的可行性，我们将策略搜索过程建模为从初始分布中采样。设 ρ_{init} 为参数流形Ω_π上的概率测度。该测度表示优化算法可访问的静态探索范围。

定义 3.3（任务可行性度量）。 可行性$\mu(\tau)$定义为在ρ_{init}下可行集的度量：

$$\mu(\tau) = \int_{\Omega_\pi} \mathbb{I}(\theta \in \Phi_\tau) \, d\rho_{init}(\theta) \quad (2)$$

含义。如果$\mu(\tau) > \delta_{min}$，则认为任务具有物理基础。相反，如果 $\mu(\tau) \approx 0$，则该任务存在可行性差距。在实践中，这表现为执行失败：尽管经过穷尽式训练，机器人强化学习智能体仍然从未收敛或很难收敛，其原因可能是物理上不可行（$\Phi_\tau = \emptyset$），也可能是生成的策略流形 Ω_π 未能与解空间相交。这会导致生成的任务不可行。

3.3. 问题陈述：可行性感知任务生成

我们将生成过程建模为从任务生成器 $\mathcal{G}$ 在任务空间上生成的任务分布 $P_\mathcal{G}$ 中进行采样。核心挑战在于，高多样性生成器 $P_\mathcal{G}$ 的支持集由未经验证的任务提案构成，通常远远超出物理基础区域，而在该区域内 $\mu(\tau) > \delta_{min}$。

定义 3.4（可行性差距）。可行性差距 定义为生成任务中可行性得分未达到最低阈值的概率质量。对于任意细化映射 $\mathcal{A}: \tau \mapsto \tau$，我们定义

$$L_\mathcal{G}(\mathcal{A}) \triangleq \mathbb{E}_{\tau\sim P_\mathcal{G}}\left[\mathbb{I}(\mu(\mathcal{A}(\tau)) \le \delta_{min})\right]. \tag{3}$$

具体而言，预对齐 可行性差距对应于恒等映射 $\mathcal{A} = \mathrm{Id}$。

该定义提供了可执行性的操作性概念：在诱导的策略流形 Ω_π 和初始化度量 ρ_{init}下，只有当任务的可行策略集具有不可忽略的测度（即 $\mu(\tau) > \delta_{min}$）时，该任务才可执行。

FATE 的目标。因此，**FATE** 的目标并非仅仅从 $P_\mathcal{G}$中采样，而是构建一种精炼机制，在弥合这一差距的同时，不削弱生成多样性，也不偏离用户意图。形式上，我们寻求一个映射 $\mathcal{A}: \tau \to \tau$，将任务假设投影到可行区域。最优操作器是以下约束最小化问题的解

$$\min_{\mathcal{A}} \quad L_\mathcal{G}(\mathcal{A}) \tag{4a}$$

$$\text{s.t.} \quad \mathbb{E}_{\tau\sim P_\mathcal{G}}[D(\tau, \mathcal{A}(\tau))] \le \epsilon_{sem}, \tag{4b}$$

其中，式 (4a) 最小化可行性差距（定义 3.4）。$D: \mathcal{T}\times\mathcal{T} \to \mathbb{R}_{\ge 0}$ 表示两个任务实例之间的语义距离度量，其值越小，表示语义一致性越高。因此，式 (4b) 强制要求投影后的任务 $\mathcal{A}(\tau)$ 在容差 ϵ_{sem}范围内与原始意图保持语义一致。

4. 方法论

本节将介绍公式(4)中定义的优化目标的构造性实现。

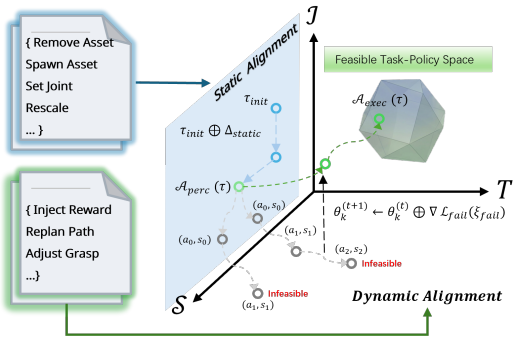


图 2. 分层可行性对齐过程的可视化。优化轨迹在联合任务空间 $\tau = (\mathcal{I}, \mathcal{S}, \Pi)$ 中演化。**（左）静态对齐（$\mathcal{A}_{perc}$）：**前置审计器通过离散修复（例如 RESCALE）纠正**场景**（$\mathcal{S}$）与**指令**（$\mathcal{I}$）之间的不兼容问题，将 τ_{init} 投影到感知有效的流形上。**（右）动态对齐（$\mathcal{A}_{exec}$）：**针对**策略**（Π），系统利用语义反馈和动态 API（例如 INJECT REWARD）优化求解器配置，引导任务从失败状态（红色轨迹）转向**可行空间**（绿色多面体）。

4.1. 可行性对齐算子

为弥合可行性差距，我们正式构造了一个映射 $\mathcal{A}: \mathcal{T}\to\mathcal{T}$，称为<对齐 id='3'>可行性对齐操作器</对齐>。不同于随机重采样，该操作器充当构造性流形投影，将一个无物理依据的假设 $\tau \sim P_\mathcal{G}$ 转换为物理有效的实例τ^*，同时最小化语义偏差：

$$\tau^* = \mathcal{A}(\tau) \triangleq \operatorname*{arg\,min}_{\mu(\tau')>\delta_{min}} D(\tau', \tau). \tag{5}$$

鉴于物理约束的层次性，我们将 $\mathcal{A}$ 结构性地分解为感知（$\mathcal{A}_{perc}$）和执行（$\mathcal{A}_{exec}$）子操作器的顺序组合。这确保任务首先映射到几何有效的子空间，随后再投影到动态交集上。

$$\mathcal{A} = \mathcal{A}_{exec} \circ \mathcal{A}_{perc} \tag{6}$$

我们将对齐操作器 $\mathcal{A}$ 表述为一种**构造性残差注入**。我们的目标并非进行随机重采样，而是合成一个任务残差 $\Delta\tau$，将当前任务 τ 投影到由度量 $\mu(\tau)$ 定义的可行区域中（定义 3.3）。

形式上，在每次迭代中，该操作器都会寻求一个最优更新 $\Delta\tau^*$，以最大化可行性似然：

$$\Delta\tau^* = \operatorname*{arg\,max}_{\Delta\tau\in\mathcal{D}} \mu(\tau \oplus \Delta\tau) \tag{7}$$

其中，$\oplus$ 表示残差的符号化应用，$\mathcal{D}$ 表示允许的编辑空间。由于计算

精确的 $\mu(\tau)$ 难以实现，因此我们使用特定阶段的替代指标来实现这一目标：在第一阶段（第 4.3 节）最大化**静态有效性分数** $\hat{}\mu_{static}$，并在第二阶段（第 4.4 节）最小化**执行失败率**。

4.2. 通过执行原语进行策略规范

为实例化受约束策略流形 Ω_π，FATE 将全局指令 $\mathcal{I}$ 分解为一系列原子原语 $\Pi = \{\pi_1, \ldots, \pi_K\}$。我们采用混合求解器策略，以满足导航阶段和操作阶段各自不同的拓扑要求：

几何感知导航（模型预测控制）。对于非接触阶段（例如，移动到目标位置），我们采用模型预测控制。求解器会优化轨迹，在严格遵守根据场景 $\mathcal{S}$ 推导出的碰撞约束的同时，尽可能缩短与子目标之间的距离。该求解器提供几何参数（例如障碍物膨胀半径），用于确定运动的保守程度。

高接触操作（强化学习）。对于交互阶段（例如抓取或打开物体），我们采用软演员-评论家（SAC）。为了促进智能体在稀疏奖励环境中的学习，多模态大语言模型会合成一个由权重 θ_{rew} 参数化的基于势函数的塑形函数。这样，系统便能通过塑造优化地形，引导智能体到达语义关键点。

审计器接口。至关重要的是，这些原语并非静态黑箱，而是参数化模块 $\pi(\cdot \mid \theta)$。它们开放了可调的配置空间（包括模型预测控制约束和强化学习奖励权重），作为**动态对齐算子**($\mathcal{A}_{exec}$)的控制接口。此设计使审计器能够主动调整求解器的逻辑来解决执行失败——实质上执行语义策略梯度更新——而不是仅仅重新采样任务。

4.3. 阶段 I：通过感知投影进行静态对齐

第一阶段 $\mathcal{A}_{perc}$ 充当感知守门器，将初始任务假设投影到感知有效区域。我们提出，静态一致性——即语义指令 $\mathcal{I}$ 与场景配置 $\mathcal{S}$ 之间的对齐——是动态可执行性的严格必要条件。

审计：检测幻觉。为在不进行高成本仿真的情况下估算可行性度量 $\mu(\tau)$（定义 3.3），前置审计器采用静态可行性检查 $\hat{}\mu_{static}$。它充当逻辑过滤器，用于验证四项一致性条件：(i) 可达性；(ii) 可供性匹配；(iii) 物理合理性；以及 (iv) 形态兼容性。任何一项条件未通过，都意味着该任务在根本上

算法 1 FATE：可行性感知任务生成 1：输入：随机种子 s，资产库 $\mathcal{L}$2：输出：可行任务 $\tau = (\mathcal{I}, \mathcal{S}, \Omega_\pi)$ 3：# 1. 初始化：构建 τ_{init}4：　$\mathcal{I} \leftarrow$采样指令($\mathcal{L}$) 5：　$\mathcal{S} \leftarrow$ 初始化场景（默认姿态，原始比例） 6：$\Omega_\pi \leftarrow$ 初始化策略规范（原语 $\Pi = \{\pi_1, \ldots, \pi_K\}$） 7：　$\tau \leftarrow (\mathcal{I}, \mathcal{S}, \Omega_\pi)$ { 初始可行性：未知 }8：# 2. 主循环：交错对齐 9：对于 $k = 1, \ldots,$ MaxIter 执行 10：# 阶段 A：静态对齐（以场景 $\mathcal{S}$为目标） 11：$valid, info \leftarrow$ AnteAuditor($\mathcal{I}, \mathcal{S}$) 12：如果非 $valid$则 13：{修复几何与语义违规}14：$\mathcal{S} \leftarrow$ 静态修复($\mathcal{S}, info$) 15：$\tau \leftarrow (\mathcal{I}, \mathcal{S}, \Omega_\pi)$ 16：继续 17：结束如果 18：# 阶段 B：动态对齐（以策略 Ω_π为目标） 19：$trace \leftarrow$ 运行仿真(τ) 20：$success, info \leftarrow$ 步内审计器 ($trace$) 21：如果非 $success$ 则 22：{优化求解器参数与原语配置}23：$\Omega_\pi \leftarrow$ 动态修复($\Omega_\pi, info$) 24：$\tau \leftarrow (\mathcal{I}, \mathcal{S}, \Omega_\pi)$ 25：继续 26：结束如果 27：# 可行性已确认 28：返回 τ29：结束循环 30：返回失败

在物理上不可能（$\mu(\tau) \approx 0$），从而立即触发修复过程。

修复：场景投影。一旦检测到违规，操作器便会合成静态残差 $\Delta\tau_{static}$，以修正场景或指令。我们将最优静态修复定义为使感知有效性代理指标 $\hat{}\mu_{static}$ 最大化的修复：

$$\Delta\tau^*_{static} = \underset{\Delta\tau \in \mathcal{D}_{static}}{\arg\max} \ \hat{\mu}_{static}(\tau \oplus \Delta\tau) \qquad (8)$$

更新 $\tau \leftarrow \tau \oplus \Delta\tau^*_{static}$ 对应于向几何有效流形的离散投影。实际上，$\Delta\tau_{static}$ 由两部分组成：**场景重构**（例如，调整资源 $M_j \to M'_j$ 的尺寸或移动无法触及的物体）和**指令细化**（例如，放宽过于具体的动词）。通过求解 (8)，$\mathcal{A}_{perc}$ 可确保提案在仿真之前具有物理合理性。

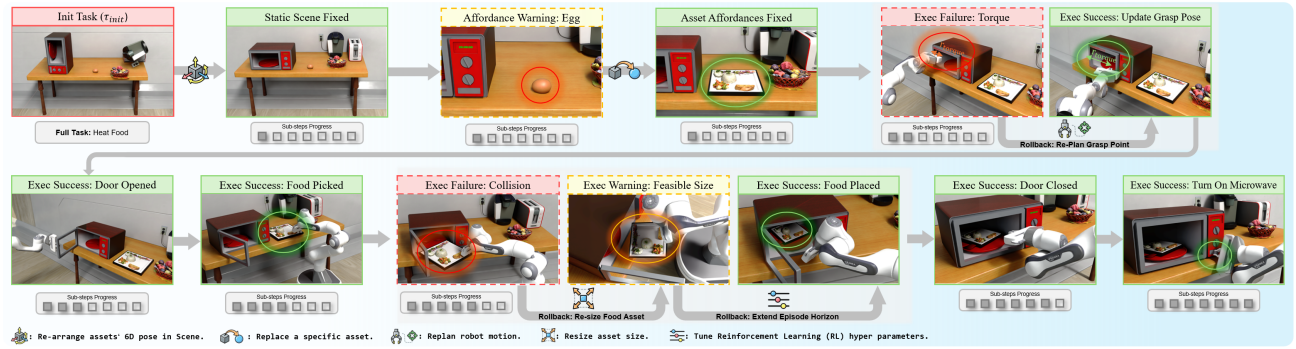


图 *3.* **"加热食物"任务的演化轨迹**。 从混乱的初始化状态（τ_{init}）开始，FATE 通过**静态修复** 重新排列物体并纠正其功能属性（用一份餐食替换一个鸡蛋）。在执行过程中，系统会自主检测并修复动态故障：通过**运动重新规划**解决抓取失败，并通过**资产缩放**修复插入碰撞。灰色弧形箭头展示了我们的**子步骤回滚** 机制，该机制会调用特定的修复 API（底部图标），迭代优化任务，直至生成成功的策略。

4.4. 第二阶段：通过语义策略梯度实现动态对齐

尽管静态对齐可确保几何有效性，但无法预见复杂接触动力学或控制器不稳定性所导致的故障。第二阶段 $\mathcal{A}_{exec}$通过确保策略规范 Π在运行时收敛，来弥合可行性差距 。

审计：监测偏差。 步内审计器会监测原语序列 $\{\pi_k\}$的执行情况。如果智能体未能在分配的时限内实现预期的状态变化，或者求解器（模型预测控制/强化学习）无法找到有效解，则会将任务标记为**动态偏差**。常见的失效模式包括高扭矩抓取失败、碰撞诱发的死锁、稀疏奖励停滞，或潜在的几何不兼容（例如插入公差过紧），这些问题只有在物理交互过程中才会显现。

修复：执行细化。 动态故障表明，策略规范 $\Omega_\pi \subset \tau$ 对于当前场景而言定义不当。为解决这一问题，FATE 会触发回滚，并根据失效轨迹 ξ_{fail}合成动态残差 $\Delta\tau_{dynamic}$。 任务更新如下：

$$\tau_{new} \leftarrow \tau_{old} \oplus \Delta\tau_{dynamic} \tag{9}$$

此处，残差 $\Delta\tau_{dynamic}$ 专门针对策略子空间 Ω_π。这一更新相当于一种**语义策略梯度**，审计器会修改物理场景（例如，通过缩放物体以便于插入的**几何放宽** ），或调整求解器逻辑（例如，通过注入密集奖励塑形项进行**逻辑调优**）。通过这些迭代更新，FATE 将发散的轨迹拉回可行区域 Φ_τ。

命题 4.1 (层级对齐的收敛性)。静态对齐与动态对齐算子的顺序运用会严格扩大可行任务集。形式上，*FATE* 算子 $\mathcal{A} = \mathcal{A}_{exec} \circ \mathcal{A}_{perc}$ 确保

$$|\Phi_{\mathcal{A}(\tau)}| \geq |\Phi_\tau|. \tag{10}$$

证明见附录 B。

推论。 命题 4.1 保证，语义多样性高但结构稳定性低的任务会被主动投射到可行性流形 $\mu(\tau) > \delta_{min}$上。

5. 实验

我们的实验验证了三个核心主张：**(1)生成能力**，确保提案分布具有较高的语义多样性和视觉多样性；**(2) 可行性感知**，评估具身大脑辨别可行性边界的准确度；**(3) 精化效能**，量化投影算子弥合可行性差距的成效。

5.1. 实验设置

尽管本框架与具体框架无关，我们仍在 NVIDIA Isaac Sim 中实现系统，以利用精确的接触物理为具身大脑提供视觉统计输入。我们使用 RidgebackFranka（安装在全向底座上的 7 自由度机械臂）评估长时程任务，这些任务需要协调导航与操作，并使用来自 Objaverse 和 PartNet-Mobility 的资源。我们采用混合执行策略：**BIT*** 用于无碰撞导航，**软演员-评论家（SAC）** 用于高接触操作。策略网络和 Q 网络均为多层感知机，隐藏层大小为 $[256, 256]$,，并通过 **Adam** 进行优化。

5.2. 生成多样性

指标与基线。 为量化生成多样性，我们采用语义和视觉指标。我们使用 **Self-BLEU-4**和**S-BERT 相似度** 衡量语言冗余，分数越低表示多样性越高。为捕捉结构差异，我们通过 ViT 和 CLIP 嵌入在不同场景间的余弦相似度评估**视觉多样性** ，确保形态的广泛性。我们将方法与以下基线进行比较：**(1) 专家设计的基准**（例如 RLBench、ManiSkill2、Behavior-100），代表由人工精心整理的标准；以及 **(2) GenSim-V2**，一种缺乏可行性反馈的先进开环基线。

结果。 表 1 表明，与开环生成相比，FATE 保持了较强的语言和视觉多样性。我们将此归因于闭环修复机制：该机制避免反复采样“安全”但不可行的模板，转而推动生成器搜索多样且仍具备物理基础的任务实现方式。

5.3. 任务可行性

指标与基线。 为量化可行性差距的缩小程度，我们采用**可行任务率（FTR）**，根据验证流程评估任务 τ 是否可行（即满足 $\mu(\tau) > \delta_{min}$）。我们将 FTR 分解为两个子指标：第一，**场景有效率（SVR）**通过 BLIP-2 语义连贯性和几何检查来衡量感知对齐；第二，**执行有效率（EVR）**使用仿真预言机，并通过模型预测控制和强化学习求解器验证物理滚动仿真是否成功，以评估动态可行性。我们将 FATE 与**原始基线（无审计）**下限以及最先进的**GenSim-V2**进行基准测试。**结果**。 表 2 展示了两个关键观察。首先，静态场景有效性虽有必要，但并不足够：许多失败只有在执行过程中才会显现并被发现，因此单纯的感知筛选仍会留下较大的可行性差距。其次，通过执行中审计和定向修复闭合循环，可显著提升端到端产出，这表明可行性问题主要源于少数可纠正的错位（例如可达性、碰撞、奖励/时域设置），而非任务多样性不足。

5.4. 审计准确率

指标与基线。 FATE 框架的可靠性取决于审计器准确近似可行性指标 $\mathbb{I}[\mu(\tau) > \delta_{min}]$ 的能力。审计器通过检查以下三个特定方面是否存在问题来检测不可行任务：**语义问题**、**几何约束**和 **动态问题**，图1直观展示了相关示例。为量化这种可靠性，我们使用人工标注的真实数据集，通过**F1分数** 衡量**审计器效能**来评估其表现。该指标直接反映判别器识别这些不同故障类型的准确度，同时尽量减少误报。**结果**。 表3证实，具身大脑是一个高精确率过滤器。前置审计器的 F1分数分别为**0.95**（语义）和**0.86** （几何），能够在仿真前有效剔除可供性违规等静态幻觉。动态类别更具挑战性（F1 **0.83**），因为物理不可能性与算法不稳定性之间存在歧义。不过，保持 0.85 的精确率可确保步内审计器采取保守策略，避免漏报，并为自我纠正闭环的收敛提供可靠反馈。

5.5. 修复效能

指标定义。 为衡量主动修复循环的贡献，我们定义了**修复成功率（RSR）**。该指标计算初始不可行的任务假设（即 $\mu(\tau) \leq \delta_{min}$）经修改后成功映射到可行区域的条件概率。我们在三个不同阶段评估 RSR：**预修复** （静态场景调整）、**基元修复**（运行时逻辑纠正）以及**强化学习修复** （学习参数调优）。较高的 RSR 表明，该系统并非仅仅充当筛选器（丢弃不可行数据），而是在主动将提案分布投影到可行性流形上。**结果**。表4显示，大多数不可行提案并非根本无法实现。FATE 最终能够修复其中的大部分失败案例。剩余的棘手案例主要源于学习导致的脆弱性，这表明改进强化学习侧是主要瓶颈。总体而言，这些结果支持了我们的主张：可行性感知生成能够提升生成任务的质量和可执行性。

表4量化了修复机制的有效性。较高的**总体修复成功率（89.5%）**表明，大多数初始可行性较低（即 $\mu(\tau) \leq \delta_{min}$）的生成提案本质上合理，只是存在错位。FATE按顺序应用分层调整算子——首先通过预修复解决静态（场景层面）违规（成功率为93.5%），随后通过步内修复处理动态（执行层面）故障——成功挽救了这些任务。值得注意的是，即使面对需要强化学习修复的最具挑战性的故障，系统也能挽救76.3%的案例。

5.6. 消融研究

为剖析对齐算子的具体贡献，我们评估了五种不同语义类别下的可行任务率。这种细分揭示了静态与动态阶段的互补性。

对于“将食材放入锅中”这类几何敏感型任务

表 *1.* **生成多样性分析**。 我们将 FATE 生成任务的语义和视觉多样性与专家设计的基准测试及开环基线进行比较。FATE 在确保物理可行性的同时，实现了与人工整理的基准测试相当的多样性。

	RLBench	ManiSkill2	Meta-World	Behavior-100	GenSim-V2	FATE（本文方法）
任务数量	106	20	50	100	70	**352**
语义多样性（语言）						
任务描述 - Self-BLEU-4（↓）	0.317	0.674	0.322	0.299	0.378	**0.276**
任务描述 - S-BERT 相似度 (↓)	0.200	0.194	0.263	0.210	0.288	**0.158**
视觉与构型多样性（场景）						
场景图像 - ViT相似度 (↓)	0.375	0.332	0.517	0.389	0.717	**0.189**
场景图像 - CLIP相似度 (↓)	0.864	0.828	0.867	0.833	0.932	**0.754**

表*2.* **量化可行性差距**。 我们将原始基线生成任务的有效性与FATE进行比较。我们将总产出（FTR）分解为静态场景有效性（SVR）和动态执行有效性（EVR）。

	场景有效性（SVR）*(静态有效性) (↑)*	执行有效性 (EVR) *(动态有效性) (↑)*	可行任务率 (FTR) *(端到端产出) (↑)*
基线			
原始基线（无审计）	41.5%	30.4%	12.6%
GenSim-V2	58.2%	51.2%	29.8%
FATE（本文方法）			
FATE-Static（事前审计）	93.8%	69.7%	65.4%
FATE（完整流水线）	**97.5%**	**94.5%**	**92.1%**

表*3*。 **审计器的诊断性能**。 我们在人工标注测试集上报告了FATE审计器在三类错误类别上的精确率、召回率和 F1分数。

诊断类别	精确率	召回率	F1分数
语义	0.96	0.94	0.95
几何	0.88	0.85	0.86
动态	0.85	0.81	0.83
总体	**0.90**	**0.87**	**0.88**

或“在微波炉中加热食物，”仅应用**静态对齐 ($\mathcal{A}_{perc}$)** 便能显著提升效果，与原始基线相比，成功率提高了 60% 以上。在这些场景中，主要的失败原因是几何不兼容（例如，锅比食材小），而感知投影算子无需耗费大量计算进行仿真，就能高效解决这一问题。

相反，对于“打开滑动窗”，这类富接触任务，静态对齐的效果并不理想（20.6%），逊于**动态对齐**（$\mathcal{A}_{exec}$）（41.8%）。这是因为滑动机制涉及复杂的摩擦和关节动力学，无法通过静态方式验证。动态对齐算子能够捕捉这些物理交互故障，但要实现高可靠性，则需要**完整FATE流程**（≈ 88%）。通过依次应用 $\mathcal{A}_{exec} \circ \mathcal{A}_{perc}$，系统可确保场景先满足几何可行性，再进行动态优化，从而覆盖多种开放式任务。

表*4*。**修复效能分析**。我们报告了三个阶段的修复成功率（RSR）。「检测到的故障」表示被识别为位于流形之外的任务，而「成功修复」则表示成功投影到流形上的任务。

修复阶段	检测到的故障（输入）	成功修复（输出）	修复成功率（%）
修复前（静态）	215	201	93.5%
原始修复（运行时）	42	39	92.9%
强化学习修复（学习）	76	58	76.3%
总计	**333**	**298**	**89.5%**

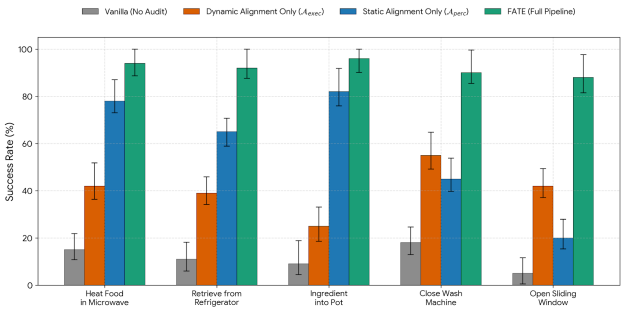


图 *4.* **任务级成功率分析**。 我们评估了五项任务中的表现，以解耦对齐算子的贡献。静态对齐（$\mathcal{A}_{perc}$）模块在几何敏感型任务（例如将食材放入锅中）中表现出色，而动态对齐（$\mathcal{A}_{exec}$）模块对于富接触任务（例如滑动窗口）至关重要。**FATE（绿色）**结合了两者，以实现稳健的性能。

6. 结论与讨论

在这项工作中，我们提出了 FATE，将任务合成重新表述为一个闭环约束优化过程。通过集成具身大脑进行分层审查与修复，FATE 使生成的假设与可行性流形保持一致，减少物理幻觉，从而生成语义多样且严格扎根于现实的课程。

尽管取得了显著进展，局限性依然存在。首先，视觉语言模型固有的随机性意味着细微的不一致仍可能绕过我们的迭代过滤器。其次，我们的框架侧重于几何与运动学约束，对细粒度动力学（例如摩擦、阻尼）的建模不够严格。

因此，依赖精确接触物理的任务仍可能遭遇超出当前修复范围的执行失败。

未来的工作将通过纳入真实世界故障日志，使物理参数根据真实情况进行校准，从而解决这些问题。此外，为缓解计算延迟和随机性，我们计划将审计能力提炼到一个轻量级内部世界模型中。这一演进旨在将 FATE 建立为自我纠正数据引擎，使其严格以仿真和真实世界物理为基础。

影响声明

本研究通过闭合语言驱动的任务提议与具身可行性验证之间的环路，推进了开放式机器人课程生成。该方法大幅减少了执行失败和仿真浪费，能够降低数据生成的计算成本，并提高下游策略学习的可靠性，从而有望促进研究可复现性，加速开发可在人类环境中运行的辅助机器人和服务机器人。与此同时，能够扩展任务合成规模的方法也可能降低开发有害或未经授权的自动化系统的门槛；而仿真中的可行性并不能保证真实世界中的安全性，尤其是在存在未建模动力学、具身不匹配或审计错误的情况下。我们的实验仅限于仿真，与接触动力学和学习控制器脆弱性相关的失效模式可能仍然存在。因此，我们将该系统定位为数据引擎和评估工具，而非可部署控制器，并建议在领域特定验证、人工监督和安全约束下负责任地使用。

参考文献

阿默斯特大学。LuciBot：从生成视频中自动学习机器人策略。*arXiv*预印本 *arXiv:2503.09871*，2025年。

Bai, S., Chen, K., Liu, X., Wang, J., Ge, W., Song, S., Dang, K., Wang, P., Wang, S., Tang, J. 等。Qwen2. 5-vl 技术报告。*arXiv*预印本 *arXiv:2502.13923*，2025年。字节跳动。

AgentGen：通过环境和任务生成增强基于大语言模型的智能体的规划能力。*arXiv*预印本 *arXiv:2408.00764*，2025年。卡内基梅隆大学。

SEEA-R1：面向自我进化具身智能体的树结构强化微调。*arXiv*预印本 *arXiv:2506.21669*，2025年。香港中文大学（深圳）。Gen2Real：利用生成视频实现无需演示的灵巧操作。*arXiv*预印本 *ar Xiv:2509.14178*，2025年。

Deitke, M., Schwenk, D., Salvador, J., Weihs, L., Michel, O., VanderBilt, E., Schmidt, L., Ehsani, K., Kembhavi, A. 和 Farhadi, A. Objaverse：一个包含带注释的三维物体的宇宙。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第 13142–13153 页，2023年。

Fu, X., Hu, Y., Li, B., Feng, Y., Wang, H., Lin, X., Roth, D., Smith, N. A., Ma, W.-C. 和 Krishna, R. Blink：多模态大语言模型看得见，但无法感知。载于欧洲计算机视觉会议，第148–166页。施普林格，2024。

Garrett, C.、Mandlekar, A.、Wen, B.和Fox, D. SkillMimicGen：面向高效技能学习与部署的自动化演示生成。*arXiv*预印本 *arXiv:2410.18907*，2024。

He, Z.、Liu, Y.、Zheng, J.-s.、Li, X.、Yao, J.-G.、Qin, B.、Xuan, R.和Yang, X. Flagevalmm：面向全面多模态模型评测的灵活框架。载于第*63*届计算语言学协会年会论文集（第*3*卷：系统演示），第51–61页，2025。Jia, M.、Qi, Z.、Zhang, S.、Zhang, W.、Yu, X.、He, J.、Wang, H.和Yi, L. OmniSpatial：面向视觉语言模型的全面空间推理基准。*arXiv*预印本 *arXiv:2506.03135*，2025。Jiang, Z.、Xie, Y.、Lin, K.、Xu, Z.、Wan, W.、Mandlekar, A.、Fan, L. J.和Zhu, Y. DexMimicGen：通过模仿学习实现双手灵巧操作的自动化数据生成。载于*2025*年*IEEE*国际机器人与自动化会议（*ICRA*），第16923–16930页。IEEE，2025。Kirillova, A.、Lyapustin, E.、Antsiferova, A.和Vatolin, D. Erqa：面向视频超分辨率的边缘复原质量评估。*arXiv*预印本 *arXiv:2110.09992*，2021。

Knightnemo等。Vid2world：将视频扩散模型打造为交互式世界模型。*arXiv*预印本 *arXiv:2505.14357*，2025。Laboratory, S. A. ReGen：通过逆向设计实现生成式机器人仿真。*OpenReview*，2025。Ma, J.、Liang, W.、Wang, H.-J.、Zhu, Y.、Fan, L.、Bastani, O.和Jayaraman, D. DrEureka：语言模型引导的仿真到现实迁移。RSS，2024。Mo, K.、Zhu, S.、Chang, A. X.、Yi, L.、Tripathi, S.、Guibas, L. J.和Su, H. PartNet：面向细粒度分层部件级三维物体理解的大规模基准测试。载于*IEEE*计算机视觉与模式识别会议（*CVPR*），2019年6月。

英伟达。MimicGen：一种利用人类演示实现可扩展机器人学习的数据生成系统。*arXiv*预印本 *arXiv:2310.17596*，2024。

英伟达. DreamGen：通过视频世界模型解锁机器人学习中的泛化能力。*arXiv* 预印本 *arXiv:2505.12705*，2025。

Qiu, L.、Chen, Y.、Ge, Y.、Ge, Y.、Shan, Y. 与 Liu, X. EgoPlan-Bench2：面向真实场景中多模态大语言模型规划的基准测试。*arXiv* 预印本 *arXiv:2412.04447*，2024。

Song, C. H.、Blukis, V.、Tremblay, J.、Tyree, S.、Su, Y. 与 Birchfield, S. RoboSpatial：面向机器人学，向二维与三维视觉语言模型教授空间理解。收录于计算机视觉与模式识别会议论文集，第 15768–15780 页，2025。

Team, B. R., Cao, M., Tan, H., Ji, Y., Chen, X., Lin, M., Li, Z., Cao, Z., Wang, P., Zhou, E. 等。RoboBrain 2.0 技术报告。*arXiv* 预印本 *arXiv:2507.02029*，2025年。UMass-Embodied-AGI。利用世界模型进行空间推理的测试时扩展。*arXiv* 预印本 *arXiv:2507.12508*，2025年。University, S. FactorSim：通过因子化表示实现生成式仿真。*NeurIPS*，2024。University, T. DemoGen：面向数据高效视觉运动策略学习的合成示范生成。*arXiv* 预印本 *arXiv:2502.16932*，2025a。University, T. GenSim2：利用多模态大语言模型和推理大语言模型扩展机器人数据生成。*arXiv* 预印本 *arXiv:2410.03645*，2025b。Wang, L., Ling, Y., Yuan, Z., Shridhar, M., Bao, C., Qin, Y., Wang, B., Xu, H. 和 Wang, X. GenSim：通过大语言模型生成机器人仿真任务。*arXiv* 预印本 *arXiv:2310.01361*，2023a。Wang, Y., Xian, Z., Chen, F., Wang, T.-H., Wang, Y., Fragkiadaki, K., Erickson, Z., Held, D. 和 Gan, C. RoboGen：通过生成式仿真释放自动化机器人学习的无限数据。*arXiv* 预印本 *arXiv:2311.01455*，2023b。Weisz, J., Huang, Y., Lier, F., Sethumadhavan, S. 和 Allen, P. RoboBench：迈向可持续的机器人学系统基准测试。载于*2016 IEEE* 国际机器人与自动化会议（*ICRA*），第 3383–3389 页。IEEE，2016。Zhu, F., Wu, H., Guo, S., Liu, Y., Cheang, C. 和 Kong, T. IRASim：面向机器人操作的细粒度世界模型。载于*IEEE/CVF* 国际计算机视觉会议论文集，第 9834–9844 页，2025a。

Zhu, N., Dong, Y., Wang, T., Li, X., Deng, S., Wang, Y., Hong, Z., Geng, T., Niu, G., Huang, H., 等。Cvbench：评估复杂多模态理解与推理中的跨视频协同。*arXiv* 预印本 *arXiv:2508.19542*，2025b。

A. 迭代修复的理论分析

本节将给出**命题 A.4**的详细数学证明，确立 FATE 修复机制的线性收敛性。该分析将审计器（EmbodiedBrain）与任务流形之间的交互建模为基于梯度的动力系统优化。

A.1. 设置与定义

令 $\mathcal{T}$表示连续任务参数空间。为分析收敛性质，我们基于定义 3.3 中引入的可行性度量 $\mu(\tau)$，定义一个**不可行性势函数** $\mathcal{J}$： $\mathcal{T} \to \mathbb{R}_{\geq 0}$

$$\mathcal{J}(\tau) \triangleq \max\left(0, \delta_{min} - \mu(\tau)\right) \tag{11}$$

我们的目标是最小化这一势函数。优化过程遵循更新规则.....

A.2. 正则性假设

为保证收敛，我们采用非凸优化中的标准假设，并针对我们的"近似预言机"（基于大语言模型的审计器）场景进行了调整。

假设 A.1 （L-光滑性）。 损失函数 $\mathcal{J}$可微，且其在流形上的梯度满足 L-Lipschitz 连续性。也就是说，对于任意 τ, τ'，存在常数 $L > 0$，使得：

$$\|\nabla\mathcal{J}(\tau) - \nabla\mathcal{J}(\tau')\| \leq L\|\tau - \tau'\| \tag{12}$$

这意味着标准的二次上界不等式成立：

$$\mathcal{J}(\tau') \leq \mathcal{J}(\tau) + \langle\nabla\mathcal{J}(\tau), \tau' - \tau\rangle + \frac{L}{2}\|\tau' - \tau\|^2 \tag{13}$$

假设 A.2 （审计器对齐条件）。 审计器不会计算精确梯度 $\nabla\mathcal{J}(\tau_k)$（这需要进行昂贵的蒙特卡洛模拟）。相反，它会提供一个与真实梯度严格对齐的语义方向 d_k。我们假设存在常数 $c_1, c_2 > 0$，使得：

$$\langle d_k, \nabla\mathcal{J}(\tau_k)\rangle \geq c_1\|\nabla\mathcal{J}(\tau_k)\|^2 \quad \text{(Sufficient Descent)} \tag{14}$$

$$\|d_k\| \leq c_2\|\nabla\mathcal{J}(\tau_k)\| \quad \text{(Bounded Magnitude)} \tag{15}$$

假设 A.3 （Polyak-Łojasiewicz（PL）条件）。 在可行流形 $\mathcal{F} = \{\tau \mid \mathcal{J}(\tau) = 0\}$的吸引盆内，损失满足参数为 $\nu > 0$ 的 PL 条件：

$$\frac{1}{2}\|\nabla\mathcal{J}(\tau)\|^2 \geq \nu\mathcal{J}(\tau) \tag{16}$$

A.3. 线性收敛证明（命题 A.4）

命题 A.4 (迭代修复的线性收敛)。 在假设 *A.1–A.3* 下，对于足够小的步长 η，迭代修复更新规则 $\tau_{k+1} = \tau_k - \eta d_k$ 线性收敛至可行流形 $\mathcal{F} = \{\tau \mid \mu(\tau) \geq \delta_{min}\}$。

证明。我们首先将更新规则 $\tau_{k+1} = \tau_k - \eta d_k$ 代入二次上界 **(13)** ：

$$\mathcal{J}(\tau_{k+1}) \leq \mathcal{J}(\tau_k) + \langle\nabla\mathcal{J}(\tau_k), -\eta d_k\rangle + \frac{L}{2}\| - \eta d_k\|^2 \tag{17}$$

重新整理各项，并应用内积的标量性质：

$$\mathcal{J}(\tau_{k+1}) \leq \mathcal{J}(\tau_k) - \eta\langle\nabla\mathcal{J}(\tau_k), d_k\rangle + \frac{L\eta^2}{2}\|d_k\|^2 \tag{18}$$

接下来，我们应用 **审计器对齐假设 (A.2)**。我们使用内积的下界以及 d_k 范数的上界：

$$\mathcal{J}(\tau_{k+1}) \leq \mathcal{J}(\tau_k) - \eta c_1 \|\nabla \mathcal{J}(\tau_k)\|^2 + \frac{L\eta^2}{2}(c_2 \|\nabla \mathcal{J}(\tau_k)\|)^2 \tag{19}$$

提取因子 $\|\nabla \mathcal{J}(\tau_k)\|^2$：

$$\mathcal{J}(\tau_{k+1}) \leq \mathcal{J}(\tau_k) - \left(\eta c_1 - \frac{L\eta^2 c_2^2}{2}\right) \|\nabla \mathcal{J}(\tau_k)\|^2 \tag{20}$$

为确保损失下降，梯度范数前的系数必须为正。这对步长η施加了约束：

$$\eta c_1 - \frac{L\eta^2 c_2^2}{2} > 0 \implies \eta < \frac{2c_1}{Lc_2^2} \tag{21}$$

假设 η 满足此条件，我们定义下降系数 $\mathcal{C}_\eta = \eta c_1 - \frac{L\eta^2 c_2^2}{2} > 0$。现在，应用**PL 条件**(A.3)，代入 $\|\nabla \mathcal{J}(\tau_k)\|^2 \geq 2\nu \mathcal{J}(\tau_k)$：

$$\mathcal{J}(\tau_{k+1}) \leq \mathcal{J}(\tau_k) - \mathcal{C}_\eta \cdot (2\nu \mathcal{J}(\tau_k)) \tag{22}$$

化简表达式：

$$\mathcal{J}(\tau_{k+1}) \leq \mathcal{J}(\tau_k)(1 - 2\nu \mathcal{C}_\eta) \tag{23}$$

设 $\rho = 2\nu(\eta c_1 - \frac{L\eta^2 c_2^2}{2})$。对于足够小的步长 η，我们可以确保 $0 < \rho < 1$。因此，得到线性递推关系：

$$\mathcal{J}(\tau_{k+1}) \leq (1 - \rho)\mathcal{J}(\tau_k) \tag{24}$$

通过归纳法，经过 k 步后：

$$\mathcal{J}(\tau_k) \leq (1 - \rho)^k \mathcal{J}(\tau_0) \tag{25}$$

由于 $(1 - \rho) < 1$，$\lim_{k\to\infty}(1 - \rho)^k = 0$，这意味着 $\lim_{k\to\infty} \mathcal{J}(\tau_k) = 0$。因此，任务 τ_k 线性收敛到满足 $\mu(\tau) \geq \delta_{min}$ 的可行区域。 □

B. 分层对齐的收敛性证明

在本节中，我们给出**命题 4.1**的详细证明。我们证明 FATE 算子 $\mathcal{A} = \mathcal{A}_{exec} \circ \mathcal{A}_{perc}$会严格保持或扩大可行策略集的体积。

B.1. 测度论表述

令 $(\Omega_\pi, \Sigma, \rho_{init})$ 为策略参数的测度空间。对于任意任务 τ，可行集定义为：

$$\Phi_\tau = \{\theta \in \Omega_\pi \mid \mathbb{E}_{\xi \sim (\pi_\theta, \mathcal{S})}[\Psi_{\mathcal{I}}(\xi)] \geq 1 - \epsilon\} \tag{26}$$

可行性体积是测度 $V(\tau) = \rho_{init}(\Phi_\tau)$。我们通过对齐的两个阶段分析 $V(\cdot)$ 的单调性。

B.2. 单调性证明

阶段 1：静态对齐（$\mathcal{A}_{perc}$）。令 $\tau_{init} = (\mathcal{I}, \mathcal{S}, \Omega_\pi)$ 为初始任务提案。我们将静态有效集$\Theta_{static}(\mathcal{S}, \mathcal{I}) \subset \Omega_\pi$ 定义为不违反基本物理定律的策略集合（例如，尝试抓取运动学上不可达的物体）。

引理 B.1 (静态必要性)。 对于任意策略 θ，如果 $\theta \in \Phi_\tau$（动态可行），那么 $\theta \in \Theta_{static}$（必须静态有效）。因此，$\Phi_\tau \subseteq \Theta_{static}$。

如果初始任务存在静态冲突（例如严重的相互穿透），那么 $\Theta_{static} = \emptyset$，这意味着 $\Phi_{\tau_{init}} = \emptyset$ 和V（τ_{init}）$= 0$。静态对齐操作器 $\mathcal{A}_{perc}$ 应用修复 Δ_{static} 以确保静态兼容性，从而创建一个非空的静态有效集 $\Theta'_{static} \neq \emptyset$。虽然这不能保证动态可行性，但它将体积从严格为零的状态转变为可能为正的状态。因此：

$$V(\mathcal{A}_{perc}(\tau_{init})) \geq V(\tau_{init}) = 0 \tag{27}$$

阶段 2：动态对齐（$\mathcal{A}_{exec}$）。设 $\tau_{static} = \mathcal{A}_{perc}(\tau_{init})$ 为静态有效的任务。动态对齐算子采用“回滚与修复”机制。我们将 $\mathcal{A}_{exec}$ 的执行建模为生成一系列经过修改的任务/求解器配置 $\mathbf{T} = \{\tau^{(0)}\tau^{(1)}\ldots,\tau^{(T)}\}$，其中 $\tau^{(0)}= \tau_{static}$。关键在于，如果策略在迭代过程中找到的任一精炼配置下执行成功，FATE 系统就会接受该任务为“可行”。设 $\Phi^{(k)}$ 为与配置 $\tau^{(k)}$相关联的可行策略集。对齐后任务 $\mathcal{A}_{exec}(\tau)$ 的有效可行集是修复轨迹上各可行集的并集：

$$\Phi_{total} = \bigcup_{k=0}^{T} \Phi^{(k)} \tag{28}$$

由于初始配置对应于 $k = 0$，因此有如下显然的包含关系：

$$\Phi^{(0)} \subseteq \bigcup_{k=0}^{T} \Phi^{(k)} = \Phi_{total} \tag{29}$$

根据测度的单调性，若 $A \subseteq B$，则 $\rho_{init}(A) \leq \rho_{init}(B)$。因此：

$$V(\tau^{(0)}) \leq \rho_{init}(\Phi_{total}) \tag{30}$$

假设审计器提供非破坏性更新（即不会在不增加新的可行区域的情况下，单纯缩小后续步骤的可行空间），则系统可访问的有效解空间会随着每次迭代而扩大。

结论。综合阶段 1 和阶段 2 的结果：

$$V(\mathcal{A}(\tau_{init})) = \rho_{init}(\Phi_{total}) \geq V(\tau_{static}) \geq V(\tau_{init}) \tag{31}$$

因此，分层对齐过程保证可行解空间的测度严格非递减，从而验证了命题 4.1。

□

C. FATE 流水线的实现细节

> **注意：** 本附录详细介绍了第 4 节中定义的**可行性对齐算子** $\mathcal{A}$ 的构造性实现。我们将说明如何把理论投影步骤（$\mathcal{A}_{perc}$ 和 $\mathcal{A}_{exec}$）实例化为具体的算法循环，并由**具身大脑**控制（详见附录 D）。

C.1. 初始化：多样化任务提议

在对齐阶段开始之前，我们首先从生成分布 $\tau_{init} \sim P_{\mathcal{G}}$（如公式 3 所定义）中采样。目标是在严格应用可行性约束之前，最大限度地提高语义多样性。

语义候选检索。我们使用来自 Objaverse（Deitke 等人，2023年）和 PartNet-Mobility（Mo 等人，2019年）数据集的资源。为确保提案分布的多样性，具身大脑基于高层标签（例如微波炉、烤箱、食物）进行广泛的语义搜索。在此阶段，我们放宽严格的几何过滤，以允许新颖的物体组合。

场景假设生成。系统输出初步场景图 $\mathcal{S}_{init}$ 和指令 $\mathcal{I}$。随后将其解析为可用于仿真的通用场景描述（USD）加载器脚本，以初始化任务元组 $\tau_{init} = (\mathcal{I}, \mathcal{S}_{init}, \Omega_{\pi}^{init})$。

C.2. 第一阶段：静态对齐算子（$\mathcal{A}_{perc}$）

此阶段实现了**感知投影算子** $\mathcal{A}_{perc}$（第 4.3 节）。其目标是最大化静态可行性代理 $\hat{\mu}_{static}$，注入随机感知偏差 Δ_{static}。

诊断分类体系（估计 $\hat{\mu}_{static}$）。前置审计器根据第 4.3 节定义的四个可行性维度评估任务元组：

- **可供性匹配：**检查语义类别是否支持所请求的动作（例如，将“加热”与“微波炉”匹配）。
- **几何可达性：**检查场景图 $\mathcal{S}$ 是否包含放置在机器人工作空间之外的物体。
- **物理合理性：**检查初始配置 $\mathcal{S}_{init}$中是否存在严重的相互穿透或不稳定放置。
- **形态兼容性：**检查物体尺度是否兼容（例如，“食材能否放入锅中”）。

修复动作空间（实现 $\oplus\Delta_{static}$）。更新步骤 $\tau' \leftarrow \tau \oplus \Delta_{static}$ 通过一组原子场景操作 API（表 5）实现。这些原语直接针对 $\hat{\mu}_{static}$ 中的违规进行处理。

表 *5*。 **静态修复 API**（$\mathcal{A}_{perc}$）。 这些原语实现离散更新 Δ_{static}，以解决第 4.3 节定义的感知幻觉。

API 签名	功能与理论映射
交换资源(tgt id, query)	替换一个物体，以最大限度地实现**可供性匹配**。用于当前对象语义无法满足 $\Psi_{\mathcal{I}}$ 的情况。
生成资产(query, pose)	在场景图 $\mathcal{S}$ 中实例化缺失节点。当指令 $\mathcal{I}$ 暗示存在 $\mathcal{S}_{init}$ 中没有的物体时使用。
移除资产(tgt id)	修剪场景图。用于移除违反**几何可达性**的**杂乱物体**或造成碰撞冲突的杂乱物体。
设置关节(tgt id, idx, val)	将构型状态 q_j 强制设为特定值。用于解决逻辑不一致的情况（例如，“打开一扇门”，但门已经打开）。
TRANSFORM POSE(tgt id, Δ)	应用刚体变换。用于解决相互穿透问题（**物理合理性**）。
RESCALE(tgt id, factor)	缩放物理流形 M_j。直接解决**形态兼容性**（例如调整物体大小以适配容器）。

C.3. 第二阶段：动态对齐算子（$\mathcal{A}_{exec}$）

此阶段实例化**动态对齐算子** $\mathcal{A}_{exec}$（第 4.4 节）。重点是优化策略规范 Ω_π 和求解器超参数 θ，以确保可行策略集 Φ_τ 非空。

层次分解与检查点机制。 任务被分解为原语 $\Pi = \{\pi_1, \ldots, \pi_K\}$。为支持**语义策略梯度**（公式 9）的迭代优化，我们采用状态检查点机制。这使系统能够回滚并应用更新 $\theta^{(t+1)} \leftarrow \theta^{(t)} \oplus \nabla_{sem}$，而无需重置整个环境。

动态干预 API（实现 ∇_{sem}）。当步内审计器检测到**动态偏差**时，会生成修复命令。对于模型预测控制原语，修复通常会调整几何约束（膨胀半径）。对于强化学习原语，Brain 会主动生成**奖励景观**（表6）。

D. 具身大脑：架构与专门化

在 FATE 框架中，审计机制的效能从根本上取决于底层模型区分物理约束与语义幻觉的能力。我们采用专门版本的

算法 2 静态对齐循环 ($\mathcal{A}_{perc}$)

输入： 初始任务 $\tau = (\mathcal{I}, \mathcal{S})$
$\mathcal{S}_{curr} \leftarrow \mathcal{S}$
当 为真 **时执行**
  // 估计静态可行性代理指标 $\hat{}\mu_{static}$
  $Critique,\ IsValid \leftarrow$ AnteAuditor($\mathcal{I},\ \mathcal{S}_{curr}$)
  如果 $IsValid$ **则**
    跳出
  如果结束
  // 合成残差 Δ_{static}
  $Repair_Op \leftarrow$Brain.生成修复($Critique$)
  // 应用投影 $\mathcal{S} \leftarrow \mathcal{S} \oplus \Delta_{static}$
  $\mathcal{S}_{curr} \leftarrow$ 模拟器.$apply$($Repair\ Op,\ \mathcal{S}_{curr}$)
  如果 $MaxIter$ 已到达 **则**
    返回 失败
  结束如果
结束循环
返回 $\mathcal{S}_{curr}$

RoboBrain 2.0（Team et al., 2025），这是一种拥有 70 亿参数的视觉语言模型，经过专门微调，以提升空间推理和可执行规划能力。本附录概述了该模型的架构基础、训练课程的组成，以及对其具身能力的定量验证。

D.1. 模型架构

审计器基于 RoboBrain 2.0 架构构建，该架构源自 Qwen2.5-VL 基础模型（Bai 等，2025）。该系统由三个主要组件构成，旨在弥合视觉感知与逻辑推理之间的差距。感知前端采用基于 Qwen-VL 视觉主干网络初始化的视觉变换器（ViT）编码器，确保能够从高分辨率输入中稳健地提取特征。这些视觉特征通过多层感知机（MLP）投影器映射到语言嵌入空间，从而实现精确的模态对齐。核心推理引擎是一个 Qwen2.5-7B 大语言模型，用于处理对齐后的视觉词元，以生成语义批评和修复代码。为满足长时程机器人任务的计算需求，训练基础设施采用了包括参数分片和注意力加速在内的深度学习优化策略，从而能够处理长度高达 16k 词元的视频上下文。

D.2. 训练数据构建

为使具身大脑具备可行性审计所需的直觉，我们采用了一个经过精心筛选、包含约 145,000 个样本的混合数据集。我们专门设计了数据分布，以弥补通用视觉语言模型中常见的行动差距，并平衡三个关键认知维度：具身行动、空间感知和推理。课程中占比最大的部分为具身行动，占数据集的 45%。这部分汇集了 RoboBench（Weisz 等，2016）、EgoPlan-Bench2（Qiu 等，2024）和 RoboSpatial-Home（Song 等，2025）等基准测试的数据，旨在培养任务分解、导航和可供性检测能力。另有同等比例的 45% 用于空间感知，采用 Blink（Fu 等，2024）、OmniSpatial（Jia 等，2025）和 CVBench（Zhu 等，2025b）等数据集，为视觉定位和几何理解提供严格监督。最后 10% 的数据通过 ERQA 推理数据集（Kirillova 等，2021）保留模型的高层因果逻辑。这种策略性的数据构成确保审计器不仅能标记物体，还能理解物理场景中的空间关系与交互可供性。

D.3. 实证验证

微调后的模型在 EmbodiedVerse基准测试（He 等，2025）上进行了评估，以量化其相较于基线 Qwen2.5-VL 的提升。结果表明，专门设计的训练方案显著提升了对 FATE审计流程 最为关键的指标。具体而言，模型的相对

表*6.* **动态修复 API**（$\mathcal{A}_{exec}$）。这些原语实例化了语义策略梯度 ∇_{sem}，通过修改求解器景观来捕获可行解。

API 签名	**功能与理论映射**
强化学习专属优化（策略流形塑形）	
INJECT REWARD(func-str)	用 Brain 合成的代码覆盖奖励函数。此操作可实现**奖励景观合成**，让视觉语言模型重塑优化曲面，引导探索朝着 Φ_τ 进行。
设置强化学习时域(steps)	调整最大回合长度。这会扩展其时间搜索空间 Ω_π，确保有充足的时间收敛。
执行动力学与规划器调优	
重新规划路径（膨胀半径 r）-	修改模型预测控制求解器的配置空间约束。用于解决碰撞诱发的死锁。
调整抓取（偏移量，　　向量）	修改末端执行器接近姿态。用于纠正高扭矩抓取失败或奇异点。
调节阻抗（刚度）	调节控制器增益，确保高接触交互期间的稳定性（例如打开重型门）。
跳过子步骤()	若转换条件已满足（例如，门因意外接触而开启）。
重置机器人关节(cfg)	从运动学奇异点中恢复，使智能体回到 Ω_π 的有效区域。

形状估计绝对提升了 34.15%，相较于基线提升了 80.01%。该指标对于前置审计器检测形态不兼容至关重要，例如判断物体是否过大而无法放入容器。此外，模型的目标分解绝对提升了 5.00%。这一提升验证了步内审计器将高级指令分解为可执行原语，并识别规划序列中逻辑断层的能力。更重要的是，这些提升并未以灾难性遗忘为代价，通用感知任务（如计数）的表现提升了 2.17%，足以证明这一点。这些性能指标证实，具身大脑具备必要的空间精度和规划前瞻性，能够在 FATE循环 中充当可靠的可行性审计器。