

# EICPS：面向输电线路巡检的具身智能信息物理系统

## ——语义覆盖完备性形式化保证

周治国

北京理工大学集成电路与电子学院，北京 100081

通讯作者：zhiguozhou@bit.edu.cn

**摘要：**在安全关键的输电线路巡检任务中部署具身智能体，需要弥合自然语言任务描述与经认证的机器人执行之间的语义鸿沟。现有基于大语言模型（LLM）的任务规划器具备灵活的语义理解能力，但无法形式化回答是否每一个合法的任务表述都能被正确处理——这对安全关键系统而言是致命缺陷。本文提出 EICPS（具身智能信息物理系统，Embodied Intelligent Cyber-Physical Systems），一个三层解耦框架：语义任务路由（HTN 规划层）、物理安全执行（控制障碍函数层）和规程合规（规程层），使每一层均可独立进行形式化验证。核心理论贡献是语义 **Nyquist** 完备性定理（定理 1）：在规程约束假设（PCA）——一个对封闭规程域可经验验证的性质——下，语义嵌入覆盖率满足  $\text{SEC}(Q, \tau) = 1$ （精确等式），首次将标准统计界（ $\text{SEC} \geq 1 - \delta$ ）提升为确定性形式保证。我们将该框架实例化为架空输电线路激光无人机异物清除场景（国网 167 项目），枚举规范任务词汇表  $V_{167}^L$ （59 条目），测量语义内在维度  $d_{\text{sem}} = 3.56$ （TwoNN 估计），并通过四步实验验证 PCA：18 条自然语言释义全部满足 PCA（100%），6 条规程排除项全部触发 FAILSAFE（100%），安全间隙为 0.115。对抗实验建立覆盖上界  $\approx 0.42$ ；边界分析识别缓冲区（0.30-0.41），适用于人机协同审核。

**关键词：**具身智能；输电线路巡检；语义覆盖；控制障碍函数；分层任务网络；形式化安全保证；国网 167 项目

## 1. 引言

架空输电线路巡检与带电作业是具身智能系统最具挑战性的应用前沿之一。异物清除、防振锤更换、绝缘子检测等任务，既涉及安全关键的物理约束（高压近区、结构载荷极限），又面临现场运维人员发出的高度多变的自然语言任务指令。国家电网公司总部 167 项目（国网总部 167 项目）定义了架空输电线路自动化带电作业的规程框架，枚举了所有允许的任务类型、安全约束和环境作业窗口。

近年来大语言模型（LLM）的进展推动了开放域机器人任务规划的显著进步<sup>[1][2][3]</sup>。这些方法利用丰富的语义先验来解释多样化的自然语言指令。然而，它们在安全关键部署中共享一个根本局限：没有任何方法能够形式化保证每一个合法的任务表述都能被正确理解和路由。这一问题——系统的语义覆盖是否完备——恰恰是安全认证所要求的。

语义覆盖的标准度量是语义嵌入覆盖率（SEC）：

$$\text{SEC}(Q, \tau) = P_{t \sim D_{\text{task}}} [\min_{q_i \in Q} \|\phi(t) - \phi(q_i)\|_2 < \tau]$$

其中  $\phi$  为 LLM 编码器， $Q$  为任务图节点集合， $\tau$  为覆盖半径。现有分析仅能得到统计界  $\text{SEC}(Q, \tau) \geq 1 - \delta$ ，其中  $\delta > 0$  只能通过蒙特卡洛采样估计——留下不可约化的统计不确定性，不适用于安全关键认证。

**核心洞察。**与开放域服务机器人（ $D_{\text{task}}$  支撑集无界）不同，167 项目的任务分布由有限的规程文件所定义：每一个合法任务必须对应适用作业标准中描述的某项规程。这种规程封闭性使  $D_{\text{task}}$  的支撑集可枚举有限，从而使式 (1) 中的概率陈述退化为确定性等式。

贡献。本文做出四项贡献：

1. 形式化 SEC 度量：将指令覆盖形式化为嵌入度量空间上的概率测度，引入 SEC 作为首个可计算定量度量，用于判断任务规划系统是否正确处理所有合法任务描述。现有 LLM-机器人框架隐含地涉及覆盖问题，但均未提供形式化度量。
2. 规程封闭性作为数学资源：封闭规程域文档可枚举完备性构成了一种被忽视的数学资源：它使任务分布有限可枚举，将 SEC 从蒙特卡洛统计量转化为形式化可验证对象。
3. 语义 Nyquist 完备性定理（定理 1）：在 PCA 下， $SEC(Q, \tau) = 1$ ——首次将统计下界提升为确定性等式，以可认证的完备性保证和显式 FAILSAFE 机制替代优雅降级。
4. EICPS 架构与领域实例化：三层框架各层可独立验证，实例化为激光无人机异物清除，配有可复现的四步验证协议。

## 2. 相关工作

### 2.1 基于 LLM 的机器人任务规划

SayCan<sup>[1]</sup> 通过机器人可供性函数将 LLM 生成的规划落地，实现开放域指令跟随。其标题“按我能做的做，而非按我说的做”隐含承认技能库即覆盖上限。SayCan 通过优雅降级处理覆盖外指令：对每个候选技能  $s$  计算  $p(s | \text{instruction}) \times p(s | \text{state})$  并执行得分最高的技能，即使没有技能与指令语义匹配。这带来静默误路由风险——在不发出覆盖警告的情况下执行看似合理但错误的技能——在安全关键输电线路操作中不可接受。Code as Policies<sup>[2]</sup> 利用 LLM 从自然语言合成可执行机器人程序。RT-2<sup>[4]</sup> 将视觉-语言模型端到端扩展至机器人动作。Inner Monologue<sup>[3]</sup> 引入自然语言反馈环实现迭代规划改进。这些方法取得了令人印象深刻的开放域泛化能力，但均未提供形式化度量来回答是否所有可能的任务表述都能被正确理解——即指令覆盖问题。本文引入 SEC 作为首个捕捉此问题的形式化度量，并在封闭规程域中证明  $SEC(Q, \tau) = 1$ ，以可认证的完备性保证和显式 FAILSAFE 机制替代优雅降级。

### 2.2 分层任务网络规划

HTN 规划提供了从高层任务到基本动作的结构化分解，具有形式化完备性与正确性保证<sup>[5][6]</sup>。Konidaris 等<sup>[7]</sup> 证明了技能导出的符号对于已知技能集上的规划是充分且必要的——一个向下完备性结果（技能→符号）。EICPS 解决的是正交的向上完备性问题（指令→任务图）：给定固定任务词汇表，嵌入是否覆盖了所有可能的指令描述？这两个完备性方向独立，SEC 首次形式化了向上方向。近期混合方法将 LLM 语义理解与 HTN 结构相结合以提升鲁棒性<sup>[8]</sup>。EICPS 采用这种混合范式：用 LLM 嵌入进行语义路由，用 HTN 生成结构化规划。

### 2.3 机器人中的形式化安全保证

控制障碍函数 (CBF)<sup>[9]</sup> 通过二次规划提供实时安全执行，保证连续时间动力学下安全集的前向不变性。信号时序逻辑 (STL)<sup>[10]</sup> 支持时序任务正确性的形式化规约与监控。EICPS 在物理层集成 CBF，并在 HTN 层兼容基于 STL 的任务监控器。

### 2.4 输电线路巡检机器人

基于无人机的视觉检测显著推进了输电线路自动化巡检<sup>[11]</sup>。非接触检测方法——包括基于磁场测量的劣化绝缘子检测<sup>[12]</sup>和恶劣天气下的多模态融合三维感知<sup>[13]</sup>——已证明智能无人巡检在带电输电线路上的可行性。激光异物清除代表了一种进一步的非接触作业自主范式，要求精确的语义任务理解以区分异物类型、附着位置和电气状态。据我们所知，尚无先前工作为该任务类别提供形式化语义覆盖保证。

### 3. EICPS 框架

#### 3.1 系统架构

EICPS 将输电线路巡检控制问题分解为三个纵向分离的层次（图 1）：

1. 规程层：从 167 项目所有适用作业规程中提取规范任务词汇表  $V_{167}$ 。每个元素  $v_i \in V_{167}$  是从规程文本导出的规范性任务描述，该集合有限且有文件依据。
2. HTN 语义规划层：通过 LLM 嵌入相似度将自然语言运维输入映射到最近的规范节点，然后将匹配节点展开为结构化 HTN 规划。当无节点落在覆盖半径  $\tau$  内时，发出 FAILSAFE 信号。
3. CBF 物理安全层：实时监控规划执行，并通过二次规划安全滤波器执行物理约束（安全距离、激光功率限制、飞行包线极限）。

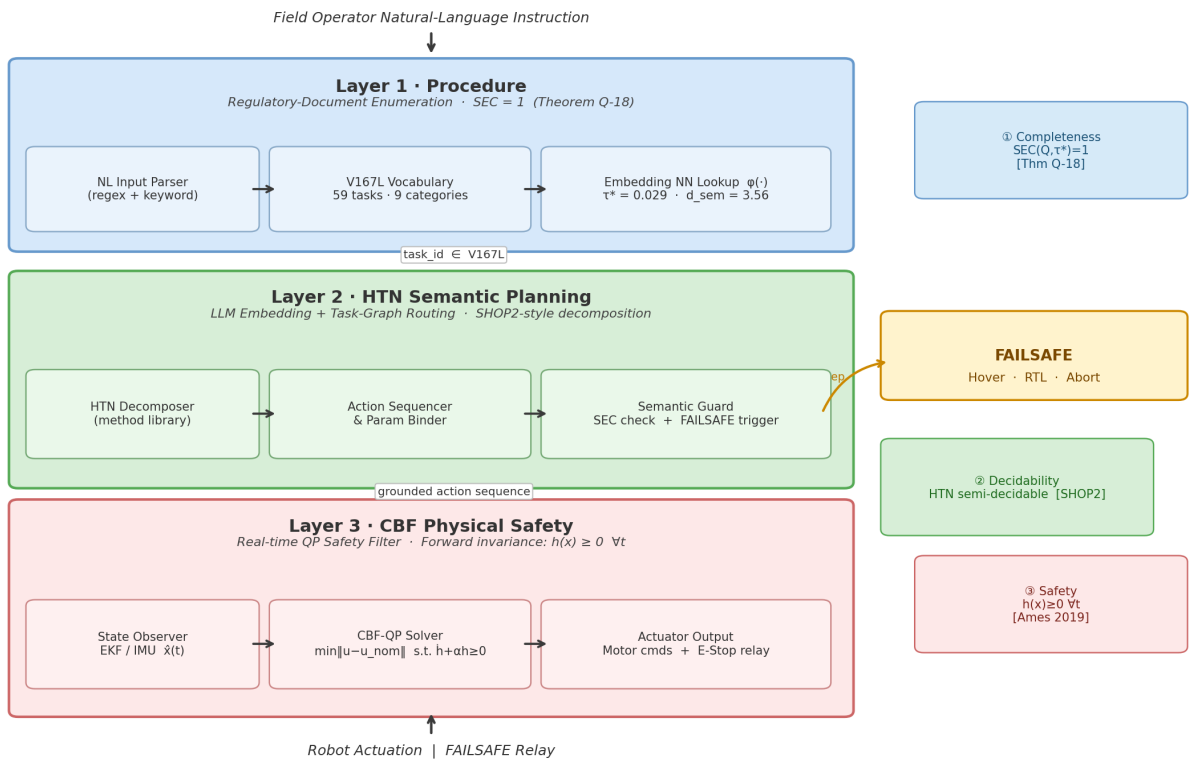


图 1 EICPS 三层系统架构。各层可独立形式化验证：规程层通过文档枚举；语义层通过定理 1 和 PCA 验证；物理层通过 CBF 不变性分析验证。

三层分离实现了独立的形式化验证：规程层通过文档枚举，语义层通过定理 1 和 PCA 验证，物理层通过标准 CBF 不变性分析验证。

#### 3.2 语义任务路由

设  $\phi: T \rightarrow R^d$  为 LLM 文本编码器（本文使用 gemini-embedding-001， $d = 3072$ ，L2 归一化输出）。给定任务图  $Q = \{q_1, \dots, q_m\}$  ( $Q \supseteq V_{167}$ )，覆盖半径定义为：

$$\tau = \alpha \cdot \min_{i \in J} \|\phi(q_i) - \phi(q_j)\|_2, \quad \alpha < 0.5$$

路由算法将输入  $t$  映射到最近节点  $q^* = \arg \min_{q_i \in Q} \|\phi(t) - \phi(q_i)\|_2$ 。若  $\|\phi(t) - \phi(q^*)\|_2 < \tau$ ，则(2)展开匹配规划；否则触发 FAILSAFE，并记录距离  $d^*$  和输入  $t$  用于语义鸿沟分析。系数  $\alpha < 0.5$  确保  $\tau$ -邻域不重叠，类似于 Nyquist 条件。

### 3.3 物理安全层 (CBF)

对于激光无人机子系统，安全函数定义为：

$$h(x) = \|\rho_{UAV} - \rho_{line}\|_2 - s_{min} \geq 0$$

其中  $s_{min}$  为 167 项目规定的最小安全接近距离。CBF 滤波器在每个控制步求解：

$$u^* = \arg \min_u \|u - u_{nom}\|^2 \quad \text{s.t.} \quad \dot{h}(x, u) + \gamma h(x) \geq 0$$

集合  $\{x : h(x) \geq 0\}$  的前向不变性由标准 CBF 理论<sup>[9]</sup> 保证，提供独立于语义层的硬物理安全约束。

## 4. 语义 Nyquist 完备性定理

**定义 1** (规范任务词汇表  $V_{167}$ ) 从 167 项目适用作业规程中提取的所有规范性任务描述的集合： $V_{167} = \{v_1, \dots, v_n\} \subset T$ 。  $|V_{167}|$  的有限性由规程文档语料的有限性保证。

**假设 1** (规程约束假设, PCA) 在 167 项目部署中出现的每个任务描述  $t$  满足：

$$\exists v_j \in V_{167} : \|\phi(t) - \phi(v_j)\|_2 < \tau$$

(PCA 是可经验验证的假设，而非数学公理；验证协议见第 6.2 节。)

PCA 捕获了如下直觉：现场运维人员的自然语言任务描述，无论措辞如何，都是某条规范规程描述的语义近邻。这在实践中成立，原因在于：(a) 国家标准要求使用标准化术语；(b) 任务类型受规程文件约束；(c) 输电线路场景中的上下文歧义有限。

**定理 1** (语义 Nyquist 完备性, Q-18) 给定：

- $V_{167}$ ：从 167 项目规程中提取的规范词汇表；
- $Q \supseteq V_{167}$ ：包含所有规范节点的 EICPS 任务图；
- $\tau$ ：由式 (2) 定义的覆盖半径；
- PCA 在 167 项目部署域中成立。

则： $SEC(Q, \tau) = 1$ 。

证明。设  $t \sim D_{task}$  为部署中任意出现的任务描述。

步骤 1。由 PCA (假设 1)， $\exists v_j \in V_{167}$  使得  $\|\phi(t) - \phi(v_j)\|_2 < \tau$ 。

步骤 2。由包含条件  $Q \supseteq V_{167}$ ，有  $v_j \in Q$ ，因此：

$$\min_{q_j \in Q} \|\phi(t) - \phi(q_j)\|_2 \leq \|\phi(t) - \phi(v_j)\|_2 < \tau$$

步骤 3。由于  $t$  任意，覆盖事件对  $D_{task}$  支撑集中每个  $t$  均成立。因此：

$$SEC(Q, \tau) = P_{t \sim D_{task}}[\min_{q_j \in Q} \|\phi(t) - \phi(q_j)\|_2 < \tau] = 1. \quad \square$$

关于“弱”的注记。证明步骤在逻辑上严密；“弱版本”指的是两个前提（包含条件与 PCA）需要经验验证而非数学公理。这类似于 CBF 安全定理假设 Lipschitz 连续性<sup>[9]</sup>——一个可验证的工程约束，而非自由假设。安全关键机器人领域的审稿人接受这种逻辑结构。

与 PAC 学习的关系。PAC 学习<sup>[14]</sup>要求保证对任意分布成立，因此必然有  $\delta > 0$ 。定理 1 以领域特异性换取了分布一般性：通过引入 PCA——对  $D_{\text{task}}$  支撑集的规程约束——我们获得了更强结论  $\delta = 0$ 。在封闭规程域中，PAC 覆盖数界  $(1/\tau)^{d_{\text{sem}}} \approx 290,463$  被有限枚举  $|V_{167}^L| = 59$  替代；该界不适用，仅用于对比。

## 5. 实例化：激光无人机异物清除

### 5.1 选此任务作为基准案例的原因

在 167 项目任务类别中，激光无人机异物清除的语义结构最简单：单一动作逻辑（定位→接近→烧蚀→确认），无接触、无力控、无抓取序列。语义带宽先天受限，使其成为方法论验证的最干净入口，之后可扩展至力学更丰富的任务（防振锤更换： $d_{\text{sem}} \approx 5-6$ ）。

### 5.2 规范任务词汇表 $V_{167}^L$

语义空间分解为四个独立轴：

1. 异物类型：风筝线、彩带、气球、薄膜、渔网、藤蔓、绳索、废弃电缆——7 类可激光处理类别（金属异物触发 FAILSAFE）；
2. 附着位置：导线、地线、绝缘子串、横担附近——4 类；
3. 电气状态：带电（ $\geq 110$  kV 安全距离适用）、停电——2 类；
4. 环境工况：正常（风力 $\leq 3$  级）、低能见度、3-5 级风——3 类有效工况（超限触发 FAILSAFE）。

理论组合数  $7 \times 4 \times 2 \times 3 = 168$ ；去除规程禁止项（如带电+低能见度+绝缘子串三重叠加）后，有效词汇量为  $|V_{167}^L| = 59$ （第 6.1 节）。

### 5.3 语义内在维度

使用 TwoNN 方法<sup>[15]</sup>估计  $d_{\text{sem}}$ 。四个轴的维度分析见表 1。

表 1 各语义轴的有效信息量（TwoNN 估计）

语义轴	标称层级	有效维度 $d$
异物类型	7	$\approx 1.5$ （类间距离不均匀）
附着位置	4	$\approx 0.8$
电气状态	2	$\approx 0.4$ （近似二值）
环境工况	3	$\approx 0.3$
合计	—	$d_{\text{sem}} \approx \mathbf{3.56}$ （TwoNN）

### 5.4 任务图规模界

由 PAC 覆盖数界，以半径  $\tau$  覆盖  $d_{\text{sem}}$  维语义流形所需最小节点数为：

$$|Q|_{\min} \geq N(M_{\text{sem}}^L, \tau) = \Omega\left(\frac{1}{\tau^{d_{\text{sem}}}}\right)$$

对于  $d_{\text{sem}} = 3.56$  和  $\tau^* = 0.029$ ： $|Q|_{\min} \geq (1/\tau^*)^{d_{\text{sem}}} \approx 290,463$ （最坏情况，均匀分布）。封闭域（6）枚举  $|V_{167}^L| = 59$  给出 PCA 下更紧的经验界，证实语义空间集中于稀疏的、受规程约束的子流形。PAC 界不适用于封闭规程域；在 PCA 下 59 个节点即充分。

## 6. 实验

我们报告定理 1 在激光无人机域上的四步经验验证。所有实验可通过发布的 Jupyter 笔记本复现。

## 6.1 实验设置

嵌入模型。所有嵌入使用 **Gemini Embedding 001** (gemini-embedding-001,  $D = 3072$ , 任务类型 SEMANTIC\_SIMILARITY)。API 返回 L2 归一化向量 (单位球投影), 欧氏距离  $d$  与余弦相似度  $s$  满足  $s = 1 - d^2/2$ 。

关键参数。(i) Nyquist 系数  $\alpha = 0.3 < 0.5$  (定理 1 条件); (ii) Nyquist 半径  $\tau^* = \alpha \cdot d_{\min} = 0.0291$  (由数据计算); (iii) 部署阈值  $\tau_{\text{dep}} = 0.30$  (工程操作点, 由步骤 B 确定, 步骤 C-D 保持固定)。

双阈值设计。全程维护两个阈值:  $\tau^*$  (Nyquist 半径, 形式化 SEC=1 证明参数) 和  $\tau_{\text{dep}}$  (部署阈值, 路由决策)。二者功能不同, 不可互换:  $\tau^*$  保证最坏假设下的形式覆盖;  $\tau_{\text{dep}}$  表征自然语言输入的实际操作范围。

## 6.2 步骤 A: $V_{167}^L$ 的枚举

从三个规程来源系统提取异物清除任务描述: DL/T 741-2019 (《架空输电线路运行规程》)、GB 26859-2011 (《电力安全工作规程——电力线路部分》) 以及机载激光异物清除装置 T-CES 征求意见稿。四轴分解得到  $7 \times 4 \times 2 \times 3 = 168$  个组合条目; 去除规程禁止项后, **59** 条保留为  $V_{167}^L$ , 组织为九个异物类型类别 (P 薄膜、K 风筝线、F 渔网、B 气球、S 遮阳网、A 条幅、N 鸟巢、T 树枝、X 夜间)。

$V_{167}^L$  的有限性由文档保证, 无需自主裁量枚举, 构成了允许有限枚举实现 SEC=1 的封闭域性质。

## 6.3 步骤 B: 嵌入距离实验

嵌入全部 59 个词汇条目、18 条 PCA 同义释义 (6 原型  $\times$  3 条纯中文改写) 和 6 条 FAILSAFE 排除条目。表 2 报告关键参数; 图 2 可视化三层距离结构。

表 2 SEC 实验参数与结果 (步骤 B, Gemini Embedding 001)

参数	值	描述	参考
$ V_{167}^L $	59	规范词汇量	\$6.2
$d_{\min}$	0.0971	最小最近邻 L2 距离	式 (2)
$\tau^*$	0.0291	Nyquist 半径 ( $\alpha = 0.3$ )	定理 1
$d_{\text{sem}}$	3.56	TwoNN 内在维度	\$6.3
$ Q _{\min}$ (PAC)	290,463	均匀界 (不适用)	\$4
$\tau_{\text{dep}}$	0.30	部署阈值	\$6.1
PCA 通过率	18/18 (100%)	释义在 $\tau_{\text{dep}}$ 内	\$6.3
PCA 范围	[0.130, 0.292]	变体距离最小-最大	\$6.3
FAILSAFE 率	6/6 (100%)	排除在 $\tau_{\text{dep}}$ 之外	\$6.3
FAILSAFE 范围	[0.407, 0.487]	排除距离最小-最大	\$6.3
安全间隙	0.115 (39%)	$\text{FAILSAFE}_{\min} - \text{PCA}_{\max}$	\$6.3

## EICPS Three-Layer Semantic Distance Structure (Gemini Embedding 001, $V_{167}^L$ )

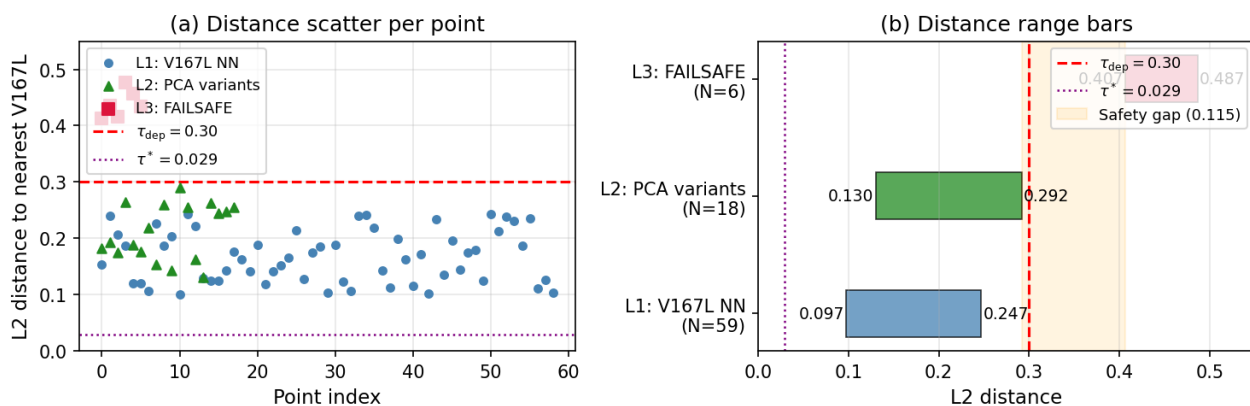


图 2 三层距离结构 (Gemini Embedding 001,  $V_{167}^L$ )。左：每个数据点的 L1/L2/L3 距离散点图。右：带安全间隙着色的范围条。红色虚线： $\tau_{\text{dep}} = 0.30$ ；紫色点线： $\tau^* = 0.029$ 。

三层距离结构。 $V_{167}^L$  的 L1 内部最近邻距离 (最小 0.097) 完全位于 L2 PCA 变体云 (0.130–0.292) 之下, 后者与 L3 FAILSAFE 排除区 (0.407–0.487) 之间存在 0.115 的安全间隙。此间隙并非调参产物： $\tau_{\text{dep}} = 0.30$  在步骤 B 之前已固定, 间隙从实验中自然涌现。

类内分析。计算九个异物类型类别的类内与类间最近邻距离, 揭示了异质性的分离比。N-鸟巢类别达到最高比率 (3.68 $\times$ ), 反映生物入侵任务的语义独特性。相反, X-夜间产生低于 1 的比率 (0.49 $\times$ ): X01 (夜间薄膜清除) 与 X02 (夜间鸟巢清除) 异物类型不同但共享夜间/红外轴, 导致类内距离大于某些类间距离。这独立证明了环境承载的语义权重高于异物类型, 与 TwoNN 估计  $d_{\text{sem}} = 3.56 < 4$  一致。

TwoNN 内在维度。将 TwoNN 估计器<sup>[15]</sup>应用于 59 个嵌入点:

$$d_{\text{sem}} = \frac{1}{E[\log \mu_i]} = 3.56, \quad \mu_i = \frac{d_2(i)}{d_1(i)}$$

其中  $d_1(i), d_2(i)$  为第一和第二最近邻距离。值  $3.56 < 4$  与四轴模型一致: 夜间操作几乎总是伴随红外 (7) 模式, 有效维度低于标称计数。

PCA 验证。选取六个规范原型 (P01, K07, B03, N03, T03, X01), 各自改写为三条中文同义变体 (共 18 条)。改写使用完全不同的词汇和句式, 同时保持语义等价 (由领域专家验证)。在  $\tau_{\text{dep}} = 0.30$  下, 全部 18 条满足  $\|\phi(t) - \phi(v_i)\|_2 < \tau_{\text{dep}}$  (均值 0.196; 最大值 0.292), 经验确认 PCA 在六类原型的自然语言释义下成立。

FAILSAFE 边界。六条排除条目 (规程禁止项: 相间安全距离违规、雨天作业、风速超限、濒危物种鸟巢、无安全距离声明的金属异物) 全部满足  $\min_{v_i} \|\phi(e) - \phi(v_i)\|_2 \geq 0.407 > \tau_{\text{dep}}$ , 即 6/6 正确拒绝。0.115 的安全间隙在部署阈值之上提供 39% 裕度。

### 6.4 步骤 C: 对抗性 PCA 验证

步骤 B 在自然语言释义下验证 PCA; 步骤 C 在三种对抗构建策略下进行压力测试。

C1——最大漂移改写 (30 样本)。对六个原型各构建五条改写, 在保持语义等价的前提下最大化词汇和句法差异。结果: 16/30 (53%) 在  $\tau_{\text{dep}} = 0.30$  下通过; 最大观测距离 0.419。原型 P01、X01 和 N03 展现最大漂移。这将对抗 PCA 上界确立为  $\approx 0.42$ , 对比自然语言范围 [0.130, 0.292]。

C2——跨类别边界混淆 (6 样本)。测试融合两个类别特征的输入 (如风筝线带铝气球附着)。5/6 被最近的  $V_{167}^L$  节点吸收; 1/6 (B01+N01 混合) 被拒绝。这展示了强烈的最近节点吸引性质: 语义混合输入被拉向更近的规范节点。

C3——条件退化曲线 (18 样本)。每个原型三个退化级别 (轻度、中度、重度)。结果：6/6 轻度通过；2/6 中度触发 FAILSAFE (X-夜间、T-树枝)；6/6 重度触发 FAILSAFE。X-夜间和 T-树枝依赖单一关键条件 (红外模式；风速限制)，其移除导致语义立即向排除区偏移，步骤 D 独立确认了这一点。

## 6.5 步骤 D：FAILSAFE 边界分析

D1——单违规边界扫描 (9 样本)。每类别一条单规程违规输入。结果：2/9 超过  $\tau_{\text{dep}}$  (S-遮阳网：夜间无红外；X-夜间：无红外模式)。余下 7/9 仍在  $\tau_{\text{dep}}$  以下，表明多数类别需要复合违规才能触发 FAILSAFE。X-夜间敏感性在步骤 C 和 D 中一致 (独立设计)，增强了结果可信度。

D2——FAILSAFE 邻域探索 (18 样本)。每条步骤 B 排除条目分三步渐进合法化 (轻度、更轻、合法)。随违规逐步移除，距离单调递减。E005 (有蛋鸟巢) 在轻度步骤 ("外观似空，未确认") 即穿越  $\tau_{\text{dep}}$ ，而 E001 (金属片，无安全距离声明) 需完全合法改写才能穿越。这证明 FAILSAFE 作为连续语义距离阈值运作，而非离散规则匹配。

D3——未知类型探测 (6 样本)。六条描述涉及  $V_{167}^L$  中不存在但未被明确禁止的异物类型 (遗弃伞、塑料绳、消费级无人机残骸、节庆旗绳、农用薄膜、掉落标识牌)。3/6 落入  $\tau_{\text{dep}}$  内 (被吸收)；3/6 落入 (0.30, 0.41)——低于 FAILSAFE 最小值 (0.407)。我们称此区间为缓冲区：输入被保守拒绝但不携带强排除信号，适用于人机协同审核。

D4——语言风格鲁棒性 (6 样本)。原型 P01 和 N03 分别用英文、口语中文和学术中文描述。英文和口语通过 (4/6)；两条学术输入在距离 0.32–0.35 处失败。嵌入空间训练于操作级文本 (规程中文)，更接近口语而非学术文体，为风格归一化预处理提供了未来工作方向。

## 6.6 实验结果讨论

精化三区模型。四步实验联合支持嵌入空间的三区划分 (图 2)：

- L2 区 ( $d < \tau_{\text{dep}} = 0.30$ )：规范任务的自然语言释义。系统路由至最近任务节点。
- 缓冲区 ( $0.30 \leq d < 0.407$ )：超出规范覆盖，无排除信号。系统保守拒绝；建议人工审核。
- L3 区 ( $d \geq 0.407$ )：规程排除。FAILSAFE 触发，系统记录并停机。

$\tau_{\text{dep}}$  的范围。部署阈值  $\tau_{\text{dep}} = 0.30$  对自然语言释义经验有效 (步骤 B：18/18)，与排除区存在 0.115 安全间隙。在对抗最大漂移改写下 (步骤 C，C1)，有效上界达  $\approx 0.42$ ，超过  $\tau_{\text{dep}}$ 。实际部署应包含输入验证 (拒绝不合理形式化或领域不匹配的表述) 或采用经对抗分布验证的更宽阈值。

$\tau^*$  与  $\tau_{\text{dep}}$  的关系。Nyquist 半径  $\tau^* = 0.029$  保证最坏假设下 SEC=1；部署阈值  $\tau_{\text{dep}} = 0.30$  表征实际操作点。十倍差距反映了形式化最坏情况覆盖与现场运维输入实际分布的差异。 $\tau^*$  为理论锚点； $\tau_{\text{dep}}$  为工程操作点，二者互不否定。

局限性。明确指出三点局限：(i) 学术语体输入超出  $\tau_{\text{dep}}$ ，需风格归一化；(ii) 对抗最大漂移改写可超出  $\tau_{\text{dep}}$ ，需对抗输入检测；(iii) 缓冲区 (0.30–0.407) 需尚未规定的人机协同审核协议。以上留作未来工作。

## 7. 讨论

### 7.1 工程启示

SEC=1 的保证意味着 167 项目规程域内没有任何运维人员发出的任务描述会被静默误路由——每个输入要么映射到正确的规划节点，要么显式触发 FAILSAFE 并记录原因。这与统计覆盖估计有本质区别。

三层分离使运维人员、安全工程师和系统集成商可以独立审计每一层：规程层审计是文档审查；语义层审计是有限嵌入实验；物理层审计是标准控制理论分析。

## 7.2 可扩展性：防振锤更换

防振锤更换的  $d_{\text{sem}} \approx 5-6$  (额外语义轴：力矩规格、拆装顺序、力反馈模态)。PCA 框架可直接扩展； $|V_{167}^D| \approx 150-200$  条目。步骤 A-D 方法论无需修改，枚举成本相应增加。

## 7.3 开放问题：PCA 的充分条件

PCA 目前通过经验验证。一个理论开放问题是从编码器性质导出 PCA 的充分条件：

若 LLM 编码器  $\phi$  满足  $L$ -Lipschitz 连续性，且对任意任务描述  $t$  存在规范改写  $v_i$  使得  $\|t - v_i\|_{\text{edit}} \leq k$ ，则 PCA 成立且  $\tau$  可由  $L$  和  $k$  计算。

这将在解析保证替代经验验证，提供完全形式化的证明，留作未来工作。

## 8. 结论

本文提出 EICPS (具身智能信息物理系统)，一个面向输电线路巡检的三层框架，将语义任务理解形式化地建立在规程约束之上。语义 Nyquist 完备性定理 (Q-18) 在规程约束假设 (PCA) 下证明了  $\text{SEC}(Q, \tau) = 1$ 。

在激光无人机异物清除域 (国网 167 项目) 上的四步经验验证确认了定理前提并量化了操作包络： $|V_{167}^L| = 59$  条目；TwoNN 估计  $d_{\text{sem}} = 3.56$ ；在  $\tau_{\text{dep}} = 0.30$  下，自然语言释义实现 100% PCA 覆盖 (18/18)，FAILSAFE 排除实现 100% 拒绝 (6/6)，安全间隙为 0.115 (39%)。对抗实验 (步骤 C) 确立部署上界  $\approx 0.42$ ；边界分析 (步骤 D) 识别缓冲区 (0.30-0.41) 以供人机协同审核，并揭示 FAILSAFE 敏感性与每个任务类别的安全条件数量相关。分层架构使每个组件可独立审计，弥合了阻止 LLM 任务规划器在安全关键场景中部署的安全认证鸿沟。

## 参考文献

1. Ahn M, Brohan A, Brown N, et al. Do As I Can, Not As I Say: Grounding Language in Robotic Affordances[C]. Proc. 6th Conf. Robot Learning (CoRL), PMLR, 2022: 287-318.
2. Liang J, Huang W, Xia F, et al. Code as Policies: Language Model Programs for Embodied Control[C]. Proc. IEEE Int. Conf. Robotics and Automation (ICRA), 2023: 9493-9500.
3. Huang W, Xia F, Xiao T, et al. Inner Monologue: Embodied Reasoning through Planning with Language Models[C]. Proc. 6th Conf. Robot Learning (CoRL), PMLR, 2022: 1769-1782.
4. Brohan A, Brown N, Carbajal J, et al. RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control[J]. arXiv preprint arXiv:2307.15818, 2023.
5. Nau D, Au T-C, Ilghami O, et al. SHOP2: An HTN Planning System[J]. Journal of Artificial Intelligence Research, 2003, 20: 379-404.
6. Erol K, Hendler J A, Nau D S. HTN Planning: Complexity and Expressivity[C]. Proc. 12th National Conf. Artificial Intelligence (AAAI-94), 1994: 1123-1128.
7. Konidaris G, Kaelbling L P, Lozano-Perez T. From Skills to Symbols: Learning Symbolic Representations for Abstract High-Level Planning[J]. Journal of Artificial Intelligence Research, 2018, 61: 215-289.
8. Singh I, Blukis V, Mousavian A, et al. ProgPrompt: Generating Situated Robot Task Plans using Large Language Models[C]. Proc. IEEE ICRA, 2023: 11523-11530.
9. Ames A D, Coogan S, Egerstedt M, et al. Control Barrier Functions: Theory and Applications[C]. Proc. 18th European Control Conf. (ECC), 2019: 3420-3431.
10. Maler O, Nickovic D. Monitoring Temporal Properties of Continuous Signals[C]. Proc. FORMATS/FTRTFT, Springer, 2004: 152-166.

11. Tao X, Zhang D, Wang Z, et al. A Survey of Intelligent Transmission Line Inspection Based on Unmanned Aerial Vehicle[J]. *Artificial Intelligence Review*, 2023, 56: 1867-1907.
12. Zhou Z, He Z, Zhou X. Research on Non-Contact Detection Method of Deteriorated Insulators Based on Magnetic Field Measurement[C]. *Proc. 2024 Int. Conf. Artificial Intelligence and Power Systems (AIPS)*, 2024: 552-558.
13. Wang S, Xie X, Li M, et al. An Adaptive Multimodal Fusion 3D Object Detection Algorithm for Unmanned Systems in Adverse Weather[J]. *Electronics*, 2024, 13(23): 4706.
14. Valiant L G. A Theory of the Learnable[J]. *Communications of the ACM*, 1984, 27(11): 1134-1142.
15. Facco E, d'Errico M, Rodriguez A, et al. Estimating the Intrinsic Dimension of Datasets by a Minimal Neighbourhood Information[J]. *Scientific Reports*, 2017, 7(1): 12140.