面向弹性多智能体系统的逻辑驱动语义通信

Tamara Alshammari，学生会员，*IEEE* 和 Mehdi Bennis，会士，*IEEE*

arXiv:2601.06733v1 [cs.MA] 11 Jan 2026

摘要——**6G网络的到来正在加速大规模、去中心化多智能体系统（MAS）的自主化、智能化和互联互通。尽管这一演进带来了高度自适应的行为，但也加剧了系统对环境变化、故障、错误信息和对抗行为等压力源的脆弱性。现有关于去中心化系统和多智能体系统韧性的研究大多聚焦于容错性或鲁棒性等孤立方面，未能提供多智能体韧性的系统化统一定义。这一缺口限制了设计能够在动态条件下持续感知、适应和恢复的系统的能力。本文提出了以两个互补维度为基础的MAS韧性形式化定义：认知韧性，即智能体恢复并维持对环境的准确认知；以及行动韧性，即智能体利用这些知识，在受到干扰时协调行动并维持目标。我们通过时序认知逻辑对韧性进行形式化，并以可恢复时间（扰动后重新建立所需属性的速度）和持久时间（恢复后准确信念和目标导向行为能够维持多久）对其进行量化。我们设计了智能体架构，并开发了实现认知韧性和行动韧性的去中心化算法。我们还提供了形式化验证保证，表明我们的规范相对于度量界限是可靠的，并支持有限时域验证，从而能够进行设计时认证和轻量级运行时监控。通过一个关于突发网络压力因素下分布式多智能体决策的案例研究，我们证明了所提出的方法始终优于基线方法。总体而言，我们的形式化验证分析和仿真结果表明，所提出的框架能够实现具有韧性的知识驱动型决策和持续运行，为下一代通信基础设施中的弹性去中心化多智能体系统奠定基础。**

索引词—**韧性、语义通信、多智能体系统、去中心化系统、6G及后续网络、可恢复性、持久性、认知逻辑、时序逻辑、模态逻辑、克里普克结构、克里普克语义。**

I. 引言

6G的到来预示着无处不在的自主性、分布式智能，以及设备、用户和基础设施之间的密集互联，从而形成能够协同感知、推理和适应的大规模去中心化多智能体系统（MAS）。然而，正是这种复杂性加剧了系统对各种压力源的暴露——从环境变化和组件故障，到错误信息级联和对抗性行为——使**韧性** 成为未来网络的一项首要性能指标 [1]–[5]。

然而，实现这种韧性不能依赖传统的集中式架构。随着 6G 系统向大规模和异构化方向发展，对集中式控制的依赖变得越来越不切实际。超密集设备部署、间歇性连接、严格的时延约束，以及隐私或管理边界，都限制了全局编排的可行性 [6]。集中式架构还会引入单点故障，并且天然容易受到通信时延、拥塞和攻击的影响 [7]–[8]。相比之下，去中心化多智能体系统 将智能与决策分布到各个智能体上，使其能够在本地进行感知、推理和行动，同时通过有限的通信进行协调 [9]–[10]。这种去中心化能够更快地适应局部扰动。因此，对于未来无线网络而言，任何严谨的韧性概念都必须建立在去中心化运行的基础上，使智能体即使在部分可观测性和动态网络条件下，也能维持功能与协同一致性。

研究动机。这一挑战并不局限于无线系统，还反映出一项更广泛的多学科难题：如何严谨地定义韧性 [11]–[19]。在无线系统中，韧性已成为下一代（6G）系统的核心关注点 [1], [20], [21], ，NSF 韧性与智能下一代系统（RINGS I–II）项目等倡议 [22] 以及人们认识到通信基础设施如今已成为关键社会公共设施、易受自然灾害和人为诱发的破坏影响 [21]，都凸显了这一趋势。在若干新兴的 6G 情景中，对快速实现分布式韧性的需求尤为突出。例如，在韧性车队编队中，道路状况意外变化、突发干扰或感知故障可能会暂时降低协调性能，但安全运行要求系统迅速恢复到标称行为。同样，用于灾害响应的分布式传感器网络可能遭遇传感器故障、信号阻塞或环境冲击，导致感知质量暂时下降，因此系统需要重新组织并恢复效能。类似地，无人机集群和机器人团队可能遇到动态障碍物、不断变化的传播环境或意外负载变化，导致运行效率暂时受损，因此需要快速且持续地恢复。然而，网络化系统中的韧性仍缺乏有原则的统一定义，导致对其理解不一，工程实践也参差不齐 [1]–[5]。韧性常与鲁棒性或可靠性混为一谈，因此有必要将它们区分开来。在 5G 中备受关注的可靠性，量化的是满足性能目标的概率，并关注时延和速率等指标的尾部风险统计数据 [23], [24]。鲁棒性方法通常依赖预先定义的不确定性模型，旨在优化不利条件下的性能。这通常通过极小极大优化、随机建模或风险敏感博弈等框架进行形式化，在这些框架中，系统的设计目标是

所有作者均来自芬兰奥卢大学无线通信中心（电子邮箱：{firstname.lastname}@oulu.fi）。

抵御最坏情况或概率场景 [25], [26]。两者都不可或缺，但都假设压力源已得到充分的先验描述。相比之下，韧性认为意外事件必将发生，并强调系统在线检测、定位和适应的能力，从而使系统即使面对模型失效和不断变化的条件，仍能实现其目标 [20]‚[25]。韧性定义方面这一缺口长期存在，根源在于缺乏严谨形式化和量化网络化系统韧性所需的成熟数学基础，从而限制了设计阶段的分析和运行时保证 [20], [21], [25]。

挑战与贡献。弥合这一差距不仅需要正式定义，还需要明确多智能体系统如何在不确定性下进行适应并维持功能的操作原则。事实上，面对未知压力源时，协调响应取决于持续更新共享态势图景，智能体在其中感知、通信，并随着情势变化相互核验，从而使决策依据不断更新的世界内部模型，而非静态计划 [20], [27]。在本研究中，**我们提出多智能体系统韧性的一种新的系统级定义**，并从两个相互关联的维度加以界定：(i) 认知韧性，指智能体检测知识缺口、采取认知行动弥补缺口、从其关于环境的内部模型（信念）与环境实际状态之间的偏差中恢复，并在恢复后维持准确内部模型的能力；(ii)行动韧性，指智能体根据更新后的知识迅速调整外部行动策略的能力，即实施恢复协议，重新建立当前信念与物理动作之间的有效一致性，并在恢复后维持最优外部行动策略。为使这一定义具有可操作性，我们引入直接对应认知循环与行动循环的指标：认知可恢复时间（多智能体系统在扰动后恢复准确、共享且一致的内部模型所需的时间）、认知持久时间（多智能体系统在认知恢复后维持准确内部模型的时长）、行动可恢复时间（多智能体系统在认知恢复后恢复最优外部行动策略所需的时间），以及行动持久时间（多智能体系统在行动恢复后维持这些最优策略的时长）。据我们所知，现有 6G/多智能体系统韧性文献并未提出这种针对网络化系统、具有可量化系统级指标的双重韧性形式化定义；相关研究通常仅探讨孤立机制，或采用宽泛且缺乏统一性的韧性概念 [1], [3], [20], [25]。

为了对这种系统级韧性进行形式化定义，并使其在多智能体系统中可量化，我们需要一种形式语言，用于：(i) 表示每个智能体对外部世界所知、不知以及认为可能的内容；(ii) 推理共同知识，以此刻画所有智能体都知道的内容，将其作为群体对齐的概念；(iii) 刻画知识/信念[1]如何在感知、通信和压力源的作用下演化；以及 (iv) 规定个体智能体层面和群体层面最优动作策略的执行。事实上，时序认知逻辑提供了用于个体和共同知识的算子并为其提供严谨的语义和更新机制，以建模信息演化 [28] [29]。至关重要的是，这些语义还支持形式化验证，从而能够在设计时或运行时检查多智能体系统是否在指定假设下满足韧性属性。这一能力对于认证恢复协议、知识更新规则和协调策略在部署前是否符合韧性要求，以及在动态环境中进行运行时监控，至关重要。

具体而言，我们使用克里普克结构实例化时序认知逻辑，为表示和演化多智能体知识提供语义基础 [30]–[34]。直观来说，克里普克模型列出所有可能世界，即环境中合理的配置。它会标注每个世界中哪些事实成立，并为每个智能体指定基于其可用信息而无法区分的世界。若某个事实在智能体认为可能的所有世界中都成立，该智能体便“知道”此事实；当所有智能体对某一事实达成一致时，就形成了共同知识。随着系统在感知、通信、故障或攻击影响下运行，模型本身也会更新：准确观测或信息丰富的消息会排除错误的可能性（收紧信念），而错误信息、传感器故障或对抗性行动则可能重新引入模糊性或矛盾（不准确的信念）。按时间顺序排列这些模型，自然就能描述系统运行过程中知识如何变化，而这正是时序认知逻辑所形式化的内容 [28]–[30], [35]。最后，这种基于克里普克模型的视角与AI原生6G中的语义通信直接相关：当一条消息能够降低接收方的认知模糊性时，它便具有价值 [36]–[43]。这种观点与韧性自然契合，因为降低模糊性能够加快认知恢复、稳定共享态势感知、促进行动恢复并提高持久性。换言之，优先降低与任务相关知识的不确定性的通信策略，能够直接增强认知韧性和行动韧性，使系统在压力源作用下实现更迅速、协同性更高的适应。

为此，我们将主要贡献概述如下：

- 我们提出了**多智能体韧性的形式化定义**，并基于两个互补维度引入可量化指标：基于认知的和基于行动的（**第四节**）。
- 我们在**时间认知逻辑**框架内**形式化**这些指标，并通过两个参数对其进行**量化**：**可恢复时间**和**持久时间**，它们分别刻画系统在遭受扰动后恢复期望属性的速度，以及系统维持准确信念和最佳性能的时长（**第四节**）。
- 为使这些指标能够付诸实践，我们设计了一种**支持语义通信的智能体架构**，其中包含相互耦合的内部组件和外部组件。内部组件利用克里普克结构对每个智能体的信念进行建模，同时通过逻辑消息交换和直接观测实现信念更新与对齐（**第二节**）。
- 我们开发了能够在去中心化多

[1]本文中，知识、信念和认知的 这几个术语交替使用。

智能体系统中同时实现认知韧性和行动韧性的**V 节**。

- 我们提供**形式化验证保证**，证明我们的韧性规范相对于所定义的度量界限是**可靠的**，并支持**有限时域验证**。这使得我们既能进行设计时认证，也能在扰动发生后进行轻量级运行时监控（**定理 1**和**推论 1**）。

- 最后，我们通过一个在突发网络压力因素下进行分布式决策的案例研究，**展示了该框架的有效性**，表明其**优于现有基线方法（第七节）**。

论文结构。本文其余部分安排如下。第二节介绍系统模型和所提出的智能体架构。第三节介绍时序认知逻辑框架和克里普克结构。第四节定义所提出的韧性度量，包括认知维度和行动维度，并从可恢复性和持久性方面对其进行定量描述。V 节详细介绍为实现认知韧性和行动韧性而设计的去中心化算法。第六节对所提出的度量进行形式化验证分析。第七节介绍并讨论数值仿真和性能评估。最后，第八节总结全文。

II. 系统模型与智能体架构

通信模型。考虑由 n 个智能体组成的系统，其通信图表示为无向图 $G = (N, E)$，其中 N 是智能体集合，基数为n，E 是表示智能体之间通信链路的边集合。N 中智能体 i 与 j 之间的一条边表示为无序对 $ij = ji \in E$。智能体 i 的邻居由集合 $V_i = \{j \in N$ 表示：$ij \in E$ 和 $j \neq i\}$。

智能体架构。在此通信层的基础上，下面将具体说明智能体的架构。受 [44], [45], 引入的信息转移系统（ITS）单智能体框架启发，每个智能体被建模为由两个相互耦合的子系统组成：外部物理环境 和内部信息处理单元，二者通过输入—输出信道进行交互。我们从两个关键方面扩展单智能体ITS框架。首先，内部处理单元 现包含一个结构化内部模型，该模型利用认知逻辑和克里普克语义捕获智能体的信念，并配备预测器和认知策略，用于更新这些信念并维持与环境的一致性。其次，我们引入一种协作式多智能体ITS表述，使智能体能够通过通信共同细化其内部模型。相比之下，原始单智能体ITS仅依靠交互轨迹历史来实现其内部模型。为使这些扩展更加具体，下面将转向介绍所提出架构的详细内容，该架构如图1所示。各组件定义如下：

外部环境。令 X 表示外部（物理）状态空间。令 U_i^{ext}表示其外部动作空间。联合动作空间定义为 $U_{\text{ext}} = \prod_{i=1}^{n} U_i^{\text{ext}}$，因此联合外部动作 $\boldsymbol{u}^{\text{ext}}$是一个 n-元组 $\boldsymbol{u}^{\text{ext}} = (u_1^{\text{ext}}, u_2^{\text{ext}}, \ldots, u_n^{\text{ext}}) \in U^{\text{ext}}$。外部动力学由以下内容描述：

$$f : X \times U^{\text{ext}} \times D \to X,$$

其中 D 表示扰动（如故障或攻击等压力源）。给定外部状态 $\boldsymbol{x}_t \in X$、联合外部动作$\boldsymbol{u}_t^{\text{ext}} \in U^{\text{ext}}$和扰动 $d_t \in D$后，下一个外部状态为 $\boldsymbol{x}_{t+1} = f(\boldsymbol{x}_t, \boldsymbol{u}_t^{\text{ext}}, d_t)$。在不存在扰动时，假定默认扰动 $d_o \in D$，以恢复标称动力学；即 f：$X \times U^{\text{ext}} \to X$。需要指出的是，该系统是分散式的：在每个时间步 t，智能体i 在本地选择其外部动作 $u_{i,\,t}^{\text{ext}}$，而无需集中式协调。

感知与通信。在每个时间步 t，每个智能体 i都会收集两类输入：

- 物理观测 $\boldsymbol{y}_{i,\,t} = s_i(\boldsymbol{x}_t) \in Y$ 其中 Y表示观测空间，s_i：$X \to Y$ 是智能体 i的传感器映射；
- 从邻近智能体收到的一组符号消息（$m_{j,\,t}$），$j \in V_i$ 这些消息传递了关于其当前信念的信息。

这些消息的精确逻辑形式将在引入底层认知语言后，于第三节中定义。

内部处理单元。内部处理单元封装了智能体的信念表征、预测能力和认知适应。它由三个关键组件构成：内部模型、预测器和认知策略（见图 1b）。下面详细介绍这三个组件。

内部模型。每个智能体都会维护一种内部表征，用以描述其当前对环境的信念或认为可能的情况。我们用 $\iota_{i,t} \in I_i$ 表示智能体 i 在时刻 t 的内部状态，其中 I_i 是该智能体可接受的信息（认知）状态集合。目前，我们抽象地将 $\iota_{i,t}$ 视为对智能体关于外部环境所知、不知以及认为可能的情况的结构化概述。在第三节-B 中，我们通过采用克里普克风格模型来具体化这一概念，即以此构造 I_i。在该解释下，$\iota_{i,t}$ 对应于一种可达结构，用于编码智能体认为可能的情境。

预测、认知策略和内部转移映射。为评估其内部模型与环境之间的一致性，智能体 i 使用以下方式预测下一次观测

$$\varrho_i : I_i \times U_i^{\text{ext}} \to Y, \qquad \hat{\boldsymbol{y}}_{i,t+1} = \varrho_i\big(\iota_{i,t}, \boldsymbol{u}_{i,t+1}^{\text{ext}}\big).$$

如果预测残差 $\Upsilon\big(\boldsymbol{y}_{i,t+1}, \hat{\boldsymbol{y}}_{i,t+1}\big)$ 在最近 L 个步骤内至少有 e_{th} 个时间步超过预定义阈值 ε_1，则预测观测 $\hat{\boldsymbol{y}}_{i,t+1}$ 与实际观测 $\boldsymbol{y}_{i,t+1}$ 之间可能出现矛盾，其中$\Upsilon(\cdot,\cdot)$ 是 Y 上的任意距离度量。智能体会根据任何已产生的矛盾，以及通过直接观测和/或邻近智能体发送的消息收到的任何新信息，更新其内部认知状态$\iota_{i,t}$。这些更新被建模为认知行动 $u_{i}^{\text{epi}}{}_{,t+1} \in U^{\text{epi}}$，其作用对象是

表 I 系统模型、时序—认知框架与韧性度量中使用的符号（第 II—IV 节）。

符号	描述
N	智能体集合，其中 $\|N\| = n$。
$G = (N, E)$	无向通信图。
E	（无向）通信链路集合。
V_i	智能体 i 的邻居集合：$\{ j \in N : ij \in E, \ j \neq i \}$。
X	外部（物理）状态空间；$\boldsymbol{x}_t \in X$ 在时间 t。
D	扰动空间；$d_t \in D$ 在时间 t；d_o 默认（标称）。
U_i^{ext}	智能体 i 的外部动作空间；$U^{\text{ext}} = \prod_{i=1}^{n} U_i^{\text{ext}}$.
$\boldsymbol{u}^{\text{ext}} = (u_1^{\text{ext}}, u_2^{\text{ext}}, \ldots, u_n^{\text{ext}})$	联合外部动作；$u_i^{\text{ext},t}$ 是智能体 i 在时间 t 的动作。
$f : X \times U^{\text{ext}} \times D \to X$	外部转移映射。
Y	观测空间；$s_i : X \to Y$；$\boldsymbol{y}_{i,t} = s_i(\boldsymbol{x}_t)$。
M_i	智能体 i 可从邻居 j接收的（有限）消息序列集合；$m_{j,t}$。
I_i	内部（认知）状态空间；$\iota_{i,t} \in I_i$。
$\hat{\boldsymbol{y}}_{i,t+1}$	时间 $t{+}1$ 的预测观测。
U^{epi}	认知行动（细化、修订、探索、广播、保持）的集合；$u_i^{\text{epi},t+1} \in U^{\text{epi}}$。
ϱ_i	预测映射。
π_i^{epi}	认知策略。
ψ_i	内部转移。
π_i^{ext}	外部策略。
T	离散时间域 $\{0, 1, 2, \ldots\}$。
Ω	原子命题集合；$\mathcal{L}_n^{\text{time}}(\Omega)$ 时序认知语言。
$\mathcal{G}_{[0,\beta)}$, $\mathcal{U}_{[0,\alpha]}$	时序算子“全局”和“直到”。
$\mathcal{K}_i$, $\mathcal{P}_i$	知识算子和可能性算子。
$\mathcal{E}_N\varphi$	智能体集合 N 上的共同知识（用于韧性公式）。
$\mathcal{M}_t$	克里普克结构 $(W, \{R_{i,t}\}_{i\in N}, v)$。
r	运行 $r : T \to W$；w_t 时间 t 的实际世界。
$\mathcal{R}_{\text{system}}$	系统级韧性规范，定义为 $\mathcal{R}_{\text{system}} \equiv \mathcal{R}_{\text{epistemic}} \wedge \mathcal{R}_{\text{action}}$。
$\mathcal{R}_{\text{epistemic}}$	认知韧性规范；参见（2）。
$\mathcal{R}_{\text{action}}$	行动韧性规范；参见（5）。
t_v	环境变化（压力源）发生的违例时刻。
$t_{\text{rec}}^{\text{epi}}$	认知恢复时间：智能体在认知上恢复的第一个时间步。
$\Delta t_{\text{rec}}^{\text{epi}}$	认知可恢复时间间隔：认知恢复所需的时间。
$t_{\text{dur}}^{\text{epi}}$	认知持久性的结束时间：拥有准确信念的最后一个时间步。
$\Delta t_{\text{dur}}^{\text{epi}}$	认识持久性区间：智能体拥有准确信念的时间区间。
$t_{\text{rec}}^{\text{act}}$	行动恢复时间。
$\Delta t_{\text{rec}}^{\text{act}}$	行动可恢复性区间：智能体恢复其策略的第一个时间步。
$t_{\text{dur}}^{\text{act}}$	行动持久性的结束时间：执行最优策略的最后一个时间步。
$\Delta t_{\text{dur}}^{\text{act}}$	行动持久性区间：智能体执行最优策略的时间区间。
α_1	允许的最大认识持久性恢复时间上界（用于（2））。
β_1	最低所需认识持久性持续时间（用于（2））。
α_2	允许的最大行动恢复时间上界（用于（5））。
β_2	最低所需动作持续时间（用于（5））。
π_i^{opt}	原子命题：智能体 i 当前在当前环境下采取最优行动；$\pi^{\text{opt}} \equiv \bigwedge_{i\in N} \pi_i^{\text{opt}}$.

认知状态$\iota_{i,t}$，以生成修订后的状态$\iota_{i,t+1}$。典型的认知行动汇集于集合 U^{epi} 中，包括：

1）细化：根据新信息排除不兼容的可能性。2）修订：调整内部状态，重新引入能够适应变化条件或证据并解决矛盾的新可能性。3）探索：主动探查环境，以减少不确定性。4）广播：与邻居共享选定的信息。5）保持：在适当情况下保持内部状态不变。

这些认知行动由每个智能体 i 根据认知策略 π_i^{epi} 选择，该策略根据智能体当前的认知状态$u_i^{\text{epi},t+1}$ 确定合适的认知行动

状态 $\iota_{i,t}$、实际观测 $\boldsymbol{y}_{i,t+1}$、预测观测 $\hat{\boldsymbol{y}}_{i,t+1}$和接收到的消息 $(m_{j,t+1})_{j\in V_i}$。令 M_i 表示智能体 i能够接收的（有限）消息序列集合（其形式化结构将在第三节中介绍）。认知策略是一个映射

$$\pi_i^{\text{epi}} : I_i \times Y \times Y \times M_i \to U^{\text{epi}},$$

用于选择动作

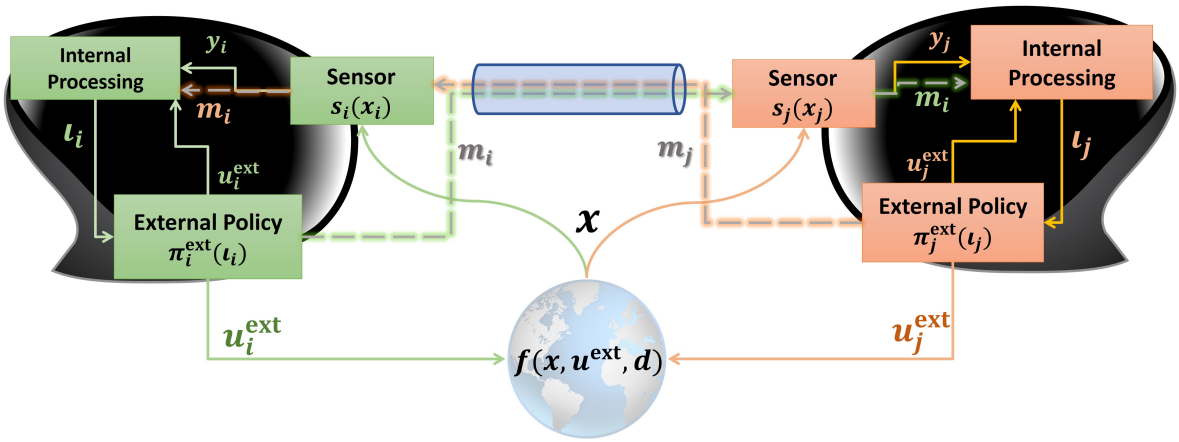
$$u_{i,t+1}^{\text{epi}} = \pi_i^{\text{epi}}\big(\iota_{i,t}, \boldsymbol{y}_{i,t+1}, \hat{\boldsymbol{y}}_{i,t+1}, M_{i,t+1}\big).$$

内部转移映射

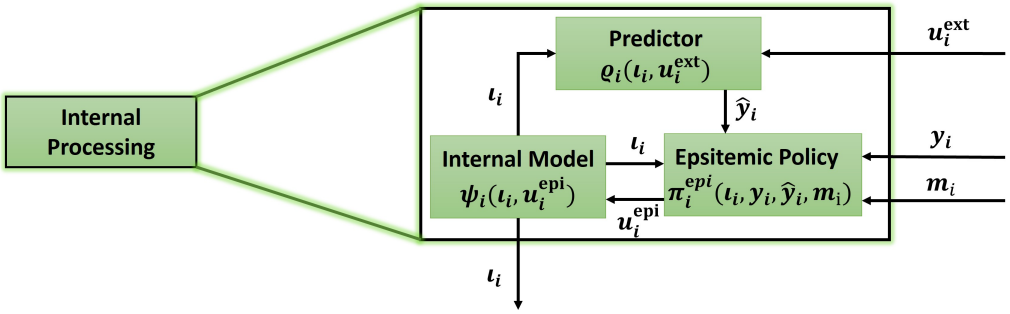
$$\psi_i : I_i \times U^{\text{epi}} \to I_i$$

随后生成下一个内部状态，$\iota_{i,t+1} = \psi_i(\iota_{i,t}, u_i^{\text{epi},t+1})$。

最后，为闭合信念与行为之间的循环，我们将外部策略定义如下。



(a) 提出的智能体架构。



(b) 智能体 i的详细结构 “内部处理” 单元，位于 所提出的 智能体架构。

图1. 所提出的智能体架构，展示了两个智能体如何相互通信以细化各自的内部模型。虚线通信箭头表示通信不一定持续进行，而是可以按需发生。智能体 j 的“内部处理”单元与 (b) 中所示的单元结构相似。

外部策略。 最后，智能体 i 通过外部策略将其当前认知状态转换为物理动作

$$\pi_i^{\text{ext}} : I_i \to U_i^{\text{ext}}, \qquad \boldsymbol{u}_{i,t+1}^{\text{ext}} = \pi_i^{\text{ext}}(\iota_{i,t}).$$

在确立智能体架构及其组成部分后，我们现在介绍用于表示每个智能体内部模型的形式化工具。具体而言，我们介绍时序认知逻辑框架和克里普克语义，它们构成了智能体内部认知状态与信念更新的结构基础。

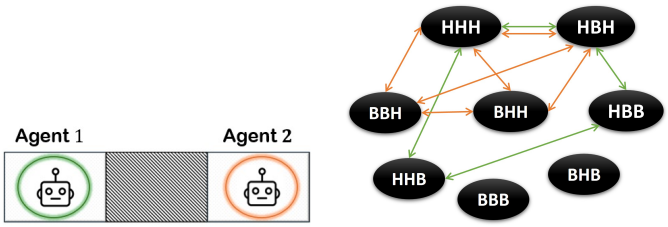
III. 认知逻辑与语义通信

在上一节中，我们介绍了智能体的内部模型，并描述了它们的认知状态如何通过预测、观测和通信而演变。现在，我们使用时序认知逻辑 和克里普克语义中的工具，对这些内部表征和通信机制进行形式化。时序认知逻辑提供了一个符号框架，用于推理智能体随着时间推移所知道、不知道或认为可能的事物；而克里普克结构则提供语义基础，使这些概念获得具体解释。在这一框架下，每个智能体 i的内部模型对应于克里普克结构的一个时序智能体切片，它编码了该智能体从信息视角出发对特定时刻世界的认识。智能体之间的通信因此被建模为不只是原始数据交换，而是逻辑陈述的传递，这些陈述的真值是相对于不断演变的克里普克结构进行评估的；这样捕捉的是消息的语义内容，而非其句法形式。

本节首先回顾时序认知逻辑的必要背景，介绍用于表达知识和信念随时间演变的公式的逻辑语言。随后，我们介绍克里普克结构作为知识的形式模型，并说明克里普克语义如何为这些结构中的逻辑公式赋予意义。接下来，我们基于克里普克结构定义智能体的内部认知空间，最后描述如何将智能体通信视为基于逻辑的语义交互。

A. 背景

在多智能体系统中，每个智能体对环境的信念如何动态演变，与环境的客观真相同样重要。经典命题逻辑将陈述判断为真或假，但无法涵盖（i）不同智能体之间的信息视角差异（认知层面），或（ii）这些视角如何随时间变化



(a) 环境设置。 (b) 克里普克结构。

图 2. 示例环境及其对应的克里普克结构。在 (a) 中，智能体 1（绿色）位于单元格 1，智能体 2（橙色）位于单元格 2。在 (b) 中，节点表示可能世界，绿色和橙色的边分别表示智能体 1 和智能体 2 的可达关系。为清晰起见，省略了自反箭头。

（时间层面）。为解决这些局限，我们采用一种时序认知逻辑，将多智能体认知算子与时序算子相结合。这种逻辑建立在一种形式语言之上，用于表达有关知识与信念的公式，并通过克里普克结构和克里普克语义进行解释。

原子命题与语言。 设 Ω 是一个非空的原子命题集合，离散时间域为 $T = \{0, 1, 2, \ldots\}$。我们还采用约定 $\mathbb{N} = \{0, 1, 2, \ldots\}$。尽管 T 和 $\mathbb{N}$ 表示相同的底层集合，但我们使用 T 表示绝对时间索引（$t \in T$），使用 $\mathbb{N}$（分别为 $\mathbb{N}_{>0}$）表示持续时间/步数范围，以便区分时间戳与长度。基于这些约定，时序—认知语言 $\mathcal{L}_n^{\text{time}}(\Omega)$ 由以下语法生成，其中每个表达式 φ 表示语言 [28] [29]的一个合式公式：

$$\varphi ::= \top \mid p \mid \neg\varphi \mid \varphi_1 \vee \varphi_2 \mid \mathcal{G}_{[0,\beta)}\,\varphi \mid \varphi_1 \mathcal{U}_{[0,\alpha]} \varphi_2 \mid \mathcal{K}_i \varphi,$$

其中，符号 $::=$ 表示通过语法构造进行定义，在形式文法中通常用于规定如何构造合式公式。它读作“定义为”或“可构造为”，竖线 $\mid$ 用于分隔不同的构造方式。此外，符号 $\top$ 表示永真公式，$p \in \Omega$，$\neg$ 表示否定，$\vee$ 表示析取，$i \in N$，而 $\alpha, \beta \in \mathbb{N}_{>0}$ 表示步数范围。合取 $\varphi_1 \wedge \varphi_2$ 和蕴含 $\varphi_1 \to \varphi_2$ 等标准联结词并非文法中的原始符号，而是通过否定 $\neg$ 和析取 $\vee$ 定义如下：$\varphi_1 \wedge \varphi_2 \equiv \neg(\neg\varphi_1 \vee \neg\varphi_2)$，以及 $\varphi_1 \to \varphi_2 \equiv \neg\varphi_1 \vee \varphi_2$。上述时态和认知算子在时间 $t \in T$ 处的解释如下：

- **时序全局算子** $\mathcal{G}_{[0,\beta)}\varphi$：表示 φ 在有界区间 $[t, t+\beta)$ 上全局成立。
- **时序直到算子** $\varphi_1 \mathcal{U}_{[0,\alpha]} \varphi_2$：表示 φ_1 持续成立，直到 φ_2 变为真；其中 φ_2 必须在某个时刻 $t'' \in [t, t+\alpha]$ 发生，并且 φ_1 在所有 $t' \in [t, t'')$ 上都成立）。
- **认知知识算子**[2] $\mathcal{K}_i\varphi$：表示智能体 i 在时刻 t 知道 φ；也就是说，在所有与智能体 i 在时刻 t 的信念相一致的**所有**可能情景中，公式 φ 都成立。

由知识算子 $\mathcal{K}_i\varphi$，我们可以推导出其对偶算子，即可能性算子 $\mathcal{P}_i\varphi \equiv \neg\mathcal{K}_i\neg\varphi$。该算子表示智能体 i 在时刻 t 认为 φ 可能成立；也就是说，存在**某些**情景，智能体 i 在时刻 t 认为这些情景可能发生，且在这些情景中 φ 为真。此外，设 $N_0 \subseteq N$ 为智能体的一个子集。我们可以推导出共同知识算子 $\mathcal{E}_{N_0}\varphi$，它表达了“群体 N_0 中的每个人都知道 φ”这一概念，定义如下：

$$\mathcal{E}_{N_0}\varphi = \bigwedge_{i \in N_0} \mathcal{K}_i\varphi.$$

克里普克结构。 为了赋予语言 $\mathcal{L}_n^{\text{time}}(\Omega)$ 中的公式以含义，并刻画智能体信念随时间的演变，我们采用如下时变克里普克结构 [30], [31]

$$\mathcal{M}_t = \big(W, \{R_{i,t}\}_{i \in N}, v\big),$$

其中：

- W 是一个非空的可能世界集合，其中每个世界代表环境的一种不同情形（配置）。
- 对于每个智能体 $i \in N$ 和时刻 $t \in T$ $R_{i,t} \subseteq W \times W$ 是智能体 i 在时刻 t 的认知可达关系。直观地说，如果 $(w, w') \in R_{i,t}$，这意味着 w' 是从 w 可访问的世界，那么在时刻 t，智能体 i 无法区分 w' 和 w；二者都与 i 的信息相符。因此，智能体可访问的世界集合编码了其知识中的不确定性。
- $v: W \to 2^{\Omega}$ 是赋值函数，它为每个世界 $w \in W$ 指定在 w 中为真的原子命题集合。

克里普克语义。 给定克里普克结构 $\mathcal{M}_t = (W, \{R_{i,t}\}_{i \in N}, v)$，时序认知语言 $\mathcal{L}_n^{\text{time}}(\Omega)$ 中公式 φ 的真值通过评价点上的满足关系 定义。评价点是一个三元组 $(\mathcal{M}_t, r, t)$，其中 $r: T \to W$ 是一次运行；也就是说，它是一个时间索引轨迹 $\langle r(0)\ r(1)\ \ldots\rangle$，且 $w_t = r(t)$ 是时刻 t [46]的实际世界。直观而言，在所有可容许轨迹中，r 是系统实际遵循的轨迹（例如，由外部动力学和扰动所诱导的轨迹）；因此，在每个时刻 t，它都确定当前世界 w_t。我们用 $(\mathcal{M}_t, r, t) \models \varphi$ 表示 φ 在时刻 t 沿着 r 于 $\mathcal{M}_t$ 克里普克语义 中为真。满足关系由以下子句归纳定义；这些子句共同构成了我们时序认知语言的克里普克语义。

- $(\mathcal{M}_t, r, t) \models \top$ 始终成立。
- $(\mathcal{M}_t, r, t) \models p$ 当且仅当 $p \in v(w_t)$。
- $(\mathcal{M}_t, r, t) \models \neg\varphi$ 当且仅当 $(\mathcal{M}_t, r, t) \not\models \varphi$。

[2]虽然句法 $\mathcal{K}_i\varphi$ 中没有明确的时间索引，但在我们的时间设定中，公式 φ 是在当前时刻 t 进行评估的。直观来说，“智能体 i 在时刻 t 知道 φ”意味着，根据 i 在 t 可获取的信息，i 仍认为可能的所有情形都会使 φ 为真。我们将在定义 Kripke 结构和 Kripke 语义时介绍使这一概念（包括其对时间的依赖）精确化的形式满足关系。

- $(\mathcal{M}_t,\ r,\ t) \models \varphi_1 \vee \varphi_2$ 当且仅当 $(\mathcal{M}_t,\ r,\ t) \models \varphi_1$ 或 $(\mathcal{M}_t,\ r,\ t) \models \varphi_2$。
- $(\mathcal{M}_t,\ r,\ t) \models \mathcal{G}_{[0,\beta)}\varphi$ 当且仅当对于所有满足 $t \le t' < t+\beta$ 的 t'，都有 $(\mathcal{M}_{t'},\ r,\ t') \models \varphi$。
- $(\mathcal{M}_t,\ r,\ t) \models \varphi_1 \mathcal{U}_{[0,\alpha]} \varphi_2$ 当且仅当存在 $t'' \in [t, t+\alpha]$，使得 $(\mathcal{M}_{t''},\ r,\ t'') \models \varphi_2$，并且对于所有 $t' \in [t, t'')$，都有 $(\mathcal{M}_{t'},\ r,\ t') \models \varphi_1$。
- $(\mathcal{M}_t,\ r,\ t) \models \mathcal{K}_i\varphi$ 当且仅当对于所有满足 $(w_t,\ w') \in R_{i,\ t}$ 的 $w' \in W$，

$$(\mathcal{M}_t, r[t \mapsto w'], t) \models \varphi,$$

其中，$r[t \mapsto w']$ 是与 r 在所有时刻都一致的运行，唯一不同之处是它将 t 映射到 w'。

在本文中，我们将通过一个连续示例来阐明所讨论的概念。

示例 1（双智能体网格世界的克里普克结构）。考虑两个智能体，它们位于一个由单元格组成的 3×1 网格$c \in \{1,2,3\}$ 中（见图 *2a*）。智能体1（绿色）占据单元格1，智能体2（橙色）占据单元格3。每个智能体 i 只能观察当前单元格的颜色 c。每个单元格都标记有两个互斥的原子命题之一：

$$H_c \ \text{(“cell } c \text{ is white”)}, \qquad B_c \ \text{(“cell } c \text{ is black”)}.$$

所有原子命题的集合为Ω。环境可能处于四种配置（即可能世界）之一：

$$W = \{\, w^{(HHH)},\, w^{(HBH)},\, w^{(BHH)},\, w^{(BBH)},\ w^{(HHB)},\, w^{(HBB)},\, w^{(BHB)},\, w^{(BBB)} \}$$

其中，上标表示单元格1、2、3的颜色。例如，$w^{(HHB)}$ 表示单元格1 和2 为白色，单元格3 为黑色。赋值函数 v 列出了每个世界中为真的原子命题：

$$v(w^{(HHH)}) = \{H_1, H_2, H_3\},\ v(w^{(HBH)}) = \{H_1, B_2, H_3\},$$
$$v(w^{(BHH)}) = \{B_1, H_2, H_3\},\ v(w^{(BBH)}) = \{B_1, B_2, H_3\},$$
$$v(w^{(HHB)}) = \{H_1, H_2, B_3\},\ v(w^{(HBB)}) = \{H_1, B_2, B_3\},$$
$$v(w^{(BHB)}) = \{B_1, H_2, B_3\},\ v(w^{(BBB)}) = \{B_1, B_2, B_3\}.$$

在此示例中，假设实际世界为 $w^{(HBH)}$。图 *2b* 展示了对应的克里普克结构：节点表示可能世界，边表示可达关系 $\{R_i\}_{i=1}^2$，这些关系为每个智能体i指定其可达世界集合（即与智能体 i的观测相一致的配置）。智能体1的可达世界为 $\{w^{(HHH)}, w^{(HBH)}, w^{(HBB)}, w^{(HHB)}\}$，因为只有这些世界与其观测相一致。由于 H_1 在智能体1的所有 可达世界中都成立，$\mathcal{K}_1 H_1$ 成立。由于智能体 1 不知道单元格2 和3的颜色，因此它认为命题 $\{H_2,\ B_2,\ H_3,\ B_3\}$是可能的，因为这些命题在其某些可达世界中成立。换言之，$\mathcal{P}_1 B_2$、$\mathcal{P}_1 H_2$、$\mathcal{P}_1 H_3$和 $\mathcal{P}_1 B_3$ 成立。同样，由于智能体2观测到 $c = 3$，其可达世界为 $\{w^{(HBH)}, w^{(BBH)}, w^{(HHH)}, w^{(BHH)}\}$，因此 $\mathcal{K}_2 H_3$、$\mathcal{P}_2 H_1$、$\mathcal{P}_2 B_1$、$\mathcal{P}_2 H_2$和 $\mathcal{P}_2 B_2$ 成立。

B. 内部认知空间

在第二节中，我们将每个智能体的内部状态$\iota_{i,\ t}$ 介绍为一种抽象表示，描述智能体对外部环境的认知或认为可能的情况。现在，我们将在上述克里普克框架中为这一概念提供形式化基础。

对于每个智能体 $i \in N$，其内部认知空间定义为

$$I_i \ = \ \big\{\, R \subseteq W \times W \ \big|\ R \text{ satisfies chosen frame conditions} \,\big\},$$

其中，$R_{i,t} \in I_i$ 表示智能体 i 在时刻 t的认知可达关系。此外，框架条件是指施加于可达关系、用于约束认知行为的结构属性。常见示例包括自反性（$\forall w : (w,w) \in R_{i,t}$，表示实际世界始终是可能的）和传递性（$(w,w') \in R_{i,t}$且$(w',w'') \in R_{i,t}$蕴含$(w,w'') \in R_{i,t}$，表示不可区分性具有传递性）。直观而言，$(w,w') \in R_{i,t}$ 表示，基于时刻 t掌握的信息，当实际世界为 w时，智能体 i 认为世界 w'是可能的。因此，$R_{i,t}$ 刻画了智能体在该时刻对不同世界的不确定性或不可区分性。由此，智能体在时刻 t 的内部状态由其可达关系$\iota_{i,t} = R_{i,t} \in I_i$给出，该关系编码了与智能体当前知识相一致的所有世界。在这种表示下，I_i 为建模智能体如何推理并随时间更新知识奠定了结构基础。

内部认知状态更新。回顾第二节，智能体 i 会通过认知行动 u_i^{epi}，U^{epi} $t+1 \in$，根据出现的矛盾以及通过观测和/或消息收到的新信息更新其内部状态。在 Kripke 表示下，这相当于更新可达关系 $\iota_{i,\ t} = R_{i,\ t} \in I_i$，具体如下：

$$R_{i,t+1} \ = \ \psi_i\big(R_{i,t},\, u_{i,t+1}^{\text{epi}}\big),$$

其中，$\psi_i :\ I_i \times U^{\text{epi}} \to I_i$ 是内部转移映射。典型动作对 ψ_i 的实例化如下：

- 细化：移除与新证据不一致的世界的可达关系，从而实现单调递减$R_{i,\ t+1} \subseteq R_{i,\ t}$。
- 修订：在与证据一致的世界之间重新连接可达关系（可以添加新的可达世界或移除现有世界），同时遵循最小变化原则。
- 保持/广播/探索：分别保持$R_{i,\ t}$ 不变、通知邻居，或触发后续证据获取。

更新 ψ_i 必须满足：(i)保留框架条件（例如自反性、传递性），以确保$R_{i,\ t+1} \in I_i$；(ii) 恢复一致性，使其与当前证据（在 t+1处的观测以及收到的消息）相符；以及 (iii) 满足最小变化原则（除确保一致性所需之外，不进行不必要的可达关系添加或移除）。这样，第二节中的抽象更新策略便具体实现为对 Kripke 可达关系的操作。

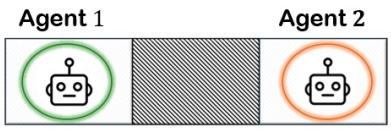
备注1（形式逻辑与操作算法之间的直观映射）。 为了直观地联系

第三节-*A*中引入的认知概念与本文介绍并在后续*V*节中使用的操作流程，我们简要概述克里普克模型在实际决策中的作用。每个世界代表一种候选环境条件，而可达关系则编码了智能体根据迄今为止的观测仍认为哪些世界是可能的（称为可及世界）。当新信息到来，且该信息并未使任何先前不可及的世界变得可能，而是排除了一些当前可及的世界时，智能体会准确移除那些不再与数据相符的世界（细化）。然而，如果智能体基于所有当前可及世界作出的预测都无法解释新观测到的数据，就会产生预测误差，促使智能体收集更多证据，以区分不同世界（探索）。随着此类证据不断累积，不相容的世界被逐步排除，智能体可能会发现必须重新组织其可能世界集合；也就是说，先前不可及的世界再次变得合理。在这种情况下，可达关系将被重构（修订），智能体也会将由此产生的信念变化告知邻居（广播）。当既无须细化也无须修订时，智能体只需保持当前的可能世界集合不变（保持）。由此，抽象的克里普克语义为后续各节算法所实现的检测、证据收集、修订和通信行为提供了概念基础。

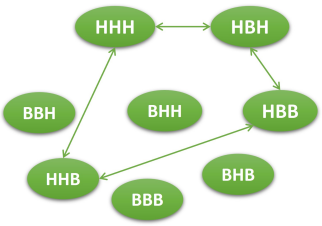
备注 2（作为压力源的碰撞）。尽管本文未对碰撞进行显式建模，但该框架能够自然地将其纳入其中。在这一形式体系中，其他智能体的行为被视为环境动力学的一部分，因此通过可达世界集合编码在每个智能体的认知状态中。如果发生碰撞（例如，两个智能体选择了不兼容的动作），由此产生的观测将与所有此类碰撞本不应发生的世界所作出的预测相矛盾。这种不一致被视为预测误差。随后，智能体会更新其可达关系，仅保留与观测到的交互一致的世界，并相应地重新调整其外部策略。这样，潜在的动作碰撞便作为环境压力因素，由本节提出的同一架构加以处理。

C. 通过认知逻辑和克里普克结构实现语义通信

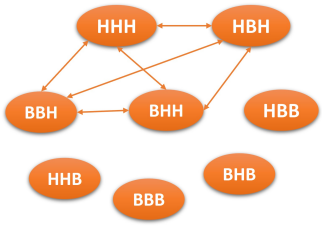
传统通信系统主要关注通过信道准确传输符号，并优化比特错误率和吞吐量等指标。这些系统将信息视为句法实体，其意义或对接收方的相关性在很大程度上被忽略。相比之下，语义通信将重点从符号保真度转向意义保真度，优先确保成功传达预期意义和与任务相关的信息。实现语义通信需要：（i）发送智能体传输能够保留意义的压缩表示；（ii）接收智能体



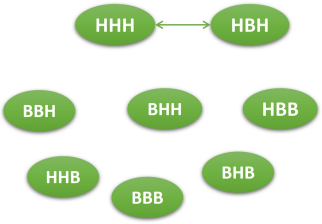
（a）环境设置。



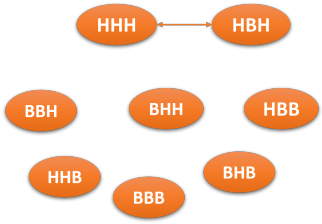
（b）智能体 1 在$t = 0$处的 $\mathcal{M}_0$ 切片（通信之前）。



（c）智能体 2 在 $\mathcal{M}_0$ 时的切片 $t = 0$（通信之前）。



（d）智能体 1 的切片 $\mathcal{M}_1$ 在 $t = 1$（通信之后）。

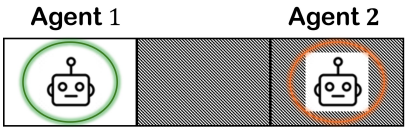


（e）智能体 2 在 $\mathcal{M}_1$的切片在 $t = 1$（通信之后）。

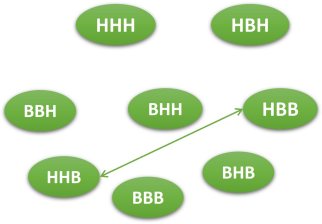
图 3. 示例环境及共享 Kripke 结构中每个智能体对应的切片，分别展示通信前后。智能体的切片是特定于智能体 i 的三元组（W，R_i，t，v）。节点表示可能世界，边表示特定于智能体的可达关系。为简洁起见，省略了世界上的自反自环。

在适当的语境中解读这些消息；（iii）通信由当前任务或目标驱动，而不是不加区分地传输所有可用数据。我们使用时序认知语言 $\mathcal{L}_n^{\text{time}}(\Omega)$及其语义所对应的共享 Kripke 结构，以一种有原则的方式实现语义通信。具体而言：（i）发送方传输来自 $\mathcal{L}_n^{\text{time}}(\Omega)$的逻辑公式，作为发送方认知状态的（可能不完整的）快照；也就是说，这些公式表示发送方知道什么或认为哪些情况有可能。（ii）接收方通过在共享 Kripke 结构中对这些公式求值来解读其语义（意义）；该结构为正确解读提供共同语境。（iii）通信由任务驱动，例如，当任务是实现韧性时，智能体仅在环境发生变化时进行通信，以便其他智能体据此更新其信念和策略。

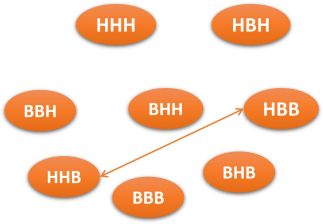
示例 2（通过通信细化认知状态）。现在，我们通过引入智能体之间的通信来扩展示例 *1*。通信之前，每个智能体的认知状态由其在共享 *Kripke* 结构 $\mathcal{M}_t$上的可达关系 $R_{i,t}$表示。图 *3b–3c* 展示了该 *Kripke* 结构中各智能体对应的切片，



（a）环境变化后的环境设置化。



（b）智能体 1 的 $\mathcal{M}_0$ 修订后切片。



（c）智能体 2 的 $\mathcal{M}_0$ 修订后切片。

图4。环境发生环境变化（压力源）后的示例，以及每个智能体执行修订认知行动后共享 Kripke 结构中相应的每个智能体切片。智能体 i 的切片是三元组（W、R_i、t、v）。节点表示可能世界，边表示特定于智能体的可达关系。为清晰起见，省略了世界上的自反自环。

其中仅显示智能体在局部可见的内容（即其自身的可达关系），而非共享结构中的全部关系。在每个切片中，节点对应可能世界，边表示智能体的可达关系，也就是根据其局部观测，对该智能体而言无法区分的世界对。由这些关系连接的世界称为可达世界。由于每个智能体只能直接观测一个单元格，因此仍有许多世界是可达的。每个切片都采用颜色编码，以匹配相应的智能体。在通信过程中，智能体1 向智能体2发送公式$\mathcal{K}_1 H_1$ ，智能体2向智能体1 发送公式 $\mathcal{K}_2 H_3$。通信之后，每个智能体通过纳入从另一个智能体收到的信息来细化其认知状态。形式上，这对应于认知行动（细化），它通过删除相应的可达关系，移除与共享消息不一致的可达世界。图 *3d–3e* 所示的结果可达关系严格缩小：每个智能体的不确定性都降低了，且两个智能体更新后的 *Kripke* 切片均与真实世界 w^{HBH}一致。现在，所有 $\mathcal{K}_1 H_1 \wedge \mathcal{K}_1 H_3$、 $\mathcal{K}_2 H_1 \wedge \mathcal{K}_2 H_3$、$\mathcal{P}_1 H_2 \wedge \mathcal{P}_1 B_2$ 和 $\mathcal{P}_2 H_2 \wedge \mathcal{P}_2 B_2$ 均成立。

示例3（压力源下的认知状态修订）。 现在考虑一种影响示例*1*和示例*2*所述情境的环境变化（压力源）。图*4a*展示了变化后的环境。具体而言，单元格 3 由白色变为黑色。起初，智能体 1 没有收到相互矛盾的观测，因此执行认知行动保持。与此同时，智能体 2在本地检测到其预测（基于自身知识）与单元格 3的观测颜色之间持续不一致。这触发了认知行动 探索：智能体 2 收集判别性观测，以确定当前有效的环境条件。收集到足够的证据后，智能体 2 需要更新其认知状态，使其知识与新证据保持一致。这对应于执行认知行动修订，该行动会更新智能体 2的认知状态，使得 $\mathcal{K}_2 B_3$ 成立，而 $\mathcal{K}_2 H_3$ 不再成立。具体而言，智能体 2移除通往世界 HHH 和 HBH 的可达链接（图*3e*），并仅添加通往满足 B_3的世界的链接，即HHB 和 HBB （图*4c*）。随后，智能体2通过公式 $\mathcal{K}_2 B_3$将此次更新告知智能体 1 ；该公式充当认知行动广播，触发智能体1进行信念修订。作为回应，智能体 1 执行修订，移除违反 B_3 的先前可达世界（图*3d*），并使其认知状态重新与新的公开共享信息保持一致（即，$\mathcal{K}_1 B_3$ 现在成立，而 $\mathcal{K}_1 H_3$ 不再成立）。智能体 1 最终的可达关系如图*4b*所示。

IV. 多智能体系统中的韧性度量

在危急情境下，及时的信息和共享理解是有效响应的关键。在我们的智能体架构中，智能体 i 的外部策略 π_i^{ext} 取决于其认知状态 $\iota_{i,\ t}$，该策略以此为条件。因此，韧性要求智能体**检测**压力源，在压力源出现后**恢复**其认知状态和策略，并在恢复后**保持**准确的认知状态和最优策略。为捕捉这些方面，我们沿着两个互补维度对韧性进行形式化：

- **认知韧性**：检测知识缺口、弥合对环境的信念（认知状态）与环境实际状态之间的差异，并在恢复后维持准确信念的能力。
- **行动韧性**：使外部动作与更新后的信念重新对齐，并在恢复后维持最优外部策略的能力。

这些维度不仅体现了个体能力，也反映了多智能体系统的集体动态。在这类系统中，韧性受到共同知识形成的影响；共同知识是一种共享且不断演进的理解，能够促成协同适应与恢复。鉴于此，我们现在正式介绍适用于多智能体系统的韧性度量，随后定义其两个维度。

定义 1 (系统韧性)。令 $\mathcal{R}_{\text{epistemic}}$、$\mathcal{R}_{\text{action}} \in \mathcal{L}_n^{time}(\Omega)$为分别表示认知韧性规范和行动韧性规范的时序认知公式。那么，系统整体韧性可由以下公式表示：

$$\mathcal{R}_{\text{system}} \equiv \mathcal{R}_{\text{epistemic}} \wedge \mathcal{R}_{\text{action}}. \tag{1}$$

这表示，当且仅当认知韧性属性和行动韧性属性都成立时，才能实现系统的整体韧性。

A. 认知韧性

设 $\varphi_1 \in \mathcal{L}_n^{\text{time}}(\Omega)$ 和 $\varphi_2 \in \mathcal{L}_n^{\text{time}}(\Omega)$ 是互斥公式（即 $\varphi_1 \to \neg\varphi_2$ 和 $\varphi_2 \to \neg\varphi_1$），用于描述环境的某个方面。初始时，实际世界中保持 φ_1 ，且所有智能体都正确地相信 φ_1 为真。也就是说，

在初始时刻 t_0，对于每个智能体 $i \in N$ 以及每个满足 $(w,\ w')$ 的 $w' \in W$，都有 $\in \iota_{i,\ t_0}$，且 $(\mathcal{M}_{t_0},\ r,\ t_0) \models \varphi_1$；等价地，

$$(\mathcal{M}_{t_0}, r, t_0) \models \bigwedge_{i \in N} \mathcal{K}_i \varphi_1 \quad \text{(i.e., } \mathcal{E}_N\, \varphi_1).$$

在某个时间 t_v，环境发生变化（形成压力源或扰动），使得 φ_1 不再成立，而 φ_2变为真。为实现认知韧性，智能体必须:

(i) 在时间界限 α_1 内将其认知状态恢复为正确的信念，这意味着智能体得知 φ_2 现已成立（可恢复性）。也就是说，令恢复时间为 $t_{\text{rec}}^{\text{epi}}$，其中 $t_{\text{rec}}^{\text{epi}} - t_v \le \alpha_1$；对于每个智能体 $i \in$$N$以及每个 $w' \in W$，若有（$w,\ w'$）$\in \iota_{i,\ t_{\text{rec}}^{\text{epi}}}$，则有（$\mathcal{M}_{t_{\text{rec}}^{\text{epi}}},\ r,\ t_{\text{rec}}^{\text{epi}}$）$\models \varphi_2$；等价地，

$$(\mathcal{M}_{t_{\text{rec}}^{\text{epi}}}, r, t_{\text{rec}}^{\text{epi}}) \models \bigwedge_{i \in N} \mathcal{K}_i \varphi_2 \quad \text{(i.e., } \mathcal{E}_N\, \varphi_2).$$

(ii) 将恢复后的信念至少保持一段持续时间 β_1（持久性）。

上述条件可以用我们的时序认知语言更简洁地表达。具体而言，认知韧性要求所有智能体保持对旧事实 φ_1 的知识，直到在 α_1 个时间单位内得知新事实 φ_2，并且该知识随后至少持续 β_1 个时间单位。我们将其形式化如下:

定义 2 (认知韧性)。 令 α_1 表示最大允许恢复时间，β_1 表示最短允许持久时长。令 $\mathcal{R}_{\text{epistemic}}$ 表示一个刻画系统认知韧性的公式。于是，我们将认知韧性定义为

$$\mathcal{R}_{\text{epistemic}} \equiv \Big(\neg \mathcal{E}_N \varphi_2\Big)\, \mathcal{U}_{[0,\alpha_1]}\, \Big(\mathcal{G}_{[0,\beta_1)}\, \big(\mathcal{E}_N \varphi_2\big)\Big). \quad (2)$$

其中:

- $\mathcal{U}_{[0,\alpha_1]}$ 是有界直到算子，确保智能体最初未意识到压力源时，能在 α_1 时间单位内恢复其认知状态;
- $\mathcal{G}_{[0,\beta_1)}$ 是有界全局算子，确保智能体恢复其认知状态后，这些认知状态至少维持 β_1 时间单位。

持久性意味着，恢复后的认知状态足够准确，能够至少持续 β_1 个时间步，且智能体不会检测到任何矛盾；前提是假设环境在此期间没有变化。这凸显了一个基本的可恢复性–持久性 权衡：快速修订信念的智能体可能依据有限证据进行修订，从而增加更新错误和频繁修订的风险（即持久性较低）。相反，等待更确凿证据的智能体恢复速度可能较慢，但更有可能采纳正确且稳定的信念（即持久性较高）。为了量化这两个因素并使（2）具有可操作性，我们现在提取两个标量，以具体体现其两个方面：认知可恢复性 和认知持久性。

定义 3 (认知可恢复性)。 设 t_v 为环境变化（压力源）发生时的违规时间，并设 φ_2 为新的环境条件。那么，认知可恢复时间间隔 $\Delta t_{\text{rec}}^{\text{epi}}$ 定义为

$$\Delta t_{\text{rec}}^{\text{epi}} = \inf\{\, t \mid t \in T,\ t > t_v \ \textit{and}\ (\mathcal{M}_t, r, t) \models \mathcal{E}_N \varphi_2 \} - t_v, \quad (3)$$

该表达式将认知可恢复时长定义为违例时刻 t_v 之后的最小时间区间，在此之后，智能体的认知状态满足以下属性: N 中的每个智能体都知道 φ_2。我们约定，$\Delta t_{\text{rec}}^{\text{epi}} = +\infty$ 如果该集合为空（即在 t_v之后认知准确性从未恢复）。随后，恢复后的认知属性首次成立的时刻表示为: $t_{\text{rec}}^{\text{epi}} = t_v + \Delta t_{\text{rec}}^{\text{epi}}$.

定义 4（认知持久性）。 设 t_v 为环境变化发生的违例时刻。则认知持久性区间 $\Delta t^{\text{epi}}{}_{\text{dur}}$ 定义为

$$\Delta t_{\text{dur}}^{\text{epi}} = \inf\{t \mid t \in T, t > t_{\text{rec}}^{\text{epi}}\ \&\ (\mathcal{M}_t, r, t) \not\models \mathcal{E}_N \varphi_2\} - t_{\text{rec}}^{\text{epi}}, \quad (4)$$

其中，$\mathcal{E}_N\varphi_2$ 表示 N 中的每个智能体都相信 φ_2，而（$\mathcal{M}_t$, r, t）$\not\models$ $\mathcal{E}_N\varphi_2$ 表示该属性在世界 w 中的时间 t 不再成立，且智能体已经改变了其信念。换言之，$\Delta t^{\text{epi}}{}_{\text{dur}}$ 是 $t_{\text{rec}}^{\text{epi}}$ 之后的最小时间区间，在此期间，恢复后的信念持续存在，直到被违例。我们约定，$\Delta t^{\text{epi}}{}_{\text{dur}} = +\infty$ 如果该集合为空（即在$t_{\text{rec}}^{\text{epi}}$之后认知准确性从未被违例）。

虽然认知韧性体现了智能体恢复并维持关于环境的准确信念的能力，但韧性并不止步于认知层面。对于在动态环境中运行的智能体而言，知识必须转化为恰当的外部动作。即使智能体能够正确修订其信念，一个系统也不能被视为真正具备韧性，除非这些信念能够促使其及时且有效地调整行为。因此，我们转向第二个维度：行动韧性。

B. 行动韧性

令 Π_i^{ext} 表示智能体 i 的可接受的外部策略集合，并令 $\pi_{i}^{\text{ext}}{}_{,\ t} \in \Pi_i^{\text{ext}}$ 表示智能体 i 在时间 t 当前实施的外部策略。每个智能体 i 都与特定的任务相关性能目标 J_i 相关联（例如期望累积奖励或成本）。给定其有限候选集 Π_i^{ext}，该目标会诱导出一个随时间变化的个体最优外部策略集合

$$\Pi_{i,t}^{\text{opt}} := \arg \max_{\pi_i \in \Pi_i^{\text{ext}}} J_i(\pi_i; t),$$

其中包含智能体 i 在时间 t、当前环境条件下的最优策略。在逻辑层面，我们为每个智能体 i 引入一个原子命题$\pi_i^{\text{opt}} \in \mathcal{L}_n^{\text{time}}(\Omega)$，其含义为“智能体 i 当前正相对于 J_i 采取最优行动”。也就是说，当且仅当智能体 i当前的外部策略 $\pi_{i\,t}^{\text{ext}}$, 属于 $\Pi_{i\,t}^{\text{opt}}$, 时，命题π_i^{opt} 为真。群体层级谓词为合取

$$\pi^{\text{opt}} \equiv \bigwedge_{i \in N} \pi_i^{\text{opt}},$$

当且仅当每个智能体都执行其个体最优外部策略之一时，该谓词成立。不假设联合最优性或均衡概念；最优性是针对每个智能体单独评估的。

最初，在压力源于时间步 t_v 发生之前，所有智能体都根据环境条件执行最优动作。也就是说，在时间 $t' < t_v$，对于每个智能体$i \in N$ 以及满足 $(w,\ w') \in \iota_{i,\ t'}$ 的每个 $w' \in W$，都有 $(\mathcal{M}_{t'},\ r,\ t') \models \pi^{\text{opt}}$。压力源在时间 t_v 发生，并在时间 $t_{\text{rec}}^{\text{epi}}$ 导致智能体的认知状态发生变化后，智能体的外部策略可能不再是最优策略。也就是说，在 $t'' \geq t_v$，对于每个智能体 i 以及满足 $(w,\ w') \in \iota_{i,\ t''}$ 的每个世界 $w' \in W$，都有 $(\mathcal{M}_{t''},\ r,\ t'') \not\models \pi^{\text{opt}}$。为实现行动韧性，智能体必须：

(i) 在时间界限 α_2 内将其外部策略恢复为最优策略，这意味着现在满足 π^{opt} （可恢复性）。也就是说，令恢复时间为 $t_{\text{rec}}^{\text{act}}$，其中$t_{\text{rec}}^{\text{act}} - t_{\text{rec}}^{\text{epi}} \leq \alpha_2$；对于每个智能体 $i \in N$ 以及满足 $(w,\ w') \in \iota_{i,\ t_{\text{rec}}^{\text{act}}}$ 的每个 $w' \in W$，都有 $(\mathcal{M}_{t_{\text{rec}}^{\text{act}}},\ r,\ t_{\text{rec}}^{\text{act}}) \models \pi^{\text{opt}}$。(ii) 至少在持续时间 β_2 内保持最优策略（持久性）。

与认知韧性类似，上述条件可以用我们的时序—认知语言更简洁地表达。具体而言，行动韧性要求所有智能体在当前环境条件下不执行最优策略（$\neg\pi^{\text{opt}}$ 成立），直到它们在 α_1 时间单位内恢复并执行最优策略（π^{opt} 成立），且至少持续 β_2 时间单位执行最优策略。我们在以下定义中对其进行形式化。

定义5（**行动韧性**）。令 α_2 表示智能体根据恢复的集体知识调整其外部动作所允许的最大恢复时间，令 β_2表示必须保持最优动作的最短持续时间，并令 π_i^{opt} 表示智能体 $i \in N$ 当前正在执行最优动作的命题。令 $\mathcal{R}_{\text{action}}$表示系统的行动韧性。于是，我们将行动韧性定义为

$$\mathcal{R}_{\text{action}} \equiv \left(\neg\pi^{opt}\right) \mathcal{U}_{[0,\alpha_2]} \left(\mathcal{G}_{[0,\beta_2)} \left(\pi^{opt}\right)\right), \tag{5}$$

其中：

- $\mathcal{U}_{[0,\alpha_2]}$ 是有界直到算子，用于确保如果智能体最初未执行最优动作，它们将在 α_2 时间单位内恢复最优动作；
- $\mathcal{G}_{[0,\beta_2)}$ 是有界全局算子，确保一旦采用最优动作，这些动作至少维持 β_2 时间单位。

这一表述体现了可恢复性–持久性 权衡：智能体可以选择快速调整行动（导致频繁变化并降低持久性），也可以在修订策略之前进行充分探索（确保长期稳定性）。与前文相同，为了量化这两个因素并使（5）具有可操作性，我们现在提取两个标量量来体现其两个方面：行动可恢复性 和行动持久性。

定义 6 (行动可恢复性)。令 t_v 为环境变化发生时的违例时刻。那么，行动可恢复性区间 $\Delta t_{\text{rec}}^{\text{act}}$ 定义为

$$\Delta t_{\text{rec}}^{\text{act}} = \inf\{\, t \mid t \in T,\ t \geq t_{\text{rec}}^{\text{epi}} \textit{ and } (\mathcal{M}_t, r, t) \models \pi^{\text{opt}} \} - t_{\text{rec}}^{\text{epi}}. \tag{6}$$

该定义表明，$\Delta t_{\text{rec}}^{\text{act}}$ 是在 $t_{\text{rec}}^{\text{epi}}$ 之后，直到系统达到 N 中所有智能体均采取最优动作的状态所经历的最小时间区间。我们约定 $\Delta t_{\text{rec}}^{\text{act}} = +\infty$，如果该集合为空（即在 $t_{\text{rec}}^{\text{epi}}$之后从未达到最优性）。

定义 7 (行动持久性)。令 t_v 为环境变化发生时的违规时间。定义

$$t_{\text{rec}}^{\text{act}} = t_{\text{rec}}^{\text{epi}} + \Delta t_{\text{rec}}^{\text{act}},$$

为达到最优策略的时间。于是，行动持久性区间 $\Delta t_{\text{dur}}^{\text{act}}$ 表示为

$$\Delta t_{\text{dur}}^{\text{act}} = \inf\{t \mid t \in T,\, t \geq t_{\text{rec}}^{\text{act}} \textit{ and } (\mathcal{M}_t, r, t) \not\models \pi^{\text{opt}}\} - t_{\text{rec}}^{\text{act}}. \tag{7}$$

该定义表明，$\Delta t_{\text{dur}}^{\text{act}}$ 是在 $t_{\text{rec}}^{\text{act}}$ 之后，直到系统不再满足每个智能体都采取最优动作这一条件所经历的最小时间区间。换言之，它衡量了恢复后的最优动作在被违反之前维持的时长。如果该集合为空（即在 $t_{\text{rec}}^{\text{act}}$之后从未违反最优性），我们约定 $\Delta t_{\text{dur}}^{\text{act}} = +\infty$。

接下来，我们将介绍在所提出的度量下，使多智能体系统实现认知韧性和行动韧性的策略与算法。

V. 实现认知韧性与行动韧性：策略与算法

定义韧性度量后，我们现在给出温和的系统假设，在这些假设下，信息能够流动且压力源可以被检测到；随后介绍实现这些度量的策略与算法。

假设 1 (通信网络连通性与可靠性)。通信网络是一个有向图$G = (N, E)$，且为强连通：对于任意 $i, j \in N$，都存在一条有向路径 $i \rightsquigarrow j$。定义（跳数）直径

$$l_{comm} := \max_{i,j \in N} l(i,j),$$

其中，$l(i,j)$ 是从 i到 j的最短有向路径长度。我们还假设，沿 E 中的边传输的消息均能可靠地在一个时间步内送达。因此，l_{comm} 可作为向整个网络传播矛盾或信念更新所需时间的上界，从而能够形成共同知识$\mathcal{E}_N(\cdot)$。

假设2 (矛盾的及时本地检测)。设 φ_1 表示一个命题，该命题在某个时间 t_v于环境中变为假。我们假设，至少存在一个智能体 $i \in N$，能够在不超过 n_{steps} 个时间步内在本地检测到这一变化。形式化地，

$$\exists i \in N,\, \exists t_{det} \in [t_v, t_v + n_{steps}] \textit{ s.t. } (\mathcal{M}_{t_{det}}, r, t_{det}) \models \mathcal{K}_i \neg\varphi_1.$$

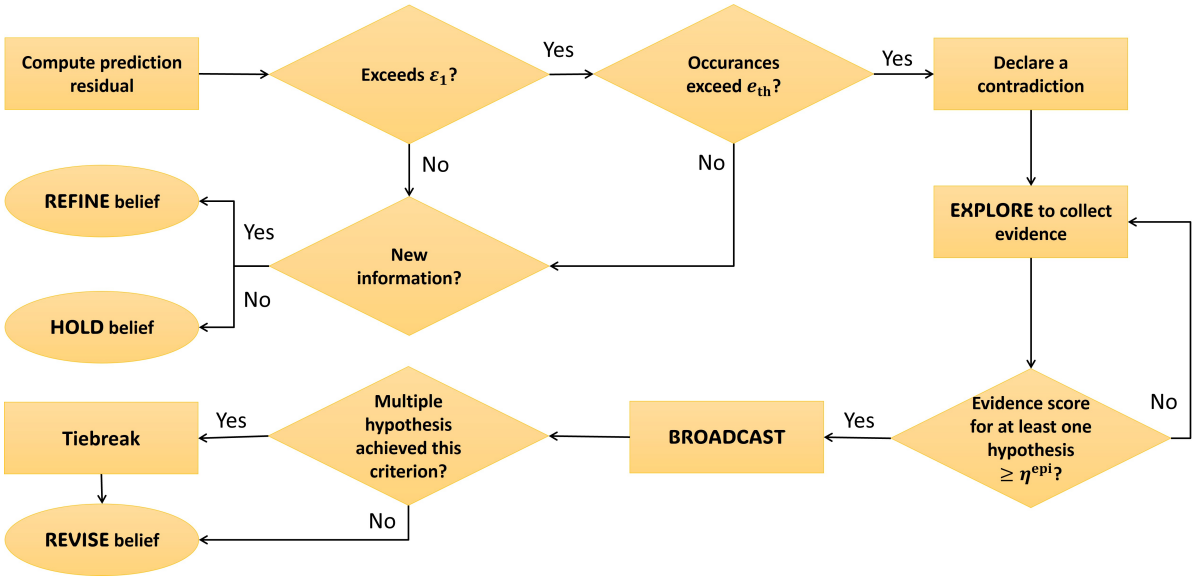


图5. 每个智能体执行的认知策略流程图，以实现认知韧性。

假设检测由内部不一致条件控制，例如预测观测与实际观测之间基于阈值的偏差。

假设2指出，至少有一个智能体能够根据自身的预测观测与实际观测之间的不匹配来察觉压力源。这是一个合理的假设：如果某个压力源对任何智能体的知识都没有可观测的影响（除了噪声之外没有持续残差），也不影响任务表现，那么它在此时就是认知静默的，无须进行“有意义的恢复”。

假设3 (最小干扰间隔)。令 $\{t_v^{(k)}\}_{k\geq 1}$ 为扰动（压力源）发生时刻的序列。存在 $\delta > 0$，使得对所有 k，

$$t_v^{(k+1)} - t_v^{(k)} \;\geq\; \delta.$$

等价地，如果变化发生在时刻 t_v，则之后发生的任何变化时刻 t'_v 都应满足 $t'_v \geq t_v + \delta$。因此，扰动发生的频率不高于每隔 δ 个时间单位一次。

假设3可避免对抗性地快速连续出现压力源，从而导致恢复在理论上不可能。然而，要满足可行性要求，扰动间隔也必须足够长，以容纳最坏情况下的恢复和持久性窗口。下面的界限刻画了这一必要的时间预算。

备注3（韧性的必要时间预算）。根据假设*3*，韧性规范可满足的一个必要条件是

$$\delta \;>\; \max\big\{\, \alpha_1 + \beta_1,\;\; \alpha_1 + \alpha_2 + \beta_2 \,\big\}.$$

事实上，在任意扰动间隔内，必须：*(i)* 完成认知恢复（至多 α_1）并至少维持 β_1；*(ii)* 在认知恢复后（至多 α_1）完成行动恢复（至多 α_2）并至少维持 β_2；因此，该间隔必须大于这两个总和。此

条件是必要但不充分的：仅满足该条件并不能保证韧性；还需要适当的认知/行动策略及其成功属性。

本节余下部分中，我们基于假设1–3进行讨论，并假设备注3中的时间预算界成立。

A. 实现认知韧性

令 $\varphi_1 \in \mathcal{L}_n^{\text{time}}(\Omega)$为表示某种环境条件的公式，该条件在违例时刻$t_v$之前成立，但在 t_v时变为不成立。令 $\Phi = \{\varphi_2, \cdots, \varphi_K\}$为一组由语言 $\mathcal{L}_n^{\text{time}}(\Omega)$生成的 $K-1$ 互斥假设。每个假设都刻画同一个环境方面，其中一个假设描述扰动后出现的新条件。由于这些假设互斥，因此它们不可能同时成立。

为了使智能体能够检测此类环境变化（即，区分先前成立的条件φ_1 何时不再成立，以及 Φ 中的新假设何时成立），我们量化智能体内部预测与实际环境之间的差异。如第二节所述，我们基于观测残差引入一种预测不一致性度量。具体而言，对于任意智能体 $i \in N$，预测残差为$\Upsilon\big(\boldsymbol{y}_{i,\ t}\ \boldsymbol{y}_i \qquad \boldsymbol{y}_{i\,t} = s_i\ \boldsymbol{x}_t \quad \boldsymbol{y}_{i\,t} = ,$
$\varrho_i\ \iota_{i\,t-1}\ u_{i\,t}^{\text{ext}} \qquad \Upsilon \qquad\qquad\qquad Y$
$\hat{},\ _t\big)$，其中，$(\,)$、$\hat{}\,.\,(\,,\ ,\ )$，且 是 上的任意距离度量（例如 ℓ_2-范数）。给定阈值 $\varepsilon_1 > 0$、窗口长度 $L \in \mathbb{N}_{\geq 1}$和超阈值计数阈值 $e_{\text{th}} \in \{1, \ldots, L\}$，定义最近 L 步内超过阈值的次数为：

$$e_{i,t} \;:=\; \sum_{t'=t-L+1}^{t} \mathbf{1}\{\, \Upsilon\big(\boldsymbol{y}_{i,t'},\, \hat{\boldsymbol{y}}_{i,t'}\big) > \varepsilon_1 \,\}. \qquad (8)$$

当且仅当$e_{i,t} \;\geq\; e_{\text{th}}$时，智能体 i 在时间 t 判定存在矛盾。该超限计数机制可确保孤立的噪声读数或短暂的预测抖动不会触发矛盾检测；只有预测值与感知观测值之间持续且具有统计意义的

偏差才会启动证据收集和认知修正。如果未检测到矛盾（即 $e_{i,t} < e_{\text{th}}$），智能体 i 会根据是否通过直接观测或消息（如有）获得了新信息，选择细化或保持。否则，智能体 i 会在最近 L 步中首次出现至少e_{th} 次超限的时间 t判定存在矛盾（即 $e_{i,t} \geq e_{\text{th}}$），也就是

$$t_{\text{det}}^{(i)} \;:=\; \min\Big\{\, t > t_v \;\Big|\; e_{i,t} \geq e_{\text{th}} \Big\}^{3}.$$

最早的网络检测时间为 $t_{\text{det}} :=$ 分钟$_{i\in N}\, t^{(i)}$.检测在时间窗 $[t_v,\; t_v + n_{\text{steps}}]$ 内检测到矛盾的智能体集合为

$$N_{\text{det}} \;:=\; \big\{\, i \in N \;\big|\; t_{\text{det}}^{(i)} \in [t_v,\; t_v + n_{\text{steps}}] \big\}.$$

根据假设2，该集合非空[4]。确认矛盾后，在时刻 $t \;\geq\; t_{\text{det}}$，每个智能体 $i \;\in\; N_{\text{det}}$选择$u_i^{\text{epi}}{}_{,\,t}$ =探索，以探索环境并收集能够区分 Φ 中各假设的证据。在此探索阶段，每个新观测都会映射为一个成对证据增量。具体而言，对于每对假设 $\varphi_k,\ \varphi_l \;\in\; \Phi$，每个智能体 $i \;\in\; N_{\text{det}}$ 都会维护一个累积成对证据分数 $\lambda^{k,}{}_{i}{}^{l}{}_{t}$，用于量化支持假设 φ_k 而非假设 φ_l的证据，定义如下：

$$\Lambda_{i,t}^{k,l} = \sum_{t'=t_{\text{det}}+1}^{t} \lambda_{i,t'}^{k,l}, \qquad k \neq l,$$

其中，逐步分数 $\lambda^{k,}{}_{i}{}^{l}{}_{t'}$ 有界，且在真实假设下具有正的均值差（例如，对数似然比、判别间隔或一致性投票）。基于这些累积分数，智能体 $i \;\in\; N_{\text{det}}$ 通过极大极小规则选择当前的最佳假设$\varphi_{\hat{k}_i\,(t)}$：

$$\varphi_{\hat{k}_i(t)} \;\in\; \arg\max_{k:\varphi_k\in\Phi} r_{i,t}^{k}, \quad \text{where} \quad r_{i,t}^{k} \;:=\; \min_{l:\varphi_l\in\Phi\setminus\varphi_k} \Lambda_{i,t}^{k,l}.$$

请注意，可以采用任意固定的平局决胜规则。有了这些最大最小值后，探索将持续进行，直到领先假设以预定义的裕度 η^{epi} 稳定地领先。智能体 i 会在所选假设与所有竞争假设之间首次达到足够分离时停止收集证据：

$$\tau_i \;:=\; \min\Big\{\, t \geq t_{\text{det}}^{(i)} \;\Big|\; r_{i,t}^{\hat{k}_i(t)} \;\geq\; \eta^{\text{epi}} \Big\}.$$

阈值 η^{epi} 的选择旨在达到期望的错误界，例如 $\Pr\{\hat{k}_i(\tau_i) \neq k^\star\} \leq \epsilon$，其中 $k^\star$ 是真实假设的索引。一旦满足分离条件，探索便终止，智能体宣布其当前最佳假设。在时间 $t = \tau_i$，智能体 i 广播一条带签名的消息

$$m_{i,\tau_i} \;=\; \big(\tau_i,\; \mathcal{K}_i\varphi_{\hat{k}_i(\tau_i)},\; r_{i,\tau_i}^{\hat{k}_i(\tau_i)}\big)$$

给所有其他智能体，生存时间 TTL 为 $= l_{\text{comm}}$。如果 N_{det} 中有多个智能体广播相互冲突的假设，网络通过在截至时间 t 收到的所有广播中，选择累积证据最高的假设来解决分歧（平局时采用任意固定规则）。选出最可信的假设后，目标便从证据收集转向信念对齐。每个智能体 $i \in N$ 将 $u_i^{\text{epi}}{}_{,\,t}$ =修订其认知状态 $\iota_{i,\,t}$ 以使所有当前可达世界都满足 $\varphi_{\hat{k}}$。随后，我们将认知恢复时间 $t_{\text{rec}}^{\text{epi}}$ 定义为第一个满足以下条件的时刻 t：所有智能体都知道 $\varphi_{\hat{k}}$:

$$t_{\text{rec}}^{\text{epi}} \;:=\; \min\{\, t > t_v \mid (\mathcal{M}_t, r, t) \models \mathcal{E}_N\varphi_{\hat{k}} \,\}.$$

因此，

$$\begin{aligned}\Delta t_{\text{rec}}^{\text{epi}}(\eta^{\text{epi}}) \;&=\; t_{\text{rec}}^{\text{epi}} - t_v \\ &= \underbrace{(t_{\text{det}} - t_v)}_{\text{detection delay}} + \underbrace{(\tau_{\max} - t_{\text{det}})}_{\text{evidence aggregation}} + \underbrace{l_{\text{comm}}}_{\text{message propagation}} \\ &\leq\; n_{\text{steps}} \;+\; (\tau_{\max} - t_{\text{det}}) \;+\; l_{\text{comm}}, \end{aligned} \tag{9}$$

其中 $\tau_{\max} :=$ max$_{i\in N_{\text{det}}}\,\tau_i$ 是最后一份报告的时间戳。请注意，为便于展示，$t_{\text{rec}}^{\text{epi}}$ 和 $\tau_{\max}$ 对 η^{epi} 的依赖关系被省略了。图 5 展示了由每个智能体执行的拟议认知策略及其相关的信息流。

满足认知可恢复性规范。 为满足认知可恢复性规范，应选择证据阈值η^{epi}，使得

$$\Delta t_{\text{rec}}^{\text{epi}}(\eta^{\text{epi}}) = t_{\text{rec}}^{\text{epi}} - t_v \;\leq\; \alpha_1. \tag{10}$$

利用广播/提交规则、公式(9)和(10)以及假设1-2，可得在至多 α_1 步内恢复的一个充分条件：

$$\tau_{\max} - t_{\text{det}} \;\leq\; \alpha_1 - n_{\text{steps}} - l_{\text{comm}},$$

（确定性地或以高概率）。这要求$\alpha_1 \geq n_{\text{steps}} + l_{\text{comm}}$。如果不满足这一条件，则在当前感知/通信机制下，配对$(\alpha_1,\ \eta^{\text{epi}})$不可行。

满足认知持久性规范。在最坏情况下的扰动间隔为 δ 时（假设3和备注3），所实现的认知持久性下界为

$$\Delta t_{\text{dur}}^{\text{epi}}(\eta^{\text{epi}}) \;\geq\; \delta - \Delta t_{\text{rec}}^{\text{epi}}(\eta^{\text{epi}}),$$

前提是变化后假设得到正确识别，并且在下一次扰动发生前始终有效。事实上，阈值η^{epi} 需要权衡正确提交的概率 $p(\eta^{\text{epi}}) := \Pr\{\hat{k}(t_{\text{rec}}^{\text{epi}}) = k^\star\}$，与恢复时间 $\Delta t_{\text{rec}}^{\text{epi}}(\eta^{\text{epi}})$之间的关系，进而影响持久时间 $\Delta t^{\text{epi}}{}_{\text{dur}}(\eta^{\text{epi}})$。一种实用的调参规则如下：

$$\begin{aligned}\max_{\eta^{\text{epi}}} \quad & p(\eta^{\text{epi}})\,\big(\delta - \Delta t_{\text{rec}}^{\text{epi}}(\eta^{\text{epi}})\big) \\ \text{s.t.} \quad & 0 \leq \Delta t_{\text{rec}}^{\text{epi}}(\eta^{\text{epi}}) \;\leq\; \min\{\alpha_1,\; \delta - \beta_1\},\end{aligned} \tag{11}$$

从而使持久性预算为非负且至少满足β_1。直观而言：收集足够的证据以确保正确性（从而信念可能得以持续），但不要收集过多，以免恢复耗尽扰动间隔。在实践中，选择 η^{epi}，以在满足标称参数的可恢复性（α_1）和持久性（β_1）的同时达到目标误差（例如，$n_{\text{steps}}, l_{\text{comm}}$）。反过来，这可归结为在可行区间 $[\eta^{\text{epi}}{}_{\min}\ \eta_{\max}^{\text{epi}}]$ 上进行一维线搜索；如果该区间为空，则规格不可行。

[3] We have the convention that $t_{\text{det}}^{(i)} = +\infty$ if the set is empty.

[4] 在实践中，选择 $n_{\text{steps}} \geq l_{\text{comm}}$ 可确保 $t_{\text{det}} - t_v \leq n_{\text{steps}}$ 前提是每个窗口内至少存在一个智能体，其变化后的残差超过阈值至少 e_{th} 次。

B. 实现行动韧性

在时间 $t_{\mathrm{rec}}^{\mathrm{epi}}$ 完成认知恢复后，智能体必须将其外部策略重新调整至新识别的条件 $\varphi_{\hat{k}}$，并保持最优性能。因此，信念修订后，每个智能体 i 都会在其有限候选集 Π_i^{ext} 中运行最佳策略识别循环。在时间 $t = t_{\mathrm{rec}}^{\mathrm{epi}},\ t_{\mathrm{rec}}^{\mathrm{epi}}+1,\ \ldots,$ 智能体 i:

1) **执行轨迹展开**。对候选策略$\pi_i \in \Pi_i^{\mathrm{ext}}$执行一次短程轨迹展开，并记录其任务成本。

2) **更新**。对于每个 $\pi_i \in \Pi_i^{\mathrm{ext}}$，维护经验平均成本 $\hat{J}_{i,\,t}(\pi_i)$ 和置信半径 $\zeta_{i,\,t}(\pi_i)$，并将其校准至置信水平 η^{act}；随着收集到更多轨迹展开 $\hat{}$，置信半径会缩小。形式上，$J_{i,\,t}(\pi_i)$ 是对第 IV 节中引入的真实性能值 $J_i(\pi_i; t)$ 的估计量，而置信区间 $[J_{i,t}(\pi_i) - \zeta_{i,\,t}(\pi_i)\ J_{i,\,t}(\pi_i)\ + \zeta_{i,\,t}(\pi_i)]$ 的构造使其以高概率包含所有 $\pi_i \in \Pi_i^{\mathrm{ext}}$ 对应的 $J_i(\pi_i; t)$。

3)**选择/停止**。探索置信下界最小的策略，并在某个策略与所有其他策略统计上可区分时立即停止；若满足以下条件，则采用策略 π_i^*

$$\hat{J}_{i,t}(\pi_i^*) + \zeta_{i,t}(\pi_i^*) \ < \ \min_{\pi_i \in \Pi_i^{\mathrm{ext}} \setminus \{\pi_i^*\}} \big[\hat{J}_{i,t}(\pi_i) - \zeta_{i,t}(\pi_i)\big],$$

即，π_i^* 的最坏可能成本低于其他任何策略的最佳可能成本。在此条件下，π_i^* 以很高的概率属于与 $J_i(\cdot; t)$ 相关联的最优集合$\Pi_{i\,t}^{\mathrm{opt}}$，。

选定这样的策略 π_i^*后，我们现在将系统何时完成行动恢复形式化定义。当智能体当前的外部策略属于集合$\Pi_{i\,t}^{\mathrm{opt}}$，时，原子命题$\pi_i^{\mathrm{opt}}$ 恰好成立:

$$(\mathcal{M}_t, r, t) \models \pi_i^{\mathrm{opt}} \qquad \Longleftrightarrow \qquad \pi_{i,t}^{\mathrm{ext}} \in \Pi_{i,t}^{\mathrm{opt}}.$$

群体层级谓词是以下合取

$$\pi^{\mathrm{opt}} \equiv \bigwedge_{i \in N} \pi_i^{\mathrm{opt}},$$

当且仅当每个智能体都执行其个体最优外部策略之一时，该谓词才成立。不假设联合最优性或均衡概念；最优性针对每个智能体单独评估。

行动恢复时间因此为

$$t_{\mathrm{rec}}^{\mathrm{act}} := \min\{\, t \geq t_{\mathrm{rec}}^{\mathrm{epi}} \mid (\mathcal{M}_t, r, t) \models \pi^{\mathrm{opt}} \,\},$$

恢复时长为

$$\Delta t_{\mathrm{rec}}^{\mathrm{act}}(\eta^{\mathrm{epi}}, \eta^{\mathrm{act}}) = t_{\mathrm{rec}}^{\mathrm{act}} - t_{\mathrm{rec}}^{\mathrm{epi}},$$

其中，为便于阅读，省略了对 η^{act} 的依赖。

满足行动可恢复性规范。为满足行动可恢复性规范，应选择置信水平 η^{act}，使得 $\Delta t_{\mathrm{rec}}^{\mathrm{act}}(\eta^{\mathrm{act}}) \leq \ \alpha_2$。如果在截止期限$\alpha_2$ 之前，根据 η^{epi} 和 η^{act}仍无法在统计上区分最佳策略，那么在当前候选集Π_i^{ext}以及噪声／轨迹展开预算下，该规范（$\alpha_2,\ \eta^{\mathrm{epi}},\ \eta^{\mathrm{act}}$）是不可行的。在这种情况下，必须调整$\alpha_2$、$\eta^{\mathrm{epi}}$和／或 η^{act}，或注入更强的先验。

满足行动持久性规范。在最坏情况下扰动间隔为 δ（假设3和备注3）时，所达到的行动持久性满足

$$\Delta t_{\mathrm{dur}}^{\mathrm{act}}(\eta^{\mathrm{epi}}, \eta^{\mathrm{act}}) \ \geq \ \delta \ - \ \Delta t_{\mathrm{rec}}^{\mathrm{epi}}(\eta^{\mathrm{epi}}) \ - \ \Delta t_{\mathrm{rec}}^{\mathrm{act}}(\eta^{\mathrm{epi}}, \eta^{\mathrm{act}}),$$

前提是正确识别变化后的假设和最优策略。使用 $\Delta t_{\mathrm{rec}}^{\mathrm{epi}}(\eta^{\mathrm{epi}}) \leq \alpha_1$ 可得

$$\Delta t_{\mathrm{dur}}^{\mathrm{act}}(\eta^{\mathrm{epi}}, \eta^{\mathrm{act}}) \ \geq \ \delta \ - \ \alpha_1 \ - \ \Delta t_{\mathrm{rec}}^{\mathrm{act}}(\eta^{\mathrm{epi}}, \eta^{\mathrm{act}}).$$

置信水平 η^{act} 在正确识别策略的概率 $p(\eta^{\mathrm{act}})$ 与恢复时间 $\Delta t_{\mathrm{rec}}^{\mathrm{act}}(\eta^{\mathrm{epi}}, \eta^{\mathrm{act}})$ 及持久时间 $\Delta t_{\mathrm{dur}}^{\mathrm{act}}(\eta^{\mathrm{epi}}, \eta^{\mathrm{act}})$ 之间进行权衡。一种实用的调参规则是

$$\begin{aligned} \max_{\eta^{\mathrm{act}}} \quad & p(\eta^{\mathrm{act}}) \left(\delta - \alpha_1 - \Delta t_{\mathrm{rec}}^{\mathrm{act}}(\eta^{\mathrm{act}})\right) \\ \text{s.t.} \quad & 0 \leq \Delta t_{\mathrm{rec}}^{\mathrm{act}}(\eta^{\mathrm{epi}}, \eta^{\mathrm{act}}) \leq \min\{\, \alpha_2,\, \delta - \alpha_1 - \beta_2 \}. \end{aligned} \tag{12}$$

因此，保证的持久性预算非负，且至少满足 β_2。直观来说：充分探索策略以确保正确性（这样选定的策略就很可能持续有效），但也不要探索过度，以免恢复耗尽扰动间隔。在实践中，η^{act} 的选择应满足目标误差，并且在标称感知参数下满足 (α_2, β_2)。这可简化为在可行区间 $[\eta_{\min}^{\mathrm{act}},\ \eta_{\max}^{\mathrm{act}}]$上进行一维线搜索；如果该区间为空，则此规范不可行。

计算复杂度。所提出的架构将认知层与动作层分离，因此可以评估每个智能体的计算开销。在认知侧，每个智能体维护一个有限假设集 Φ 及其对应的累积成对证据分数。更新残差 $\Upsilon(\boldsymbol{y}_{i,\,t},\ \hat{}\boldsymbol{y}_{i,\,t})$需要 $O(\dim Y)$ 次计算，而维护滑动窗口阈值超限计数器每步需要 $O(1)$ 次计算。在探索阶段，更新所有成对证据分数并计算极大极小统计量 $r_{i,\,t}^k$ 每个更新步需要 $O(|\Phi|^2)$次计算。在动作侧，每个智能体 i 都会针对其有限候选集 Π_i^{ext}执行最优策略识别。更新经验均值和置信半径，并评估停止条件，每个更新步需要 $O(|\Pi_i^{\mathrm{ext}}|)$次计算。总体而言，该框架每个智能体每步的最坏情况计算复杂度为

$$O\big(\dim Y \ + \ |\Phi|^2 \ + \ |\Pi_i^{\mathrm{ext}}|\big),$$

也就是说，其复杂度与观测维度和候选外部策略的数量呈线性关系，与认知假设的数量呈二次关系。该框架适用于资源受限部署，因为其主要计算项取决于假设集 Φ、外部策略集 Π_i^{ext} 和观测空间维度 Y 的基数。通过设计（例如对假设进行粗粒化、限制候选策略或缩减观测表示），可以使这些量保持较小，从而让系统设计者在模型分辨率和计算占用之间进行权衡。

VI. 韧性属性的形式化验证

我们沿着实际运行过程 r ，使用 $(\mathcal{M}_t,\ r,\ t) \models \cdot$ 评估公式，具体定义见第三节。在时间t_v 发生的扰动，是指环境发生变化，使得变化前的事实φ_1 沿着 r对所有 $t < t_v$ 均成立，而变化后的事实 φ_2沿着 r对所有 $t \geq t_v$ 均成立；智能体之间的关系$R_{i,\ t}$ 随t演变。

定理 1 (韧性规范的可靠性)。若$(\mathcal{M}_{t_v},\ r,\ t_v) \models \mathcal{R}_{\text{epistemic}} \wedge \mathcal{R}_{\text{action}}$，则沿着 r:

1）(3)中定义的认知恢复时间 $\Delta t^{\text{epi}}_{\text{rec}}$ 满足$\Delta t^{\text{epi}}_{\text{rec}} \leq \alpha_1$，而 (4) 中定义的认知持久时间满足 $\Delta t^{\text{epi dur}} \geq \beta_1$。*2*）(6) 中定义的行动恢复时间 $\Delta t^{\text{act}}_{\text{rec}}$ 满足$\Delta t^{\text{act}}_{\text{rec}} \leq \alpha_2$，而(7)中定义的行动持久时间满足 $\Delta t^{\text{act}}_{\text{dur}} \geq \beta_2$。

证明。根据有界 $\mathcal{U}$和 $\mathcal{G}$的语义，$(\mathcal{M}_{t_v}, r, t_v) \models \mathcal{R}_{\text{epistemic}}$意味着存在 $t^{\text{epi}}_{\text{rec}} \in \{t_v, \ldots, t_v + \alpha_1\}$，使得对所有 $t' \in \{t^{\text{epi}}_{\text{rec}}, \ldots, t^{\text{epi}}_{\text{rec}} + \beta_1 - 1\}$，$(\mathcal{M}_{t^{\text{epi}}_{\text{rec}}}, r, t^{\text{epi}}_{\text{rec}}) \models \mathcal{E}_N\varphi_2$且$(\mathcal{M}_{t'}, r, t') \models \mathcal{E}_N\varphi_2$。因此，$\Delta t^{\text{epi}}_{\text{rec}} \leq \alpha_1$且 $\Delta t^{\text{epi}}\text{dur} \geq \beta_1$。类似地，$(\mathcal{M}_{t_v}, r, t_v) \models \mathcal{R}_{\text{action}}$推出存在$t^{\text{act}}_{\text{rec}} \in \{t_v, \ldots, t_v + \alpha_2\}$，使得对所有 $t' \in \{t^{\text{act}}_{\text{rec}}, \ldots, t^{\text{act}}_{\text{rec}} + \beta_2 - 1\}$，$(\mathcal{M}_{t^{\text{act}}_{\text{rec}}}, r, t^{\text{act}}_{\text{rec}}) \models \pi^{\text{opt}}$且$(\mathcal{M}_{t'}, r, t') \models \pi^{\text{opt}}$。由此得出$\Delta t^{\text{act}}_{\text{rec}} \leq \alpha_2$ 和 $\Delta t^{\text{act}}_{\text{dur}} \geq \beta_2$。 □

推论 1（有界视界完备性（有限反例））。若（$\mathcal{M}_{t_v},\ r,\ t_v) \not\models \mathcal{R}_{\text{epistemic}} \wedge \mathcal{R}_{\text{action}}$，则存在一个有限时间 $t \in [t_v,\ t_v + B]$ 见证该违规，其中

$$B \ := \ \max\{\, \alpha_1 + \beta_1,\ \alpha_2 + \beta_2 \,\}.$$

等价地，要反驳 $\mathcal{R}_{\text{epistemic}} \wedge \mathcal{R}_{\text{action}}$，只需在扰动之后检查至多 B 步。

证明。考虑认知部分 $\mathcal{R}_{\text{epistemic}} \equiv (\neg\, \mathcal{E}_N\varphi_2)\, \mathcal{U}_{[0,\alpha_1]}\, (\mathcal{G}_{[0,\beta_1)}\ \mathcal{E}_N\varphi_2)$。如果 $\mathcal{R}_{\text{epistemic}}$ 在 t_v 处失效，则有以下两种情况：（i）不存在满足 $(\mathcal{M}_{t^{\text{epi}}_{\text{rec}}},\ r,\ t^{\text{epi}}_{\text{rec}}) \models \mathcal{E}_N\varphi_2$ 的 $t^{\text{epi}}_{\text{rec}} \leq t_v + \alpha_1$，因此该失效可在 α_1 步内得到见证；或（ii）存在这样的 $t^{\text{epi}}_{\text{rec}}$，但 $\mathcal{G}_{[0,\beta_1)}\ \mathcal{E}_N\varphi_2$ 失效，因此存在一个 $t' \in \{t^{\text{epi}}_{\text{rec}}, \ldots, t^{\text{epi}}_{\text{rec}} + \beta_1 - 1\}$，满足 $(\mathcal{M}_{t'},\ r,\ t') \models \neg\, \mathcal{E}_N\varphi_2$，从而可在 $\alpha_1 + \beta_1$ 步内得到见证。对 $\mathcal{R}_{\text{action}}$ 进行相同的分类讨论，可在 $\alpha_2 + \beta_2$ 步内得到见证。取两个界的最大值，得到 B。 □

形式化验证旨在设计时或运行时检查多智能体系统是否在给定假设下满足韧性属性。具体而言，根据定理 1 和推论 1，验证可归结为有界推理：只需检查扰动发生后最多 $B = \max\{\alpha_1+\beta_1,\ \alpha_2+\beta_2\}$ 步。由于验证范围限定在有限时域内，韧性规范与标准有界模型检验方法相兼容，例如基于 SAT/SMT 的展开、在长度为 B 的窗口内进行显式状态探索，以及仅跟踪最近 B 步的轻量级运行时监视器。因此，本节确立的形式化结果能够支持在部署前对恢复协议、知识更新规则和协调策略进行计算上可行的认证，并支持在动态环境中进行高效的运行时监控。

VII. 数值结果与分析

为了展示所提框架的有效性，我们研究了一个在网络压力因素突变情况下的分布式多智能体决策用例。具体而言，我们研究一组网络化智能体 N，它们在环境发生变化（压力因素）时，反复从 $A = |U^{\text{ext}}|$ 个动作中进行选择。假设所有智能体共享相同的外部行动集，即 $U_1^{\text{ext}} = \cdots = U_{\text{ext}_n} = U_{\text{ext}}$。为便于表述，将每个动作记为 a，其中 $a = \boldsymbol{u}^{\text{ext}} \in U^{\text{ext}}$；该动作会产生方差为 σ^2 、均值为$\mu_a(t)$的高斯奖励，而均值可能因环境压力因素而变化。这一模型描述了外部事件突然改变性能的运行状态，例如强干扰源启动、移动性导致的阴影衰落，或热点出现或消失。该群体的目标是在时域 T 内最大化总奖励（等价于最小化累积遗憾）。在此处考虑的设定中，这等价于每个智能体i最大化其自身的期望奖励 J_i，因为所有智能体面临相同的行动集和奖励结构。这种抽象在文献中通常称为多智能体多臂老虎机，它概括了网络化通信系统中的若干典型任务，这些任务需要快速的分布式适应。在此情形下，外部行动也称为臂。

多臂老虎机问题的经典解法通常依赖上置信界（ UCB） [47], 、汤普森采样 [48],或 ε-贪心策略 [49]等统计算法。这些方法利用动作奖励的统计估计值、置信界或后验分布，在探索与利用之间取得平衡。然而，这些策略本质上假设奖励偏差纯粹是随机且无偏的。在对抗性或非平稳条件下，这一假设并不成立。为应对非平稳性，一种标准策略是通过衰减记忆进行被动适应。折扣 UCB 就是这种方法的一个例子，它会降低陈旧数据的权重，以跟踪分段常数漂移 [50]。

在本研究中，我们将 UCB 用作所有基线方法和我们提出方法的底层算法。具体而言，对于每个动作 a，每个智能体 i都会维护两项运行统计量：(i)计数：截至当前时间步，该智能体 i 选择动作 a 的次数；(ii) 总和：截至当前时间步，该智能体 i 从动作 a 获得的奖励累计总和。这两个量是计算动作 a经验均值所需的充分统计量； $\mu_a(t)$。我们的实现采用了 [51, 第 7] 章中介绍的高斯上置信界算法。在协作基线方法中，智能体不会共享原始奖励历史；它们只与邻居交换逐动作计数和总和，并通过线性共识步骤对这些数据取平均。混合后，每个智能体将每个智能体的网络平均值乘以群体规模 n ，以近似全局计数和总和，从而得到全局化经验均值，供上置信界决策规则使用。

基线： 我们模拟了两种标准方案。在**独立折扣 UCB**中，每个智能体运行一种折扣 UCB

策略，仅使用自身的观测，并通过指数遗忘率来适应非平稳奖励。在**协作折扣 UCB**中，智能体通过无向通信图相互连接，并采用基于共识的折扣UCB，与邻居共享充分统计量，以加快适应速度。

提出的方法：我们考虑认知（Kripke）框架的两种变体。在第一种变体中，即**Light-Coop Kripke**，智能体遵循独立的UCB动态，仅在宣布其知识或信念状态发生变化时才进行通信。在第二种变体中，即**Cooperative Kripke**，智能体从一开始便采用协作UCB；它们不仅交换信念变化消息，还共同进行协作恢复，汇集证据和行动统计量，以识别新的正确假设并重新与之对齐。

仿真设置： 我们仿真 $n = 12$ 个智能体在 $T = 2500$ 个时间步的时间范围内同步运行的网络。行动集包含 $|U^{\text{ext}}| = 16$ 个动作（臂），每个臂的真实平均奖励从区间$[0.1, 1.2]$中随机选取。智能体通过无向通信图相连，该图采用环形拓扑。环境遵循分段平稳结构：每个臂的真实平均奖励在一个片段内保持不变，并在预先设定的变点处突然变化（智能体不知道该变点），以表示外部应激因素。在主要实验中，应激因素于 $t = 1400$发生，导致最优臂的排序改变。对于协作算法，智能体使用Metropolis–Hastings权重与邻居交换局部统计数据，并在每个决策时段进行一轮共识。所有方法均使用相同的随机种子初始化，以确保它们在各次试验中经历相同的环境随机实现。每种配置进行10次独立试验，报告的性能曲线显示样本均值及基于这些试验计算的95% 置信区间。为改善可视化效果，曲线经过轻度平滑处理，采用窗口长度为50个时间步的移动平均。此外，为考察不同噪声条件下的韧性，我们针对两种噪声水平进行仿真：$\sigma \in \{0.5, 1.0\}$。这使我们能够比较独立探索和协作探索在不同不确定性下的平衡表现。

我们实验中的韧性界限（α_1，β_1，α_2，β_2）并非算法超参数，而是设计者针对恢复和持久性规定的要求（见 IV 节和 V 节）。由于仿真环境中只有一个压力源，因此这些规格经过选择，确保韧性在固定时域 T 内仍然可行。相比之下，检测阈值（ε_1，L，e_{th}）决定何时将矛盾证据视为持续存在而非暂时现象。这些阈值根据变化前残差统计量选定（经实证确保虚警概率可忽略不计）。最后，认知证据阈值 η^{epi} 和行动置信度阈值 η^{act} 分别依据公式（11）和（12）中的一维规则进行调优。

我们在认知逻辑中对多臂赌博机环境建模如下。可能世界的集合为 $W = \{w_1, \dots, w_H\}$，其中每个世界 w_h 对应一个完整的臂均值向量 $\mu^{(h)} \in \mathbb{R}^A$。在任意给定时刻，恰有一个世界是真实的环境条件。

我们引入按世界逐一评估的原子命题：

- **逐臂均值原子：**

$$p_{a,h}:\ \mu_a = \mu_a^{(h)} \qquad (a = 1, \dots, A;\ h = 1, \dots, H),$$

- **可行动原子：**

$$b_a:\ a \text{ is an optimal arm}, \quad \text{i.e.,} \quad a \in \arg\max_{a'} \mu_{a'}.$$

世界 身份公式

$$\varphi_h \ = \ \bigwedge_{a=1}^{A} p_{a,h}$$

刻画世界 w_h。其可行动后果是公式

$$\psi_h \ = \ \bigvee_{a \in \arg\max_j \mu_j^{(h)}} b_a,$$

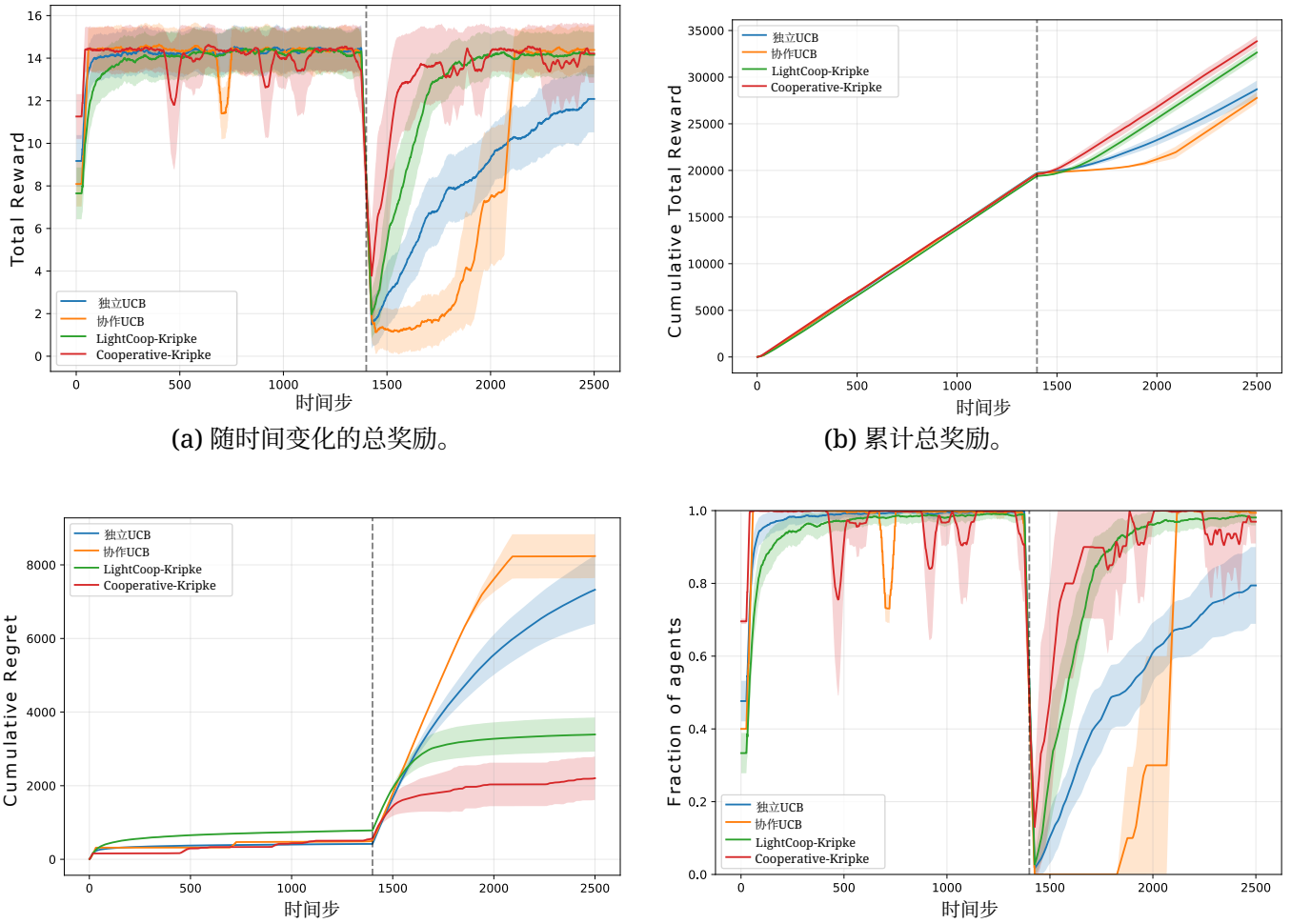
该命题指出，在世界 w_h 中所选的臂是最优的。对数似然比（LLR）提供了用于在 $\{\varphi_h\}_{h=1}^{H}$ 之间进行区分的累积成对证据分数。

在我们的仿真中，环境会在单一世界 w_h 中连续保持一段时间，并且仅在预先定义的扰动（压力源）时刻确定性地切换到另一个世界。我们将智能体 i 的抽象最优行动命题具体化为

$$(\mathcal{M}_t, r, t) \models \pi_i^{\text{opt}} \quad \Longleftrightarrow \quad a_{i,t} \in \arg\max_a \mu_a(t),$$

也就是说，智能体 i 选择的臂在时间 t 的真实世界中是最优的。在此，外部策略 $\pi_i^{\text{ext}}{}_{,t}$ 指的是直接决定智能体 i 在时间 t 所选臂的行为，使得 $a_{i,t} = \pi_i^{\text{ext}}{}_{t,}$。因此，当智能体的外部策略 $\pi_i^{\text{ext}}{}_{,t}$ 选择了当前环境条件下具有最高平均奖励的臂时，π_i^{opt} 恰好成立。在这种设定下，集合 $\Pi_{i\,t}^{\text{opt}}$，由所有在时间 t 的当前环境下达到最优的外部策略组成。因此，智能体 i 的期望奖励 J_i 由其外部策略以及随时间变化的世界序列 $\left(w_{h(t)}\right)_{t \geq 0}$ 共同决定。最后，我们通过标准指标评估性能：选择最优动作的智能体比例、总奖励和累积奖励，以及随时间变化的累积遗憾。这些共享设定为比较独立式、合作式和认知的（基于 Kripke 的）学习方案提供了统一基准，比较其在突发环境变化下的适应性、恢复速度和总体性能。

中等噪声水平。图 6 总结了所有学习模式的性能，包括独立折扣 UCB、协作折扣 UCB，以及在 $\sigma = 0.5$时提出的独立 Kripke 和 Cooperative-Kripke 框架。为简洁起见，图例中省略了术语 折扣，但该术语适用于独立UCB和协作UCB基线。对于此设置，两个 Kripke 学习模式的参数分别为 $\varepsilon_1 = 1.4$、$e_{\text{th}} = 12$、$L = 30$、$\eta^{\text{epi}} = 13$、$\alpha_1 = 550$、$\beta_1 = 600$、$\alpha_2 = 174$和 $\beta_2 = 436$。在协作基线中，智能体每个时间步执行一次共识更新。



(a) 随时间变化的总奖励。

(b) 累计总奖励。

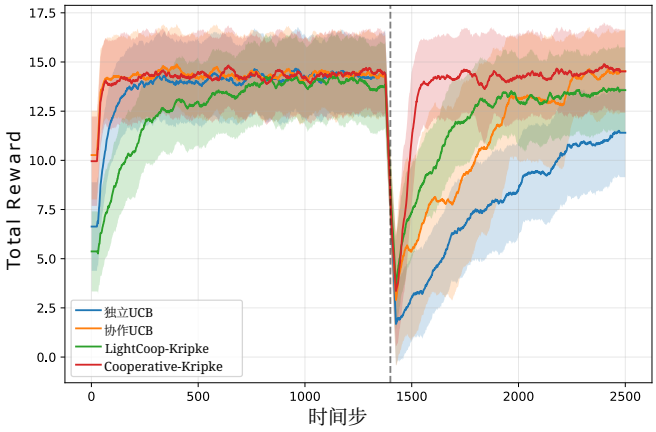
(c) 随时间变化的累积遗憾。

(d) 选择最优动作的智能体比例。

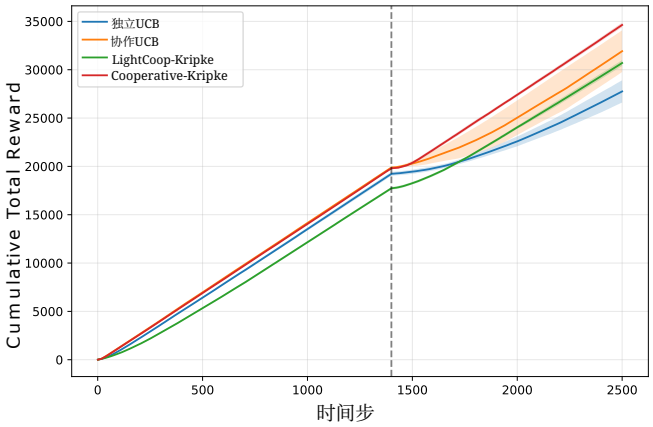
图 6. 噪声水平（$\sigma = 0.5$）下的性能。每个面板报告 10 次试验的均值及 95% 置信区间；虚线标示 $t = 1400$ 处的变点。对于折扣 UCB 基线，遗忘率设为 0。998。

在压力源发生之前（$t < 1400$），所有方法都迅速收敛到最优动作，并获得几乎相同的累积总奖励。在累积遗憾曲线中，差异变得更加明显，其中 LightCoop-Kripke 表现出最高的遗憾。这主要是因为每个智能体都独立进行探索，以解决认知矛盾，而不是像协作模式那样利用共享信息，也不像独立UCB那样更多地进行利用，从而导致累积遗憾略有增加。从图 6a 的总奖励随时间变化曲线可以看出，协作UCB和 Cooperative-Kripke 智能体在最初阶段表现出更快的收敛速度，这得益于通过共识共享统计数据。然而，协作UCB和 Cooperative-Kripke 在压力源出现之前的奖励中都出现了轻微的暂时性下降。在 Cooperative-Kripke 中，这些偏离对应于基于残差的门控机制偶尔触发的状态变化误检。由于智能体通过共识共享信息，虚警可能在时间上产生相关性，导致多个智能体同时进入证据收集阶段，并使总奖励暂时下降。尽管这些压力源出现前的偏离在视觉上较为明显，但其幅度和持续时间都很有限。在协作模式中，

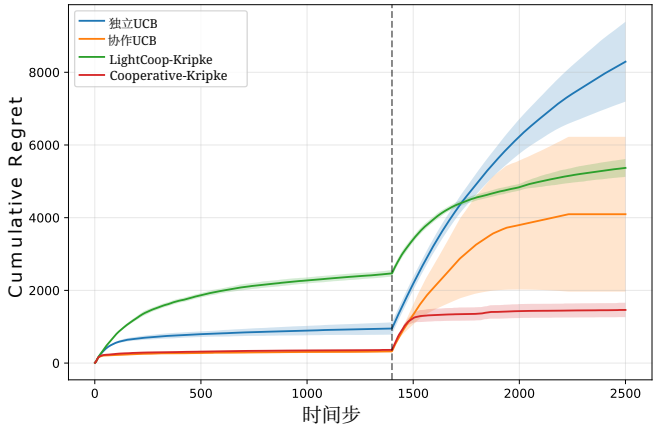
Kripke 中，每次误报警事件都会导致约 $10 - 25\%$ 的瞬时奖励下降，并持续约$40 - 50$ 个时间步。从累计奖励（图6b）和累积遗憾（图6c）可以看出，这些短暂的下降对长期表现影响微乎其微。也就是说，Cooperative-Kripke 累计获得的变化前总奖励与其他方法大致相同，并且保持了最低的累积遗憾。因此，这些暂时性下降反映的是短暂的探索性修正，而非显著的性能退化。相比之下，LightCoop-Kripke 在变化前没有出现明显的性能退化，因为其检测异步发生，不会使智能体的试探动作同步。Cooperative-UCB 中也出现了类似的下降，这是由共识诱导偏差造成的：有限的通信轮次使智能体的局部估计暂时趋同于某个次优臂，从而引发短暂的集体探索阶段。在此期间，Independent-UCB 保持稳定，因为各智能体的探索误差互不相关，并且会在总体表现中相互抵消。在 $t = 1400$ 处出现压力源后，所有学习模式的总奖励和选择最优动作的智能体比例都下降，累积遗憾也高于压力源出现前的水平



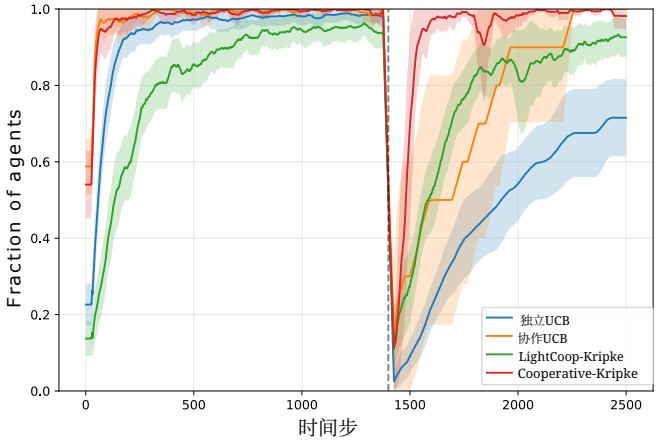
(a) 随时间变化的总奖励。



(b) 累计总奖励。



(c) 随时间变化的累积遗憾。

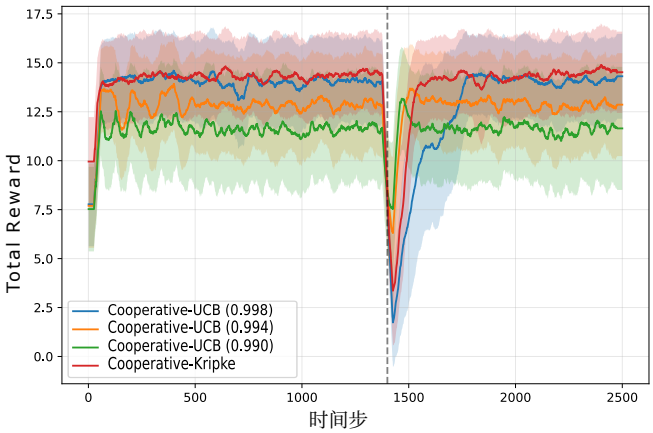


(d) 选择最优动作的智能体比例。

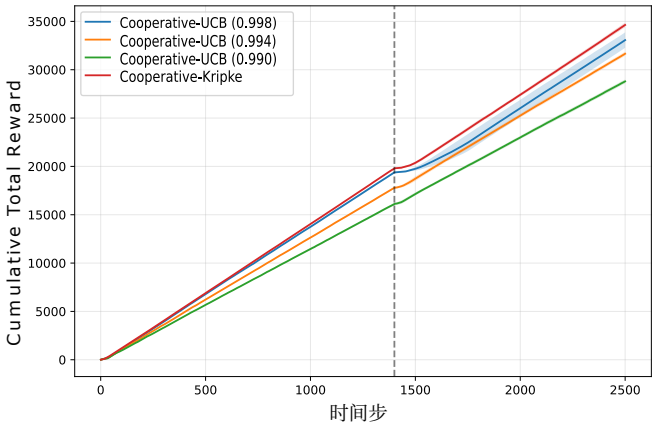
图7. 噪声水平（$\sigma = 1.0$）下的性能。每个面板报告 10 次试验的均值及 95% 置信区间；虚线标出 $t = 1400$ 处的变点。对于折扣 UCB 基线，遗忘率设为 0.998。

的水平。具体而言，独立UCB和协作UCB基线在所有指标上都出现了显著退化，因为仅靠遗忘不足以促使智能体及时重新探索并识别新的最优动作。不过，协作UCB基线的奖励下降幅度更大、持续时间更短，因此与独立UCB相比，其累积遗憾更高，累计总奖励更低。这是因为变化发生后，共享共识变量会使所有智能体暂时依赖过时的折扣全局估计值，从而延迟各自的重新探索。一旦少数智能体开始转向新的最优动作进行更新，共识机制便会迅速通过网络传播这一信息，促成总体表现的急剧恢复。相比之下，Independent-UCB智能体会异步响应：每个智能体独立进行探索，并在不同时间逐渐发现新的最优动作。这种去同步化使总体恢复更加平稳、瞬时损失更小，但整体适应速度较慢。

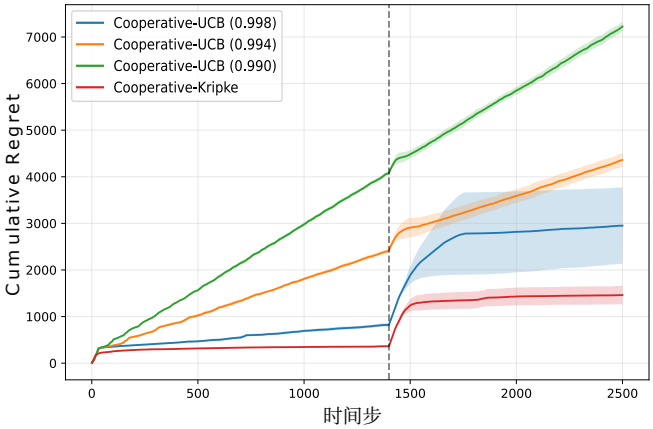
相比之下，基于克里普克的学习者展现出显著更快的认知恢复和行动恢复。LightCoop-Kripke（采用认知通信的独立学习）能够迅速检测到机制转变并开始重新调整，但是由于行动学习仍然相互独立，因此所有智能体都需要额外时间才能稳定在新的最优状态上。完全的 Cooperative-Kripke 模式整合了认知层和协作行动层，实现了最短的恢复时间窗口和变化后的最小瞬态遗憾。该模式迅速重建了对新假设的全局共识，并同步智能体行动，使其趋向最优动作。如子图6b和6c所示，其累积奖励曲线几乎沿着变化前的趋势，仅有极小的斜率下降；其累积遗憾则在所有方法中以最低水平趋于饱和。子图6d进一步证实，LightCoop-Kripke 和 Cooperative-Kripke 智能体能够更快地重新同步，而独立基线和协作基线则表现出较慢的转变。总体而言，结果验证了认知协调显著提升了突发环境变化后的恢复速度和稳定性。值得注意的是，即使通信受限（仅在认知恢复期间使用），LightCoop-Kripke 智能体的表现仍优于依赖持续通信来交换完整统计数据的基线方法。其表现也略逊于完全的 Cooperative-Kripke 框架；在该框架中，智能体协作恢复的同时还会共享统计数据。这意味着



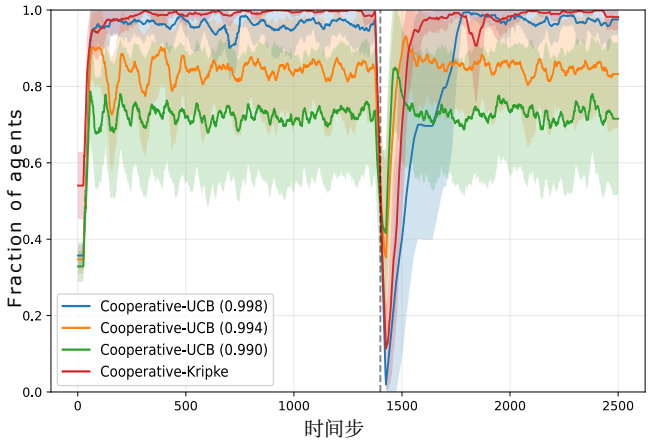
(a) 随时间变化的总奖励。



(b) 累计总奖励。



(c) 随时间变化的累积遗憾。



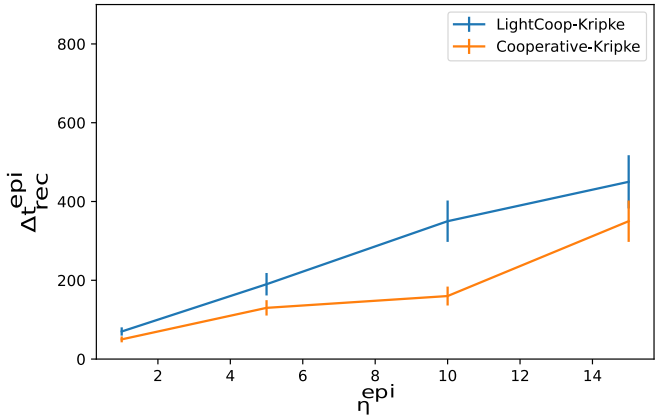
(d) 选择最优动作的智能体比例。

图 8. 噪声水平为（$\sigma = 1.0$）时的性能。每个面板报告 10 次试验的均值及 95% 置信区间；虚线标示 $t = 1400$ 处的变点。

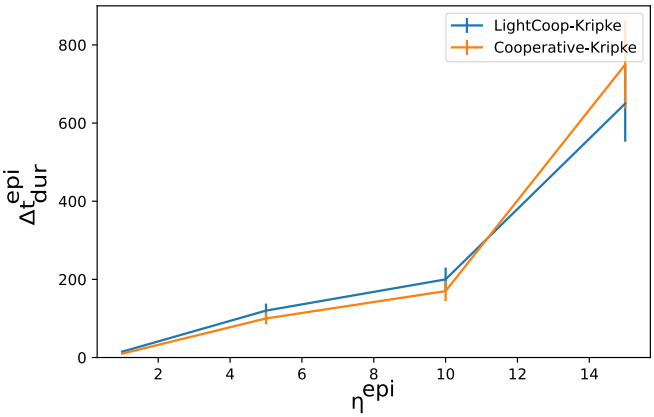
LightCoop-Kripke 在保持通信效率的同时，也能实现出色的性能（见表 II）。

高噪声水平。图 7 总结了 $\sigma = 1.0$ 时所有学习模式、独立折扣 UCB、协作折扣 UCB，以及所提出的独立 Kripke 和 Cooperative-Kripke 框架的性能。在这一设置下，两个 Kripke 学习模式的参数为 $\varepsilon_1 = 1.6$, $e_{\text{th}} = 13$, $L = 30$，以及 $\eta^{\text{epi}} = 10$, $\alpha_1 = 550$,$\beta_1 = 600$, $\alpha_2 = 174$，和 $\beta_2 = 436$。在较高噪声水平下，各方法之间的总体性能排序与中等噪声情形保持一致，但也出现了若干重要差异。值得注意的是，Cooperative-Kripke 和协作UCB 均不再表现出低方差下观察到的压力源发生前的小幅下降。噪声的增加有效平滑了小幅残余波动，降低了误检和共识引发的振荡的可能性。在压力源发生后，独立UCB 的遗憾明显更高、累积奖励明显低于协作UCB，这表明随着不确定性增加，通信变得更加有益。在这一机制下，协作帮助智能体对嘈杂的奖励观测进行平均，并加快对新均值的重新估计，从而实现更快速的集体恢复。Cooperative-Kripke 再次实现了最快且最稳定的恢复，而 LightCoop-Kripke 保持了略低于完全协作模式、但仍优于两个 UCB 基线的中等性能。这再次表明，LightCoop-Kripke 在实现通信效率的同时能够提供可比的性能（见表 II）。总体而言，这些结果表明，在高方差环境中，通信和认知协调能够带来更大的优势；在此类环境下，共享信息可以缓解噪声，并增强对突发机制转变的韧性。

遗忘率的影响：图 8 比较了不同遗忘率下 Cooperative-Kripke 框架与协作折扣 UCB 基线的表现。所有其他仿真参数均与图 7 中使用的参数相同。在所有指标上，Cooperative-Kripke 始终实现最高的累积奖励和最低的累积遗憾，展现出变化后快速且完全的恢复能力。相比之下，协作 UCB 基线的表现很大程度上取决于遗忘率。较高的遗忘率（例如 0.998）会导致适应速度变慢，因为历史样本在估计值中占主导地位，从而延迟了对



(a) 认知恢复时间与 η^{epi}的关系。



(b) 认知持久时间与 η^{epi} 的关系。

图 9. 不同 η^{epi} 水平下的认知恢复时间和持久时间，噪声水平为（$\sigma = 1.0$）。每个面板报告 10 次试验的均值及 95% 置信区间。

新奖励均值的响应。较低的遗忘率（例如 0.990）通过更积极地折扣旧信息加快恢复，但代价是恢复后的方差更高，稳态奖励水平更低，因为智能体会对近期含噪观测赋予过高权重。因此，提高遗忘率可增强响应能力，但会降低渐近精度。Cooperative-Kripke 框架通过显式检测机制转变，并仅在认知证据支持变化时重置信念，避免了这种权衡，从而既能快速恢复，又能收敛到新的最优值，同时不牺牲稳态表现。

可恢复性—持久性权衡：图 9 展示了随着证据阈值 η^{epi} 增大，认知恢复与持久性之间固有的权衡。更高的阈值会在提交新的世界假设之前施加更严格的证据要求，从而延长证据累积阶段，进而使 $\Delta t^{\mathrm{epi}}_{\mathrm{rec}}$ 增大。同时，较高的阈值也增强了对已提交假设的信心，使恢复后的持久性更强、$\Delta t^{\mathrm{epi}}{}_{\mathrm{dur}}$ 更长。这一关系体现了 V-A 节中提出的理论界限，其中 η^{epi} 有效地调节最大允许恢复时间 α_1 与最低要求持久时间 β_1 之间的平衡。在所有阈值下，LightCoop-Kripke 的恢复时间始终长于 Cooperative-Kripke，因为每个智能体都必须收集足够的局部证据，才能就变化达成共识。协作模式受益于信息汇集，因此能够更快地实现认知对齐。相反，LightCoop-Kripke 在恢复后具有略高的持久性，因为其个体决策较不容易受到偶发全网扰动的影响。在较高的 η^{epi} 下，两种方法都能稳定运行至仿真时域结束；因此，持久性排序表面上的反转，仅仅反映了较早恢复的方法能在剩余时间内维持新信念更久。总体而言，结果证实了预测的可恢复性—持久性权衡，并表明 Cooperative Kripke 提供了一个有利的运行点，能够实现快速且稳定的认知适应。此外，在我们的实验中，行动恢复时间 $\Delta t^{\mathrm{act}}_{\mathrm{rec}}$ 比认知恢复晚几个时间步（即紧接在 $t^{\mathrm{epi}}_{\mathrm{rec}}$ 之后的第一次决策）。这是因为两种 Kripke 模式下的智能体都会使用已提交的认知状态重新初始化其 UCB 学习器，导致最优动作的 UCB 指数立即占据优势。我们还观察到，行动持久性与认知持久性紧密同步，两者仅相差一个对应于短暂行动恢复阶段的小偏移。一旦群体的认知状态与正确世界保持对齐，UCB 奖励便会迅速减小，最优动作也会始终明显占优。只有当认知层因虚警重新进入证据模式，或引入新的压力源时，才会偏离最优动作。具体而言，在所有实验中，我们都使用行动集实例化 Π^{ext}_i，并通过高斯 UCB 实现最优策略识别。置信水平 η^{act} 隐含于 UCB 探索奖励以及在 $t^{\mathrm{epi}}_{\mathrm{rec}}$ 时的重新初始化中；因此，统计分离条件几乎立即得到满足。分离条件几乎立即得到满足。此外，我们的实现采用在线赌博机更新，其中提交时的重新初始化步骤等同于对每个动作执行少量虚拟展开，

在所有情形下，认知协调都是韧性的关键驱动因素：与衰减记忆基线相比，我们提出的认知策略可缩短恢复窗口，并提高恢复后稳定性。合作式 Kripke 模式通过汇集认知知识和行动统计量，实现了最快的恢复速度和最低的遗憾。与此同时，LightCoop-Kripke（事件触发变体）在通信量远低于独立式/合作式 UCB 基线的情况下，仍优于这些基线。它的表现也非常接近 Cooperative-Kripke，同时实现了更高的通信效率。我们还表明，噪声越大，合作的价值越突出；而折扣 UCB 中的遗忘率调节面临速度—稳定性权衡，认知层则在很大程度上避免了这一问题，因为它只会在证据支持时进行修正。

可扩展性：为评估可扩展性，我们在网络规模不断增大的情况下重复了图 7 中采用的相同设置 $|N| \in \{10, 150, 300\}$，同时保持环形通信拓扑。我们在 V-A 节（见公式（9））中的分析表明

表 II 环形拓扑与小世界拓扑下的性能比较。总恢复时间定义为总韧性延迟$\Delta t_{\mathrm{rec}}^{\mathrm{epi}} + \Delta t_{\mathrm{rec}}^{\mathrm{act}}$，即扰动发生后，平均奖励恢复至其稳态最大水平（在较小容差范围内）所需的时间。

数字的智能体	算法	环形拓扑			小世界拓扑		
		总计恢复time	平均奖励之后恢复	总计通信（每个智能体每步）	总计恢复time	平均奖励后恢复	总计通信（每个智能体每步）
10	独立UCB	1100	0.95	–	1100	0.95	–
	协作UCB	600	**1.20**	50，000(2)	500	1.18	100，000(4)
	LightCoop-Kripke	400	1.19	**20(8 × 10⁻⁴)**	360	1.19	**24(9.6 × 10⁻⁴)**
	快速轻量协作Kripke	350	1.19	**20(8 × 10⁻⁴)**	350	1.19	**24(9.6 × 10⁻⁴)**
	Cooperative-Kripke	**160**	**1.22**	50，010(∼ 2)	**150**	**1.20**	100，020(∼ 4)
150	独立UCB	1100	0.95	–	1100	0.95	–
	协作UCB	650	**1.20**	750, 000(30)	500	1.18	1，500，000(60)
	LightCoop-Kripke	750	1.00	**15，750(0.63)**	600	1.10	**4，860(0.19)**
	快速轻量协作Kripke	400	1.15	**15, 750(0.63)**	400	1.18	**4，860(0.19)**
	Cooperative-Kripke	**150**	**1.20**	765，750(30。6)	**120**	**1.20**	1，505，100(60。2)
300	独立UCB	1100	0.95	–	1100	0.95	–
	协作UCB	650	**1.20**	1，500，000(60)	500	**1.18**	3，000，000（120）
	LightCoop-Kripke	950	1.00	**54，000(2.16)**	800	1.10	**13，086(0。52)**
	快速轻量协作Kripke	410	1.15	**54，000(2.16)**	400	1.15	**13，086(0.52)**
	Cooperative-Kripke	**150**	**1.20**	1，581，000（63。24）	**130**	1.17	3, 017, 280(120.7)

LightCoop-Kripke 的恢复延迟 $\Delta t_{\mathrm{rec}}^{\mathrm{epi}}$ 会随着通信图的直径 l_{comm} 增大。这意味着，在环形拓扑中，LightCoop-Kripke 的恢复时间 $\Delta t_{\mathrm{rec}}^{\mathrm{epi}}$ 会随着智能体数量增加而显著增长。这一行为是环形拓扑跳数直径 $\ell_{\mathrm{comm}} = \Theta(|N|)$ 直接导致的：认知提交必须依次在网络中传播，随着图规模扩大，共同知识形成也因此变慢。相比之下，Cooperative-Kripke 基本不受规模扩大的影响，因为其持续进行的共识混合能够维持同步信念，并避免冗长的提交传播阶段。

为缓解拓扑引发的延迟，我们考虑了两项改进。首先，我们用小世界网络替代环形网络。小世界网络保留了稀疏性，但其直径相对于 $|N|$ 仅呈对数增长。与环形拓扑相比，这类拓扑呈现出更符合现实的邻居连通模式。其次，我们测试了快速轻量协作Kripke变体，该变体允许智能体在其局部对数似然比超过 η^{epi} 后立即提交，同时无需等待完整的时域确认即可进行广播；如果之后收到具有更高证据的假设，智能体将相应地进行修订。这降低了对 ℓ_{comm} 的依赖，从而能够及时实现认知对齐。

表 II 从总恢复时间 $\Delta t_{\mathrm{rec}}^{\mathrm{epi}} + \Delta t_{\mathrm{rec}}^{\mathrm{act}}$、恢复后的平均奖励（即在较小容差范围内的稳态最大水平）以及总通信成本（括号内为归一化的每个智能体每步成本）三个方面考察了五种方法的可扩展性。总通信成本衡量所有智能体在所有轮次中通过图边传输的消息总数，其中包括共识更新（如有）以及用于认知提交的广播泛洪（如有）。这反映了该算法的完整通信开销。独立UCB不产生通信开销，但始终表现出恢复最慢，稳态奖励也最低，表明其在动态环境中的韧性较差。协作UCB能够快速恢复，并在恢复后获得较高的平均奖励；随着网络规模扩大，其恢复时间仍保持平稳。不过，这要以极高的通信开销为代价；由于节点度较高，在小世界拓扑中，通信开销还会进一步增加。也就是说，在我们的设置中，环形拓扑中的每个智能体的度为 2（有两个邻居），而小世界图中的平均节点度为 4。由于每步都要与每个邻居交换共识消息，节点度翻倍会使每步通信负载大致翻倍，这解释了为什么小世界拓扑中的总通信成本约为环形拓扑的两倍。由于需要共享认知证据，Cooperative-Kripke 的通信量略高于协作UCB，但其恢复速度显著更快，而且随着智能体数量增加，恢复时间同样保持平稳，恢复后也能获得较高的平均奖励。若优先考虑通信效率，LightCoop-Kripke 则提供了颇具优势的权衡：即使在 $n = 300$ 下也能显著降低通信成本，但在环形拓扑中，其总恢复时间会随着智能体数量增加而延长（因为 l_{comm} 也会随之增加）。LightCoop-Kripke-fast 通过允许智能体在收集到足够的局部证据后立即提交，消除了等待网络直径所带来的延迟；它在保持较低通信成本的同时，使恢复时间随着智能体数量增加仍几乎保持平稳，并在恢复后获得接近较高水平的平均奖励。请注意，LightCoop-Kripke 和 LightCoop-Kripke-fast 的通信成本基本相同，因为二者使用相同的 Kripke 证据层：只有当智能体的局部对数似然差值超过认知阈值时，才会发送假设广播消息，并且每条此类广播都会沿每条边转发一次。“快速”

变体只改变智能体决定提交的时机（它们不再等待 l_{comm} 步后才更新世界模型），但不会改变广播消息的发送数量，因此总通信负载实际上保持不变。由于小世界拓扑会大幅缩短网络直径，有关新假设的信息能够更快地传播到所有智能体。因此，智能体较少发起重复广播，因为它们能更快地从其他智能体获知变化，从而使通信开销相比环形拓扑大幅降低。这些结果表明，LightCoop-Kripke 在环形拓扑中的减速并非认知协调的根本局限，而是由拓扑结构引起的效应：在实际的小世界连通性下，即使在 $n = 300$ 时，恢复时间和通信成本也都能得到良好控制；FAST变体则以最小的通信开销实现了可扩展的韧性。

VIII. 结论与未来工作

本文提出了一种新颖的形式化框架，用于量化和设计动态系统中的多智能体韧性。我们引入了一种基于Kripke的学习架构，将认知韧性（检测、共享并恢复关于世界状况的准确信息的能力）与行动韧性（在受到干扰后重新调整并维持最优行为的能力）相结合。我们的形式化方法通过可测量的恢复时间和持久时间将这两个维度联系起来，从而为韧性给出了一个有原则的定义，贯通了学习、通信与控制。通过大量仿真，我们证明了所提出的 Cooperative-Kripke 和 LightCoop-Kripke 架构优于独立折扣 UCB 和协作折扣 UCB 等经典基线。结果表明，即使在通信受限的情况下，认知协调也能显著提升恢复速度和变化后的稳定性。具体而言，Cooperative-Kripke 能够快速实现认知恢复和行动恢复，同时保持较高的累积奖励和较低的遗憾；LightCoop-Kripke 的性能略低，但通信负载大幅减少。实验还验证了可恢复性—持久性权衡：更严格的证据阈值会延长证据收集阶段，但也能在恢复后带来更长时间的稳定性。

一个很有前景的研究方向是将韧性设计问题表述为混合整数线性规划（MILP）。这种表述能够将认知转移的逻辑约束、恢复和持久性的时间预算（α_1，β_1，α_2，β_2），以及通信和计算预算编码为线性与整数约束，从而在通信和感知限制下，精确计算可行的韧性规范或最优参数选择（η^{epi}，η^{act}）。求解该混合整数线性规划，可为合成韧性智能体配置并在部署前验证韧性保证提供系统化方法。这一优化视角与基于学习的方法相辅相成，在统一框架中兼顾设计时验证和运行时韧性。

未来工作的另一个方向是纳入显式的碰撞感知动力学。在当前表述中，智能体之间的交互被视为环境的一部分，因此会进入每个智能体的认知状态。这使得潜在碰撞可以被解读为意外观测，并触发与处理其他压力源时相同的认知更新机制。然而，若通过显式碰撞模型（例如考虑竞争的奖励结构、联合动作可行性约束或协调协议）扩展该框架，便可在易发生碰撞的场景中实现主动规避，并支持更丰富的多智能体交互。最后，尽管我们的算法通过残差阈值、超限计数和证据裕度来防范噪声观测，但对不可靠通信进行建模仍是未来一项重要的扩展方向。

参考文献

[1] L. Khaloopour、Y. Su、F. Raskob、T. Meuser、R. Bless、L. Janzen、K. Abedi、M. Andjelkovic、H. Chaari、P. Chakraborty 等，“6G网络中的设计韧性：文献综述与新型赋能概念”，*IEEE Access*，2024年。[2] V. Vesterby，“复杂网络中的通用韧性模式——对Gao、Barzel和Barab´asi论文的分析”，2022年10月。 [在线]。可用地址：osf.io/hz5pr [3] R.-J. Reifert、M. Krawczyk-Becker、L. Prenzel、S. Pavlichkov、M. A. Khatib、S. A. Hiremath、M. Al-Askary、N. Bajcinca、S. Steinhorst和A. Sezgin，“迈向混合关键工业控制系统的韧性：多学科视角”，*IEEE Access*，第10卷，第124 563–124 581页，2022年。[4] A. Weissberger，“面向2030年及未来的IMT框架”，https://techblog.comsoc.org/tag/imt-framework-for-2030-and-beyond/，2023年，IEEE ComSoc技术博客，发布于2023年7月9日。[5]中国工业和信息化部，“中国工业和信息化部将优先推进6G项目，加快5G和千兆光网络部署”，https://techblog.comsoc.org/2023/01/17/chinas-miit-to-prioritize-6g-project-accelerate-5g-and-gigabit-optical-network-deployments/，2023年，IEEE ComSoc技术博客，发布于2023年1月17日。[6] M. Bennis、M. Debbah和H. V. Poor，“超可靠低时延无线通信：尾部、风险与规模”，*IEEE*学报，第106卷，第10期，第1834–1853页，2018年。[7] T. Meuser、L. Lov´en、M. Bhuyan、S. G. Patil、S. Dustdar、A. Aral、S. Bayhan、C. Becker、E. d. Lara、A. Y. Ding、J. Edinger、J. Gross、N. Mohan、A. D. Pimentel、E. Rivi`ere、H. Schulzrinne、P. Simoens、G. Solmaz和M. Welzl，“重新审视边缘人工智能：机遇与挑战”，*IEEE Internet Computing*，第28卷，第4期，第49–59页，2024年。[8] M. van Steen和A. S. Tanenbaum，《分布式系统》，第*3*版。阿姆斯特丹自由大学，2017年，初步版本。[9] L. Sun、Y. Yang、Q. Duan、Y. Shi、C. Lyu、Y.-C. Chang、C.-T. Lin和Y. Shen，“跨多样化应用的多智能体协调：综述”，2025 [ ] 2502 14743 。在线。可用地址：https://arxiv.org/abs/ 。[10] Y. Zou、H. Liu、K. Xia、S. Zhang、Y. Wu、D. Li和Z. Zuo，“异构非对称输入饱和与异步时变通信时延下多智能体系统的一致性鲁棒性”，*Information Sciences*，第647卷，第119483页，2023年。 [在线]。可用地址：https://www.sciencedirect.com/science/article/pii/S002002552301068X [11] R. Francis和B. Bekera，“工程系统与基础设施系统韧性分析的度量与框架”，*Reliability Engineering & System Safety*，第121卷，第90–103页，2014年1月。[12] V. Ramaswamy，“主题演讲——从凤凰的视角展望未来网络设计”，2023年3月。[13] M. Scheffer、J. Bascompte、W. A. Brock、V. Brovkin、S. R. Carpenter、V. Dakos、H. Held、E. H. Van Nes、M. Rietkerk和G. S ugihara，“临界转变的早期预警信号”，《自然》，第461卷，第7260期，第53–59页，2009年。[14] L. Dai、D. Vorselen、K. S. Korolev和J. Gore，“导致种群崩溃的临界点之前韧性丧失的通用指标”，《科学》，第336卷，第6085期，第1175–1177页，2012年。[15] K. Barker、J. E. Ramirez-Marquez和C. M. Rocco，“基于韧性的网络组件重要性度量”，*Reliability Engineering & SystemSafety*，第117卷，第89–97页，2013年。[16] Y.-P. Fang、N. Pedroni和E. Zio，“关键基础设施网络系统的基于韧性的组件重要性度量”，*IEEE*可靠性汇刊，第65卷，第2期，第502–512页，2016年。

[17] T.-Y. Liao、T.-Y. Hu和Y.-N. Ko，“灾害下交通网络的韧性优化模型”，Natural hazards，第93卷，第469–489页，2018年。[18] D. Hulse和C. Hoyle，“理解韧性优化架构：多层分解策略中的对齐与耦合”，Journal of Mechanical Design，第144卷，第11期，第111704页，2022年。[19] Y. Wu、I. El Hallaoui、F. Soumis等，“迈向韧性：原始大规模再优化”，Transportation Research Part E: Logistics and Transportation Review，第192卷，第103819页，2024年。[20] N. H. Mahmood、S. Samarakoon、P. Porambage、M. Bennis和M. Latva-aho，“设计韧性：面向未来无线网络的韧性框架”，2024年。 [在线]。可用地址：https: //arxiv.org/abs/2410 23203 。R.-J. Reifert、Y. Karacora、C. Chaccour、[21] A. Sezgin和W. Saad，“韧性与关键性：6G的两大支柱”，2024年。 [在线]。可用地址：https://arxiv.org/abs/2412 03661 。[22] “Rings II 报告”，https://rings-vo.org/wp-content/uploads/2024/12/NSF-RINGS-Virtual-Organization-PI-Meeting-Report July2024 Yorktown-Heights NY.pdf。[23] M. Bennis、M. Debbah和H. V. Poor，“超可靠低时延无线通信：尾部、风险与规模”，IEEE学报，第106卷，第10期，第1834–1853页，2018年。[24] M. Di Mauro、W. Cerroni、F. Postiglione、M. Tornatore和K. S. Trivedi，“虚拟化网络中的可靠性与可用性：标准、建模方法与研究挑战综述”，arXiv预印本，2025 [ ] 2503 22034 。在线。可[25] 6用地址：https://arxiv.org/abs/ 。“一旦6G被定义，韧性将至关重要”，https://spectrum. ieee.org/6g-resiliency。[26] J. Shi、X. Chen、S. Xing、A. Liu和C. Chen，“受执行器故障影响的不确定多智能体系统的鲁棒协作容错控制”，Sensors，第24卷，第8期，第2651页，2024年。[27] A. Prorok、M. Malencia、L. Carlone、G. S. Sukhatme、B. M. Sadler和V. Kumar，“超越鲁棒性：迈向韧性多机器人系统的方法分类”，arXiv预印本arXiv:2109.12343，2021年。[28] A. Lomuscio和W. Penczek，《时序认知逻辑的符号模型检测》。柏林、海德堡：Springer Berlin Heidelberg，2012年，第172–195页。 [在线]。可用地址：https://doi.org/10. 1007/978-3-642-29414-3 10[29] H. van Ditmarsch、W. van der Hoek、J. Halpern和B. Kooi，编，《认知逻辑手册》。College Publications，2015年。[30] R. Fagin、J. Y. Halpern、Y. Moses和M. Y. Vardi，《关于知识的推理》。麻省理工学院出版社，2004年。[31] S. A. Kripke，“关于模态逻辑的语义考察”，载于《模态逻辑与多值逻辑讨论会论文集》，Acta Philosophica Fennica系列，1963年，第16期，第83–94页。[32] P. Blackburn和J. van Benthem，“1 模态逻辑：语义视角”，载于《模态逻辑手册》，Studies in Logic and Practical Reasoning系列，P. Blackburn、J. Van Benthem和F. Wolter编。Elsevier，2007年，第3卷，第1–84页。 [在线]。可用地址：https://www.sciencedirect.com/science/article/pii/S1570246407800048[33] A. Bjorndahl，《经典逻辑与模态逻辑导论：知识概要》，剑桥哲学导论系列。剑桥大学出版社，2024年。[34] J. Garson，“模态逻辑”，载于《斯坦福哲学百科全书》，2024年春季版，E. N. Zalta和U. Nodelman编。斯坦福大学形而上学研究实验室，2024年。[35] H. P. van Ditmarsch、W. van der Hoek和B. Kooi，《动态认知逻辑》，Synthese Library系列。Springer，2008年，第337卷。[36] M. Bennis和S. Lahlou，“语义通信遇上系统2机器学习：抽象、组合性与涌现语言如何塑造智能”，2025 [ ] 2505 20964 。在线。[37] 可用地址：https://arxiv.org/abs/ 。H. Seo、J. Park、M. Bennis和M. Debbah，“通过上下文推理实现语义原生通信”，IEEE Transactions on Cognitive Communications and Networking，第9卷，第3期，第604–617页，2023年。[38] E. C. Strinati、P. Di Lorenzo、V. Sciancalepore、A. Aijaz、M. Kountouris、D. G¨und¨uz、P. Popovski、M. Sana、P. A. Stavrou、B. Soret、N. Cordeschi、S. Scardapane、M. Merluzzi、L. Zanzi、M. B. Renato、T. Quek、N. D. Pietro、O. Forceville、F. Costanzo和P. Li，“6G人工智能原生网络中的目标导向与语义通信：6G-GOALS 方法”，载于《2024年联合欧洲网络与通信会议暨6G峰会》（EuCNC/6G Summit），2024年，第1–6页。[39] W. Yang、H. Du、Z. Q. Liew、W. Y. B. Lim、Z. Xiong、D. Niyato、X. Chi、X. Shen和C. Miao，“面向未来互联网的语义通信：基础理论、应用与挑战”，IEEE Communications Surveys & Tutorials，第25卷，第1期，第213–250页，2023年。[40] G. Xin、P. Fan 和 K. B. Letaief，《语义通信：理论发展的综述》，*Entropy*，第26卷，第2期，2024年。 [在线]。可访问：https://www.mdpi.com/1099 4300 26 2 102 - / / / Y. Liu、X. Wang、Z.[41] Ning、M. Zhou、L. Guo 和 B. Jedari，《语义通信综述：技术、解决方案、应用与挑战》，*Digital Communications andNetworks*，第10卷，第3期，第528–545页，2024年。 [在线]。可访问：https: //www.sciencedirect.com/science/article/pii/S2352864823000925[42] S. Guo、Y. Wang、N. Zhang、Z. Su、T. H. Luan、Z. Tian、Xuemin 和 Shen，《语义通信网络综述：架构、安全与隐私》，2024年。 [在线]。可访问：https: //arxiv.org/abs/2405 01221 . T. M. Getu、G. Kaddoum 和 M. Bennis，《目标导[43] 向语义通信综述：技术、挑战与未来方向》，*IEEE Access*，第12卷，第51 223–51274页，2024年。[44] B. Sakcak、K. G. Timperi、V. Weinstein 和 S. M. LaValle，《对最低限度充分的机器人脑的数学刻画》，*The International Journal of Robotics Research*，第0, 0, 1 21, 卷，2023 [ ] 10 1177 02783649231198898 期，页码–。在线。可访问：https://doi.org/。/[45] K. Friston、J. Kilner 和 L. Harrison，《大脑的自由能原理》，*Journal of Physiology-Paris*，第100卷，第1期，第70–87页，2006年，理论与计算神经科学：理解大脑功能。 [在线]。可访问：https://www.sciencedirect.com/science/article/pii/S092842570600060X[46] J. Y. Halpern，《关于知识的推理：综述》。 [在线]。可访问：https://www.cs.cornell.edu/home/halpern/papers/knowledgesurvey.pdf[47] M. Agarwal、V. Aggarwal 和 K. Azizzadenesheli，《通信受限的多智能体多臂老虎机》，*Journal of Machine Learning Research*，第23卷，第212期，第1–24页，2022年。 [在线]。可访问：http://jmlr.org/papers/v23/21-138.html[48] S. Agrawal 和 N. Goyal，《多臂老虎机问题中汤普森采样的分析》，载于第*25*届年度学习理论会议，Proceedings of Machine Learning Research 系列，S. Mannor、N. Srebro 和 R. C. Williamson 编，第23卷。苏格兰爱丁堡：PMLR，2012年6月25日至27日，第39.1–39.26页。 [在线]。可访问：https://proceedings.mlr.press/v23/agrawal12.html[49] A. Slivkins，《多臂老虎机导论》，2024年。 [在线]。可访问：https://arxiv.org/abs/1904 07272 . A. Garivier 和 E. Moulines，《切换[50] 老虎机问题的上置信界策略》，载于算法学习理论，J. Kivinen、C. Szepesv´ari、E. Ukkonen 和 T. Zeugmann 编。柏林、海德堡：Springer Berlin Heidelberg，2011年，第174–188页。[51] T. Lattimore 和 C. Szepesv´ari，《老虎机算法》。剑桥大学出版社，2020年。