arXiv:2301.06157v1 [cs.GT] 15 Jan 2023

合作并发博弈

Julian Gutierrez[a]、Szymon Kowara[c]、Sarit Kraus[b]、
Thomas Steeples[c]、Michael Wooldridge[c]

*a*信息技术学院，莫纳什大学，澳大利亚 *b*计算机科学系，巴伊兰大学，以色列 *c*计算机科学系，牛津大学，英国

摘要

在理性验证中，目标是在假设系统中的智能体（“玩家”）选择的行动策略构成博弈论均衡的前提下，验证多智能体系统中哪些时序逻辑性质会成立。偏好通常通过假设智能体以实现个体目标为行动目的来定义，这些目标以时序逻辑公式的形式指定。迄今为止，理性验证一直采用非合作解概念进行研究，例如纳什均衡及其细化概念。这类非合作解概念假定智能体不可能达成有约束力的协议来进行合作，因此其适用范围有限。本文将理性验证拓展至合作博弈论领域所研究的合作解概念。我们聚焦于核心，因为它是最基础且研究最广泛的合作解概念。我们首先提出一种似乎适用于并发博弈情形的核心变体，并证明这一版本的核心可以用 ATL*刻画。随后，我们研究与核心相关的关键决策问题的计算复杂性，这些问题的复杂度从 PSpace 到 3-指数时间不等。

⋆本文是对 [1] 的扩展和修订版本。

*通讯作者：Julian Gutierrez。电子邮箱：julian.gutierrez@monash.edu (Julian Gutierrez), sarit@cs.biu.ac.il (Sarit Kraus), thomas.steeples@cs.ox.ac.uk (Thomas Steeples), michael.wooldridge@cs.ox.ac.uk (Michael Wooldridge)

提交至《人工智能》的预印本 *2023*年*1*月*18*日

我们还探讨了确保核心非空的充分条件，并研究核心在何种情况下具有双模拟不变性。接着，我们提出并研究了核心主要定义的若干变体，由此引出可信偏离问题，以及更强的集体稳定行为概念。最后，我们采用另一种偏好模型研究合作理性验证。在该模型中，玩家试图在允许定量信息的博弈中，最大化其在无限次对局中获得的平均收益。关键词：并发博弈；合作博弈；多智能体系统；逻辑；形式化验证

---

1. 引言

如今，Siri、Alexa 和 Cortana 等智能体已被数百万人每天使用。早在 20 世纪 90 年代初，人们就提出了这样一种构想：此类智能体应当成为主动型助手，致力于追求人类用户委托给它们的目标。当前的智能体运用了日益丰富的人工智能技术，从其界面中的自然语言理解开始，用于与人类用户进行交流。此类智能体发展的下一步，预计将是能够在追求其委托目标的过程中，不仅与人类用户互动，还能彼此互动。这类多智能体系统带来了一系列研究挑战，在过去三十年中一直是大量研究的主题。本文所要探讨的问题，是理解此类系统的动态。特别是，相互作用的智能体集合可能呈现出不理想且不可预测的动态。因此，如何验证多智能体系统可能表现出的行为，自然就成为一个问题。

一种日益流行的多智能体系统验证方法，是将系统视为博弈（即博弈论意义上的博弈），其中智能体为追求委托目标而理性且有策略地行动。在此基础上，自然会产生这样的问题：假定智能体以理性方式行动——遵循博弈论解概念，一个系统可能会表现出哪些行为。这正是理性验证范式的核心思想 [2–5]。

以往关于理性验证的研究一直采用 并发博弈 作为多智能体系统的语义模型。自首次提出以来 [6], ，并发博弈已在人工智能社区以及验证与计算机科学领域得到广泛应用 [2–4]。并发博弈 [6] 是一种有限状态环境，其中包含一组彼此独立、追求自身利益的智能体。博弈以无限序列的回合进行，每一回合中，每个智能体都会选择一个动作来执行。并发博弈中的偏好通常通过假设每个智能体都关联一个时序逻辑目标公式 [7], 来建模，而该智能体希望这个公式得到满足。博弈产生的无限对局要么满足每位玩家的目标，要么不满足；玩家会理性地行动，试图实现自己的目标。由于玩家目标是否满足可能取决于其他玩家的选择，因此玩家必须采取策略性选择，才能实现最优对局。

然而，据我们所知，此前所有研究都假设并发博弈是非合作的：玩家独自行动，玩家之间不允许达成有约束力的协议。因此，此前并发博弈研究所采用的博弈论解概念，都是非合作博弈论所研究的概念——主要是纳什均衡及其各种精炼。在这种非合作场景中，我们针对并发博弈提出的基本问题包括：某个时序逻辑性质是否成立于系统的某个 可能计算路径上，而该路径可能由玩家选择策略后产生，且这些策略构成一个

纳什均衡（E-纳什问题）；或者某个性质是否成立于所有 这类计算路径上（A-纳什问题）。这些问题可视为传统模型检测问题 [8] 的博弈论对应问题：在模型检测中，我们通常关注某个性质是否可能成立于系统的某些或所有可能计算路径上；而在理性验证中，我们关注某个性质是否成立于玩家理性选择可能产生的某些或所有计算路径上。并发博弈中相应判定问题的复杂度已得到广泛研究，目前已有少量支持理性验证的软件工具 [9–12]。

然而，非合作博弈论只是博弈论的一个分支。合作博弈论是一个广受研究但知名度较低的博弈论分支，它与非合作博弈论的不同之处在于，它允许智能体之间达成有约束力的协议。有约束力的协议使智能体能够组成团队，并享受合作带来的好处。在现实世界中，我们通过合同和其他机制（包括正式和非正式的）来促成有约束力的协议。

因此，本文旨在将理性验证的研究扩展到包含合作解概念 [13–15]。因此，我们假设存在某种（外生）机制，使并发博弈中的玩家能够达成有约束力的协议并组建团队（“联盟”），以共同实现目标（尽管我们强调，此类机制的性质不在本文讨论范围之内）。有约束力的合作与联盟形成的可能性，消除了非合作情形中出现的一些不理想均衡，并使得一系列无法通过合作实现的结果成为可能。我们重点关注核心这一关键解概念。核心背后的基本思想是，它由这样一些策略组合构成：不存在任何子集的

玩家能够通过共同偏离而获益。在传统合作博弈中（具有可转移效用 [15] 的特征函数博弈），这一理念可以得到一种简单而自然的形式化定义，因此，核心或许是合作博弈中研究最广泛的解概念。然而，核心的传统定义并不能自然地适用于我们的并发博弈情形，因为在我们的博弈中，联盟会受到外部性的影响：某个联盟能否进行有利偏离，不仅取决于该联盟的组成，也取决于其余玩家的行为。

我们先介绍并发博弈框架，然后定义适用于这类情形的核心的两种变体。第一种变体假定，如果某个玩家联盟有可采取的行动方案，且无论其余玩家采取何种行动，该方案都能使联盟受益，那么该联盟就存在有利偏离（参见经典博弈论文献中的 α-核心概念 [16]）。然而，这种“最坏情况”（极大极小）核心表述要求偏离针对其余玩家所有 行动方案都能带来收益——即便其余智能体不会理性地选择这些方案。由此引出了第二种定义：偏离只需针对其余玩家所有可信的行动方案都能带来收益；这里的“可信”是指，采取这些方案后，那些玩家的处境不会比原先更差。我们还考虑智能体具有定量偏好的博弈，并将其建模为博弈每个回合中的某种奖励。对于每种情形，我们都对相关解概念作出形式化定义，确定其若干关键计算性质，给出逻辑刻画，并在可能的情况下提供复杂性结果；这些结果涵盖从可在 PSpace 中判定的性质到可在 3-指数时间内判定的性质。我们还研究与核心相关的模型论性质：特别是核心是否保证非空，以及时序逻辑性质是否在由核心元素所诱导的对局（计算运行）上的互模拟系统中成立。

博弈（从形式化验证的角度来看，这是一种非常理想的性质）。

论文结构。 本文其余部分安排如下：

- 下一节将概述本文中会用到的逻辑与并发博弈的关键概念。

- 第 3 节定义核心及其相关的主要计算性质。第 3.4 节将探讨可信联盟形成问题及相关复杂性结果。

- 第 4 节将在并发平均收益博弈的定量情形下研究核心。

- 第 5 节将给出总结性评述与相关工作，并讨论如何使用模型检测器实现我们的概念。

2. 预备知识

2.1. 集合与序列

给定任意集合 S，我们分别用 S^*、S^ω 和 S^+ 表示由 S 中元素构成的有限序列、无限序列和非空有限序列的集合。如果 $w_1 \in S^*$ 和 w_2 是任意其他序列（有限或无限），我们用 $w_1 w_2$ 表示它们的拼接。对于 $Q \subseteq S$，我们用 S_{-Q} 表示 $S \setminus Q$，并且当 $Q = \{i\}$ 时，用 S_{-i} 表示。我们将此表示法扩展到元组 $u = (s_1, \ldots, s_k, \ldots, s_n) \in S_1 \times \cdots \times S_n$，并用 u_{-k} 表示 $(s_1, \ldots, s_{k-1}, s_{k+1}, \ldots, s_n)$；对于元素集合，也采用类似表示法，即 u_{-Q} 表示不含每个 s_k 的 u，其中 $k \in Q$。给定一个序列 w，我们用 $w[t]$ 表示该序列中位置 $t+1$ 上的元素；例如，$w[0]$ 是 w 的第一个元素。我们还使用切片表示法：用 $w[l \ldots m]$ 表示序列 $w[l] \ldots w[m-1]$，用 $w[l \ldots]$ 表示 $w[l]w[l+1] \ldots$，并用 $w[\ldots m]$ 表示 $w[0] \ldots w[m-1]$；如果 $m = 0$，则令 $w[l \ldots m]$ 为空序列，记作 ϵ。

2.2. 博弈

我们首先介绍本文余下部分将始终采用的多智能体系统模型：并发博弈结构 [6]。非正式地说，并发博弈由一组玩家、每位玩家各自的一组动作、一组系统状态，以及一个转移函数构成；该转移函数描述给定当前状态和每位玩家各自的一个动作时，博弈状态如何变化。

形式上，并发博弈结构 M 由以下元组给出：

$$M = (\mathrm{Ag}, \mathrm{St}, \{\mathrm{Ac}_i\}_{i\in\mathrm{Ag}}, s^0, \mathrm{tr}),$$

其中：

- Ag 和 St 分别是有限的非空**智能体**集合和**状态**集合——我们通常将 Ag 视为集合 $\{1, \ldots, n\}$；
- 对于每个 $i \in \mathrm{Ag}$，Ac_i 是主体 i 可用的有限非空**行动**集合。我们将每个状态 s 与该状态下可用动作的集合 $\mathrm{Ac}_i(s) \subseteq \mathrm{Ac}_i$ 相关联，并将 Ac 记作 $\mathrm{Ac}_1 \times \cdots \times \mathrm{Ac}_n$；
- $s^0 \in \mathrm{St}$ 是**初始/起始状态**；最后，
- $\mathrm{tr} : \mathrm{St} \times \mathrm{Ac} \to \mathrm{St}$ 是该博弈的**转移函数**。

由于博弈的所有其他组成部分都可以从转移函数推导出来（例如，可以先列出初始状态的转移来指定初始状态），因此 M 的大小就是其转移函数的大小，由 $|\mathrm{St}| \times |\mathrm{Ac}|^{|\mathrm{Ag}|}$给出。

给定一个并发博弈结构 M，我们可以在其中进行如下博弈：博弈从状态 s^0 开始，每个玩家 $i \in \mathrm{Ag}$ 选择一个对其可用的动作 $\mathrm{ac}_i^0 \in \mathrm{Ac}_i(s^0)$。随后，博弈转移到一个新状态，

$$s^1 = \mathrm{tr}(s^0, \mathrm{ac}_1^0, \ldots, \mathrm{ac}_n^0).$$

然后重复这一过程。我们通常用 s^i 表示序列中的 i^{th} 状态，并用 $\text{ac}^i = (\text{ac}_1^i、\ldots\ldots、\text{ac}_n^i)$ 表示序列中所执行的 i^{th} 动作向量。因此，对所有 $t \in \mathbb{N}$，都有

$$s^{t+1} = \text{tr}(s^t, \text{ac}^t)$$

一个**运行**，ρ 是一个无限序列 $\rho = s^0 s^1 s^2 \ldots\ldots$，满足对于每个 $t \in \mathbb{N}$，都存在某个 $\text{ac} \in \text{Ac}$，使得 $s^{t+1} = \text{tr}(s^t, \text{ac})$。**路径**，$\pi$ 是某个运行的有限前缀。

玩家的一种**策略** i 规定了玩家 i 如何在博弈的每个回合中选择动作，具体取决于博弈此前的历史。形式上，策略是一个函数 $\sigma_i : \text{St}^+ \to \text{Ac}_i$，满足对于每个 $\pi \in \text{St}^*$ 和 $s \in \text{St}$，$\sigma_i(\pi s) \in \text{Ac}_i(s)$。因此，对于每条路径 π，玩家的一种策略 i 都会为 i 指定一个从该路径最后一个状态出发可采取的动作。玩家 i 的策略集合记为 Σ_i.[1]

一个**策略组合** $\circledR\sigma$ 是由策略构成的元组，每位玩家各有一个：$\circledR\sigma = (\sigma_1, \ldots\ldots, \sigma_n) \in \Sigma_1 \times \cdots \times \Sigma_n$。注意，策略组合$\circledR\sigma$与一个状态 s共同决定唯一的运行 $\rho(\circledR\sigma, s)$，其中：

- $\rho(\circledR\sigma, s)[0] = s$；以及
- $\rho(\circledR\sigma, s)[t+1] = \text{tr}(\rho(\circledR\sigma, s)[t], \sigma_1(\rho(\circledR\sigma, s)[\ldots t]), \ldots, \sigma_n(\rho(\circledR\sigma, s)[\ldots t]))$, 对所有 $t \in \mathbb{N}$。

我们写作 $\rho(\circledR\sigma)$，如果 $s = s^0$。

[1]这里，我们针对状态的有限序列定义策略。也可以根据动作组合的有限序列来定义策略。由于可以从动作组合推导出状态序列，反之则不然，因此这种策略比我们使用的策略更强大，确实具有一些理想的性质，例如纳什均衡在双模拟等价下的不变性 [17]。然而，基于状态定义的策略是并发博弈结构 [6] 和理性验证 [3] 文献中的标准做法，因此为保持一致，我们采用这一模型。更多详情和进一步讨论，请参阅 [17]。

注意，将策略视为函数 $\sigma_i\colon \mathrm{St}^+ \to \mathrm{Ac}_i$ 在计算分析方面存在问题，因为此类函数的定义域是无限的。为了能够回答与计算复杂性（例如）相关的问题，我们需要用有限表示来描述必须在无限多个回合中运行的策略。为此，我们遵循并发博弈文献中的标准做法，将策略建模为**有限状态输出机**（转换器） [2, 3]。值得注意的是，对于具有 LTL 目标的玩家来说，这类策略已经足够：无需使用更强大的策略模型 [2, 3]。形式上，玩家 i 的策略是一个结构

$$\sigma_i = (Q_i, q_i^0, \delta_i, \tau_i)$$

其中：

- Q_i 是一个有限的非空**策略状态**集合；
- $q_i^0 \in Q_i$ 是**初始**策略状态；
- $\delta_i\colon \quad Q_i \times \mathrm{St} \to Q_i$ 是一个 **转移函数**；且
- $\tau_i\colon \quad Q_i \to \mathrm{Ac}_i$ 是一个 **输出函数**。

机器策略的运作方式如下：它从初始状态 q_i^0 开始，并据此选择一个动作， $\tau(q_i^0)$。随后，博弈状态按照转移函数转移到一个新状态。根据这个新的（博弈）状态以及策略的状态，策略也会按照策略转移函数转移到一个新状态。这个过程不断重复，每个时间步都会产生一个新动作。很容易看出，仅由机器策略构成的策略组合将是最终周期性的（也就是说，它最终会进入一个先前出现过的配置，并从那时起重复其行为）。

有时，我们会发现可以采用一种更简单的策略模型：**无记忆策略**。无记忆策略 $\sigma_i\colon \mathrm{St} \to \mathrm{Ac}_i$ 仅根据当前环境状态选择一个动作。尽管无记忆策略的表达能力远不及有限记忆策略，但它们仍然非常重要，因为其概念简单、易于实现，而且在许多类型的博弈中（例如双人平均收益博弈 [18], 和奇偶博弈 [19]），无记忆策略足以实现最优表现。[2]

2.3. 逻辑

总体而言，我们需要借助某些逻辑，主要出于两个原因：表达智能体的偏好，以及对这些智能体及其偏好进行推理。为此，本文将始终使用三种逻辑：线性时序逻辑 （LTL） [22], 交替时序逻辑 （ATL*） [6], 和策略逻辑 （SL） [23]。我们使用 LTL 对智能体偏好进行建模，并使用 ATL*和 SL 对我们将要研究的博弈论概念进行逻辑刻画。虽然从下文给出的语法和语义中可以更形式化地看出这一点，但我们在此指出，LTL 可以被视为 ATL* 的子集，而 ATL* 又可以被视为 SL 的子集。

2.3.1. 线性时序逻辑

LTL 是一种广泛用于对并发系统的行为进行推理的逻辑。尽管我们将在此介绍关键概念，但有关详细内容，请读者参阅任何一本标准的时序逻辑教材（例如， [24]）。

设 AP 为一个命题变量集合。那么，LTL公式的语法

[2]关于哪些类型的博弈中，无记忆策略足以让一方或双方玩家实现最优玩法的刻画，请参阅 [20, 21]。

φ 由以下语法给出：

$$\varphi := p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathrm{X}\varphi \mid \varphi \mathrm{U} \varphi,$$

其中 $p \in \mathrm{AP}$。设命题变量集合 AP 上的所有 LTL公式记为 L(AP)。如果上下文中 AP 的含义明确，我们也可以直接记为 L。我们还引入传统命题缩写 $\cdot \wedge \cdot$、$\cdot \rightarrow \cdot$、$\cdot \leftrightarrow \cdot$，其定义采用通常的方式；此外，还引入缩写 $\mathrm{F}\varphi$ 表示 $\top \mathrm{U} \varphi$，以及 $\mathrm{G}\varphi$ 表示$\neg \mathrm{F} \neg \varphi$。

通常，在模型检测的语境下，LTL公式的语义是相对于标记迁移系统 [25], 或克里普克结构 [26],定义的；但就我们的目的而言，我们根据并发博弈结构所生成的无限运行来定义它们。形式化地，设 M 是一个并发博弈结构，并设 $\lambda : \mathrm{St} \rightarrow \mathrm{P}(\mathrm{AP})$ 是一个标记函数，将状态映射到命题变量集合。于是，给定一个无限运行 $\rho \in \mathrm{St}^{\omega}$ 和一个 LTL公式 φ，如果 (M, λ, ρ) 满足 φ，我们就称其成立，并根据以下归纳定义记作 $(M, \lambda, \rho) \models \varphi$：

- 对于 $p \in$原子命题集，我们有 $(M, \lambda, \rho) \models p$当且仅当 $p \in \lambda(\rho[0])$;
- 对于 $\varphi \in$L (原子命题集)，我们有 $(M, \lambda, \rho) \models \neg\varphi$当且仅当 $(M, \lambda, \rho) \models \varphi$不成立；
- 对于 $\varphi, \psi \in$L (原子命题集)，我们有 $(M, \lambda, \rho) \models \varphi \wedge \psi$当且仅当 $(M, \lambda, \rho) \models \varphi$和 $(M, \lambda, \rho) \models \psi$同时成立；
- 对于 $\varphi \in$L (原子命题集)，我们有 $(M, \lambda, \rho) \models \mathrm{X}\varphi$当且仅当 $(M, \lambda, \rho[1 \ldots]) \models \varphi$;
- 对于 φ，$\psi \in$L (原子命题集)，有 $(M, \lambda, \rho) \models \varphi \mathrm{U} \psi$，当且仅当存在某个 $j \geq 0$，使得 $(M, \lambda, \rho[j \ldots]) \models \psi$，并且对于所有 $i < j$，都有 $(M, \lambda, \rho[i \ldots]) \models \varphi$。

为方便记号，我们将 $(G, \rho) \models \varphi$ 作为 $(M, \lambda, \rho) \models \varphi$ 的简写；如果 G 可由上下文推知，我们就只写 $\rho \models \varphi$。

2.3.2. 交替时序逻辑

ATL*是分支时序逻辑 CTL* [27]的扩展——我们将在证明中使用该逻辑，以更有效地推理智能体联盟。ATL*中的关键算子是合作模态 $\langle\langle C\rangle\rangle \phi$，它断言联盟 C 有能力确保时序性质 ϕ成立；更具体地说，存在一组适用于 C 的策略，使得如果 C 遵循这些策略，那么无论其他智能体如何行动，都可以保证 ϕ 成立。给定一个原子命题集 AP 和一个智能体集合 Ag，ATL* 公式的语言由以下语法定义：

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \langle\langle C\rangle\rangle \psi$$

$$\psi ::= \varphi \mid \neg\psi \mid \psi \vee \psi \mid \mathrm{X}\psi \mid \psi \mathrm{U} \psi$$

其中 $p \in \mathrm{AP}$ 且 $C \subseteq \mathrm{Ag}$。我们将上述语法中由 φ 生成的公式称为**ATL*状态公式**，并用 $\mathrm{L}_s(\mathrm{AP}, \mathrm{Ag})$ 表示包含这些公式的集合；将由上述语法中 ψ 生成的公式称为**ATL*路径公式**，记作 $\mathrm{L}_P(\mathrm{AP}, \mathrm{Ag})$。（之所以这样命名，是因为它们的语义分别相对于状态和路径定义。）需要强调的是，只有 ATL* 状态公式才是良构的 ATL*公式，因此，我们用 $\mathrm{L}(\mathrm{AP}, \mathrm{Ag})$ 作为 $\mathrm{L}_s(\mathrm{AP}, \mathrm{Ag})$ 的简写。当 AP 或 Ag 已知时，或者二者都已知时，我们可以省略它们。对于 $\mathrm{AP}' \subseteq \mathrm{AP}$，若 $\varphi \in \mathrm{L}(\mathrm{AP}', \mathrm{Ag})$ 对某个智能体集合 Ag 成立，则可以写作 $\varphi|_{\mathrm{AP}'}$。

除了上述用于 LTL公式的缩写外，我们还使用简写形式：$\mathrm{E}\ \varphi$ for $\langle\langle \mathrm{Ag}\rangle\rangle\ \varphi$、$\mathrm{A}\ \varphi$ for $\langle\langle \emptyset\rangle\rangle\ \varphi$以及 $[\![C]\!]\ \varphi$ for $\neg \langle\langle C\rangle\rangle \neg\varphi$。最后，我们将 ATL*公式 φ 的大小 定义为其子公式的数量。

要定义 ATL*公式的语义，实际上我们需要定义两种

语义关系，$\models_s$（用于状态公式）和 $\models_p$（用于路径公式）。因此，设 M 为某个并发博弈结构，并配有一个标记函数 $\lambda : \mathrm{St} \to \mathrm{P}(\mathrm{AP})$。那么，给定一个状态 $s \in St$ 和一个 ATL*公式 φ，我们称 (M, λ, s) 满足 φ，并根据以下归纳定义记作 $(M, \lambda, s) \models \varphi$：

- 对于 $\varphi \in \mathrm{L}_s(\mathrm{AP}, \mathrm{Ag})$，有 $(M, \lambda, s) \models \varphi$ 当且仅当 $(M, \lambda, s) \models_s \varphi$；
- 对于属于线性时序逻辑（LTL）的运算符，其归纳语义与 LTL 中的相同；
- 对于 $\varphi \in \mathrm{L}_p(\mathrm{AP}, \mathrm{Ag})$，有 $(M, \lambda, s) \models_s \langle\langle C \rangle\rangle \varphi$ 当且仅当存在一个联盟 C 的策略向量 $\mathrm{R}\sigma_C$，使得对于所有互补的策略组合 $\mathrm{R}\sigma_{\mathrm{Ag}\setminus C}$，均有 $(M, \lambda, \rho((\mathrm{R}\sigma_{\mathrm{Ag}\setminus C}, \mathrm{R}\sigma_C), s)) \models_p \varphi$ 成立；
- 对于 $\varphi \in \mathrm{L}_s(\mathrm{AP}, \mathrm{Ag})$，有 $(M, \lambda, \rho) \models_p \varphi$ 当且仅当 $(M, \lambda, \rho[0]) \models_s \varphi$；

如果上下文已明确并发博弈和标记函数，我们将简写为 $s \models \phi$。给定并发博弈结构 M 和标记函数 $\lambda : \mathrm{St} \to \mathrm{P}(\mathrm{AP})$，如果存在某个状态 $s \in \mathrm{St}$，使得 $(M, \lambda, s) \models \varphi$，我们就称 φ 为**可满足**。此外，如果对于所有状态 $s \in \mathrm{St}$，都有 $(M, \lambda, s) \models \varphi$ 当且仅当 $(M, \lambda, s) \models \psi$，我们就称 φ 与 ψ **等价**。

注意，LTL 可以看作 ATL*的子逻辑，它由所有满足以下条件的公式 $\mathrm{A}\varphi$ 构成：公式 φ 不包含“联盟”量词 $\langle\langle C \rangle\rangle$ 或 $[\![C]\!]$。因此，ATL*是一个特别有效的工具，可用于推理联盟能够实现的线性时序逻辑性质，而这正是我们在证明判定问题的复杂度界时使用它的方式。具体而言，若有一个线性时序逻辑博弈 G，我们可以将 $(G, s) \models \varphi$ 作为 $(M, \lambda, s) \models \varphi$ 的简写，并且

此外，如果根据上下文可以明确博弈 G 指的是什么，我们就省略它，直接写作 $s \models \varphi$。

这里相关的判定问题是**ATL 的模型检测问题***，我们将在下一节大量使用它：

ATL*模型检测：

给定：并发博弈 M、标记函数 λ、状态 $s \in \mathrm{St}$，以及

ATL*公式 φ。

问题：$(M, \lambda, s) \models \varphi$ 是否成立？

对于两个或更多玩家参与的博弈，该问题是 2-指数时间完全 [6]；对于单人博弈，该问题是 PSPACE完全（因为此时该问题可归约为线性时序逻辑模型检测） [28]。

2.3.3. 策略逻辑

尽管 ATL* 是用于推理联盟的强大工具，但众所周知，它无法表示某些博弈论概念，例如纳什均衡 [23]——要实现这一点，需要对策略进行显式量化。就我们的目的而言，我们将能够使用 ATL*完成大部分必要的逻辑刻画；但当我们在第 3.4 节讨论< शैली id='4'>强核心 时，同样需要具备对策略进行显式推理的能力。SL 通过两个< शैली id='6'>策略量词扩展了 LTL，即 hh$x$ ii 和 $[[x]]$，以及一个< शैली id='10'>主体绑定 算子（i, x），其中 i 是一个主体，x 是一个变量。这些算子分别可读作“存在一个策略 x”、“对于每个策略 x”以及“将主体 i 绑定到与变量 x相关联的策略”。形式上，SL公式由命题集合 AP、变量集合 Var 和主体集合 Ag 归纳构成，所使用的语法如下，其中 $p \in \mathrm{AP}$、$x \in \mathrm{Var}$ 和 $i \in \mathrm{Ag}$：

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \mathrm{X}\varphi \mid \varphi \mathrm{U} \varphi \mid \langle\langle x \rangle\rangle \varphi \mid [[x]]\varphi \mid (i, x)\varphi.$$

现在我们可以给出 SL 的语义。首先，用 Str 表示所有策略的集合，我们将赋值定义为一个偏函数，它把变量和主体映射到策略，$\chi \in \mathrm{Asg} = (\mathrm{Var} \cup \mathrm{Ag}) \rightharpoonup \mathrm{Str}$。随后，我们使用记号$\chi[i \mapsto f]/\chi[x \mapsto f]$表示这样一个赋值：它在 i/x上的取值与 f相同，并且在其他所有有定义的输入上都与 χ一致。然后，给定一个并发博弈结构 M、任意 SL 公式 φ、 M中的状态 $s \in \mathrm{St}$，以及赋值 $\chi \in \mathrm{Asg}$，关系 $M, \chi, s \models \varphi$ 定义如下：

1. 对于嵌入 SL 的 LTL公式，其归纳语义与 LTL 中的相同；

2. 对于所有公式 φ 和变量 $x \in \mathrm{Var}$，我们有：

   (a) $M, \chi, s \models \mathrm{hh}x\mathrm{ii}\varphi$ 当且仅当存在某个 $f \in \mathrm{Str}$ 使得$M, \chi[x \mapsto f]$, $s \models \varphi$；(b) $M, \chi, s \models [[x]]\varphi$ 当且仅当对于所有 $f \in \mathrm{Str}$，我们有 $M, \chi[x \mapsto f], s \models \varphi$；

3. 对于每个主体 $i \in \mathrm{Ag}$ 和变量 $x \in \mathrm{Var}$，我们有 $M, \chi, s \models (i, x)\varphi$ 当且仅当 $M, \chi[i \mapsto \chi(x)], s \models \varphi$。

对于句子 φ，也就是说，对于没有自由变量且主体为 [23], 的公式，我们称 M 满足 φ，并在此情况下记作 $M \models \varphi$，当且仅当 M、$\emptyset$、$s^0 \models \varphi$，其中 $\emptyset$ 是空赋值。我们使用以下缩写：用 $\mathrm{h}i\mathrm{i}\varphi$ 表示 $\mathrm{hh}x\mathrm{ii}(i, x)\varphi$，并用 $[i]\varphi$ 表示 $[[x]](i, x)\varphi$；直观而言，它们分别表示“存在主体 i 的一个策略，使得 φ 成立”和“对于主体 i 的所有策略，φ 都成立”。我们将此记法扩展到玩家集合，例如，用 $\mathrm{h}C\mathrm{i}\varphi$ 代替 $\mathrm{h}i\mathrm{i} \ldots \mathrm{h}\, j\mathrm{i}\varphi$，其中 $C = \{i, \ldots, j$，全称量词算子也类似。于是，$\mathrm{h}C\mathrm{i}\varphi$ 表示“联盟 C 有一个联合策略，使得 φ 成立”。

2.4. LTL 博弈

现在我们可以定义 *LTL* 博弈 [2, 29, 30]。LTL 博弈的核心思想是，每位玩家 i 都与一个 LTL 目标公式 γ_i 相关联，并希望该公式得到满足。形式化地说，一个 LTL 博弈 G 由一个结构给出

$$G = (M, \mathrm{AP}, \lambda, (\gamma_i)_{i \in \mathrm{Ag}}),$$

其中，M 是一个并发博弈结构，AP 是一个原子命题集，$\lambda : \mathrm{St} \to \mathrm{P}(\mathrm{AP})$ 是一个标记函数；对于每个 $i \in \mathrm{Ag}$，γ_i 都是一个关于 AP 的 LTL公式，用于定义该玩家对运行的偏好关系。

现在我们来说明时序目标公式 γ_i 如何在运行上诱导出**偏好关系**$\succeq_i$。首先，由于在给定的运行中，LTL 目标要么得到满足，要么未得到满足，因此我们可以确定该运行的一组“获胜者”和一组“失败者”。形式化地，令 $\mathrm{W}(\rho)$ 表示在 ρ 中实现了其目标的玩家集合，令 $\mathrm{L}(\rho)$ 表示未实现其目标的玩家集合：

$$\mathcal{W}(\rho) = \{i \in \mathrm{Ag} \mid \rho \models \gamma_i\}$$

$$\mathcal{L}(\rho) = \mathrm{Ag} \setminus \mathcal{W}(\rho).$$

现在，我们可以使用赢家与输家来定义智能体的偏好关系。直观地说，智能体 i 总是严格偏好满足其目标 γ_i的运行，而不是不满足该目标的运行；但对于两个都满足其目标的运行，它会无差异；对于两个都不满足其目标的运行，它也会无差异。形式化地，对于两个运行 ρ，ρ'以及一个玩家 i，当且仅当 i 在 ρ 下是赢家（注意，这里不依赖于 ρ'），或者 i 在 ρ 和 ρ'下都是输家时，我们有 $\rho \succeq_i \rho'$。换一种说法，当且仅当（M，λ，ρ'）$\models \gamma_i$ 蕴含（M，λ，ρ）$\models \gamma_i$时，我们有$\rho \succeq_i \rho'$。严格偏好关系 $\succ_i$ 按标准方式定义：当且仅当 $\rho \succeq_i \rho'$ 但不$\rho' \succeq_i \rho$时，有 $\rho \succ_i \rho'$。我们把验证工作留给读者：首先，这一定义与我们上述对偏好的非形式化解释相符；其次，以此定义的关系 $\succeq_i$ 确实是

偏好关系（即，二元关系 $\succeq_i$ 具有自反性、完备性和传递性）。最后需要指出的是，本文在两种不同意义下使用 L：一是表示各种逻辑的语言，二是表示某个运行的输家集合。由于这两种用法通常出现在不同语境中，且所接受的输入也不同，因此对其应作何种解释不会产生歧义。

定义偏好关系后，我们可以介绍博弈论解概念。目前，我们先讨论非合作解概念。首先，若不存在玩家 $i \in \mathrm{Ag}$ 和该玩家的策略 σ_i'，使得 $(\circledR \sigma_{-i}, \sigma_i') \succ_i \circledR\sigma$，则称策略组合 $\circledR\sigma$ 为**纳什均衡**。也就是说，如果没有玩家能够通过单方面改变自己的策略而获益（假设其他所有玩家都保持策略不变），那么 $\circledR\sigma$ 就是纳什均衡。用 $\mathrm{NE}(G)$ 表示博弈 G的纳什均衡集合。

我们强调，纳什均衡只考虑单方面偏离，即例如，由单个玩家进行的偏离。将此与**强纳什均衡** [31, 32]这一概念进行比较：如果不存在联盟 C 及策略 σ_C'，使得对所有 $i \in C$ 都有 $(\circledR\sigma_{-i}, \sigma_i') \succ_i \circledR\sigma$，则称策略组合 $\circledR\sigma$ 为强纳什均衡。因此，强纳什均衡是能够抵御多边偏离的策略组合。

最后，我们顺带提一下联盟证明的纳什均衡 [33, 34]这一概念。虽然这里不作正式定义，但非正式地说，可以将其理解为刻画玩家拥有无限但不具约束力的赛前沟通时所产生的结果。

3. 合作理性验证

3.1. 定义核心

我们希望为合作情形定义与理性验证问题 E-纳什 和 A-纳什 相对应的概念，这两个问题在 [2, 3], 中已有研究。为此，我们需要为并发博弈情形定义一个核心。核心大概是合作博弈论中最著名的解概念。与非合作场景中的纳什均衡一样，核心为博弈定义了一种稳定性概念；但纳什均衡只要求结果稳定，即不存在个体有利偏离，而核心要求结果不存在联盟有利偏离。在“标准”的合作博弈模型中（特征函数形式的可转移效用合作博弈 [15]），这种直觉很容易形式化；但在并发博弈中，存在一个重要难点，如下所述。

假设玩家联盟 $C \subseteq \mathrm{Ag}$ 正在考虑参与策略组合 $\mathrm{®}\sigma$，并且尤其试图确定，他们是否能够从 $\mathrm{®}\sigma$ 中进行合作有利偏离。现在，当他们考虑可能的有利偏离——集体策略 $\mathrm{®}\sigma_C \in \hat{\mathrm{I}}_{i\in C}\,\Sigma_i$ ——应该对 C其余玩家的行为作何假设 $\mathrm{Ag}\setminus C$? 尤其是，假设其余玩家不会作出回应（例如可能改变自己的策略），在合作场景中是不可信的。[3] 能够合作的理性玩家会理性地、以合作的方式回应这一偏离，对抗参与 C的玩家。而且，至关重要的是， C所设想的偏离是否确实有利，很可能取决于其余玩家的行为。用博弈论术语来说，我们的并发博弈情形受到外部性的影响：联盟的表现

[3]要定义强纳什均衡这一非合作解概念，就必须假设存在这类行为。

C 不仅取决于联盟 C 的组成，还取决于其余玩家的行为。

众所周知，在存在外部性时，合作解概念难以定义 [15]。尤其是，对于存在外部性的博弈，核心并没有普遍认可的定义。因此，我们对并发博弈的核心所作的第一个定义，采用了最坏情况（极大极小）推理。也就是说，当联盟 C 考虑偏离时，它要求无论其余玩家如何行动，这种偏离都有利。合作博弈中的 α-核心概念已探讨过这一思想 [16]。为了将其形式化，我们需要形式化有利偏离这一概念。设 $\mathbb{R}\sigma$ 为一个策略组合， C为一个联盟；那么，如果以下两项都成立，我们就称 $\mathbb{R}\sigma'_C$ 是从 $\mathbb{R}\sigma$ 出发的有利偏离：

1. $C \subseteq \mathrm{L}(\rho(\mathbb{R}\sigma))$; 且 2. 对所有 $\mathbb{R}\,\sigma'_{-C}$，
都有 $C \subseteq \mathrm{W}(\rho(\mathbb{R}\sigma'_C, \mathbb{R}\sigma'_{-C}))$.

换言之，如果联盟中的玩家无论联盟外的玩家采取什么策略，都认为采用 $\mathbb{R}\sigma'_C$ 比采用自己在 $\mathbb{R}\sigma$ 中各自的策略更好，那么就称 $\mathbb{R}\sigma'_C$ 是相对于 $\mathbb{R}\sigma$ 的有利偏离。博弈 G 的核心（记作 core(G)）定义为 G 中不存在任何联盟可对其进行有利偏离的结果集合。

例1. 考虑以下博弈，它说明了合作解概念与非合作解概念之间的差异：该博弈包含一个质量较差且不属于核心的纳什均衡。合作的能力使智能体能够避免这一不理想的均衡。该博弈包含两名玩家，玩家集 $\mathrm{Ag} = \{1,2\}$，以及两个变量 $\mathrm{AP} = \{p, q\}$。玩家1的行动集为 $\mathrm{Ac}_1 = \{pt, pf\}$，玩家2的行动集为 $\mathrm{Ac}_2 = \{qt, qf\}$，并满足：对于每个可达状态，如果玩家1/2执行 pt/qt，那么 p/q 就会成立；如果改为执行 pf/qf，则该条件不会成立（也就是说，玩家1“控制” p 的取值，玩家2则“控制” q的取值）。他们的

目标相同（因此该博弈是一个协调博弈）：$\gamma_1 = \gamma_2 = \mathsf{G}(p \wedge q)$。现在，考虑策略组合 $\circledR\sigma$，其中两名玩家都只是永远将各自的变量固定为假（也就是说，永远执行 pf 和 qf）。在这样的策略组合下，两名玩家的目标都无法实现。然而，该策略组合构成一个纳什均衡，因为单方面偏离无法改善局面：任何一名玩家都没有能让自己获益更多的替代策略。事实上，该博弈中存在无穷多个这样的劣质纳什均衡，其中两名玩家都无法实现自己的目标。不过，该策略组合并不属于核心，因为存在一种合作有利偏离，可转向这样的策略组合：两名玩家都永远将各自的变量固定为真（也就是说，永远执行 pt 和 qt）。事实上，核心中的每个策略组合都能让两名玩家实现各自的目标。因此，以核心取代纳什均衡，可以消除博弈中的劣质均衡，从而带来更符合社会利益的结果。

在继续之前，我们注意到，博弈论关于具有外部性的合作博弈的文献中还提出了核心的其他变体，特别是 β-核心和 γ-核心。β-核心假设，偏离联盟之外的玩家将选择策略，以最大限度地惩罚偏离联盟；随后，偏离联盟的成员将对这些惩罚策略集体作出最佳响应。因此，β-核心假设偏离联盟采取极小极大 选择。相比之下，γ-核心假设，偏离联盟之外的玩家将组成单人联盟，其中每个未偏离者针对偏离联盟的策略选择个人最佳响应策略。我们注意到，与 α-核心相比，对这些解概念的研究要少得多，而且可以说它们不那么自然，动机也不那么充分。因此，我们将其留待未来研究（另见相关工作第5节中的评论）。

3.2. 判定问题

在理性验证 [2–4] 中，我们主要关注检查给定博弈的某个解概念满足哪些时序逻辑性质；通常，在非合作场景中，我们研究博弈的纳什均衡 NE(*G*) 中成立哪些LTL公式 *G*。在此处介绍的合作场景中，我们关注的则是博弈核心中成立哪些性质。理性验证中的两个主要判定问题，是检查某个时序逻辑公式是否由博弈的某个/每个稳定策略组合满足。对于核心，这些问题定义如下——参见 [2–4]。

E-Core：

给定：博弈 G，线性时序逻辑公式 φ。

问题：是否存在 $\exists ®\sigma \in$ 核心(G)：$\rho(®\sigma) \models \varphi$?

A-Core：

给定：博弈 G，线性时序逻辑公式 φ。

问题：是否有 $\forall ®\sigma \in$ 博弈(G)：$\rho(®\sigma) \models \varphi$?

非合作场景中一个受到研究的判定问题是非空问题，它询问给定博弈的纳什均衡集合是否非空。然而，正如我们马上将要证明的，博弈的核心总是非空的，因此在合作场景中考虑相应的问题并无意义。

我们还将关注另外两个判定问题：判断给定的策略组合是否属于核心（核心成员资格），以及判断联盟的给定策略向量相对于某个策略组合是否构成有利偏离（有利偏离）。形式化地说：

核心成员资格：

给定：博弈 G，策略组合 $\text{®}\sigma$。

问题：是否有 $\text{®}\sigma \in \text{core}(G)$?

有利偏离：

给定：博弈 G、策略组合 $\text{®}\sigma$、联盟 C 和偏离 $\text{®}\sigma'_C$。

问题：$\text{®}\sigma'_C$ 是否是相对于 $\text{®}\sigma$ 的有利偏离?

以下讨论中，使用**已实现联盟**这一概念很有帮助：如果一个联盟能够实现其目标，且无论其他玩家采取何种行动，就称该联盟已实现；也就是说，他们拥有一种集体策略，能够保证实现其目标。形式上，如果存在一个针对 $C \subseteq \text{Ag}$ 的联合策略 $\text{®}\sigma_C$，使得对于 $\text{Ag} \setminus C$ 的所有联合策略 $\text{®}\sigma_{-C}$，我们都有

$$\rho((\vec{\sigma}_C, \vec{\sigma}_{-C})) \vDash \bigwedge_{i \in C} \gamma_i.$$

因此，已实现联盟拥有获胜策略，能够共同实现其目标。由于我们讨论的是合作博弈，问题在于这样的联盟是否会形成。我们有如下结论：

引理 *1. 1.* 存在博弈 G，其策略组合 $\text{®}\sigma \in core(G)$ 包含已实现联盟 $C \subseteq Ag$，使得 $C \notin W(\rho(\text{®}\sigma))$; *2.* 对于每个博弈 G、策略组合 $\text{®}\sigma \in core(G)$ 和已实现联盟 C，都有 $C \cap W(\rho(\text{®}\sigma)) \neq \emptyset$; *3.* 对于每个博弈 G 和已实现联盟 C，存在 $\text{®}\sigma \in core(G)$，使得 $C \subseteq W(\rho(\text{®}\sigma))$。

通俗地说，该引理的第一部分指出，某个联盟已实现，并不意味着该联盟中的每个玩家在核心的任意成员下都必然能实现其目标。然而，该引理的第二部分

指出，在核心的任何成员中，每个已实现联盟中的某些智能体都必须实现其目标。此外，该引理的第三部分指出，对于每个已实现联盟，核心中都存在一个策略组合，使该联盟中的每个玩家都能实现其目标。

1.1 的证明。考虑一个包含三个状态——起始状态、向上状态和向下状态——以及三名玩家（{1, 2, 3}）的博弈。令起始状态为该博弈的初始状态，并假设玩家2和玩家3各自只有一个可选动作，而玩家1有两个动作 H 和 T。因此，转移函数仅需根据玩家1的动作来定义；令 tr(start, H) = 向上状态，tr(start, T) = 向下状态。其余两个状态为汇状态，其中 tr(向上状态, H) = tr(向上状态, T) = 向上状态，tr(向下状态, H) = tr(向下状态, T) = 向下状态。进一步假设 AP = $\{p, q\}$，且 λ(起始状态) = $\emptyset$, λ(向上状态) = $\{p\}$ 和 λ(向下状态) = $\{q\}$。最后，设玩家的目标如下：$\gamma_1 = \top$,$\gamma_2 = \mathrm{X\,G}\ p$，以及 $\gamma_3 = \mathrm{X\,G}\ q$。由于博弈在执行第一个动作后必定进入两个汇状态之一，因此只需根据玩家1的第一个动作 H 或 T 来确定策略。

现在请注意，联盟 {1，3} 是一个已实现联盟：如果玩家1选择**T**，那么博弈最终会进入向下状态，并永久停留在那里，从而满足玩家1和玩家3的目标。不过请注意，玩家1选择**H** 的策略组合属于核心：玩家1和玩家2的目标都得到了满足；而由于玩家3只有一个动作，无法通过偏离来实现自己的目标。因此，我们有一个已实现联盟和一个核心成员，但该已实现联盟并非该核心成员获胜者的子集。请注意，同样的论证也可以反过来应用——{1，2} 是一个已实现联盟，而**T** 是一个核心成员，其获胜者并不包含 {1，2}。□

1.2 的证明。设 G 为一个博弈，®σ 为一个属于核心的策略组合，并且

C 是一个已实现联盟。为导出矛盾，进一步假设 C与 W (ρ(®σ)) 的交集为空。这意味着，对于每个 $i \in C$，都有ρ(®σ) $\not\models \gamma_i$。但是，由于 C 是一个已实现联盟，存在一个策略 ®σ'_C，使得对于所有反策略 ®σ_{-C}，都有 ρ((®σ_C，®σ_{-C})) $\models$Ó $_{i\in C}\ \gamma_i$。自然地，这意味着对于所有 $i \in C$，都有 ρ((®σ_C，®σ_{-C})) $\models \gamma_i$。然而，这就意味着 ®σ'_C是对 C的有利偏离，这与 ®σ 属于核心这一事实相矛盾。因此，我们得出结论： C与 W (ρ(®σ)) 的交集非空。□

1.3 的证明。设 G 是一个博弈， C 是一个已实现联盟。设 ®σ^1 = (®σ_C, ®σ_{-C}) 是任意一个策略组合，在该组合下 C 得以实现（也就是说，无论反联盟采取什么行动， C 的每个成员都能实现自己的目标）。如果 ®σ^1 是核心中的成员，那么结论成立。否则，就存在某个联盟 $D \subseteq \mathrm{Ag} \setminus C$，它有一个有利偏离 ®$\sigma'_D$。那么，在策略 ®$\sigma^2$ =(®σ_C，®σ'_D，®$^{\sigma}{}_{-(C\cup D)}$) 下， C 中的每位玩家都能实现自己的目标。和之前一样，如果 ®σ^2 是核心中的成员，那么结论成立。否则，我们可以重复这个过程，得到一个策略序列：®σ^1，®σ^2，®σ^3 .....。由于连续策略组合中的失败者每次都严格减少，这个过程只能持续有限次。根据构造，序列的最后一个成员将实现 C 中每位玩家的目标，且不存在有利偏离，因此它是核心中的成员。由此可见，存在一个核心中的策略组合，使得 C 是该策略组合获胜者的子集。□

事实上，我们对 1.3 的证明可以加以修改，从而得到一个更强的结果，即核心总是非空。这是一项非常理想的博弈论性质，因为它确保每个博弈都存在稳定策略组合，使其在实践中能够理性地实施。这与转移效用博弈中核心的传统表述形成对比 [15]。

定理 1。 对于每个博弈 G，我们都有 核心$(G) \neq \emptyset$。

证明。 与引理 1.3 的证明相同，但不要将 $\circledR\sigma^1$ 设为某个联盟能够得到满足时的策略组合，而是将其设为任意策略组合，然后照常继续。 □

由于已满足联盟有助于我们理解博弈中的联盟形成能力，我们也将关注以下关于联盟的判定问题。

已满足联盟：

给定：博弈 G，联盟 $C \subseteq \mathrm{Ag}$。

问题： C 是 G的已满足联盟吗?

在下一节中，我们将研究这些判定问题，以及核心的一些模型论性质。

3.3. 核心的刻画

本节将研究上一节所介绍的判定问题的复杂度，并确立核心的其他一些性质。我们已经在定理 1 中看到，核心总是非空；接下来我们将证明，在核心中的某个/每个结果上满足 LTL性质具有双模拟不变性 [35]。这两项结果与非合作场景下的理性验证问题形成鲜明对比：在非合作场景中，博弈的纳什均衡集合不一定总是非空 [2], ，而且在一般情形下，双模拟不变性也不成立 [17]。

我们将要考察的第一个判定问题是已实现联盟；我们通过使用 ATL^* 进行逻辑刻画，在一般情形下解决了该问题。

定理 2。 对于单人博弈，已实现联盟问题是 *PSPACE* 完全的；对于玩家人数多于一人的博弈，该问题是 *2*-指数时间完全的。

证明。 关于归属性，我们注意到，对于给定的博弈 $G = (M, \gamma_1, \ldots\ldots, \gamma_n)$ 和联盟 $C \subseteq \mathrm{Ag}$， C 已实现，当且仅当 $s^0 \models \mathrm{hh}C\mathrm{ii}\,\acute{\mathrm{O}}_{i \in C}\, \gamma_i$ 成立。根据 ATL* 模型检测的结果，可立即得出这两个上界。对于下界，我们可以将问题归约为：检查一个具有 LTL 目标的双人博弈中是否存在获胜策略（定义见 [36]，用于证明 2ExpTime 困难性），以及存在性 LTL 模型检验问题（用于证明 PSpace 困难性） [28]。请注意，因此，即使对于双人零和博弈（即具有 $\gamma_1 = \neg\gamma_2$ 的博弈），玩家人数多于一人的已实现联盟问题也是 2ExpTime 困难的。□

已实现联盟可以说明哪些稳定联盟可能形成，但不足以刻画核心，因此也不足以检验多智能体系统的 E-Core 和 A-Core 性质。为此，我们采用另一种策略，并证明这两个判定问题一般也是 2ExpTime 完全的。

定理 3. 对于单人博弈，*E-Core* 和 *A-Core* 是 *PSPACE* 完全的；对于玩家人数多于一人的博弈，它们是 *2*-指数时间完全的。

证明。 我们先考虑 E-Core。关于归属性，我们注意到，给定一个博弈 G 和一个线性时序逻辑公式 φ，当且仅当$s^0 \models \varphi_{\text{E-Core}}(G, \varphi)$成立时，$(G, \varphi) \in$ E-Core，其中 $\varphi_{\text{E-Core}}(G, \varphi)$是以下 ATL* 公式:

$$\bigvee_{W \subseteq \mathrm{Ag}} \left(\langle\!\langle \mathrm{Ag} \rangle\!\rangle \left(\varphi \wedge \bigwedge_{i \in W} \gamma_i \wedge \bigwedge_{j \in \mathrm{Ag} \setminus W} \neg\gamma_j \right) \wedge \bigwedge_{L \subseteq \mathrm{Ag} \setminus W} [\![L]\!] \bigvee_{j \in L} \neg\gamma_j \right)$$

它指出，在 G 中存在一条路径，该路径满足 φ 以及一组玩家 W （“获胜者”）的目标，并且对于任何一组未能在这条路径上实现其目标的玩家 L （“失败者”），这些玩家都无法从该路径中有利偏离。与之前一样，我们借助 ATL* 模型检测来获得上界；这里需要调用一个 2ExpTime算法

指数次——每个获胜者联盟调用一次，并在其中任意一个接受时接受。

对于下界，和已实现联盟问题一样，我们可以将如下问题归约过来：检查如 [36] 中定义的、具有 LTL 目标的双人博弈中是否存在获胜策略，以证明 2ExpTime 困难性；以及将存在性 LTL 模型检验问题归约过来，以证明 PSpace 困难性 [28]。事实上，请注意，对于单人博弈，当 $\phi = \gamma_1$ 时，以及当 $\gamma_1 = \phi, \gamma_2 = \neg\phi$ 和 $C = \{1\}$ 时，E-Core 与已实现联盟等价；这些正是上述两种归约中出现的情形。

最后，对于 A-Core，请注意，$(G, \varphi) \notin$ A-Core，当且仅当 $(G, \neg\varphi) \in$E-Core：$(G, \varphi) \notin$ A-Core 意味着并非对核心 $\mathbb{R}\sigma$ 的所有成员都有 $\rho(\mathbb{R}\sigma) \models \varphi$。将否定号移入量词后可知，存在一个核心成员使得 $\rho(\mathbb{R}\sigma) \not\models \varphi$，或者 $\rho(\mathbb{R}\sigma) \models \neg\varphi$。反向论证也可采用相同的方法。然后，由于 PSpace 和 2ExpTime 都是确定性复杂度类，我们可以得出结论：当 $|\mathrm{Ag}| = 1$ 时，A-Core 是 PSpace完全；当 $|\mathrm{Ag}| > 1$ 时，A-Core 是 2-指数时间完全，与 E-Core 的情形相同。□

现在我们研究核心成员资格和有利偏离。对于这两个问题，我们首先需要定义如何表示策略组合，因为目前策略被定义为无穷结构，它将有限历史映射为玩家行动。

有了这种策略的有限表示，我们就可以确定核心成员资格和有利偏离的复杂度。

定理 4. 对于单人博弈，核心成员资格是 *PSPACE* 完全的；对于玩家人数多于一人的博弈，则是 *2*-指数时间完全的。

证明。 对于归属性，我们首先根据 $\mathbb{R}\sigma$ =计算赢家与输家

$(\sigma_1$、……、 $\sigma_n)$，即博弈结果。这可以在 PSpace 中完成（这等价于对“乘积自动机”或“并发程序” [37]进行 LTL 模型检测）。计算出 W 后，我们可以针对每个 $L \subseteq \mathrm{Ag} \setminus W$检查其是否存在有利偏离。当且仅当 L 是一个已实现联盟时，这一条件成立。由于对于单人博弈，该问题可以在 PSpace 中检查，而对于多于一名玩家的博弈，则可以在 2ExpTime 中检查，因此两个上界立即成立。对于下界，我们再次使用引理 1 和定理 2。考虑以下博弈。设 φ 是一个可满足的 LTL公式，且 $\mathbb{R}\sigma$ 是一个不满足 φ的结果。那么，只要对于每个玩家 $j \in \mathrm{Ag} \setminus \{1\}$， $\gamma_1 = \varphi$ 且 $\gamma_j = \neg\varphi$，$(G$、$\mathbb{R}\sigma) \in$ 核心成员资格，当且仅当 $(G$、{1}) $\notin$ 已实现联盟。□

现在让我们来考虑有利偏离。这是多人博弈中唯一一个“容易”的问题：它可以在 PSpace 内解决。为此，我们还需要采用一种不同的证明策略。考虑该问题的任意输入实例 $(G$，$\mathbb{R}\sigma$，$\mathbb{R}\sigma'_C)$。我们注意到，由于 $\mathbb{R}\sigma'_C$ 是固定的，我们可以将它纳入进行博弈的竞技场，然后检查不属于 C 的玩家是否有针对 $\hat{\mathrm{O}}_{j\in C} \neg\gamma_j$ 的联合策略。根据有利偏离的定义，我们还需要检查 $\rho(\mathbb{R}\sigma) \models \acute{\mathrm{O}}_{j\in} \neg\gamma_j$ 是否成立。

换句话说，与我们迄今研究的所有其他判定问题（通常可在双指数时间内解决）不同，这个问题之所以能在 PSpace 内解决多人博弈，是因为它可以归约为一个单人博弈（由联盟 $\mathrm{Ag} \setminus C$ 给出），其 LTL目标（由 $\gamma_{\mathrm{Ag}\setminus C} = \hat{\mathrm{O}}_{j\in C} \neg\gamma_j$ 给出）定义在一个“乘积竞技场”（记作 M_C）上；该乘积竞技场由并发博弈结构 M 和我们要检查的联合策略 $\mathbb{R}\sigma'_C$ 构成。

定理 5. 有利偏离是 *PSPACE*-完全的，即使对于单人博弈也是如此。证明。检查 $\rho(\mathbb{R}\sigma) \models \acute{\mathrm{O}}_{j\in C} \neg\gamma_j$是否成立，可以在 PSpace 中完成。同样，这等价于在“乘积自动机”或“并发程序”上进行线性时序逻辑公式的模型检测。

如果该陈述不成立，那么根据定义，$\circledR\sigma'_C$ 不是有利偏离，因为 C 中至少有一位玩家已经通过 $\circledR\sigma$实现了其目标。如果该陈述成立，那么我们会检查对于不在 C中的玩家的所有联合策略 $\circledR\sigma'_{-C}$， $\rho(\circledR\sigma'_{-C}, \circledR\sigma'_C) \models \acute{O}_{j\in C}\gamma_j$是否成立。我们通过检查 $(M_C, \lambda', s^{0\prime}) \models \hat{O}_{j\in C}\neg\gamma_j$不成立是否成立，在 PSpace 中完成此操作，其中$M_C = (\mathrm{Ag}', \mathrm{Ac}', \mathrm{St}', s^{0\prime}, \mathrm{tr}')$ 是如下定义的并发博弈结构：

- $\mathrm{Ag}' = \{0\}$, $\mathrm{Ac}' = \Pi_{i\in \mathrm{Ag}\setminus C}\mathrm{Ac}_i$;
- $\mathrm{St}' = \mathrm{St} \times \Pi_{j\in C} Q_j$;
- $s^{0\prime} = (s^0, q_x^0, \ldots, q_y^0)$, such that $\sigma_z = (Q_z, q_z^0, \delta_z, \tau_z)$, $\vec{\sigma}'_C = (\sigma_x, \ldots, \sigma_y)$, and $z \in \{x, \ldots, y\}$;
- $\mathrm{tr}'((s, q_x, \ldots, q_y), (a, \ldots, b)) = (s', q'_x, \ldots, q'_y)$ such that
  - $s' = \mathrm{tr}(s, \tau(q_x), \ldots, \tau(q_y), a, \ldots, b)$, and
  - $q'_z = \delta(q_z, s)$, with $z \in \{x, \ldots, y\}$.

并且 λ' 定义为，

$$\lambda'(s, q_x, \ldots, q_y) = \lambda(s).$$

换言之， M_C 的转移方式与 M 相同，只是其行为受限于$\circledR\sigma'_C$已经定义的行为。

对于下界，我们使用 LTL 模型检测。 □

除了上述复杂度结果之外，我们还有一个模型论结果。不过，在陈述该结果之前，我们需要定义双模拟等价的概念 [35, 38]。设 M 和 M' 是两个具有相同智能体 Ag、动作 Ac 和原子命题 AP 的并发博弈结构。此外，分别用 St/St′ 和 tr/tr′ 表示它们各自的状态集合和转移函数。最后，设 λ 为

St 上的一个标记函数，且 λ' 为 St′上的一个标记函数。那么， $s^* \in \mathrm{St}$ 与 $t^* \in \mathrm{St}'$ 之间的一个双模拟是一个非空二元关系 $\sim \subseteq \mathrm{St} \times \mathrm{St}'$，满足：

- 我们有 $s^* \sim t^*$;
- 对于所有 $s \in \mathrm{St}$ 和所有 $t \in \mathrm{St}'$，若 $s \sim t$，则 $\lambda(s) = \lambda'(t)$;
- 对于所有 $s_1, s_2 \in \mathrm{St}$ 和所有 $t_1 \in \mathrm{St}'$，若 $s_1 \sim t_1$ 且对于某个 $\mathrm{ac} \in \mathrm{Ac}$，$\mathrm{tr}(s_1, \mathrm{ac}) = s_2$，则对于某个满足 $s_2 \sim t_2$ 的 $t_2 \in \mathrm{St}'$，$\mathrm{tr}'(t_1, \mathrm{ac}) = t_2$；
- 对于所有 $s_1 \in \mathrm{St}$ 和所有 $t_1, t_2 \in \mathrm{St}'$，若 $s_1 \sim t_1$ 且对于某个 $\mathrm{ac} \in \mathrm{Ac}$，$\mathrm{tr}(t_1, \mathrm{ac}) = t_2$，则对于某个满足 $s_2 \sim t_2$ 的 $s_2 \in \mathrm{St}'$，$\mathrm{tr}'(s_1, \mathrm{ac}) = s_2$。

如果两个并发博弈结构的初始状态双模拟等价，我们就称这两个并发博弈结构双模拟等价。有关并发博弈结构上的双模拟等价的更多细节，请参阅 [17]。接着，如果性质 P 满足：当 $M \sim N$时， P 在 M 上成立，当且仅当 P 在 N 上成立，那么我们称该性质具有双模拟不变性。

现在我们可以陈述我们的结果：非正式地说，检查某个LTL公式是否由核心中的某个结果满足，是一个双模拟不变性质。该结果很容易得到，并且直接源于E-Core的归属性证明。

推论 1。 设 $G = (M, \gamma_1, \ldots, \gamma_n)$是一个博弈， φ是一个*LTL*公式，且 M'是一个与 M双模拟的并发博弈结构。那么，$(G, \varphi) \in$ 属于*E-Core*，当且仅当$(G', \varphi) \in$属于*E-Core*，其中 $G' = (M', \gamma_1, \ldots, \gamma_n)$。

证明。 这源于 ATL^* 具有双模拟不变性，且核心可以在 ATL^* 中使用 $\varphi_{\text{E-Core}}$刻画，正如定理 3 的证明中所定义的那样。更具体地说，这是因为 $(M, \lambda, s^0) \models \varphi_{\text{E-Core}}(G, \varphi)$ 当且仅当 $(M', \lambda', s^{0\prime}) \models \varphi_{\text{E-Core}}(G', \varphi)$。□

3.4. 关于可信联盟形成：强核心

如上所述，我们对核心的定义假设采取最坏情况推理：偏离必须在面对所有 反制响应时都有利。这个定义具有稳健性，因为任何核心稳定结果都具有很强意义上的稳定性，但也可以认为，在某些情况下它过于 严格。具体而言，当某个联盟 C 考虑偏离 $\mathbb{R}\sigma_C$时，它当然可以假设其余玩家不会做出损害自身利益的行动。因此，有人可能会认为，偏离无需对其余玩家的所有 行为都有利，只需对那些可信的行为有利即可，也就是说，其余玩家可能会理性地选择这些行为。为使讨论更具体，请看下面的例子。

例2. 假设我们有一个双人博弈 G，其中有一个初始状态 s_0 和三个汇状态 s_1，s_2 和 s_3。每个玩家都有两个可选动作， a 和 b，初始状态的转移函数定义如下：

$$\begin{aligned}
\mathrm{tr}(s_0,(a,a)) &= s_1,\\
\mathrm{tr}(s_0,(a,b)) &= s_2,\\
\mathrm{tr}(s_0,(b,a)) &= s_3,\\
\mathrm{tr}(s_0,(b,b)) &= s_3
\end{aligned}$$

因此，该博弈如图2所示。此外，假设最终到达 s_1 的无限运行是唯一满足玩家一目标的运行，而最终到达 s_2 的运行是唯一满足玩家二目标的运行。现在，在这个博弈中，最终到达 s_1 的运行属于核心，但这是通过玩家1采用不可信（惩罚性）策略实现的。请注意，玩家2从 $(a,\ a)$ 出发唯一可能的偏离是选择 b，对此玩家1可以回应，也选择 b。尽管这种行为会阻止玩家2实现其目标，但鉴于玩家1的偏好关系，这种做法可以被视为不理性的：玩家1

当然更喜欢以 s_1 结束的运行，而非另外两种可能的运行，但除此之外并无偏好。

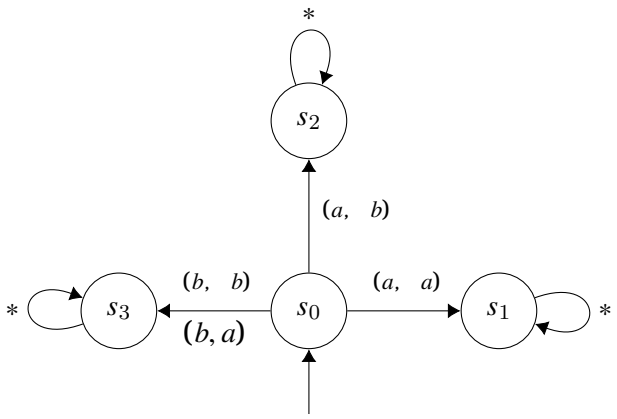


图1：一个具有不可信策略的博弈。

受这一现象启发，我们现在为核心提出一个更强的定义。更具体地说，根据这一定义，如果一个联盟 C想要偏离给定的策略组合，那么剩余玩家只有在其拥有一种反制响应时，才能可信地威胁 C；在这种反制响应中，至少有一名 C 中的玩家无法实现其目标，并且原策略组合中的每一名获胜者在新策略组合中仍然是获胜者，即，反制联盟依照其偏好关系行事。因此，在这一解概念中，我们刻画了剩余玩家愿意惩罚偏离者这一思想，但这种惩罚是有限度的：留下来的玩家更希望自己的处境不要比原来更糟。正如具有外部性的合作博弈中的所有解概念一样，我们并不声称这一解概念总是恰当的。但它刻画了不偏离联盟可能如何响应偏离的另一种有趣变化。

接着，我们重新表述有利偏离的定义。现在，如果满足以下条件，我们就称偏离 $\circledR\sigma'_C$ 是相对于 $\circledR\sigma$ 的强有利偏离

成立：

1. $C \subseteq \mathrm{L}(\circledR\sigma)$; 2. $C \subseteq \mathrm{W}(\circledR\sigma_{-C}, \circledR\sigma'_C)$; 3. 对于每个关于 $\mathrm{Ag}\backslash C$ 的联合策略 $\circledR\sigma'_{-C}$，如果 $\mathrm{W}(\circledR\sigma) \subseteq \mathrm{W}(\circledR\sigma'_{-C}, \circledR\sigma'_C)$，那么 $C \subseteq \mathrm{W}(\circledR\sigma'_{-C}, \circledR\sigma'_C)$。

有了这个定义，我们称博弈的**强核心**为 $\mathrm{core}^+(G)$，它是 G 中不存在强有利偏离的结果集合。接着，对于例2，我们看到，尽管 $\circledR\sigma_{a_1a_1}$ 属于 $\mathrm{core}(G)$，但 $\circledR\sigma_{a_1a_1}$ 并不属于 $\mathrm{core}^+(G)$，因为玩家2现在可以从 $\circledR\sigma_{a_1a_1}$ 强有利偏离到 $\circledR\sigma_{a_1b_1}$。

值得暂停片刻，思考在存在外部性的情况下如何定义核心这一问题。博弈论文献中关于这个主题的讨论非常丰富。原因在于，外部性的存在会导致稳定行为有许多不同的定义（参见例如， [13, 16, 39, 40] 其中列举了核心的许多变体）。在这里，我们提出一种定义，但绝不声称它是唯一的可能性。实质上，按照我们的定义，要使惩罚性联合策略可信，获胜者在发出威胁后必须仍然是获胜者。

现在，我们将研究前几节所定义的判定问题在强核心下的复杂度。共有四个判定问题，其定义取决于核心的性质：E-Core、A-Core、核心成员资格和有利偏离。我们将继续使用这些问题的相同名称，但需注意，本节中的结果均针对强核心。接下来我们将说明，这四个问题的复杂度与核心下相同，但需要更复杂的逻辑刻画；本文将使用策略逻辑（SL）的双交替片段给出这种刻画 [23]。[4]

[4]我们无法使用 ATL* 对强核心进行逻辑刻画。事实上，我们

定理 6。 给定一个博弈 $G = (M, \gamma_1, \ldots, \gamma_n)$ 和一个 LTL公式 φ，当且仅当 $M \models \varphi^{+}_{E\text{-}Core}(G, \varphi)$时，我们有$(G, \varphi) \in$*E-Core*，其中 $\varphi^{+}_{E\text{-}Core}(G, \varphi)$是以下 *SL* 公式:

$$\varphi^{+}_{E\text{-}Core}(G, \varphi) = \bigvee_{W \subseteq \mathrm{Ag}} \langle \mathrm{Ag} \rangle \left(\varphi \wedge \bigwedge_{i \in W} \gamma_i \wedge \bigwedge_{j \in \mathrm{Ag} \setminus W} \neg \gamma_j \wedge \bigwedge_{C \subseteq \mathrm{Ag} \setminus W} \varphi_{\mathrm{NoBD}}(G, W, C) \right)$$

并且公式 $\varphi_{\mathrm{NoSBD}}(G, W, C)$ 定义如下:

$$\varphi_{\mathrm{NoSBD}}(G, W, C) = [C] \left(\bigwedge_{j \in C} \gamma_j \rightarrow \langle \mathrm{Ag} \setminus C \rangle \left(\bigwedge_{i \in W} \gamma_i \wedge \bigvee_{j \in C} \neg \gamma_j \right) \right)$$

证明。 这个 SL 公式表示，在并发博弈结构中，存在一条路径 hAgi (...), 满足

1. 公式 φ 成立;
2. 一些玩家实现了自己的目标：Ó $_{i \in W}\, \gamma_i$；3. 其余玩家没有实现自己的目标：Ó $_{j \in \mathrm{Ag} \setminus W}\, \neg \gamma_j$；4. 失败者联盟不存在强有利偏离：Ó $_{C \subseteq \mathrm{Ag} \setminus W}\, \varphi_{\mathrm{NoSBD}}(G, W, C)$。

我们用 SL 公式 $\varphi_{\mathrm{NoSBD}}(G, W, C)$ 表达失败者联盟 C 不存在强有利偏离这一条件；该公式可分解如下：对于 C 的每个联合策略，如果 C 中的每个玩家都能获得更好的结果 $\left(\text{Ó}_{j \in C}\, \gamma_j\right)$，那么 C 之外的玩家联盟就有一个联合策略 (hAg \ Ci...)，使得原结果中的获胜者在提出威胁后仍然是获胜者 (Ó $_{i \in W}\, \gamma_i$)，并且偏离联盟中至少有一名玩家 C 未能实现其目标 $\left(\text{Ô}_{j \in C}\, \neg \gamma_j\right)$。□

此时，我们想提出几点观察。首先，检查 SL公式的复杂度是非初等的，并且取决于

认为这种多人博弈的逻辑刻画在 ATL* 中是不可能实现的。

公式的交替深度（[23]）：交替深度为 n 的 SL 公式可在 $(n+1)$-指数时间内完成检查；对于在语义上等价于 CTL* 公式的公式，则可在 PSpace 内完成检查。由于 $\varphi^{+}_{\text{E-Core}}(G, \varphi)$ 是一个含有两次交替的 SL 公式，因此可以在 3- 指数时间内完成检查（若 $|\text{Ag}| = 1$，则可以在 PSpace 内完成检查）。其次，我们还想重申，有限状态机策略（如我们在此使用的策略）可以使用 [2, 3] 中介绍的技术，通过线性时序逻辑来刻画。利用这些逻辑刻画，我们可以得出以下复杂性结果。

定理7。 对于多人博弈，*E-Core* 和 *A-Core* 属于 *3*-指数时间，而核心成员资格是 *2*-指数时间完全问题，有利偏离是 *PSPACE* 完全问题。对于单人博弈，所有问题都是 *PSPACE* 完全问题。

由于我们使用 SL 刻画了强核心（与 ATL* 不同，SL 并非双模拟不变逻辑），因此无法得出结论：强核心中的结果满足线性时序逻辑性质这一点是双模拟不变性质。我们认为事实并非如此。

在最初对核心的表述中，我们看到核心总是非空的。因此，一个自然的问题是：强核心是否总是非空？我们首先证明，对于玩家人数不超过三人的博弈，答案是肯定的：

定理 8。 对于任意二人博弈和三人博弈，G，有 核心$^{+}(G) \neq \emptyset$。

证明。 首先，考虑只有两名玩家的博弈。为导出矛盾，假设对于某个博弈 G，结果核心$^{+}(G)$ 为空。这意味着，对于每个结果，玩家1或玩家2，或者两者，都存在强有利偏离。于是，我们知道，没有任何结果能够同时满足两个目标，γ_1 和 γ_2。接下来考虑剩下的三种可能情况：只满足 γ_1 的结果（情况1）、只满足 γ_2 的结果（情况2），以及两者都不满足的结果

®γ_1（情况3）。设 $f=(f_1,f_2)$ 为一个结果，$f'1$ 为玩家1的一个偏离，f_2' 为玩家2的一个偏离，并考虑上述三种情况。在情况1中，只有玩家2会偏离。于是，结果（f_1,f_2'）只满足 γ_2。由于（f_1,f_2'）也不在核心中，因此从该结果出发，只有玩家1会偏离，转而到达另一个结果（$f'1,f'2$）。于是，结果（$f'1,f'2$）只满足 γ_1。但是，这就产生了矛盾，因为这意味着（f_1,f_2）将属于核心$^+(G)$。通过对称推理，可以证明情况2同样不可能。对于情况3，我们注意到，只有单个偏离才是可能的。然而，任何此类偏离都会转向一个只满足 γ_1 或只满足 γ_2的结果，而这两种情况都已不再可能。由于不存在其他情况，我们必须否定原先的假设，并得出结论：对于双人博弈，核心$^+(G)$并非空集。

对于三人博弈，证明过程类似，但需要仔细地逐一分析各种情况。完整证明见附录。 □

相比之下，对于四名或更多玩家参与的博弈，强核心可能为空，以下例子便说明了这一点。

例3. 考虑以下四人博弈：该博弈有一个初始状态 s^0 和六个汇状态 s^1、……、s^6。每位玩家各有两个可选动作 {0，1}。将动作 $(a_1,\ a_2,\ a_3,\ a_4)$ 记作 $a_1a_2a_3a_4$，则从初始状态出发的转移函数如下：

$$\begin{aligned}
&\mathrm{tr}(0000)=s^1, && \mathrm{tr}(0001)=s^1,\\
&\mathrm{tr}(0010)=s^2, && \mathrm{tr}(0011)=s^2,\\
&\mathrm{tr}(0100)=s^1, && \mathrm{tr}(0101)=s^3,\\
&\mathrm{tr}(0110)=s^2, && \mathrm{tr}(0111)=s^5,\\
&\mathrm{tr}(1000)=s^6, && \mathrm{tr}(1001)=s^4,\\
&\mathrm{tr}(1010)=s^4, && \mathrm{tr}(1011)=s^4,
\end{aligned}$$

$$\mathrm{tr}(1100) = s^1, \qquad \mathrm{tr}(1101) = s^3,$$
$$\mathrm{tr}(1110) = s^4, \qquad \mathrm{tr}(1111) = s^3.$$

此外，假设玩家一偏好最终到达状态 s^1、s^2、s^3 的运行，玩家二偏好最终到达 s^1、s^4、s^5 的运行，玩家三偏好 s^2、s^4、s^6，玩家四偏好 s^3、s^5、s^6。博弈如图2所示——我们用在该状态获胜的玩家标记状态，并用可能的动作元组标记边。[5]

[5]我们想指出，这个（反）例子是通过对不同规模的博弈进行有界穷举搜索自动生成的。这或许可以解释为什么这个博弈如此违反直觉，更重要的是，为什么像这样的博弈很难靠手工找到或构造出来。

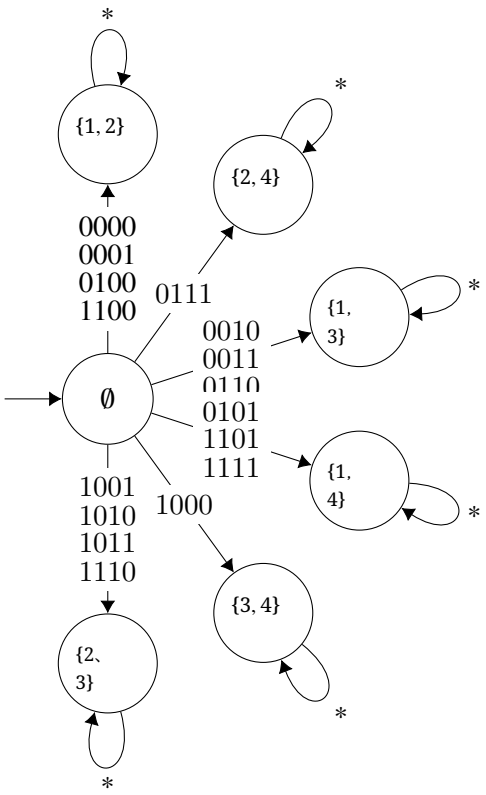


图2：一个空核博弈$^{+}$。

通过仔细逐一分析各种情况，可以验证，对于每个状态，都存在某个联盟可以进行强有利偏离。因此，该博弈的强核心为空。关于各联盟可以进行的偏离的完整分析，请参阅附录，其中给出了完整的逐一分析。

4. 平均收益博弈

到目前为止，我们研究的是具有定性偏好的博弈——每个玩家都有一个由某个时序逻辑公式给出的目标；在给定运行下，该目标要么得到满足，要么不满足。但这种方法对于偏好的描述而言过于粗略，

无法表达单个玩家偏好的强度。一种获得更丰富偏好模型的可能方法，是为每个玩家引入多个 LTL 目标，并定义一个从满足的公式集合到实数 [41–43] 的映射。然而，另一种在文献中得到广泛研究的方法，是完全绕开时序逻辑，为状态而非原子命题赋予权重。然后，我们计算运行的**平均收益**，其思想是，智能体偏好使其平均收益最大化的运行 [18, 44, 45]。本节将重新审视均值收益情形下核心的表述。

形式化地说，均值收益博弈 G 是一个元组，

$$G = (M, \{w_i\}_{i \in \mathrm{Ag}}),$$

其中 M 是一个并发博弈结构，并且对于每个 $i \in \mathrm{Ag}$，$w_i : \mathrm{St} \to \mathbb{Z}$ 是一个**权重**函数，将状态映射为整数。博弈的进行方式与线性时序逻辑设置中的方式相同，但此处智能体的偏好关系定义有所不同。设 $\beta \in \mathbb{R}^\omega$ 是一个实数无限序列。那么，β 的平均收益记为 $\mathsf{mp}(\beta)$，定义如下：

$$mp(\beta) = \liminf_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \beta_i.$$

在均值收益博弈中，一次运行 $\rho = s^0 s^1 \ldots$ 会为每个玩家导出一个权重无限序列，$w_i(s^0) w_i(s^1) \ldots$——我们用 $w_i(\rho)$ 表示该序列。为方便记号，我们将用 $\text{收益}_i(\rho)$ 表示 $\mathsf{mp}(w_i(\rho))$。据此，我们可以定义每个玩家的偏好关系：给定两次运行 ρ 和 ρ'，若 $\rho \succeq_i \rho'$，则 $\text{收益}_i(\rho) \geq \text{收益}_i(\rho')$；严格关系 $\succ_i$ 按照通常方式定义。

现在，回顾一下我们对 LTL 博弈情形下核心的定义，该定义依赖于博弈中赢家与输家的概念。在平均收益情形下，将玩家划分为赢家或输家显然毫无意义——他们可能获得范围很广的

收益谱。因此，我们需要重新审视有利偏离的概念。在平均收益情形下，对于给定的策略组合 $®\sigma$，我们称联盟 C 的有利偏离是一个策略向量 $®\sigma'_C$，使得对于所有互补的策略组合 $®\sigma'_{\mathrm{Ag}\setminus C}$，对于所有 $i \in C$，均有 $\rho(®\sigma'_C, ®\sigma'_{\mathrm{Ag}\setminus C}) \succ_i \rho(®\sigma)$。接着，如果不存在任何联盟 C 能够从 $®\sigma$ 进行有利偏离，我们就称 $®\sigma$ 属于核心。

4.1. 核心的非空性

我们首先要问，均值收益博弈的核心是否总是非空的。回顾一下在线性时序逻辑博弈中，我们发现核心必定非空；但我们发现，对于均值收益博弈而言，这通常并不成立。在证明该定理之前，我们需要一个小引理来说明.....

引理 2. 设 $\{\vec{\sigma}^j\}_{j\in\mathbb{N}}$是一个策略组合序列，并假设对于某个玩家 i，有 $\lim_{j\to\infty}$收益$_i(\rho(\vec{\sigma}^j)) = x$，其中 $x \in \mathbb{R}$。那么存在某个策略组合 $\vec{\sigma}^x$，使得收益$_i(\rho(\vec{\sigma}^x)) = x$。

证明。 首先注意，博弈最终会进入一个强连通分量。令 C 为该强连通分量的博弈图中所有简单环构成的集合。考虑一个解为 x 的线性规划——它给出了各环的比例 □

定理 9。 在均值收益博弈中，如果 $|\mathrm{Ag}| \leq 2$，那么核心非空。对于$|\mathrm{Ag}| > 2$，存在核心为空的博弈。

证明。 如果 $|\mathrm{Ag}| = 1$，很容易看出核心总是非空；我们使用 Karp 算法来确定加权图 [46]中的最大环，从而确定一名玩家能够获得的最大收益。对于双人博弈，令 $®\sigma = (\sigma_1, \sigma_2)$ 为任意策略组合。如果 $®\sigma$ 不在核心中，那么

玩家1、玩家2，或由两名玩家组成的联盟存在有利偏离。如果后一种情况成立，那么存在一个策略组合 $®\sigma' = (\sigma_1', \sigma_2')$，使得 $®\sigma' \succ_i ®\sigma$ 对所有 $i \in \{1, 2\}$ 均成立。我们重复这一过程，直到两名玩家组成的联盟不再存在有利偏离。最终必定会出现这种情况，因为：1）每位玩家的收益都不超过其最大权重；2）根据 [47], 中的定理 4，我们知道联盟能够达到的收益集合是闭集，因此任何极限点都可以达到。[6] 所以，终会出现双方无法共同进行有利偏离的时刻。此时，我们要么处于核心中，要么玩家1或玩家2存在有利偏离。如果玩家 $j \in \{1, 2\}$ 存在有利偏离，例如 σ_j，那么任何满足 $i \neq j$ 的策略组合 (σ_j, σ_i) 只要能使玩家 i 的平均收益最大化，就属于核心。因此，对于每个双人博弈，都存在一个属于核心的策略组合。

然而，对于三名或更多玩家参与的平均收益博弈，博弈的核心可能为空。下面的例子说明了这种情况。

例 4. 考虑以下三人博弈 G，其中每位玩家有两个行动，**H**、**T**，并且有四个状态： P、 R、 B、 Y。各玩家对这些状态的权重如下：

$w_i(s)$	1	2	3
P	-1	-1	-1
R	2	1	0
B	0	2	1
Y	1	0	2

如果博弈处于除 P 以外的任何状态，那么无论采取什么行动集合

[6] [47] 的结果实际上指的是双人多重平均收益博弈，而我们讨论的是多人平均收益博弈。我们马上会解释为什么这并不构成问题。

一旦采取该行动，博弈就会保持在该状态。因此，我们只需指定状态 P 的转移：

Ac	St
(**H**,**H**, **H**)	*R*
(**H**, **H**, **T**)	*R*
(**H**,**T**, **H**)	*B*
(**H**, **T**, **T**)	*P*
(**T**, **H**, **H**)	*P*
(**T**,**H**, **T**)	*Y*
(**T**,**T**, **H**)	*B*
(**T**,**T**, **T**)	*Y*

图3展示了该博弈的结构。

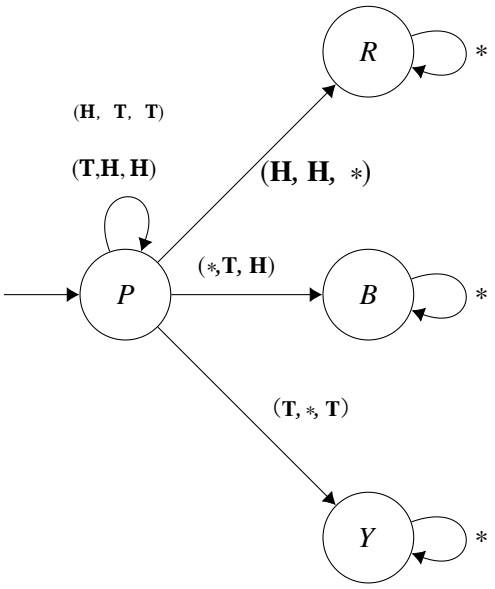


图3：一个具有空核的博弈。

请注意，策略由博弈最终到达的状态来表征。如果玩家永远停留在 P，那么他们可以共同改变策略，转移到 R、 B或 Y中的一个，并且每个人都能获得更高的收益。现在，如果博弈最终到达R，那么玩家2和玩家3可以通过采取（**T**，**H**）来偏离；无论玩家1采取什么行动，博弈都会处于状态 B，使这两位偏离玩家的境况更好。但同样地，如果博弈处于 B，那么玩家1和玩家3可以通过采取（**T**，**T**）来偏离，进入状态 Y；无论玩家2采取什么行动，他们的境况都会更好。最后，如果博弈处于 Y，那么玩家1和玩家2可以通过采取（**H**， **H**）来偏离，进入 R；无论玩家3采取什么行动，他们的境况都会更好。因此，没有任何策略组合属于核心。

□

4.2. 判定问题

现在我们转而关注均值收益设定下与核心相关的判定问题。然而，从计算视角来看，这里有一个直接的问题——给定一个潜在的有利偏离，我们如何验证：在所有 可能的反制响应下，它是否都优于现状？幸运的是，正如我们将在下一个引理中所示，在考虑针对玩家偏离的潜在反制响应时，我们可以将注意力限制在无记忆策略 上（无记忆策略的定义见第 2 节）：

引理 3。 设 G 是一个博弈， $C \subseteq \mathrm{Ag}$ 是一个联盟，且 $\vec{\sigma}$ 是一个策略组合。进一步假设$\vec{\sigma}'_C$ 是一个策略向量，使得对于所有无记忆策略向量$\vec{\sigma}'_{\mathrm{Ag}\setminus C}$，都有，

$$\rho(\vec{\sigma}'_C, \vec{\sigma}'_{\mathrm{Ag}\setminus C}) \succ_i \rho(\vec{\sigma}).$$

那么，对于所有策略向量$\vec{\sigma}'_{\mathrm{Ag}\setminus C}$（不一定是无记忆的），我们有：

$$\rho(\vec{\sigma}'_C, \vec{\sigma}'_{\mathrm{Ag}\setminus C}) \succ_i \rho(\vec{\sigma}).$$

在证明这一点之前，我们需要引入一个辅助概念：双人轮流制零和多重平均收益博弈 [48]（下文将简称为多重平均收益博弈）。非正式地说，这类博弈与双人轮流制零和平均收益博弈类似，不同之处在于，玩家1为边关联了 k 个权重函数，并试图确保所得的 k 维平均收益向量按分量大于某个阈值。形式上，一个多重平均收益博弈由一个结构给出

$$G = (V_1, V_2, v^0, E, w, z^k)$$

其中 V_1，V_2 分别是由玩家1和玩家2控制的状态集合，$V = V_1 \cup V_2$ 为状态空间，$v^0 \in V$ 为初始状态，$E \subseteq V \times V$ 为边集合，$w: E \to \mathbb{Z}^k$ 为权重函数，为每条边分配一个权重向量，$z^k \in \mathbb{Q}^k$ 为阈值向量。

博弈从初始状态 $v^0 \in V_i$ 开始，玩家 i 选择一条边（v^0，v^1）并沿此边移动到下一个状态。到达新状态后，$v^1 \in S_j$，玩家 j 选择一条边，如此反复，永远持续下去。运行按通常方式定义，而一次运行 ρ 的收益 $\mathsf{pay}(\rho)$ 只是向量（$\mathsf{mp}(w_1(\rho))$，……，$\mathsf{mp}(w_k(\rho))$）。如果对于所有 $i \in \{1, \ldots\ldots, k\}$，玩家1的 $\mathsf{pay}_i(\rho) \geq z_i$，则玩家1获胜，否则失败。与这些博弈相关的基本问题是：玩家1能否确保获胜：

多维平均收益阈值问题：

已知：多维平均收益博弈 G。

问题：玩家1是否拥有获胜策略?

如 [48], 所示，这个问题是 co-NP 完全的。虽然我们目前不需要使用这一复杂度结果，但之后会用到这一事实。还值得注意的是，在我们的多人平均收益博弈中，权重附加在状态上，而在

多重平均收益博弈中，权重附加在边上。就我们的目的而言，这种差异只是表面上的——只需将权重推到出边上，就可以把前者映射为后者；而通过增加更多状态（每条边对应一个状态），则可以把后者映射为前者。只要以正确的方式进行映射，额外开销就只有多项式级。接下来我们会隐式地使用这种映射（事实上，我们已经在定理 9 的证明中使用过了）。

介绍了这个判定问题后，我们现在可以证明引理 3。

引理 3 的证明。设 $\mathbb{R}\sigma' \text{Ag}\backslash C$ 是任意策略，并设 $i \in C$ 是任意主体。假设并非如此： $\rho(\mathbb{R}\sigma'_C, \mathbb{R}\sigma' \text{Ag}\backslash C) > _i \rho(\mathbb{R}\sigma)$。因此，我们有$\rho(\mathbb{R}\sigma) \geq_i \rho(\mathbb{R}\sigma'_C, \mathbb{R}\sigma' \text{Ag}\backslash C)$。将其视为一个双人多重平均收益博弈，其中玩家 1 的策略固定并编码在博弈结构中（即玩家1遵循 $\mathbb{R}\sigma'_C$，但无权决定），且收益阈值为 $\text{mp}(\rho(\mathbb{R}\sigma))$，那么 $\mathbb{R}\sigma' \text{Ag}\backslash C$ 就是该博弈中玩家2的获胜策略。根据 [21, 48],，如果玩家2有获胜策略，那么他们就有无记忆获胜策略。因此，存在一个无记忆策略 $\mathbb{R}\sigma'' \text{Ag}\backslash C$，使得 $\rho(\mathbb{R}\sigma) \geq_i \rho(\mathbb{R}\sigma'_C, \mathbb{R}\sigma'' \text{Ag}\backslash C)$。但这与引理的假设矛盾，因此必有 $\rho(\mathbb{R}\sigma'_C, \mathbb{R}\sigma' \text{Ag}\backslash C) > _i \rho(\mathbb{R}\sigma)$。□

现在我们可以考察合作场景下均值收益博弈的一些复杂度界。首先考虑以下与有利偏离有关的判定问题：

有利偏离：

给定：博弈 G 和策略组合$\mathbb{R}\sigma$。

问题：是否存在某个联盟，能够从 $\mathbb{R}\sigma$ 实现有利偏离？也就是说，是否存在 $C \subseteq \text{Ag}$ 和 $\mathbb{R}\sigma'_C \in \Sigma_C$，使得对于所有

$\text{®}\sigma'_{\text{Ag}\setminus C} \in \Sigma_{\text{Ag}\setminus C}$，对于所有 $i \in C$，都有：

$$\rho(\vec{\sigma}'_C, \vec{\sigma}'_{\text{Ag}\setminus C}) \succ_i \rho(\vec{\sigma})?$$

我们有：

定理 10. 如果给定的策略组合 $\vec{\sigma}$ 是无记忆的，那么有利偏离是 *NP* 完全。

证明。 首先，正确地猜测一个偏离联盟 C 以及该玩家联盟的策略组合 $\text{®}\sigma'_C$。然后，使用以下三步算法。首先，计算玩家在 C 上通过（$\text{®}\sigma$）获得的平均收益，也就是说，对每个 $j \in C$ 计算一组值 $z^*_j = \text{收益}_j(\rho(\text{®}\sigma))$——只需“运行”策略组合 $\text{®}\sigma$，即可在多项式时间内完成。接着，计算图 $G[\text{®}\sigma'_C]$，其中包含关于 $\text{®}\sigma$ 的 $\text{Ag}\setminus C$ 的所有可能行为（即策略组合）——这一构造与定理 5 的证明中所用的构造类似，也就是固定 $\text{®}\sigma'_C$ 时的博弈，并且可以在多项式时间内完成。最后，我们询问 $G[\text{®}\sigma'_C]$ 中的每条路径 ρ 是否都满足 $\text{收益}_j(\rho) > z^*_j$，对每个 $j \in C$ 都如此——对于这一步，我们可以使用 Karp 算法 [46]，在多项式时间内对每个 $j \in C$ 回答这个问题。如果 $G[\text{®}\sigma'_C]$ 中的每条路径都具有这一性质，则接受；否则拒绝。

对于困难性，我们从 3SAT 归约，采用 [28] 中构造的一个小变体。令 $P = \{x_1, \ldots\ldots, x_n\}$ 为一个原子命题集合。给定一个关于 P — 的布尔公式 $\varphi = \text{Ó}_{1 \le c \le m} C_c$（合取范式），其中每个 $C_c = l_{c1} \vee l_{c2} \vee l_{c3}$，且每个文字 $l_{ck} = x_j$ 或 $\neg x_j$，其中 $1 \le k \le 3$，对于某个 $1 \le j \le n$ —，我们构造 $M = \left(\text{Ag}, \text{St}, s^0, (\text{Ac}_i)_{i \in \text{Ag}}, \text{tr}\right)$，这是一个 m 玩家并发博弈结构，定义如下，并在图 4 中示出：

- $\text{Ag} = \{1, \ldots, m\}$;
- $\text{St} = \{x_v \mid 1 \le v \le n\} \cup \{x'_v \mid 1 \le v \le n\} \cup \{y_0, y_n, y^0, y^*\}$;

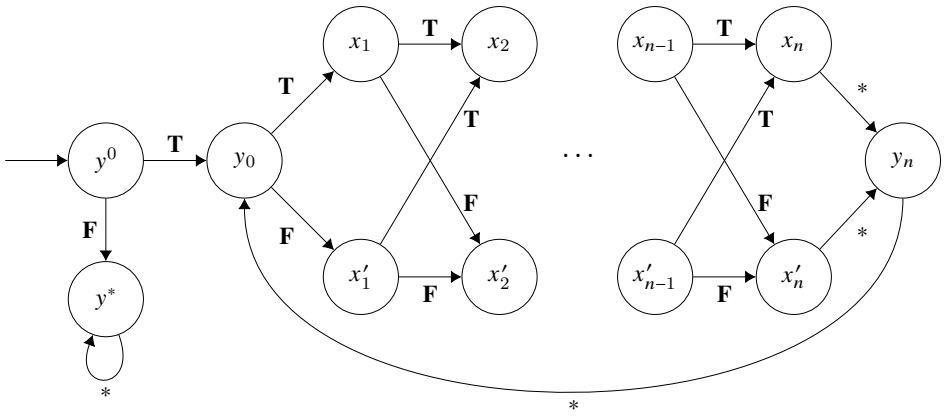


图 4：从 3SAT 归约得到的并发博弈结构。

- $s^0 = y^0$;

- $\mathrm{Ac}_i = \{t,\ f\}$，对每个 $i \in \mathrm{Ag}$ 均成立，且 $\mathrm{Ac} = \mathrm{Ac}_1 \times \cdots \times \mathrm{Ac}_m$;

- 关于 tr，参见图 4，其中 $T = \{(t_1, \ldots, t_m)\}$ 且 $F = \mathrm{Ac} \setminus T$。

有了 M，我们使用以下权重函数构造一个均值收益博弈：

- $w_i(x_v) = 1$如果 x_v 是 C_i 中的一个文字，且 $w_i(x_v) = 0$否则，对所有 $i \in \mathrm{Ag}$ 以及 $1 \le v \le n$

- $w_i(x_v) = 1$ 若 $\neg x_v$ 是 C_i 中的文字，且对所有 $i \in \mathrm{Ag}$ 和 $1 \le v \le n$，否则为 $w_i(x_v) = 0$

- $w_i(y_0) = w_i(y_n) = w_i(y^0) = w_i(y^*) = 0$，对所有 $i \in \mathrm{Ag}$

接着，我们考虑博弈 G 在 M 上，以及任意满足 $\sigma(s^0) = y^*$ 的（无记忆策略）策略组合。对于这类策略组合中的任意一个，每位玩家的平均收益都是 0。然而，如果 φ 可满足，那么在 M中存在一条从 y_0

到 y_n的路径，使得沿着这条路径，每位玩家都有一个收益不为 0 的状态。因此，大联盟 Ag 有动机偏离，因为无限次地沿着这条路径移动，会使每位玩家的平均收益严格大于 0。注意两点。首先，只有大联盟 Ag 同意，博弈才能在 y^0之后到达 y_0。否则，博弈必然会从此永远停留在 y^*。其次，由于我们考虑的是无记忆策略，起初从 y_0 到 y_n所经过的路径，与之后将无限次重复经过的路径相同。由此可得，当且仅当 φ 可满足时，才存在有利偏离（必然是 Ag 的偏离）；否则，博弈中至少有一位玩家没有偏离的动机（因为其平均收益仍将为 0）。因此，我们得出结论：$(G, \sigma) \in$ 存在有利偏离，当且仅当 φ可满足。□

由定理 10 可知，判断是否不存在任何玩家联盟相对于给定策略组合存在有利偏离是 co-NP 完全的，因此我们有：

推论 2. 如果给定的策略组合$\vec{\sigma}$ 是无记忆的，那么核心成员判定问题是 *co-NP* 完全的。

我们还可以利用有利偏离得到以下结论：

定理 11. 如果将策略上的存在量化限制为只考虑无记忆策略，那么 *E-Core* 属于 Σ_2^P。

证明。 给定一个博弈 G，我们猜测一个策略组合 σ，并检查 (G,σ) 是否不是有利偏离的一个实例。前者可以在多项式时间内完成，而后者可以借助有利偏离问题的 co-NP 预言机求解。因此，我们得到一个在 NP $^{\text{co-NP}} = \Sigma_2^P$ 中运行的过程。

□

由于核心和有利偏离的定义中存在交替量词，我们认为 E-Core 问题是 Σ_2^P 完全的，但我们尚未能找到相关证明。请注意，这表明它与均值收益博弈中纳什均衡的对应问题形成对比；后者属于 NP [45]。更重要的是，该结果还表明，对联盟偏离类型的（复杂度）依赖性较弱：只要能够借助 NP 或 co-NP 预言机检查这些偏离，就可以在同一复杂度类中考虑不同类型的有利偏离。

4.3. 已满足联盟

我们现在将已实现联盟的概念加以推广。为此，我们引入**下界**的概念。设 $C \subseteq \text{Ag}$ 是博弈 G 中的一个联盟，且令 $®z_C \in \mathbb{Q}^C$。如果存在一个针对 C 的联合策略 $®\sigma_C$，使得对于针对 $\text{Ag} \setminus C$的所有策略 $®\sigma_{-C}$，对于每个 $i \in C$，都有支付$_i(\rho(®\sigma_C, ®\sigma_{-C})) \geq z_i$，那么我们称 $®z_C$ 是 C 的下界。基于这个定义，我们可以证明以下结果，它用下界刻画了核心：

引理 4。设 ρ 是 G中的一次运行。当且仅当对于每个联盟 $C \subseteq \text{Ag}$ 以及 C 的下界$\vec{z}_C \in \mathbb{Q}^C$，都存在某个$i \in C$，使得 $z_i \leq$支付$_i(\rho)$时，存在$\vec{\sigma} \in$ 核心(G)，满足 $\rho = \rho(\vec{\sigma})$。

证明。 为了证明从左到右的方向，假设存在一个核心 $®\sigma \in \text{core}(G)$ 的成员，并进一步假设存在某个联盟 $C \subseteq \text{Ag}$ 和关于 C 的下界 $®z_C \in \mathbb{Q}^C$，使得对于每个 $i \in C$，都有 $z_i >$收益$_i(\rho)$。由于 $®z_C$ 是 C 的下界，并且对于每个 $i \in C$，都有 $z_i >$收益$_i(\rho)$，因此存在一个关于 C 的联合策略 $®\sigma_C$，使得对于 $\text{Ag} \setminus C$ 的所有策略 $®\sigma_{-C}$，对于每个 $i \in C$，都有 收益$_i(\rho(®\sigma_C, ®\sigma_{-C})) \geq z_i >$收益$_i(\rho)$。由此可知，$(G, ®\sigma)$ 是有利偏离，这进一步意味着 $®\sigma$ 不可能处于 G 的核心中——这与我们的初始假设矛盾。

为了证明从右到左的方向，假设存在 ρ 属于 G，使得对于每个联盟 $C \subseteq \text{Ag}$ 和关于 C 的下界 $\circledR z_C \in \mathbb{Q}^C$，都存在 $i \in C$，使得$z_i \leq$收益$_i(\rho)$。于是，我们只需令 $\circledR\sigma$ 为任意一个满足 $\rho = \rho(\circledR\sigma)$ 的策略组合。现在，任取联盟 $C = \{j, \ldots\ldots, k\} \subseteq \text{Ag}$，并令 $\circledR\sigma'_C$ 为 C 相对于 $\circledR\sigma$ 的任意可能偏离。要么 $\circledR z'_C = (\text{收益}_j(\rho(\circledR\sigma_{-C}, \circledR\sigma'_C)), \ldots\ldots, \text{收益}_k(\rho(\circledR\sigma_{-C}, \circledR\sigma'_C)))$ 是 C 的下界，要么不是。

如果前一种情况成立，根据假设，我们知道存在 $i \in C$，使得收益$_i(\rho(\circledR\sigma_{-C}, \circledR\sigma'_C)) \leq$收益$_i(\rho)$。因此，$i$不会有动机与 $C \setminus \{i\}$ 一起偏离 $\circledR\sigma$，因而联盟 C 无法从 $\circledR\sigma$ 中进行有利偏离。

$\circledR$ 另一方面，如果 z'_C 不是 C的下界，那么根据下界的定义，我们知道并非存在 $\circledR\sigma'_C$，使其为 C的联合策略，且对于 $\text{Ag} \setminus C$ 的所有策略 $\circledR\sigma'_{-C}$，都有 $\text{pay}_i(\rho(\circledR\sigma'_C, \circledR\sigma'_{-C})) \geq \text{pay}_i(\rho(\circledR\sigma_{-C}, \circledR\sigma'_C))$，对每个 $i \in C$ 都成立。也就是说，存在 $i \in C$ 和 $\text{Ag} \setminus C$ 的 $\circledR\sigma'_{-C}$，使得 $\text{pay}_i(\rho(\circledR\sigma'_C, \circledR\sigma'_{-C})) < \text{pay}_i(\rho(\circledR\sigma_{-C}, \circledR\sigma'_C))$。现在我们将选择 $\circledR\sigma'_{-C}$，使得此外，对于某个 i，$\text{pay}_i(\rho) \geq \text{pay}_i(\rho(\circledR\sigma'_C, \circledR\sigma'_{-C}))$。

令 $\circledR z''_C = (\text{pay}_j(\rho(\circledR\sigma^j_{-C}, \circledR\sigma'_C)), \ldots, \text{pay}_k(\rho(\circledR\sigma^k_{-C}, \circledR\sigma'_C)))$，其中 $\text{pay}_i(\rho(\circledR\sigma^i_{-C}, \circledR\sigma'_C))$ 定义为 $\min_{\circledR\sigma'_{-C} \in \Sigma_{-C}} \text{pay}_i(\rho(\circledR\sigma'_{-C}, \circledR\sigma'_C))$。也就是说，$\circledR\sigma^i_{-C}$是 $\text{Ag} \setminus C$ 的一种策略，它能在假设 C 正在采用联合策略 $\circledR\sigma'_C$ 的情况下，为 i 确保最低的平均收益。根据构造，$\circledR z''_C$ 是 C 的一个下界——因为每个 $z''_i = \text{pay}_i(\rho(\circledR\sigma^i_{-C}, \circledR\sigma'_C))$ 都是 i 在 C 采用 $\circledR\sigma'_C$ 时，无论联盟 $\text{Ag} \setminus C$ 如何行动，都能为自己确保的最大平均收益值——因此，根据假设，我们知道对于某个 $i \in C$，有 $\text{pay}_i(\rho(\circledR\sigma^i_{-C}, \circledR\sigma'_C)) \leq \text{pay}_i(\rho)$。因此，与之前一样，$i$ 没有动机与 $C \setminus \{i\}$ 一起偏离 $\circledR\sigma$，所以联盟 C 无法通过偏离 $\circledR\sigma$ 获得更高收益。由于 C 和 $\circledR\sigma'_C$ 是任意选取的，我们得出 $\circledR\sigma \in \text{core}(G)$，由此证明了从右到左的方向，并完成了证明。□

有了这个引理，我们想要确定给定的向量 $\circledR z_C$ 是否确实是一个下界，更重要的是，如何高效地完成这一判断。也就是说，我们要研究以下判定问题：

下界：

给定：博弈 G、联盟 $C \subseteq \mathrm{Ag}$ 和向量 $\circledR z_C \in \mathbb{Q}^{\mathrm{Ag}}$。

问题：$\circledR z_C$ 是否是 C 在 G 中的下界?

利用前面介绍的多维平均收益阈值问题，我们可以显然地证明以下定理：

定理 12. 下界问题是 *co-NP* 完全。

证明。 我们通过将多维平均收益阈值问题分别规约到下界问题以及从下界问题规约到多维平均收益阈值问题，以显然的方式证明其归属性和困难性。首先，我们通过将下界问题规约到多维平均收益阈值问题，证明下界问题属于 co-NP。假设给定一个实例 $(G, C, \circledR z_C)$，我们要判断它是否属于下界问题。我们可以构造一个双人多重平均收益博弈 $G'=(V_1, V_2, v^0, E, w', z^k)$。这里，$V_1 = \mathrm{St}$、$V_2 = \mathrm{St} \times \mathrm{Ac}_C$ 和 $v^0 = s^0$。此外，G'、E 的边集合定义如下，

$$\begin{aligned} E = &\{(s, (s, \mathrm{ac}_C)) \mid s \in \mathrm{St}, \mathrm{ac}_C \in \mathrm{Ac}_C\} \\ &\cup \{((s, \mathrm{ac}_C), \mathrm{tr}(s, (\mathrm{ac}_C, \mathrm{ac}_{\mathrm{Ag}\setminus C}))) \mid \mathrm{ac}_{\mathrm{Ag}\setminus C} \in \mathrm{Ac}_{\mathrm{Ag}\setminus C}\}, \end{aligned}$$

权重函数 $w'\colon E \to \mathbb{Z}^{|C|}$ 由以下两种模式定义：

$$\begin{aligned} w'_i(s, (s, \mathrm{ac}_C)) &= w_i(s); \\ w'_i((s, \mathrm{ac}_C), \mathrm{tr}(s, (\mathrm{ac}_C, \mathrm{ac}_{\mathrm{Ag}\setminus C}))) &= w_i(s). \end{aligned}$$

最后，我们令 $z^{|C|}$ 为 $\circledR z_c$。

非正式地说，该博弈中的两个玩家是 C 和 $\mathrm{Ag} \setminus C$，向量权重函数由聚合 C 的权重函数得到，阈值为 $®z_C$。现在，如果在这个博弈中，玩家1有获胜策略，那么存在某个策略 $®\sigma_C$，使得对于玩家2的所有策略 $®^{\sigma_{\mathrm{Ag}\setminus C}}$，$\rho(®\sigma_C, ®^{\sigma_{\mathrm{Ag}\setminus C}})$ 都是玩家1的获胜运行。但这意味着，对于所有 $i \in C$，收益$_i(\rho(®\sigma_C, ®^{\sigma_{\mathrm{Ag}\setminus C}})) \geq z_i$。然而，很容易验证这意味着 $®z_C$ 是 G中 C 的下界。反过来，如果玩家1没有获胜策略，那么对于所有策略 $®\sigma_C$，都存在某个策略 $®^{\sigma_{\mathrm{Ag}\setminus C}}$，使得 $\rho(®\sigma_C, ®^{\sigma_{\mathrm{Ag}\setminus C}})$ 不是获胜运行。这反过来意味着，对于某个 $j \in C$，收益 $_j(\rho(®\sigma_C, ®^{\sigma_{\mathrm{Ag}\setminus C}})) < z_j$，这意味着 $®z_C$ 不是 G中 C 的下界。另请注意，这种构造可以在多项式时间内完成，从而得到 co-NP 上界。

对于下界，我们采取相反方向，从多维平均收益阈值问题归约而来。假设我们想确定一个实例 G 是否属于多维平均收益阈值问题。那么，我们构造一个并发平均收益博弈 G'，其中有 $k+1$ 个玩家，且 G' 的状态与 G的状态完全一致。在这个博弈中，只有 1^{st} 和 $(k+1)^{\mathrm{th}}$ 玩家会对博弈的策略性质产生影响。如果博弈处于 V_1中的某个状态，玩家一可以决定下一步转移到哪个状态。否则，如果博弈处于 V_2中的某个状态，那么 $(k+1)^{\mathrm{th}}$玩家将采取行动。注意，我们只允许采取符合 G中允许的行动的移动。

现在，在 G' 中，前 k 个玩家的权重函数分别对应于 G 中玩家1的 k 个权重函数。最后一个玩家可以使用任意权重函数。做好这些准备后，我们要问：z^k 是否是 $\{1, \ldots, k\}$ 的下界。按照与上述类似的推理，很容易验证，当且仅当 G 是多维平均收益阈值问题的一个实例时，z^k 才是所构造的并发平均收益博弈中 $\{1, \ldots, k\}$ 的下界。此外，这一归约可以在多项式时间内完成，因此我们可以得出结论：下界

是 co-NP 完全的。 □

本节一个值得注意的遗漏是，我们没有给出一般情形下 E-Core 问题复杂度的任何界。上界之所以仍难以确定，原因之一在于：在多重平均收益博弈中，玩家2可以采用无记忆策略实现最优行动，而玩家1可能需要无限记忆 [21, 48]。鉴于我们所研究的并发多智能体设定中的核心与多重平均收益博弈中的获胜策略密切相关，这给 E-Core 问题带来了计算方面的顾虑。此外，在 [47], 中，作者研究了多重平均收益博弈的帕累托前沿，并提出了一种算法，用于构造给定博弈的可实现值表示，但该过程所需时间随输入规模呈指数增长。同一篇论文还证明了 多面体值问题 是 Σ_2^p -完全的。这两个问题似乎都与核心密切相关，我们希望未来能够利用这些结果，加深对 E-Core 的认识。

5. 讨论与相关工作

本节中，我们总结了一些结论，简要讨论了相关工作和核心相关问题，并推测如何使用模型检测技术实现我们的方法。

5.1. 合作博弈中的联盟形成

合作博弈论文献中已有关于具有外部性的联盟形成问题的研究 [16, 39, 40]。这些研究探讨了核心的几种可能定义。例如， α-核心采用悲观的方法，要求偏离联盟 S 的所有成员无论其他可能形成的联盟如何行动，都能从偏离中获益。我们首先提出的

核心定义沿用了这一方法。相比之下， β-核心采用乐观的方法，要求偏离联盟 S 的成员至少能在其余玩家的一种联盟形成方案中获益。此外， γ-核心 [49, 50]假定偏离后形成的联盟结构将包含偏离联盟 S，而其余联盟结构由所有单人联盟构成。现在， S 的“价值”被定义为其在纳什均衡中的收益；该均衡由 S 与其他单独行动的玩家共同构成，其中 S 的成员针对其余玩家各自的最佳响应策略，采取联合最佳响应策略。众所周知，α-和 β-特征函数会产生较大的核心 [51]，这与我们的观察一致：按照我们的第一个定义，核心永不为空。联盟形成在多智能体系统中十分重要 [52]。然而，尽管具有外部性的联盟形成是多智能体系统中广受研究的问题，针对具有外部性的多智能体联盟形成中的稳定性概念，相关研究却不多 [53]。相反，在人工智能和多智能体系统领域，大多数研究都聚焦于结构本身的形成 [54]。请注意，从合作博弈论的角度来看，我们的博弈属于非可转移效用博弈（即 NTU 博弈） [15, p.71]。对于某些类型的 NTU 博弈，还存在其他定义核心的方法，例如将其转换为常规的（可转移效用）博弈（参见例如 [55, p.238]）。这种方法或许也可以用于我们的情形。

5.2. 并发博弈的理性验证

在假设此类系统中的智能体采取理性行为的前提下，对多智能体系统时序逻辑性质进行形式化验证的研究至今已有近十年；例如，参见 [2–4, 43, 56]。然而，据我们所知，这些研究均考虑了非合作场景，

即使允许联盟权力，例如强纳什均衡中的情形。尽管如此，即便在此类非合作情境中，检查博弈的稳定结果是否满足某项时序逻辑性质，其复杂度仍是2ExpTime 完全问题；即使对于只能组成平凡联盟的双人零和博弈，情况也是如此。值得肯定的是，在理性验证框架中，合作博弈似乎具有更好的模型论性质：对于我们对核心的第一个定义（对应于合作博弈文献中的 α-核心这一概念），始终可以保证存在一个核心中的见证（因为核心永不为空）；该见证在互模拟系统之间保持不变，并且可以在实践中通过 ATL* 模型检测技术进行检查。这些技术受到诸如 MCMAS [10], 这样的自动化形式化验证工具的支持。MCMAS 能够执行 ATL*模型检测（通过实现涵盖 ATL* [57] 的 SL[1G],），并允许使用 ISPL（解释系统编程语言，即 MCMAS 的建模语言，它基于解释系统形式主义 [58]）对并发博弈结构进行规范。

致谢。Sarit Kraus部分获得了以色列科学基金会（资助编号：1958/20）的支持。Thomas Steeples谨此感谢EPSRC自主智能机器与系统博士培训中心（EP/L015897/1）以及Ian Palmer纪念奖学金提供的支持。Michael Wooldridge获得了UKRI图灵人工智能世界领先研究员奖学金（资助编号：EP/W002949/1）的支持。Sarit Kraus和Michael Wooldridge还获得了欧盟TAILOR项目（资助编号：992215）的支持。

参考文献

[1] J. Gutierrez、S. Kraus、M. J. Wooldridge，《合作并发博弈》，载于：AAMAS，IFAAMAS，2019年，第1198–1206页。

[2] J. Gutierrez、P. Harrenstein、M. Wooldridge，《迭代布尔博弈》，《信息与计算》242（2015年），53–79页。doi:10.1016/j.ic.2015.03.011。

[3] J. Gutierrez、P. Harrenstein、M. Wooldridge，从模型检测到均衡检测：用于理性验证的反应式模块，《人工智能》248 (2017) 123–157。doi:10.1016/j.artint.2017.04.003。

[4]M. Wooldridge、J. Gutierrez、P. Harrenstein、E. Marchioni、G. Perelli、A. Toumi，理性验证：从模型检测到均衡检测，载于：AAAI 会议论文集，AAAI出版社，2016，第 4184–4191 页。

[5] A. Abate、J. Gutierrez、L. Hammond、P. Harrenstein、M. Kwiatkowska、M.Na-jib、G. Perelli、T. Steeples、M. J. Wooldridge，理性验证：多智能体系统的博弈论验证，《应用智能》51 (9) (2021) 6569–6584。doi:10.1007/s10489-021-02658-y。

[6] R. Alur、T. A. Henzinger、O. Kupferman，《交替时序逻辑》，ACM期刊 49 (5) (2002) 672–713。doi:10.1145/585265.585270。

[7]E. A. Emerson，《时序逻辑与模态逻辑》，载于：《形式模型与语义》，Elsevier，1990，第995–1072页。doi:10.1016/b978-0-444-88074-1.50021-4。

[8] E. M. Clarke、O. Grumberg、D. A. Peled，《模型检测》，麻省理工学院出版社，2002年。

[9] J. Gutierrez、M. Najib、G. Perelli、M. Wooldridge，《EVE：用于时序均衡分析的工具》，载于：ATVA，《计算机科学讲义》第11138卷，施普林格，2018年，第551–557页。

[10] A. Lomuscio、H. Qu、F. Raimondi，MCMAS：用于验证多智能体系统的开源模型检测器，《国际软件工具技术与转化杂志》19 (1) (2017) 9–30。doi:10.1007/s10009-015-0378-x。

[11] R. Brenguier，PRALINE：一种用于计算并发博弈中纳什均衡的工具，载于：CAV，《计算机科学讲义》第 8044 卷，施普林格，2013，第 890–895 页。

[12] M. Kwiatkowska、G. Norman、D. Parker、G. Santos，《{PRISM}-games 3.0：具有并发性、均衡和时间的随机博弈验证》，载于：S. K. Lahiri、C. Wang（编），《计算机辅助验证——第32届国际会议》，CAV 2020，美国加利福尼亚州洛杉矶，2020年7月21日至24日，论文集第二部分，《计算机科学讲义》第12225卷，施普林格，2020年，第475–487页。doi:10.1007/978-3-030-53291-8_25。

[13] M. J. Osborne、A. Rubinstein，《博弈论教程》，麻省理工学院出版社，1994。

[14] M. Maschler、E. Solan、S. Zamir，《博弈论》，剑桥大学出版社，2013。doi:10.1017/cbo9780511794216。

[15] G. Chalkiadakis、E. Elkind、M.Wooldridge，《合作博弈论的计算方面》，Morgan-Claypool，2011。

[16] M. Uyanık，关于非连续博弈的 α-核心的非空性：可转移效用与不可转移效用，《经济理论杂志》158（2015）213–231。doi:10.1016/j.jet.2015.04.005。

[17] J. Gutierrez、P. Harrenstein、G. Perelli、M. Wooldridge，纳什均衡与双模拟不变性，载于：CONCUR，LIPIcs 第 85 卷，Schloss Dagstuhl，2017，第 17:1–17:16 页。

[18] A. Ehrenfeucht、J. Mycielski，平均收益博弈的位置策略，《国际博弈论杂志》8（2）（1979）109–113。doi:10.1007/bf01768705。

[19] E. A. Emerson、C. S. Jutla，《树自动机、μ演算与确定性》（扩展摘要），载于：第32届计算机科学基础年度研讨会

，圣胡安，波多黎各，1991年10月1日至4日，IEEE计算机学会，1991年，第368–377页。doi:10.1109/SFCS.1991.185392。

[20] H. Gimbert、W. Zielonka，无需记忆即可最优游玩的博弈，载于：M. Abadi、L. de Alfaro（编），CONCUR 2005——并发理论，第16届国际会议，CONCUR 2005，2005年8月23日至26日，美国加利福尼亚州旧金山，会议论文集，《计算机科学讲义》第3653卷，施普林格，2005年，第428–442页。doi:10.1007/11539452_33。

[21] E. Kopczynski，无限博弈的半位置确定性，载于：M. Bugliesi、B. Preneel、V. Sassone、I. Wegener（编），自动机、语言与编程，第33届国际研讨会，ICALP 2006，意大利威尼斯，2006年7月10日至14日，会议论文集，第二部分，《计算机科学讲义》第4052卷，施普林格，2006年，第336–347页。doi:10.1007/11787006_29。

[22] A. Pnueli，《程序的时序逻辑》，载于：第 18 届年度计算机科学基础研讨会（sfcs 1977），IEEE，1977 年，第 46–57 页。doi:10.1109/sfcs.1977.32。

[23] F. Mogavero、A. Murano、G. Perelli、M. Y. Vardi，《策略推理》，《ACM 计算逻辑汇刊》15（4）（2014）：1–47。doi:10.1145/2631917。

[24] C. Baier、J.-P. Katoen，《模型检测原理》，麻省理工学院出版社，2008 年。

[25] R. M. Keller，《并行程序的形式化验证》，《ACM 通讯》19（7）（1976）：371–384，出版地：美国纽约州纽约市；出版商：计算机协会。doi:10.1145/360248.360251。

[26] S. A. Kripke，《正规模态命题演算中的模态逻辑语义分析》

，《数学逻辑与数学基础杂志》9 (5-6) (1963) 67–96。doi:10.1002/malq.19630090502。

[27] E. A. Emerson、J. Y. Halpern，《“有时”与“并非从不”再探》，ACM期刊 33 (1) (1986) 151–178。doi:10.1145/4904.4999。

[28] A. P. Sistla、E. M. Clarke，《命题线性时序逻辑的复杂度》，ACM期刊 32 (3) (1985) 733–749。doi:10.1145/3828.3837。

[29] A. Pnueli、R. Rosner，《关于反应式模块的综合》，载于：第十六届ACM 程序设计语言原理年度研讨会会议记录，美国得克萨斯州奥斯汀，1989年1月11日至13日，ACM出版社，1989年，第179–190页。doi:10.1145/75277.75293。

[30] D. Fisman、O. Kupferman、Y. Lustig，《理性综合》，载于：J. Esparza、R. Majumdar（编），《系统构造与分析的工具和算法》，第16届国际会议，TACAS 2010，作为软件理论与实践联合欧洲会议 ETAPS 2010 的一部分举行，塞浦路斯帕福斯，2010年3月20日至28日。会议论文集，第6015卷《计算机科学讲义》，施普林格，2010年，第190–204页。doi:10.1007/978-3-642-12002-2_16。

[31] R. J. Aumann，《一般合作型 n人博弈中的可接受点》，载于：《博弈论贡献》第 IV 卷，《数学研究年刊》第 40 号，普林斯顿大学出版社，新泽西州普林斯顿，1959 年，第 287–324 页。

[32] R. J. Aumann，《完全信息博弈中的可接受点》，《太平洋数学杂志》10 (1960) 381–417。

[33]B. D. Bernheim、B. Peleg、M. D. Whin- ston，《联盟稳健纳什均衡i.概念》，《经济理论杂志》

42（1）（1987）1–12。doi:https://doi.org/10.1016/0022-0531(87)90099-8。URL https://www.sciencedirect.com/science/article/pii/0022053187900998

[34]B. D. Bernheim, M. D. Whinston，联盟稳健纳什均衡 II：应用，Jour- nal of Economic Theory 42 (1) (1987) 13–29. doi:https://doi.org/10.1016/0022-0531(87)90100-1. URL https://www.sciencedirect.com/science/article/pii/0022053187901001

[35] M. Hennessy, R. Milner，非确定性与并发的代数定律，ACM期刊 32 (1) (1985) 137–161. doi:10.1145/2455.2460.

[36]R. Alur, S. La Torre，LTL片段的确定性生成器与博弈，ACM Trans. Comput. Log. 5 (1) (2004) 1–25. doi:10.1145/963927.963928.

[37] O. Kupferman, M. Y. Vardi, P. Wolper，分支时间模型检测的自动机理论方法，ACM期刊 47 (2) (2000) 312–360。doi:10.1145/333979.333987。

[38] R. Milner，《通信与并发》，Prentice-Hall, Inc.，美国，1989。

[39] S.-S. Yi，具有外部性的稳定联盟结构，Game. Econ. Behav. 20 (2) (1997) 201–237。doi:10.1006/game.1997.0567。

[40] M. Finus, B. Rundshagen，正外部性 {NTU}-联盟博弈中核心稳定性的非合作基础，SSRNElectronic Journal (2003)。doi:10.2139/ssrn.406400。

[41] M. Mavronicolas、B. Monien、K. W. Wagner，《加权布尔公式博弈》，载于：X. Deng、F. C. Graham（编），《互联网与网络经济学》，

第三届国际研讨会，WINE 2007，美国加利福尼亚州圣迭戈，2007年12月12日至14日，会议论文集，《计算机科学讲义》第4858卷，施普林格，2007年，第469–481页。doi:10.1007/978-3-540-77105-0_51。

[42] S. Almagor、O. Kupferman、G. Perelli，《定量目标博弈中可控纳什均衡的综合》，载于：第二十七届国际人工智能联合会议会议论文集，IJCAI'18，国际人工智能联合会，2018年，第 35–41 页，会议地点：瑞典斯德哥尔摩。doi:10.24963/ijcai.2018/5。

[43]O. Kupferman、G. Perelli、M. Y. Vardi，《具有理性环境的综合》，《数学与人工智能年刊》78 (1) (2016) 3–20。
doi:10.1007/s10472-016-9508-8。

[44]U. Zwick、M. Paterson，《图上平均收益博弈的复杂度》，《理论计算机科学》158 (1-2) (1996年) 343–359。
doi:10.1016/0304-3975(95)00188-3。

[45] M. Ummels, D. Wojtczak，《随机多玩家博弈中纳什均衡的复杂度》，《逻辑方法与计算机科学》7 (3)（2011年9月）。
doi:10.2168/lmcs-7(3:20)2011.

[46] R. M. Karp，《有向图中最小环平均值的刻画》，Universitext。23 (3)（1978年）309–311。doi:10.1016/0012-365x(78)90011-0.

[47]R. Brenguier、J.-F. Raskin，多维平均收益博弈的帕累托曲线，载于：D. Kroening、C. S. Pasareanu（编），计算机辅助验证—— 27^{th} 国际会议，CAV 2015，美国加利福尼亚州旧金山，2015年7月18日至24日，会议论文集，第二部分，第9207卷

《计算机科学讲义》，施普林格，2015年，第251–267页。
doi:10.1007/978-3-319-21668-3_15。

[48] Y. Velner、K. Chatterjee、L. Doyen、T. A. Henzinger、A. Rabinovich、J.-F. Raskin，《多重平均收益博弈与多重能量博弈的复杂度》，《信息与计算》241（2015）177–196。doi:10.1016/j.ic.2015.03.001。

[49] P. Chander 等，《具有正外部性的博弈的核心》，CORE DP 2010/4（2010）。

[50] P. Chander，《伽马核与联盟形成》，《国际博弈论杂志》35（4）（2007）539–556。doi:10.1007/s00182-006-0067-9。

[51] D. Ray、R. Vohra，《均衡约束协议》，《经济理论杂志》73（1）（1997）30–78。doi:10.1006/jeth.1996.2236。

[52]O. Shehory、S. Kraus，通过智能体联盟形成进行任务分配的方法，《人工智能》101 (1-2) (1998) 165–200。
doi:10.1016/s0004-3702(98)00045-9。

[53] D. Mutzari、J. Gan、S. Kraus，多防御者安全博弈中的联盟形成，《人工智能》会议论文集 35 (6) (2021) 5603–5610。

[54]T. Rahwan、T. Michalak、M. Wooldridge、N. R. Jennings，具有正外部性或负外部性的多智能体系统中的任意时间联盟结构生成，《人工智能》186 (2012) 95–122。doi:10.1016/j.artint.2012.03.007。

[55]P. Borm、Y. Ju、D. Wettstein，《具有联盟外部性的博弈中的理性谈判》，《经济理论杂志》157（2015）236–254。
doi:10.1016/j.jet.2015.01.011。

[56] O. Kupferman、Y. Lustig，《什么触发了一种行为？》，载于：计算机辅助设计形式化方法（FMCAD'07），《计算机科学讲义》第 6015 卷，IEEE，2007 年，第 190–204 页。doi:10.1109/famcad.2007.41。

[57] P. Cermák, A. Lomuscio, A. Murano，《针对单目标策略逻辑规范验证和综合多智能体系统》，载于：B. Bonet、S. Koenig（编），《第二十九届 AAAI 人工智能会议论文集》，2015年1月25日至30日，美国得克萨斯州奥斯汀，AAAI出版社，2015，页码 2038–2044。

[58] R. Fagin、J. Y. Halpern、Y. Moses、M. Y. Vardi，《关于知识的推理》，麻省理工学院出版社，1995年。doi:10.7551/mitpress/5803.001.0001。

附录A. 证明例3具有空强核心

我们先前声称存在强核心为空的四人博弈，并在例3中给出了这样一个博弈的示例。下面我们给出一个表格，列出每种动作组合下某个玩家集合的强有利偏离。记号 $\{2 \mapsto 0, 4 \mapsto 1\}$ 表示联盟 $\{2,4\}$ 有一个强有利偏离，其中玩家2选择动作0，玩家4选择动作1。请注意，联盟的强有利偏离并不要求所有玩家都改变其相对于现状的动作。

动作组合	获胜玩家集合	强有利偏离
0000	{1，2}	$\{3 \mapsto 1\}$
0001	{1，2}	$\{3 \mapsto 1\}$
0010	{1，3}	$\{2 \mapsto 1,\ \ 4 \mapsto 1\}$
0011	{1，3}	$\{2 \mapsto 1,\ \ 4 \mapsto 1\}$
0100	{1，2}	$\{3 \mapsto 1\}$
0101	{1，4}	$\{2 \mapsto 0\}$
0110	{1，3}	$\{2 \mapsto 1,\ \ 4 \mapsto 1\}$
0111	{2，4}	$\{1 \mapsto 1\}$
1000	{3，4}	$\{1 \mapsto 0\}$
1001	{2，3}	$\{1 \mapsto 0\}$
1010	{2，3}	$\{1 \mapsto 0\}$
1011	{2，3}	$\{1 \mapsto 0\}$
1100	{1，2}	$\{3 \mapsto 1\}$
1101	{1，4}	$\{2 \mapsto 0\}$
1110	{2，3}	$\{1 \mapsto 0\}$
1111	{1，4}	$\{2 \mapsto 0\}$

该表说明所给博弈的强核心为空。验证偏离玩家在所有使原获胜玩家获胜的反制响应下都能获胜，这项工作并不难，但有些繁琐。

附录 B. 定理 8 的证明

我们将通过假设存在一个强核心为空的三人博弈G，再证明两个相互矛盾的事实来证明该定理，这便可立即推出结论。这两个论断是：

1. 存在某种策略 ®σ，在 G 上恰好对 γ_1、 γ_2、 γ_3 中的两个进行建模。

2. 不存在一种策略 ®σ 在 G 上恰好模拟 γ_1、 γ_2、 γ_3 中的两个。

我们在全文中使用以下引理：

引理 5. 设 G 是一个强核心为空的三人博弈。那么，不存在满足三位玩家所有目标的策略组合。

证明。 如果 G 存在一种满足所有玩家目标的策略，那么该策略就属于强核心，这就产生了矛盾。□

我们需要证明的第一个结论，在两个结论中更容易证明：

引理 6. 设 G 是一个强核心为空的三人博弈。那么，存在某种策略 ¯σ 在 G 上恰好模拟 γ_1、 γ_2、 γ_3 中的两个。

证明。 根据主张 1，我们知道， G 上的每个策略必须恰好满足零个、一个或两个目标，而不能满足三个。反设 G 只有满足 γ_1、 γ_2、 γ_3中零个或一个目标的策略。首先注意，如果 G 只有满足 γ_1、 γ_2、 γ_3中零个目标的策略，那么这些策略都不可能存在强有利偏离，因此都属于核心$^+$。因此， G 必须有一个恰好满足 γ_1、 γ_2、 γ_3中一个目标的策略。将这样的策略记为 ®σ = (σ_1、 σ_2、 σ_3)，不失一般性，我们可以假设它只满足 γ_1。现在，由于根据假设，®σ 不属于核心$^+$(G)，因此必定存在某个联盟 $C \subseteq$ {2、3}，能够从该策略进行强有利偏离。由于强有利偏离中，联盟的每个成员都会获胜，并且我们假设G 的策略满足的目标少于两个，因此该联盟必为单人联盟。同样，不失一般性，我们可以假设偏离者是玩家2，其偏离为 σ_2'。根据假设， σ_1、 σ_2'、 σ_3 只满足 γ_2。

但同样，(σ_1， σ_2'， σ_3) 必须存在针对玩家1或玩家3的强有利偏离，使其策略分别仅满足 γ_1 或仅满足 γ_3。如果前一种情况成立，

那么这将与以下事实矛盾： σ_2' 是从 ®σ 出发的强有利偏离。因此，我们必然处于后一种情况，并且存在针对玩家3的强有利偏离 σ_3' 。

新策略（σ_1， σ_2'， σ_3'）必须再次存在一个强有利偏离。如果它存在一个使 γ_1 仅成立的、针对 1 的强有利偏离 σ_1' ，那么联合策略（σ_1'， σ_3'）将是对 σ_2' 的惩罚。否则，如果存在一个使 γ_2 仅成立的、针对 2 的强有利偏离 σ_2''，那么（σ_1， σ_2''）将构成对 σ_3' 的惩罚。无论哪种情况，都会得出矛盾，从而完成证明。□

在证明第二个结论之前，我们还需要用到另外两个技术引理。第一个如下：

引理 7。 设 $\{a, b, c\}$ 是 $\{1,2,3\}$ 的一个排列。那么，不存在 G 的策略 $(\sigma_a, \sigma_b, \sigma_c)$、$(\sigma_a, \sigma_b, \sigma_c')$、$(\sigma_a', \sigma_b, \sigma_c')$ 使得以下条件全部成立：

- $(\sigma_a, \sigma_b, \sigma_c)$ 满足 γ_a 和 γ_b ，但不满足 γ_c
- $(\sigma_a, \sigma_b, \sigma\prime c)$ 仅满足 γ_c
- $(\sigma\prime a, \sigma_b, \sigma\prime c)$ 仅满足 γ_a
- $\sigma' c$ 是 c 从 $(\sigma_a、\ \sigma_b、\ \sigma_c)$ 变为 $(\sigma_a、\ \sigma_b、\ \sigma' c)$的有利偏离
- $\sigma' a$ 是 a 从 $(\sigma_a、\ \sigma_b、\ \sigma' c)$ 变为 $(\sigma' a、\ \sigma_b、\ \sigma' c)$的有利偏离

证明。 为导出矛盾，我们假设引理的陈述不成立，且所有陈述均成立。不失一般性，我们还可以假设$a = 1$、 $b = 2$和 $c = 3$。

现在，第三个策略 $(\sigma_1', \sigma_2, \sigma_3')$ 必须对 2、3 或两者都有强有利偏离。如果它存在一个偏离 (σ_2', σ_3'')（可能还包括 $\sigma_2' = \sigma_2$），转向一个策略

满足 γ_3 但不满足 γ_1，那么这就会是对 σ_1' 的惩罚。因此，新策略中可能的获胜者集合只有 {2}、{1,2} 和 {1,3}。我们依次考虑这些情况。

如果在新策略下只有玩家2获胜，那么强有利偏离必定是 σ_2' 仅由2进行，转向策略（σ_1'，σ_2'，σ_3'）。

同样，这个策略必须对1、3或两者都有有利偏离。设该偏离为（σ_1''，σ_3''），其中可能包含 $\sigma_1'' = \sigma_1'$ 或 $\sigma_3'' = \sigma_3'$，但不能两者兼有。考虑由此产生的策略（σ_1''，σ_2'，σ_3''）的以下情形：

- 如果只有 γ_1满足，或恰好 γ_1 和 γ_3 满足，那么（σ''_1，σ''_3）是对 σ'_2的惩罚；这与假设矛盾。

- 如果恰好 γ_1 和 γ_2 满足，那么由于3输掉了，我们必然有$\sigma''_3 = \sigma'_3$（该偏离仅由1进行）。在这种情况下，联合策略（σ''_1，σ'_2）可以将（σ_1，σ_2，σ'_3）转变为（σ''_1，σ'_2，σ''_3），并构成对 σ'_3的惩罚；这与假设矛盾。

- 如果只有 γ_3，或者只有 γ_2 和 γ_3 恰好得到满足，那么无论哪种情况，1 都会输，因此我们必须有 $\sigma''_1 = \sigma'_1$，而联合策略 (σ'_2, σ''_3) 可以将我们从 $(\sigma'_1, \sigma_2, \sigma'_3)$ 带到 $(\sigma''_1, \sigma'_2, \sigma''_3)$，并构成对 σ'_1 的惩罚；这就产生了矛盾。

现在假设新策略中的获胜者是 1 和 2。由于 1 在 $(\sigma_1', \sigma_2, \sigma_3')$ 中已经是获胜者，因此强有利偏离必定再次是 σ_2' 仅由 2 进行，转向策略 $(\sigma_1', \sigma_2', \sigma_3')$。

但是现在，我们可以看出，联合策略 (σ_1', σ_2') 是对 σ_3' 的惩罚（1 和 2 都仍是获胜者）；这就产生了矛盾。

最后，假设新策略中的获胜者是 1 和 3。由于 1 在（σ_1'，σ_2，σ_3'）中已经是获胜者，因此这次的强有利偏离必定仅由 3 做出，转向策略（σ_1'，σ_2，σ_3''）。

唯一的失败者 2 必定能从该策略出发进行强有利偏离 σ_2'。现在，请看：

- 如果在所得策略（σ_1'，σ_2'，σ_3''）中恰好满足 γ_1 和 γ_2，那么联合策略（σ_1'，σ_2'）将是对 σ_3'的惩罚。

- 如果在（σ_1'，σ_2'，σ_3''）中恰好满足 γ_2 和 γ_3，那么联合策略（σ_2'，σ_3''）将是对 σ_1'的惩罚。

因此，在（σ_1'，σ_2'，σ_3''）中，2必定是唯一的获胜者。

同样地，该策略必定允许1、3中的一方或双方进行强有利偏离。设该偏离为（σ_1''，σ_3'''），其中 $\sigma_1'' = \sigma_1'$ 或 $\sigma_3''' = \sigma_3''$可能成立，但不能同时成立。对于所得策略（σ_1''，σ_2'，σ_3'''），考虑以下情形：

- 如果只有 γ_1，或者恰好 γ_1 和 γ_2 成立，那么我们必然有$\sigma_3''' = \sigma_3''$，而联合策略（σ_1''，σ_2'）将是对 σ_3'' 的惩罚，这就产生了矛盾。

- 如果恰好 γ_1 和 γ_3 成立，那么（σ_1''，σ_3'''）将是对σ_2'的惩罚，这就产生了矛盾。

- 如果只有 γ_3，或者恰好 γ_2 和 γ_3 成立，那么我们必然有 $\sigma_1'' = \sigma_1'$，而联合策略（σ_2'，σ_3'''）将是对 σ_1'的惩罚，这就产生了矛盾。

因此，三种可能的获胜者集合都会导致矛盾，从而得出我们的结论。□接下来给出两个技术引理中的第二个：

引理 8。令 $\{a, b, c\} = \{1、2、3\}$ 为集合 Ag。不存在策略$(\sigma_a, \sigma_b, \sigma_c)$、$(\sigma_a, \sigma_b, \sigma_c')$ 属于 G，且满足以下所有条件：

- $(\sigma_a, \sigma_b, \sigma_c)$ *satisfies* γ_a *and* γ_b *but not* γ_c

- $(\sigma_a, \sigma_b, \sigma\prime_c)$ 仅满足 γ_c

- σ'_c 是对 c 从$(\sigma_a,\ \sigma_b,\ \sigma_c)$ 到$(\sigma_a,\ \sigma_b,\ \sigma'_c)$的强有利偏离

证明。为得出矛盾，反设结论不成立。不失一般性，我们可以再次假设 $a = 1$、$b = 2$ 和 $c = 3$。

第二个策略 $(\sigma_a,\ \sigma_b,\ \sigma'_c)$ 必定存在对 1、2 或两者都有利的强有利偏离。然而请注意，偏离到仅满足 γ_1 或仅满足 γ_2 的策略会与引理 7 矛盾。因此，新策略必须恰好有两个获胜者。此外，如果新策略中的获胜者恰好是 1 和 2，那么该偏离将为 σ'_3 带来惩罚。因此，不失一般性，我们可以假设 $(\sigma_a,\ \sigma_b,\ \sigma'_c)$ 存在一个强有利偏离 σ'_1，使 1 偏离到恰好满足 γ_1 和 γ_3 的策略。

现在，$(\sigma'_1, \sigma_2, \sigma'_3)$ 中唯一的失败者 2 必须存在一个强有利偏离σ'_2。然而，如果在 $(\sigma'_1, \sigma'_2, \sigma'_3)$ 中恰好满足 γ_1 和 γ_2，那么 (σ'_1, σ'_2) 就会对 σ'_3 构成一种惩罚。如果恰好满足 γ_2 和 γ_3，那么 (σ'_2, σ'_3) 就会对 σ'_1 构成一种惩罚。因此，2 必须是唯一的获胜者。

再一次，从 $(\sigma'_1, \sigma'_2, \sigma'_3)$ 偏离到仅满足 γ_1 或仅满足 γ_3 的策略所构成的强有利偏离，会与引理 7 矛盾。因此，1、3 或两者必定存在一个强有利偏离，转向恰有两名获胜者的策略。现在：

- 如果获胜者是 1 和 3，那么该偏离将对$\sigma\prime_2$构成一种惩罚。

- 如果获胜者是 1 和 2，那么偏离必定是仅由 1 采取的 $\sigma\prime\prime_1$（因为 3 输了）。但是，联合策略（$\sigma\prime\prime_1$, $\sigma\prime_2$）将对初始的 $\sigma\prime_3$构成惩罚。

- 如果获胜者是 2 和 3，那么偏离必定是仅由 3 采取的 σ''_3 。联合策略（σ'_2, σ''_3 ）将对 σ'_1构成惩罚。

这就穷尽了所有情况，而每种情况都会导致矛盾，并完成了引理 8 的证明。□

有了这两个技术引理，我们就可以证明所述主张的第二部分了。

引理 9. 设 G 是一个强核心为空的三人博弈。那么，在 ®σ 上不存在一个策略组合 G 能够恰好刻画 γ_1、 γ_2、 γ_3中的两个。证明。 为导出矛盾，假设结论不成立。不失一般性，设 (σ_1、 σ_2、 σ_3) 是一个恰好满足 γ_1 和 γ_2 的策略。它必然存在一个强有利偏离 σ'_3，使唯一的失败者玩家 3 获利。此外，根据引理 8，他不可能在新策略中独自获胜。因此，不失一般性，我们可以假设 1 和 3 是 (σ_1、 σ_2、 σ'_3) 中的获胜者。

同样，根据引理 8，新策略必然存在一个强有利偏离σ'_2 ，由玩家 2 转向一种恰好有两名玩家获胜的策略。然而，如果 1 和 2 是获胜者，那么 (σ_1、σ'_2) 就会构成对 σ'_3 的惩罚。因此，在 (σ_1、 σ'_2、 σ'_3) 中必须恰好满足 γ_2 和 γ_3 。

但现在，再次根据引理 8，必定存在一个强有利偏离 σ'_1，使玩家 1 转向另一种策略，且恰有两名获胜者。然而，如果获胜者是 1 和 2，那么（σ'_1， σ'_2）就会构成对 σ'_3 的惩罚。否则，如果获胜者是 1 和 3，那么（σ'_1， σ'_3）就会构成对 σ'_2 的惩罚。至此，证明完毕。□

最后，将两个引理结合起来即可得到所需结果。

定理 13. 如果 G 是一个三人博弈，那么它具有非空的强核心。证明。 结果可由引理6和引理9结合得出。

□