arXiv:2301.10697v1 [cs.GT] 25 Jan 2023

具有前缀无关目标的并发博弈中的子博弈最优策略

Benjamin Bordais、Patricia Bouyer 和 St´ephane Le Roux

大学生 伊韦特河畔吉夫，法国巴黎萨克雷大学、法国国家科学研究中心、巴黎萨克雷高等师范学院、LMF，邮编 91190

摘要

我们研究有限图上具有前缀无关目标的双人胜负随机并发博弈。我们刻画了子博弈最优策略，并利用这一刻画证明了若干记忆转移结果：1）对于给定的（前缀无关）目标，如果每个具有子博弈几乎必胜策略 的博弈也具有位置型策略，那么每个具有子博弈最优 策略的博弈也具有位置型策略；2）假设该（前缀无关）目标具有中性色。如果每个具有子博弈几乎必胜策略的回合制 博弈也具有位置型策略，那么每个具有有限选择 （待定义概念）子博弈最优策略的博弈也具有位置型策略。

我们收集或设计了一些例子，以从多个方面说明我们的结果是紧的。我们还将结果应用于 B¨uchi、co-B¨uchi、奇偶和均值收益目标，从而得到更简洁的表述。

1 引言

有限图上的轮流制双人胜负（随机）博弈，在广义的模型检测领域得到了深入研究 [21, 1]。在各种设置中，这类博弈都具有良好的最优性性质。对本文而言，最重要的是， [15] 证明了有限轮流制随机博弈在前缀无关目标下满足以下结果：（1）每个博弈都有确定性最优策略；（2）从每个值为1的状态出发，都存在最优策略，即几乎必然获胜的策略；（3）如果从每个博弈的每个值为1的状态出发，都存在使用某个固定记忆量的最优策略，那么每个博弈都有一个使用相同记忆量的最优策略。这些结果具有以下两种一般形式之一：

- 在所有博弈中（从所有良好状态出发），都存在一种良好策略。
- 如果从所有博弈的良好状态 出发都存在一种良好策略，那么从所有状态出发也是如此。

与这些轮流制（随机）博弈相比，并发版本的建模能力更强：这在实践中非常有用，因为现实世界系统本质上是并发的 [17]。博弈在有限图上进行，规则如下：在每个玩家状态下，两位玩家分别随机且独立地从有限个动作中选择一个。这会产生一个自然状态，它随机抽取下一个玩家状态；随后，每位玩家再次选择一个动作，如此反复。每个玩家状态都标有一个颜色，而胜负取决于由（随机生成的）无限玩家状态序列所对应的无限颜色序列。遗憾的是，即使对于简单的胜利条件和简单的图结构，例如有限图，这些并发博弈通常也无法良好运行：

- 可达性目标：存在一个没有最优策略的博弈 [14];
- B¨uchi 目标：存在一个值为 1 的博弈，而所有有限记忆策略的值均为 0 [13];
- Co-B¨uchi 目标：尽管总是存在位置型 ε-最优策略， [9], 但存在一个拥有最优策略、却不存在有限记忆最优策略的博弈 [5];
- 奇偶 [13] 与均值收益 [11] 目标：存在一个博弈，其中有子博弈几乎必然获胜的策略，但所有有限记忆策略的值均为 0。

本文聚焦于并发随机有限博弈。因此，为了考虑上述差异，我们的结果的一般形式会更加复杂。这些结果将以如下形式给出一般性陈述：

- 每个拥有不错策略的博弈，也拥有更好的策略。
- 如果所有拥有不错策略的特殊博弈都有更好的策略，那么所有博也都是如此 *es.*

上述“良好”“更好”和“特殊”的强度调节构成了大部分难点。下面介绍我们关于具有前缀无关目标的有限双人胜负并发随机博弈的主要贡献：

1. 我们给出了子博弈最优策略的刻画；这类策略在每段历史之后都具有最优性（定理 1）：玩家 A 的策略是子博弈最优策略，当且仅当：1）它是局部最优的；2）对于每个玩家 B 的确定性策略，在每段历史之后，如果访问过的状态具有相同的正值，那么玩家 A 以概率 1 获胜。我们利用这一刻画证明了下文的所有结果。

2. 我们证明了从子博弈几乎必然获胜策略到子博弈最优策略的记忆转移结果：

   (a) 定理 2：如果每个具有子博弈几乎必然获胜策略的博弈也具有位置型策略，那么每个具有子博弈最优策略的博弈也具有位置型策略。(b) 推论 1：每个具有子博弈最优策略的 B¨uchi 或 co-B¨uchi 博弈都具有位置型策略。（而奇偶博弈可能需要无限记忆 [13]。）

   Not该传递结果**2a** 在附录中从位置型推广到有限记忆 .

3. 如果一种策略只使用有限多个动作分布，我们就称其具有有限选择。注意，有限记忆策略（分别地，确定性策略）显然具有有限选择。

   (a) 定理 4：在给定的博弈中，如果存在有限选择最优策略，那么就存在有限选择子博弈最优策略。(b) 定理 5：假设目标具有中性色。如果每个具有子博弈几乎必胜策略的轮流制博弈也具有位置型策略，那么每个具有有限选择子博弈最优策略的博弈也具有位置型策略。(c) 推论 2：每个具有有限记忆子博弈最优策略的奇偶博弈或平均收益博弈，也具有位置型策略。

注意，**3a** 和**3b** 在删去“有限选择”一词后便不成立 [5]。证明**3b** 运用了**3a**。**3b** 和**2a** 的风格（以及证明）相似，但正如所强调的，**3b**中的前提和结论都被弱化了。

相关工作。本文很大一部分内容致力于将以下文献中的结果推广到并发博弈： [15] 关于如何在轮流制博弈中将记忆从几乎必然获胜策略转移到最优策略。请注意， [15]中使用的证明技术有所不同，无法适用于我们更一般的设定。在他们的证明中，双方玩家都同意对自然状态的偏好，并按照这一偏好进行博弈。在我们的证明中，我们将图切分为价值区域（即具有相同值的状态集合），并证明在每个切片中采用几乎必然获胜策略就足够了；随后，我们将这些（部分）策略拼接起来，从而得到适用于整个图的子博弈最优策略。

切片技术已在并发博弈的背景下用于 [9]。作者聚焦于奇偶目标，并建立了从极限必胜策略到近乎最优策略的记忆转移结果。作为应用，他们证明，对于余 B¨奇目标，由于位置策略足以实现极限必胜，因此也足以实现近乎最优。他们的构造大量利用了奇偶目标的特有性质。

我们还提到 [7], ，其重点同样是具有前缀无关目标的并发博弈。特别是，作者证明了一个（非常有用的）结果：如果所有状态的值都为正，那么它们的值都为 1。（请注意，本文给出了这一结果的一个加强版本（定理 3），这也可视为对 [15]中所证明结果的改编。随后，这一结果被用于非零和博弈的另一种情境中。

最后， [2, 4, 5], 近期也开展了一些关于并发博弈的研究，其目标如下：两个玩家在玩家状态下的局部交互由二维表给出；这些表可以抽象为博弈形式，其中（输出）变量是局部交互中的事项（同一个变量可以标记多个事项）。这一系列研究旨在给出这些博弈形式的（内在）性质，从而确保在图博弈中使用它们时，最优策略存在。例如， [4], 提出了一种称为 RM 的博弈形式性质；该性质保证，如果图中仅使用 RM 博弈形式，那么对于每个可达性目标，玩家 A 总能拥有该目标的最优策略。此性质刻画了在可达性目标方面表现良好的博弈形式，因为任何不满足 RM 性质的博弈形式都可以嵌入某个（小型）图博弈，使得玩家 A 没有最优策略。这一研究方向与本文的目标确实大不相同。

论文结构。第2 节介绍符号，第3 节回顾博弈形式的概念，第4 节介绍我们的形式体系，第5 节给出子博弈最优性的充要条件，第6 节展示了并发博弈中从子博弈几乎必胜到子博弈最优的记忆转移，第7 节将上一节的结果推广到存在子博弈有限选择策略的情形。

2 预备知识

考虑一个非空集合 Q。我们分别用 Q^*、 Q^+ 和 Q^ω 表示由 Q 中的元素构成的有限序列、非空有限序列和无限序列。对于 $n \in \mathbb{N}$，我们用 Q^n （分别用 $Q^{\leq n}$）表示由 Q中的 n 个（分别是不超过 n 个）元素构成的序列集合。对于所有$\rho = q_1 \cdots q_n \in Q^n$和 $i \leq n$，我们用 ρ_i表示元素 $q_i \in Q$，用 $\rho_{\leq i} \in Q^i$表示有限序列 $q_1 \cdots q_i$。对于子集 $S \subseteq Q$，我们用 $Q^* \cdot S^\omega \subseteq Q^\omega$表示最终停留在其中的无限路径集合

，用 $(Q^* \cdot S)^\omega \subseteq Q^\omega$ 表示无限次访问集合 S的无限路径集合。

非空有限集 Q 上的离散概率分布 是一个函数 $\mu: Q \to [0,1]$，满足 $\sum_{x \in Q} \mu(x) = 1$。概率分布 $\mu: Q \to [0,1]$的支撑集 Supp(μ)是该分布中非零元素构成的集合：Supp(μ) $= \{q \in Q \mid \mu(q) \in (0, 1]\}$。$Q$ 上所有分布构成的集合记作 $\mathcal{D}(Q)$。

3 博弈形式

我们回顾博弈形式的定义——非正式地说，它是带有变量的二维表——以及标准形博弈的定义，即结果为介于 0 和 1 之间的值的博弈形式。

定义 1（博弈形式与标准形博弈）一个博弈形式 （简称 *GF*）是一个元组 $\mathcal{F} = \langle$动作$_A$,动作$_B$, O $\varrho\rangle$，其中动作$_A$ （分别为动作$_B$）是玩家A （分别为B）可用动作构成的非空有限集，O 是非空结果集，而 ϱ : 动作$_A$ $\times$动作$_B$ $\to$ O是一个函数，为每一对动作指定一个结果。当结果集O等于 $[0,1]$, 时，我们称 $\mathcal{F}$为一个标准形博弈。对于结果的赋值 $v \in [0,1]^O$，记号 $\langle \mathcal{F}, v\rangle$ 指代标准形博弈 $\langle$动作$_A$, 动作$_B$, $[0,1]$, $v \circ \varrho\rangle$。

我们使用博弈形式来表示两个玩家之间的交互。玩家 A（分别地，玩家 B）可用的策略是动作的凸组合，这些动作分别列于表格的各行（各列）。在标准形博弈中，玩家 A 试图使结果最大化，而玩家 B 则试图使其最小化。

定义 *2*（标准形博弈的结果）考虑一个标准形博弈 $\mathcal{F} =\langle Act_A、Act_B, [0,1], \varrho\rangle$。集合 $\mathcal{D}(Act_A)$（分别地，$\mathcal{D}(Act_B)$）是玩家 *A*（分别地，玩家 *B*）可用的策略集合。对于一对策略 $(\sigma_A, \sigma_B) \in \mathcal{D}(Act_A) \times \mathcal{D}(Act_B)$，策略 (σ_A, σ_B) 的结果 $out_{\mathcal{F}}(\sigma_A, \sigma_B)$（属于 $\mathcal{F}$）定义如下：

$$\mathsf{out}_{\mathcal{F}}(\sigma_A, \sigma_B) := \sum_{a \in \mathsf{Act}_A} \sum_{b \in \mathsf{Act}_B} \sigma_A(a) \cdot \sigma_B(b) \cdot \varrho(a,b) \in [0,1]$$

定义 *3*（标准形博弈的值与最优策略）考虑一个标准形博弈 $\mathcal{F} = \langle Act_A、Act_B, [0,1], \varrho\rangle$ 以及玩家 *A* 的一个策略 $\sigma_A \in \mathcal{D}(Act_A)$。策略 σ_A 的值记作 $val_{\mathcal{F}}(\sigma_A)$，等于：$val_{\mathcal{F}}(\sigma_A) := inf_{\sigma_B \in \mathcal{D}(Act_B)} out_{\mathcal{F}}(\sigma_A, \sigma_B)$，玩家 *B* 的情况类似，只需将 *inf* 换成 *sup*。当 $sup_{\sigma_A \in \mathcal{D}(Act_A)} val_{\mathcal{F}}(\sigma_A) = inf_{\sigma_B \in \mathcal{D}(Act_B)} val_{\mathcal{F}}(\sigma_B)$ 时，它定义了博弈 $\mathcal{F}$ 的值，记作 $val_{\mathcal{F}}$。

策略 $\sigma_A \in \mathcal{D}$(行动$_A$) 确保 值$_{\mathcal{F}} =$ 值$_{\mathcal{F}}(\sigma_A)$ 的策略称为 最优。玩家 A 的所有最优策略组成的集合记为 Opt$_A$ $(\mathcal{F}) \subseteq \mathcal{D}$(行动$_A$)，玩家 B的情形也类似。冯·诺依曼极小极大定理 *[22]* 保证了最优策略（对双方玩家而言）的存在。

下文中，标准形博弈中的策略称为 GF-策略，以免与并发（图）博弈中的策略混淆。

4 并发博弈与最优策略

4.1 并发博弈场与策略

我们给出在有限图上进行的并发博弈场的定义。

定义 *4*（有限随机并发博弈场）着色并发博弈场 $\mathcal{C}$ 是一个元组 $\langle Q, (A_q)_{q\in Q}, (B_q)_{q\in Q}, D\ \delta, dist, K, col\rangle$，其中 Q 是非空有限状态集；对于所有$q \in Q$，A_q（分别为 B_q）是玩家 *A*（分别为玩家 *B*）在状态 q下可用的非空有限动作集；*D* 是有限自然状态集；$\delta : \bigcup_{q\in Q}(\{q\} \times A_q \times B_q) \to D$ 是转移函数；$dist : D \to \mathcal{D}(Q)$ 是分布函数。此外，*K* 是非空有限颜色集，$col : Q \to K$ 是着色函数。

下文中，除非另有说明，博弈场 $\mathcal{C}$表示元组 $\langle Q, (A_q)_{q\in Q}, (B_q)_{q\in Q}, \mathrm{D}\ \delta, \mathrm{dist}, \mathrm{K}, \mathrm{col}\rangle$。并发博弈是在并发竞技场的基础上加入获胜条件得到的：该条件指定了对玩家 A 获胜（对玩家 B 失败）的无限路径集合。

定义 5（有限随机并发博弈）有限并发博弈是一个二元组 $\langle \mathcal{C}, W\rangle$，其中 $\mathcal{C}$ 是有限并发着色博弈场，且 $W \subseteq \mathsf{K}^\omega$ 是博雷尔集。集合 W 称为目标，因为它对应于对玩家A获胜的着色路径集合。

本文仅考虑一种特定类型的目标：前缀无关目标。非正式地说，这类目标是指满足以下性质的目标 W：无限路径 ρ 属于 W，当且仅当它的任意一个后缀都属于 W。更正式地说：

定义 6（前缀无关目标）对于非空有限颜色集K，并且$W \subseteq \mathsf{K}^\omega$, W 称为前缀无关（简称 *PI*），如果对于所有 $\rho \in \mathsf{K}^\omega$ 和 $i \geq 0$,$\rho \in W \Leftrightarrow \rho_{\geq i} \in W$。

下文中，我们将具有前缀无关目标的并发博弈称为 PI 并发博弈。我们的主要结果将针对任意前缀无关目标陈述，但我们会将其应用于一些特定目标，即奇偶目标，以及相关的特殊情况 B¨uchi 和 co-B¨uchi。

定义 *7*（奇偶、*B¨uchi*、*co-B¨uchi* 目标）设 $K \subset \mathbb{N}$ 是一个非空有限整数集。考虑一个以 *K* 为颜色集的并发竞技场 $\mathcal{C}$。对于一条无限路径 $\rho \in Q^\omega$，我们用 $col(\rho)_\infty \subseteq \mathbb{N}$ 表示在 ρ中无限次出现的颜色集合：$col(\rho)_\infty := \{n \in \mathbb{N} \mid \forall i \in \mathbb{N}, \exists j \geq i, col(\rho_j) = n\}$。于是，相对于颜色函数 *col* 的奇偶性目标是集合 $W^{\mathsf{Parity}}(col) := \{\rho \in Q^\omega \mid max\ col(\rho)_\infty$ 为偶数 $\}$。*B¨uchi*（分别为 *co-B¨uchi*）目标对应于满足 $K := \{1, 2\}$（分别为 $K := \{0, 1\}$）的奇偶性目标。

策略定义为函数：给定博弈的历史记录（即已出现的状态序列），函数会为玩家可用的动作分配一个分布。

定义 8（策略）考虑一个并发博弈 $\mathcal{C}$。玩家 A 的策略是一个函数 $\mathsf{s_A} : Q^+ \to \mathcal{D}(A)$，满足 $A := \bigcup_{q\in Q} A_q$，使得对于所有 $\rho = q_0 \cdots q_n \in Q^+$，都有$\mathsf{s_A}(\rho) \in \mathcal{D}(A_{q_n})$。我们用 $\mathsf{S}^{\mathsf{A}}_{\mathcal{C}}$ 表示博弈场 $\mathcal{C}$ 中玩家 A的所有策略的集合。玩家 B的定义类似。

我们希望定义给定两个策略（每位玩家各一个）时博弈的结果。首先，定义在给定状态下，给定两个 GF-策略时从状态 q 转移到另一状态 q' 的概率。

定义 9（概率转移）考虑一个并发竞技场 $\mathcal{C}$、一个状态 $q \in Q$ 以及两个策略 $(\sigma_\mathsf{A}, \sigma_\mathsf{B}) \in \mathcal{D}(A_q) \times \mathcal{D}(B_q)$。再考虑另一个状态 $q' \in Q$。当玩家在 *q* 中分别采取 σ_A 和 σ_B时，从 q 转移到 q' 的概率记为 $\mathbb{P}^{q,\ q'}(\sigma_\mathsf{A}, \sigma_\mathsf{B})$，其值等于：

$$\mathbb{P}^{q,q'}(\sigma_\mathsf{A}, \sigma_\mathsf{B}) = \sum_{a\in A_q} \sum_{b\in B_q} \sigma_\mathsf{A}(a) \cdot \sigma_\mathsf{B}(b) \cdot \mathsf{dist} \circ \delta(q, a, b)(q')$$

现在，我们来定义给定两个策略时，任意有限路径以及由此任意博雷尔集的发生概率。

定义 10（给定两个策略的概率分布）考虑一个并发博弈场 $\mathcal{C}$，以及$\mathsf{s_A}$、$\mathsf{s_B} \in \mathsf{S}_{\mathcal{C}}^{\mathsf{A}} \times$、$\mathsf{S}_{\mathcal{C}}^{\mathsf{B}}$，它们分别是玩家A 和B的任意两个策略。我们用$\mathbb{P}^{\mathsf{s_A},\ \mathsf{s^B}}:\ Q^+ \to \mathcal{D}(Q)$ 表示这样一个函数：给定已经观察到的状态序列，返回博弈场下一状态上的概率分布。也就是说，对于所有有限路径 $\pi = \pi_0 \ldots \pi_n \in Q^+$ 和 $q \in Q$，有：

$$\mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(\pi)[q] = \mathbb{P}^{\pi_n,q}(\mathsf{s_A}(\pi), \mathsf{s_B}(\pi))$$

那么，给定一对策略 (s_A, s_B)，从状态 $q_0 \in Q$ 出发的有限路径 $\pi = \pi_0 \cdots \pi_n \in Q^+$ 的发生概率等于 $\mathbb{P}^{\mathcal{C},q_0 s_A, s_B}(\pi) = \Pi_{i=0}^{n-1} \mathbb{P}^{s_A, s^B(\pi_{\leq i})}[\pi_{i+1}]$，当 $\pi_0 = q_0$ 成立时；否则为 0。对于任意有限路径 $\pi \in Q^*$，圆柱集 $Cyl(\pi)$ 的概率为 $\mathbb{P}^{\mathcal{C},q_0 s_A, s_B}[Cyl(\pi)] = \mathbb{P}^{s_A, s^B(\pi)}$。按照通常的方式，这就确定了任意博雷尔集的概率；我们用$\mathbb{P}^{\mathcal{C},q_0 s_A, s_B} : Borel(Q) \to [0,1]$ 表示相应的概率测度。注意，在 Q^ω 中，颜色序列属于博雷尔集 $W \subseteq K^\omega$ 的无限路径集合本身也是一个博雷尔集（它是连续函数对博雷尔集的原像）。

策略值和博弈值的后续定义如下。

定义 *11*（策略值和博弈值）设 $\mathcal{G} = \langle \mathcal{C}, W \rangle$ 是一个 *PI* 并发博弈，并考虑玩家 *A* 的一个策略 $s_A \in S_{\mathcal{C}}^{\mathsf{A}}$。函数 $\chi_{\mathcal{G}}[s_A] : Q \to [0,1]$ 给出策略 s_A 的值，其定义为：对所有 $q_0 \in Q$，都有 $\chi_{\mathcal{G}}[s_A](q_0) := inf_{s_B \in S_{\mathcal{C}}^{\mathsf{B}}} \mathbb{P}^{\mathcal{C},q_0 s_A, s_B}[W]$。函数 $\chi_{\mathcal{G}}[A] : Q \to [0,1]$ 给出玩家 *A* 的值，其定义为：对所有 $q_0 \in Q$，都有 $\chi_{\mathcal{G}}[A](q_0) := sup_{s_A \in S_{\mathcal{C}}^{\mathsf{A}}} \chi_{\mathcal{G}}[s_A](q_0)$。类似地定义函数 $\chi_{\mathcal{G}}[B] : Q \to [0,1]$，用于给出玩家 *B* 的博弈值，只需交换上确界和下确界。

根据 *Martin* 关于 *Blackwell* 博弈确定性的结果 *[19]*，对于所有并发博弈$\mathcal{G} = \langle \mathcal{C}, W \rangle$，两位玩家的值函数相等，由此定义值函数$\chi_{\mathcal{G}}$：$Q \to [0,1]$ 博弈的 $\chi_{\mathcal{G}} := \chi_{\mathcal{G}}[\mathsf{A}] = \chi_{\mathcal{G}}[\mathsf{B}]$。

我们定义价值区域：值相同的状态子集。

定义 12（价值区域）在 *PI* 并发博弈 $\mathcal{G}$ 中，$V_{\mathcal{G}}$ 指博弈中出现的值的集合：$V_{\mathcal{G}} := \{\chi_{\mathcal{G}}[q] \mid q \in Q\}$。此外，对于所有 $u \in V_{\mathcal{G}}$，$Q_u \subseteq Q$ 指其值相对于 $\chi_{\mathcal{G}}$ 为 u 的状态集合：$Q_u := \{q \in Q \mid \chi_{\mathcal{G}}(q) = u\}$。

在并发博弈中，每个状态都会出现博弈形式，用于描述玩家在该状态下的交互。此外，将每个状态映射到其在博弈中的值的赋值，可以通过凸组合提升为对自然状态的赋值。进而，这自然地诱导出一种方式，用于定义每个状态出现的标准形博弈。

定义 13（局部交互，赋值提升）在 *PI* 并发博弈 $\mathcal{G}$ 中，赋值 $\chi_{\mathcal{G}}$：$Q \to [0,1]$ 给出博弈的值，而提升 $\nu_{\mathcal{G}}$：$\mathsf{D} \to [0,1]$ 满足：对所有 $d \in \mathsf{D}$，都有 $\nu_{\mathcal{G}}(d) := \sum_{q \in Q} \chi_{\mathcal{G}}(q) \cdot$ 距离函数$(d)(q)$（回忆一下，距离函数：$\mathsf{D} \to \mathcal{D}(Q)$ 是分布函数）。

令 $q \in Q$。状态 q 处的局部交互是博弈形式 $\mathcal{F}_q = \langle A_q, B_q, \mathsf{D}\ \delta(q, \cdot, \cdot) \rangle$。因此，状态 q 处的标准形博弈为 $\mathcal{F}_q^{\mathsf{nf}} := \langle \mathcal{F}_q, \nu_{\mathcal{G}} \rangle$。

请注意，标准形博弈 $\mathcal{F}_q^{\mathsf{nf}}$的值与状态 q 的值相等。

命题 1 对于 *PI* 并发博弈 $\mathcal{G}$，对所有状态 $q \in Q$，都有 $\chi_{\mathcal{G}}(q) = \mathsf{out}^{\mathcal{F}_q^{\mathsf{nf}}}$。

4.2 关于策略的更多内容

在本小节中，我们定义几类后续将要讨论的策略。接下来这一节中，我们固定一个 PI 并发博弈 $\mathcal{G}$。首先，我们考虑最优策略，即能够实现博弈值的策略。我们还考虑正向最优策略，即从所有可能状态出发，其值都为正的策略。下面的定义 14 对此作出形式化定义。

定义 14（（正向）最优策略）玩家 A 的策略 $\mathsf{s}_A \in \mathsf{S}^A_C$ 从状态 $q \in Q$ 出发是（分别是正向）最优的，当且仅当 $\chi_{\mathcal{G}}(q) = \chi_{\mathcal{G}}[\mathsf{s}_A](q)$（分别是当 $\chi_{\mathcal{G}}(q) > 0 \Rightarrow \chi_{\mathcal{G}}[\mathsf{s}_A](q) > 0$）。如果该条件从所有状态 $q \in Q$ 出发都成立，则称其为（分别是正向）最优策略。

请注意，我们所采用的最优策略定义有时也称为一致最优性，因为它要求策略从博弈的每个状态出发都成立。然而，这一定义没有说明在观察到某些状态序列之后会发生什么。现在我们希望定义一种策略概念，使其在观察到任意有限状态序列后，从任何可能到达的位置出发都是最优的。这对应于子博弈最优策略。要定义这类策略，我们需要引入剩余策略的概念。

定义 15（剩余策略与子博弈最优策略）考虑玩家 A 的一个策略 s_A。对于所有有限序列 $\rho \in Q^+$，策略 s_A 的剩余策略$\mathsf{s}^\rho{}_A$ 是策略 $\mathsf{s}^\rho{}_A\colon Q^+ \to \mathcal{D}(A)$，满足对所有 $\pi \in Q^+$，均有 $\mathsf{s}^\rho{}_A(\pi) := \mathsf{s}_A(\rho \cdot \pi)$。

玩家 A 策略 s_A 是子博弈最优 的，当且仅当对于所有 $\rho = \rho' \cdot q \in Q^+$，剩余策略$\mathsf{s}^\rho{}_A$ 从 q 出发是最优的，也就是说 $\chi_{\mathcal{G}}[\mathsf{s}^\rho{}_A](q) = \chi_{\mathcal{G}}(q)$。

请注意，特别是，子博弈最优策略也是最优策略。如果存在此类策略，我们希望它们尽可能简单，例如希望它们是位置型的，也就是说，它们只依赖于博弈的当前状态。

此外，一旦固定了玩家 A 策略，就会得到一个（无限）马尔可夫决策过程。在这种情形下，ε-最优策略可以从确定性策略中选取（例如参见 [10, 定理 1]中的解释），其中确定性策略是指：在任意有限状态序列之后，都以概率 1 执行某个特定动作。这两个概念都在下面的定义16中作了形式化定义。

定义 16（位置型、确定性策略）玩家 A 策略s_A 是位置型的，当且仅当对于所有状态 $q \in Q$ 和路径 $\rho \in Q^+$，都有 $\mathsf{s}_A(\rho \cdot q) = \mathsf{s}_A(q)$。

玩家 B 策略 s_B 是确定性的，若对于所有有限序列 $\rho \cdot q \in Q^+$，都存在$b \in B_q$ 使得 $\mathsf{s}_B(\rho \cdot q)(b) = 1$。

5 子博弈最优性的充要条件

本节给出玩家A 策略为子博弈最优的充要条件对，并在定理1中予以正式表述。这里给出的论证与 [5], 第 4 节中的论证有些相似，该节讨论的是同一问题在位置策略情形下的限制版本。

第一个条件是局部条件：它规定了策略在博弈的每次局部交互中，如何在标准形博弈中表现。如命题1所述，在每个状态 q，标准形博弈 $\mathcal{F}^{\mathsf{nf}}_q$的博弈值等于状态 q 的值（由赋值 $\chi_{\mathcal{G}} \in [0,1]^Q$给出）。这意味着，对于所有以该状态$q$为终点的状态有限序列 $\rho \in Q^+$，GF 策略 $\mathsf{s}_A(\rho)$都需要在标准形博弈 $\mathcal{F}^{\mathsf{nf}}_q$中对剩余策略 $\mathsf{s}^\rho{}_A$ 达到最优，才能使策略从 q起达到最优。具有这一性质的策略称为局部最优策略。

这是子博弈最优性的必要条件。（然而，正如第 6 节所述，它既不是最优性的必要条件，也不是充分条件。）

定义 17（局部最优策略）考虑一个 *PI* 并发博弈 $\mathcal{G}$。玩家A 策略 s_A 是局部最优的，如果对于所有 $\rho = \rho' \cdot q \in Q^+$，GF-策略 $\mathsf{s}_\mathsf{A}(\rho)$ 在标准形博弈 $\mathcal{F}_q^{\mathsf{nf}}$ 中是最优的。也就是说——回想一下，$\nu_\mathcal{G} \in [0,1]^{\mathsf{D}}$ 会将赋值 $\chi_\mathcal{G} \in [0,1]^Q$提升到自然状态——对于所有 $b \in B_q$:

$$\chi_\mathcal{G}(q) \leq \mathsf{out}_{\mathcal{F}_q^{\mathsf{nf}}}(\mathsf{s}_\mathsf{A}(\rho), b) = \sum_{a \in A_q} \mathsf{s}_\mathsf{A}(\rho)(a) \cdot \nu_\mathcal{G} \circ \delta(q,a,b)$$

引理 1（证明见附录 B.1）在 *PI* 并发博弈中，子博弈最优策略是局部最优的。

请注意，这一点已针对位置策略在 [5]中证明。

一般而言，局部最优性并不能确保子博弈最优性。不过，对于所有玩家 B 的确定性策略，它确实能确保博弈几乎必然最终进入某个值区域，即对于某个 Q_u，存在某个 $u \in V_\mathcal{G}$。

引理 2（证明见附录 B.2）考虑一个 *PI* 并发博弈 $\mathcal{G}$ 以及玩家 A 局部最优策略 s_A。对于所有玩家B 的确定性策略，以几乎必然的概率，无限次出现的状态具有相同的值。也就是说:

$$\mathbb{P}^{\mathsf{s}_\mathsf{A},\mathsf{s}_\mathsf{B}}[\bigcup_{u \in V_\mathcal{G}} Q^* \cdot (Q_u)^\omega] = 1$$

证明（概要）首先，如果到达了一个值为 *1* 的状态（即 Q_1 中的状态），那么所有以正概率可达的状态的值也都是 *1*（即都属于 Q_1），因为策略 s_A 是局部最优的。现在令 $u \in V_\mathcal{G}$ 为 $V_\mathcal{G}$ 中小于 *1* 的最高值，并考虑这样一类无限路径：集合 Q_u 被无限次访问，但博弈并未最终停留在该集合中，即集合 $(Q^* \cdot (Q \setminus Q_u))^\omega \cap (Q^* \cdot Q_u)^\omega \subseteq Q^\omega$。由于策略 s_A 是局部最优的（且 $V_\mathcal{G}$ 是有限的），可以证明存在一个正概率 $p > 0$，使得已知 Q_u 被离开时，到达 Q_1 的条件概率至少为 p。因此，如果 Q_u 被无限次离开，那么几乎必然会访问集合 Q_1（并且此后不再离开）。由此可知，事件 $(Q^* \cdot (Q \setminus Q_u))^\omega \cap (Q^* \cdot Q_u)^\omega$ 的概率为 *0*。这意味着，几乎必然地，如果集合 Q_u 被无限次访问，那么从某个时刻起便不再离开它。然后可以用相同的论证处理 $V_\mathcal{G}$ 中小于 u 的最高值，依此类推。总之，对于所有 $u \in V_\mathcal{G}$，如果集合 Q_u 被无限次访问，那么几乎必然会从某个时刻起不再离开该集合。 □

局部最优性确保在每一步，到达各状态的期望值都不会变差（如果玩家 B 没有采取最优策略，甚至可能有所改善）。通过传递这一性质，我们得到：给定玩家 A 的局部最优策略以及玩家 B 的确定性策略，值 u 在 $V_\mathcal{G}$ 上的凸组合，以最终落入值区域 Q_u的概率为权重，从状态 q 出发时，至少等于其值 $\chi_\mathcal{G}(q)$。这一结论在下文的引理 3中给出。

引理 3（证明见附录 B.3）对于一个 *PI* 并发博弈 $\mathcal{G}$，玩家 A 的局部最优策略 s_A、玩家 B 的确定性策略 s_B 以及状态 $q \in Q$:

$$\chi_\mathcal{G}(q) \leq \sum_{u \in V_\mathcal{G}} u \cdot \mathbb{P}_q^{\mathsf{s}_\mathsf{A},\mathsf{s}_\mathsf{B}}[Q^* \cdot (Q_u)^\omega]$$

注意，如果玩家 B 采用子博弈最优策略，那么这个不等式就是等式。

证明（概要）首先，令 $\mathbb{P}^{s_A, s_q}{}_B$ 表示为 $\mathbb{P}$。可以通过归纳法证明，对于所有 $i \in \mathbb{N}^*$，性质 $\mathcal{P}(i)$：$\chi_{\mathcal{G}}(q) \leq \sum_{\pi \cdot q' \in q \cdot Q^i} \chi_{\mathcal{G}}(q') \cdot \mathbb{P}(\pi \cdot q') = \sum_{u \in V_{\mathcal{G}} \setminus \{0\}} u \cdot \mathbb{P}[q \cdot Q^{i-1} \cdot Q_u]$ 均成立。此外，由于根据引理 2，该博弈几乎必然会进入某个值区域，因此可以证明，当 n 足够大时，n 步后处于 Q_u 的概率（即 $\mathbb{P}[q \cdot Q^{n-1} \cdot Q_u]$）与最终进入 Q_u 的概率（即 $\mathbb{P}[Q^* \cdot (Q_u)^\omega]$）可以任意接近。随后，我们可以应用 $\mathcal{P}(n)$，得到所需的不等式。 □

回想一下，我们考虑一对条件来刻画策略是否为子博弈最优。第一个条件是局部最优性。总而言之，我们已经看到，策略局部最优这一事实保证：从任意状态 q出发，博弈结束时所在价值区域的期望值至少为 $\chi_{\mathcal{G}}(q)$。然而，局部最优性并不能保证在博弈结束于某个特定价值区域的条件下，W 的概率是多少，附录 C.1中的例子便说明了这一点。第二个条件正是在此发挥作用。为了说明这个条件，我们需要用到下面的引理4：它是 Levy 的 0-1 定律的推论。

引理 4（见附录 B.4）设 $\mathcal{M}$ 是一个具有 *PI* 目标的可数马尔可夫链。若存在一个 $q \in Q$ 使得 $\chi_{\mathcal{M}}(q) < 1$，则 $\inf_{q' \in Q} \chi_{\mathcal{M}}(q') = 0$。

现在考虑一名玩家A 子博弈最优策略 s_A 以及一名玩家 B 确定性策略。让我们考虑当博弈最终稳定在 Q_u 对于某个 $u \in V_{\mathcal{G}} \setminus \{0\}$时会发生什么。反设存在一条有限路径，在此之后，给定博弈最终稳定在 W 的情况下，Q_u 的概率小于 1。那么，存在该路径的一条延续，以Q_u 结束，并且在该路径下，W 的概率小于 u。事实上，已有研究表明，对于 PI 目标，在可数马尔可夫链中（双方玩家的策略一旦确定后，我们得到的正是这种链），如果存在一个值小于 1 的状态，那么该马尔可夫链中各值的下确界为 0（引理 4所陈述的正是这一点）。按照上述反证假设，将存在一条有限路径，从该路径开始，玩家 A 的策略 s_A 不是最优的。这与它是子博弈最优策略这一事实相矛盾。因此，除了局部最优性假设之外，子博弈最优性的第二个必要条件是：对于所有有限路径、所有玩家 B 确定性策略以及所有正值 $u \in V_{\mathcal{G}} \setminus \{0\}$，$W$ 并最终稳定在 Q_u 的概率等于最终稳定在 Q_u的概率。我们得到以下定理。

定理 1（证明见附录 B.5）考虑一个并发博弈 $\mathcal{G}$，其具有一个 *PI* 目标W，以及玩家 A 的策略 $s_A \in S^A_{\mathcal{C}}$。策略 s_A 是子博弈最优的，当且仅当：

- 它是局部最优的；
- 对于所有 $\rho \in Q^+$、玩家 B 的所有确定性策略 s_B，以及所有值 $u \in V_{\mathcal{G}} \setminus \{0\}$，都有 $\mathbb{P}_{s^\rho{}_A, \rho}[W \cap Q^* \cdot s^\rho{}_B(Q_u)^\omega] = \mathbb{P}_{s^\rho{}_A, \rho}[Q^* \cdot s^\rho{}_B(Q_u)^\omega]$。

证明（概要）引理 *1* 指出，局部最优性是必要条件；并且如上文所述，第二个条件对于子博弈最优性也是必要的。至于这些条件足以保证子博弈最优性，则可直接由引理 *2* 和 *3* 推出，同时还需用到这样一个事实：在马尔可夫决策过程（*MDP*）中，确定性策略能够达到与任意策略相同的值（固定玩家 A 的策略后即可得到这一事实），如第 *4.2* 小节所述。 □

可以探讨一下，在所考虑的策略 s_A 为位置型策略这一特殊情况下会发生什么。如上所述，[5][1]已经给出了这样的刻画。总体而言，我们得到

[1]该证明 仅针对特定类别的目标进行了阐述 。

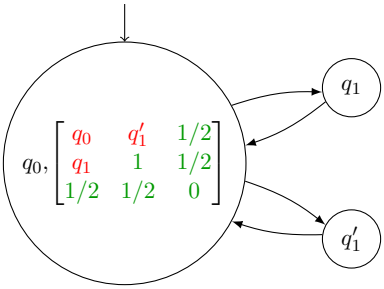


图 1：一个 co-B¨uchi 博弈。

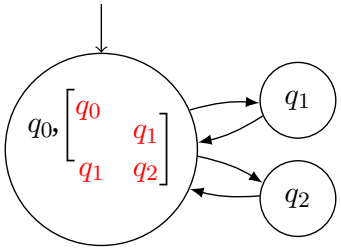


图 2：一个奇偶博弈。

一个类似的结果，只是第二个条件被替换为：在由位置策略 s_A 诱导的马尔可夫决策过程中，将博弈限制在终端分量上时会发生什么。

6 从子博弈几乎必然获胜到子博弈最优性

在 [15, 定理 4.5]，中，作者在 PI 回合制博弈中证明了一项迁移结果：在每个博弈中每个值为 1 的状态下进行最优博弈所需的记忆量，也足以在每个博弈中进行最优博弈。该结果不能直接推广到并发博弈。首先，尽管 PI 回合制博弈中总存在最优策略（如同一篇论文[15, 定理 4.3] 中所证明），但有些 PI 并发博弈不存在最优策略。其次，在 co-B¨uchi 并发博弈中，进行最优博弈可能需要无限记忆，而在回合制设定中，几乎必然获胜策略可以从位置策略中找到。图 1 中的博弈展示了这一点，其中 $\mathrm{col}(q_0) = 0$，且 $\mathrm{col}(q_1) = \mathrm{col}(q_1') = 1$。状态 q_0 处局部交互中的绿色值，是到达相应状态时博弈的值（博弈会立即结束）。若未到达绿色值，玩家 A 的目标是只有限次看到状态 q_1 和 q_1'。文献 [5] 已经论证，该博弈的值为 $1/2$，且玩家 A 存在一个最优策略，但该策略需要无限记忆。为了实现最优，玩家 A 必须以概率 $1 - \varepsilon_k$ 选择顶行，以概率 ε_k 选择中间行；对于 $\varepsilon_k > 0$，当 k 趋近于 ∞ 时，该概率快速趋近于 0（其中 k 表示步数）。ε_k 必须选择得足够合适，使得如果玩家 B 始终以概率 1 选择左列，那么以概率 1 只有限次看到状态 q_1。此外，一旦访问状态 q_1'，玩家 A 就切换到一个位置策略：以足够小的概率 ε_k' 选择底行（其中 k 表示状态 q_1' 被看到之前的步数），并以概率 $(1 - \varepsilon_k')/2$ 选择两条顶行。

因此，即使假设最优策略存在，从几乎必然获胜到最优的记忆转移在并发博弈中仍不成立。不过，可以注意到，尽管上述策略是最优的，却不是子博弈最优的。事实上，当策略切换时，剩余策略的值为 $1/2 - \varepsilon_k' < 1/2$。实际上，该博弈中不存在子博弈最优策略。事实上，如果我们假设不仅最优策略存在，而且子博弈最优策略也存在，那么记忆转移就会成立。

本节的目标有两个：首先，我们确定子博弈最优策略存在的充要条件[2]。其次，我们建立上述记忆转移关系，将记忆量与进行子博弈最优以及

[2]请注意，这与我们在上一节中的做法不同：在那里，我们为某个特定策略确立了其为子博弈最优策略的充要条件。而在这里，给定一个博弈，我们考察该博弈满足哪些充要条件时，子博弈最优策略才存在。

几乎必然获胜联系起来。在陈述本节的主要定理之前，我们首先介绍位置性子博弈几乎必胜目标的定义，即对于这类目标，可以在位置策略中找到子博弈几乎必胜策略。

定义18（位置性子博弈几乎必胜目标）考虑一个 *PI* 目标 $W \subseteq \mathsf{K}^\omega$。若满足以下条件，则称其为位置性子博弈几乎必胜目标（简称 *PSAW*）：在所有存在子博弈几乎必胜策略的并发博弈 $\mathcal{G} = \langle \mathcal{C}, W \rangle$ 中，都存在一个位置策略。

定理2考虑一个非空有限颜色集 K 和一个 *PI* 目标 $\emptyset \subsetneq W \subseteq \mathsf{K}^\omega$。考虑一个目标为 W的并发博弈 $\mathcal{G}$ 。那么，以下三个断言等价：

a. 存在一个子博弈最优策略；

b. 存在一个局部最优的最优策略；

c. 存在一个局部最优的正值最优策略。

此外，若上述条件成立且目标 W 是 *PSAW*，则存在一个子博弈最优位置策略。

首先请注意，这种等价关系是针对策略是否存在而言的，而不是针对策略本身而言的。特别是，任何子博弈最优策略既是最优策略，也是局部最优策略；然而，局部最优的最优策略未必是子博弈最优策略。附录C.1给出了一个例子。其次，显然，命题 a 推出命题 b （由定理 1得出），而命题 b 推出命题 c （根据正向最优策略的定义）。本节余下部分将非正式地解释导向该定理证明的构造，即证明命题 c 推出命题a的构造。记忆转移是该定理证明方式的直接结果。完整证明见附录 C。本节余下部分中，我们固定一个 PI 并发博弈 $\mathcal{G} = \langle \mathcal{C}, W \rangle$。

思路如下。正如定理 1所述，子博弈最优策略是局部最优的，并且当博弈最终进入值区域 Q_u 时，几乎必然获胜，其中$u \in V_\mathcal{G} \setminus \{0\}$为某个正值。因此，我们的思路是在导出博弈 $\mathcal{G}_u$中考虑子博弈几乎必胜策略：该博弈是对博弈 $\mathcal{G}$ 在 Q_u 上的一种“限制”（稍后会详细说明）。然后，我们可以将针对所有 $u \in V_\mathcal{G} \setminus \{0\}$定义的这些子博弈几乎必胜策略拼接起来，从而构造出一个子博弈最优策略。不过，其中存在一些事项：

1. 博弈 $\mathcal{G}_u$ 中的状态值应全部等于 1；
2. 此外， $\mathcal{G}_u$ 中必须存在一个子博弈几乎必胜策略；
3. 将此子博弈几乎必胜策略置于整个博弈 $\mathcal{G}$ 中考量时，它应当是局部最优的。

请注意，我们在此使用的方法不同于 [15] 为证明轮流制博弈中的记忆转移所采用的方法。

我们先来处理问题 3。通过适当定义博弈 $\mathcal{G}_u$ ，可以确保其中的几乎必然获胜策略在 $\mathcal{G}$ 中都是局部最优的。具体而言，我们通过规定玩家 A 在 $\mathcal{G}_u$ 中唯一可能采用的策略在博弈 $\mathcal{G}$ 中都是局部最优的，来实现这一点。为此，我们构造博弈 $\mathcal{G}_u$ ，其状态空间为 Q_u （加上辅助状态），但其在状态 $q \in Q_u$下的动作集合 $A_{\mathcal{F}_q^{\mathsf{nf}}}$使得策略集合$\mathcal{D}(A_{\mathcal{F}_q^{\mathsf{nf}}})$恰好对应于原始标准形博弈中的最优策略集合

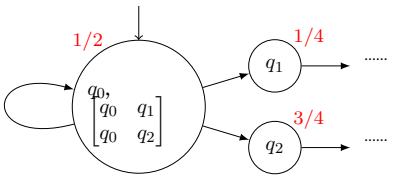


图 3：一个具有 $A_{q_0} = \{a_1, a_2\}$ 的并发博弈。

$$\begin{matrix} a_1 \\ a_2 \end{matrix}\begin{bmatrix} q_0 & q_1 \\ q_0 & q_2 \end{bmatrix}$$

图 4：状态 q_0 处的局部交互 $\mathcal{F}_{q_0}$ 。

$$\begin{matrix} a_1 \\ a_2 \end{matrix}\begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{3}{4} \end{bmatrix}$$

图 5：一般形式的博弈 $\mathcal{F}^{\mathsf{nf}}_{q_0}$。

$$\begin{matrix} \frac{a_1+a_2}{2} \\ a_2 \end{matrix}\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{4} \end{bmatrix} \qquad \begin{matrix} \frac{a_1+a_2}{2} \\ a_2 \end{matrix}\begin{bmatrix} q_0 & \frac{q_1+q_2}{2} \\ q_0 & q_2 \end{bmatrix}$$

图 6：博弈 $\mathcal{F}^{\mathsf{opt}}_{q_0}$ 与图 7：博弈形式 $\mathcal{F}^{\mathsf{opt}}_{q_0}$仅有最优策略。

形博弈 $\mathcal{F}^{\mathsf{nf}}_q$，同时保持玩家 A 的动作集合 $A_{\mathcal{F}^{\mathsf{nf}}_q}$为有限集。这得益于下面的命题 2 ：对于每个标准形博弈 $\mathcal{F}^{\mathsf{nf}}_q$，在状态 $q \in Q_u$处，都存在一个有限的最优策略集合 $A_{\mathcal{F}^{\mathsf{nf}}_q}$，使得 $\mathcal{F}^{\mathsf{nf}}_q$中的最优策略恰好是 $A_{\mathcal{F}^{\mathsf{nf}}_q}$中策略的凸组合。这是一个众所周知的结果，例如文献 [20]中就论证了这一点。

命题 2 考虑一个标准形博弈 $\mathcal{F}^{\mathsf{nf}} = \langle \mathsf{Act_A}, \mathsf{Act_B}, [0,1], \delta \rangle$，其中 $|\mathsf{Act_A}| = n$且 $|\mathsf{Act_B}| = k$。存在一个最优策略集合 $A_{\mathcal{F}^{\mathsf{nf}}} \subseteq \mathsf{Opt_A}(\mathcal{F}^{\mathsf{nf}})$，使得 $|A_{\mathcal{F}^{\mathsf{nf}}}| \leq n+k$且 $\mathcal{D}(A_{\mathcal{F}^{\mathsf{nf}}}) = \mathsf{Opt_A}(\mathcal{F}^{\mathsf{nf}})$。

证明（概要） 可以写出一个不等式组 $n+k$（附带一些额外等式），其解集恰好是最优GF策略的集合 $\mathsf{Opt_A}(\mathcal{F}^{\mathsf{nf}})$。由于解空间实际上是一个至多有 $n+k$ 个顶点的多面体，依据不等式组的标准论证即可得出结论。 □

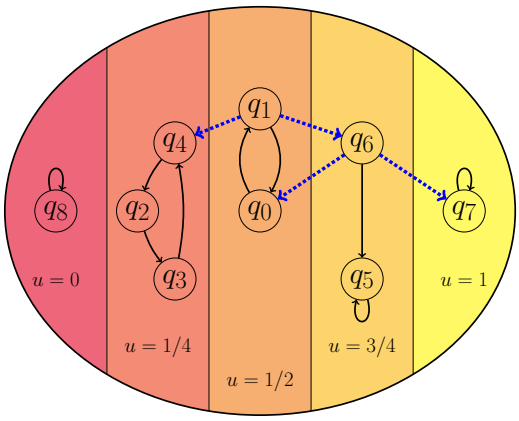
我们通过一个例子说明这一构造：图 3 展示了并发博弈的一部分，以及玩家在状态 q_0 处交互方式的变化，见图 4、5、6 和7。

博弈 $\mathcal{G}_u$ 与博弈具有相同的目标 W。由于我们希望 $\mathcal{G}$中的所有状态的值都为 1（请回想问题 $\mathcal{G}_u$ （回顾问题 1），我们将构造博弈 $\mathcal{G}_u$，使得任何一条通向不属于 Q_u 的状态的边，在 $\mathcal{G}$ 中都会转而通向一个 PI 并发博弈 $\mathcal{G}_W$ （具有相同的目标 W），且其中所有状态的值均为 1。博弈 $\mathcal{G}_W$ 可以（例如）是一个包含 K 中所有颜色的团，且由玩家A 单独行动。博弈 $\mathcal{G}_W = \langle \mathcal{C}_W, W \rangle$ 和博弈 $\mathcal{G}_u$的正式定义见附录 C.2。

图 8 和9展示了这一构造的示意图。蓝色虚线箭头是更改博弈时需要重新指向的箭头。通过这样的定义，我们在此前提到的问题1 （关于值等于 1）上取得了一些进展：博弈 $\mathcal{G}_u$ 中所有状态的值均为正（当 u为正时）。

引理 5（证明见附录 C.3） 考虑博弈 $\mathcal{G}_u$，其中某个正值 $u \in V_{\mathcal{G}} \setminus \{0\}$，并假设在 $\mathcal{G}$中存在一个局部最优的正值最优策略。那么，对于所有状态 q 在 $\mathcal{G}_u$中，状态 q 在 $\mathcal{G}_u$ 中的值为正： $\chi_{\mathcal{G}_u}(q) > 0$。

证明（概要）考虑状态 $q \in Q_u$，以及在 $\mathcal{G}$ 中从 q 出发正最优的玩家 A 局部最优策略 s_{A}。那么，策略 s_{A} （限制在 Q_u^+ 上）可以视为 $\mathcal{G}_u$ 中的策略（它必须在 $\mathcal{G}_W$ 中定义，但这可以直接做到）。注意，只有因为策略 s_{A} 是局部最优的（根据 $\mathcal{G}_u$ 的定义），这才有可能。对于 $\mathcal{G}_u$ 中的 B 玩家策略 s_{B}，考虑策略 s_{A} 和 s_{B} 在博弈 $\mathcal{G}_u$ 和 $\mathcal{G}$ 中的情形。博弈要么永远停留在 Q_u 中，此时 $\mathcal{G}_u$ 和 $\mathcal{G}$ 中发生的情况完全相同；要么最终离开 Q_u，到达值为 1 的状态 $\mathcal{G}_u$。因此，采用策略 s_{A} 和 s_{B} 时，从 q 开始的博弈 $\mathcal{G}_u$ 的值至少与采用相同策略时从 q 开始的博弈 $\mathcal{G}$ 的值一样高。因此，状态 q 在 $\mathcal{G}_u$ 中的值为正。 □



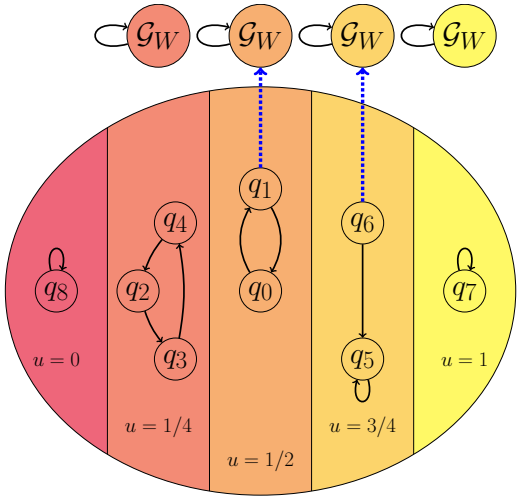


图8：PI 并发博弈及其值区域示意图。图9：经过上述修改后的 PI 并发博弈。

事实证明，引理 5 足以同时处理这两个事项**1** 和 **2** 。事实上，正如下文定理 3所述，有一个一般性结论：在 PI 并发博弈中，如果所有状态的值都为正，那么所有状态的值都为 **1**，并且存在一个子博弈几乎必胜策略。

定理 3（证明见附录 D）考虑一个 *PI* 并发博弈 $\mathcal{G}$，并假设所有状态值均大于或等于 $c > 0$，即对所有 $q \in Q$，$\chi_{\mathcal{G}}(q) \geq c$。那么，在 $\mathcal{G}$ 中存在一个子博弈几乎必胜策略。

注1 这个定理可以看作是对 *[7]* 中定理 *1* 的加强。事实上，该定理 *1* 指出，如果所有状态的值都为正，那么它们的值都为 *1*（随后这一结论被推广到具有可数多个状态的博弈）。定理 *3* 更强，因为它确保了（子博弈）几乎必然获胜策略的存在。尽管附录*D*中给出了详细证明，但请注意，该定理已在 *[15]* 中针对 *PI* 回合制博弈陈述并证明。不过，他们的论证也可以原封不动地用于并发博弈。在附录中，我们给出了采用相同构造（即重置策略）的证明，但对该构造为何能够证明定理给出了不同的论证。

现在，我们可以将所有博弈中定义的策略片段 $\mathsf{s}_{\mathsf{A}}^{u}$ 拼接成一个策略 $\mathsf{s}_{\mathsf{A}}[(\mathsf{s}_{\mathsf{A}}^{u})_{u \in V_{\mathcal{G}} \setminus \{0\}}]$。通俗地说，粘合策略会在 Q_u^+ 上模仿相应的策略，并在离开一个值区域、进入另一个值区域时切换策略。

定义19（拼接策略）考虑一个 *PI* 并发博弈 $\mathcal{G}$，并针对所有值 $u \in V_{\mathcal{G}} \setminus \{0\}$，设有一个博弈 $\mathcal{G}_u$ 中的策略 $\mathsf{s}_{\mathsf{A}}^{u}$。然后，我们将这些策略拼接成策略$\mathsf{s}_{\mathsf{A}}[(\mathsf{s}_{\mathsf{A}}^{u})_{u \in V_{\mathcal{G}} \setminus \{0\}}]$：$Q^+ \to \mathcal{D}(A)$，简写为$\mathsf{s}_{\mathsf{A}}$，使得对于所有以状态 $q \in Q$ 结束的 ρ：

$$\mathsf{s}_{\mathsf{A}}(\rho) := \begin{cases} \mathsf{s}_{\mathsf{A}}^{u}(\pi) & \text{if } u = \chi_{\mathcal{G}}(q) > 0 \text{ for } \pi \text{ the longest suffix of } \rho \text{ in } Q_u^+ \\ \text{is arbitrary} & \text{if } \chi_{\mathcal{G}}(q) = 0 \end{cases}$$

如下面的引理 6 所述，定义 19 中描述的构造会将 $\mathcal{G}_u$ 中的几乎必胜策略转化为 $\mathcal{G}$中的子博弈最优策略。

引理 6（证明见附录 C.4）对于所有 $u \in V_{\mathcal{G}} \setminus \{0\}$，设 $\mathsf{s}_{\mathsf{A}}^{u}$ 是 $\mathcal{G}_u$中的子博弈几乎必胜策略。粘合策略 $\mathsf{s}_{\mathsf{A}}[(\mathsf{s}_{\mathsf{A}}^{u})_{u \in V_{\mathcal{G}} \setminus \{0\}}]$记作 s_{A}，是$\mathcal{G}$中的子博弈最优策略。

证明（概要）我们应用定理 *1*。首先，根据定义博弈 $\mathcal{G}_u$ 时所作的策略限制，在所有值为 u >*0* 的 Q_u 中，策略 s_{A} 都是局部最优的（在状态 $q \in Q_u$ 下，每个标准形 $\mathcal{F}_q^{\mathsf{nf}}$ 博弈中只考虑最优策略）。此外，在值为 *0* 的标准形博弈中，任意策略都是最优的（状态 Q_0 对应的标准形博弈正是如此）。其次，如果博弈最终停留在某个值区域 Q_u 中，其中某个 $u > 0$，那么从那时起，策略 s_{A} 就模仿策略 s_{A}^{u}；后者在 $\mathcal{G}_u$ 中是子博弈几乎必胜策略。因此，给定博弈最终停留在 Q_u 中，发生 W 的概率为 *1*。这对所有 $u \in V_{\mathcal{G}} \setminus \{0\}$ 都成立，所以定理 *1* 的第二个条件成立。 □

现在，我们已经具备证明定理 2 所需的所有要素。

证明（定理2） 我们考虑 *PI* 并发博弈 $\mathcal{G}$，并假设存在一个局部最优的正值最优策略。那么，根据引理*5*，对于所有正值$u \in V_{\mathcal{G}} \setminus \{0\}$，$\mathcal{G}_u$ 中的所有状态都具有正值。由此，根据定理 *3*，对于每个满足 $\mathcal{G}_u$ 的 $u \in V_{\mathcal{G}} \setminus \{0\}$，都存在一个子博弈几乎必胜策略。接着，根据引理 *6*，将这些策略拼接起来，便可得到一个子博弈最优策略。

定理的第二部分讨论如何将位置性从子博弈几乎必然获胜转移到子博弈最优。这一结论可由以下事实推出：若对于所有 $u \in V_{\mathcal{G}} \setminus \{0\}$，所有策略 $\mathsf{s}_{\mathsf{A}}^{u}$ 都是位置型的，那么粘合策略 $\mathsf{s}_{\mathsf{A}}[(\mathsf{s}_{\mathsf{A}}^{u})_{u \in V_{\mathcal{G}} \setminus \{0\}}]$ 也具有位置性。 □

现在，我们将定理 2 的结果应用于两类特定目标：B¨uchi 和 co-B¨uchi 目标。请注意，对于 B¨uchi 目标，这一结果已为人所知，并已在 [5]中证明。

推论 1 考虑一个具有 *B ¨uchi*（分别为 *co-B ¨uchi*）目标的并发博弈，并假设存在一个局部最优的正值最优策略。那么就存在一个子博弈最优位置策略。

请注意，也可以针对任意记忆骨架，而不只是位置策略，证明从子博弈几乎必然获胜到子博弈最优的记忆转移。这只会带来少量细微的困难。相关内容见附录 C.5。

应用于轮流制设定。第6 节旨在将一项关于轮流制博弈的既有结果推广到并发博弈的背景中。这需要调整相关假设。然而，事实上，我们可以相当直接地从定理 2 推出关于轮流制博弈的原始结果。具体而言，只需证明：对于所有有限轮流制博弈 $\mathcal{G}$，对于所有值 $u \in V_{\mathcal{G}} \setminus \{0\}$，都存在一个局部最优策略，该策略从 Q_u中的所有状态出发都是正最优的。这一点在附录 C.6 中证明。

7 有限选择策略

在本节中，我们将介绍一种新型策略，即有限选择策略。首先来说明我们为何考虑这类策略。再次考虑图 ¨1中的余 Büchi 博弈。回想一下，我们最先描述的最优策略是以递增的概率选择顶行，以递减的概率选择中间行；之后，一旦玩家 B 选择第二列，就切换到一个位置策略，以正概率（但概率足够小）选择底行。请注意，切换策略至关重要。事实上，如果玩家 A 不切换，玩家 B就可能在某个时刻选择中间列，从而无限期地看到状态 q_1'，且其概率非常高

事实上，这种情况下发生的事情相当违反直觉：玩家 B 一旦切换，就会有无穷多次以正概率到达值为 1 的结果。然而，如果玩家 B 等待足够长的时间才选择中间行，那么最终到达该结果的概率可以任意小。这是因为到达该结果的概率 ε_k 会随着 k 趋于 ∞而（迅速）趋于 0。事实上，这种最优策略具有“无限选择”，因为它可能规定无穷多种不同的概率分布。

在本节中，我们考察有限选择策略，即在每个状态下只能使用有限多个 GF-策略的策略。

定义 20（有限选择策略）设 $\mathcal{G}$ 是一个并发博弈。若玩家 A 在 $\mathcal{G}$ 中的策略s_A 满足：对于所有 $q \in Q$，集合 $S_q^{\mathsf{s}_\mathsf{A}} := \{\mathsf{s}_\mathsf{A}(\rho \cdot q) \mid \rho \in Q^+\} \subseteq \mathcal{D}(A_q)$ 都是有限的，则称其具有有限选择。

注意，位置型策略（甚至有限记忆策略）和确定性策略都是有限选择策略的例子。

有趣的是，我们可以将有限选择策略与子博弈最优策略的存在联系起来。一般而言，即使存在最优策略，也不一定存在子博弈最优策略（如图 1中的博弈所示）。然而，在下文的定理 4中，我们将说明，如果进一步假设所考虑的最优策略具有有限选择，那么就存在一个子博弈最优策略（它也具有有限选择）。

定理 4（证明见附录 E.1）考虑一个 *PI* 并发博弈 $\mathcal{G}$。如果存在一种有限选择最优策略，那么就存在一种有限选择子博弈最优策略。

证明（概要）考虑这样一种有限选择最优策略 s_A。特别要注意，存在一个常数 $c > 0$，使得对于所有 $\rho \cdot q \in Q^+$、所有 $a \in A_q$，都有：$s_\mathsf{A}(\rho \cdot q)(q) > 0 \Rightarrow s_\mathsf{A}(\rho \cdot q)(q) \geq c$。我们按如下方式构造子博弈最优策略 s'_A：对于所有 $\rho = \rho' \cdot q \in Q^+$，若剩余策略 s^ρ *A* 是最优的，则令 $s'_\mathsf{A}(\rho) := s_\mathsf{A}(\rho)$；否则令 $s'_\mathsf{A}(\rho) := s_\mathsf{A}(q)$（即重置策略）。显然，策略 s'_A 具有有限选择。我们希望应用定理 *1* 来证明它是子博弈最优的。可以看出，它是局部最优的（依据所选择的策略重置准则）。现在考虑某个 $\rho \in Q^+$，其终止于状态 $q \in Q$，以及另一个状态 $q' \in Q$。假设剩余策略 s^ρ *A* 是最优的，但剩余策略 $s^{\rho \cdot q'}$ *A* 不是。与局部最优性是子博弈最优性的必要条件的原因类似（见命题 *1*），可以证明，任何以正概率从 ρ 到达q' 的玩家 *B* 的动作 $b \in B_q$，都满足 $\chi_\mathcal{G}(q) < out^{\mathcal{F}^{\mathsf{nf}}_q}(s_\mathsf{A}(\rho)\ b)$。因此，若玩家 *B* 选择动作 b，则从 ρ 出发，有正概率到达一个值不同于 $u = \chi_\mathcal{G}(q)$ 的状态。若这种情况无限次发生，则几乎必然会到达一个值不同于 u 的状态。换言之，若从未离开一个值区域，那么几乎必然地，策略 s'_A 只会有限次重置。

现在考虑某个$\rho \in Q^+$，玩家B 的确定性策略s_B 以及一个值 $u \in V_\mathcal{G} \setminus \{0\}$。根据我们上面的论述，事件 $Q^* \cdot (Q_u)^\omega$（分别是 $W \cap Q^* \cdot (Q_u)^\omega$）的概率，与将其和策略$\mathsf{s}'_\mathsf{A}$ 只进行有限次重置这一事实取交集时的概率相同。此外，如果策略从某个时刻起不再重置，并且所有状态的值都相同$u > 0$，那么可得 W 的概率为 *1*（因为 W 是 *PI*）。于是，我们可以应用定理 *1*得出结论。□

有限选择策略还有另一个值得关注的原因。在上一节中，我们将定理2中的记忆转移应用于 B¨uchi 和 co-B¨uchi 目标。我们没有将其应用于其他目标——尤其是奇偶性目标。事实上，一般而言，与

[3]这是因为策略 s_A 具有有限选择：观察到值不同的状态的概率至少为 c 与所有自然状态中最小正概率的乘积。

轮流制博弈的情况不同，在奇偶博弈中，要几乎必然获胜就需要无限记忆。这种情况出现在图 2 （已在 [13]中介绍）中，玩家A 的目标是无限多次看到 q_1 ，同时只有限次看到 q_2 。下面描述一个玩家A 的子博弈几乎必胜策略。以概率 $1-\varepsilon_k$ 走顶行，以概率 $\varepsilon_k > 0$ 走底行，其中 ε_k 在 k 趋于 ∞ 时趋于 0（图ε_k中博弈所用的（1 ）在这里同样适用），其中 k 表示状态 q_0 被看到的次数。这种策略是子博弈几乎必胜策略，但并非有限选择策略。事实上，可以证明，在该博弈中，玩家A 的所有有限选择策略的值都是 0。

有趣的是，定理 2 中的记忆转移，在定理 5中经过了调整，采用了在轮流制博弈中足够的记忆——针对那些具有“中性色”的 PI 目标——前提是我们另外假设所考虑的子博弈最优策略具有有限选择。首先，我们来定义“中性色”的含义，然后定义轮流制版本的 PSAW。

定义 21（具有中性色的目标）考虑一个颜色集合K 以及一个 *PI* 目标 $W \subseteq \mathsf{K}^\omega$。若存在某个（中性）颜色 $k \in \mathsf{K}$ ，使得对于所有$\rho = \rho_0 \cdot \rho_1 \cdots \in \mathsf{K}^\omega$，都有 $\rho \in W \Leftrightarrow \rho_0 \cdot k \cdot \rho_1 \cdot k \cdots \in W$，则该目标具有中性色。

定义 22（轮流制博弈中的 PASW 目标）考虑一个 *PI* 目标 $W \subseteq \mathsf{K}^\omega$。若在所有存在子博弈几乎必胜策略的轮流制博弈 $\mathcal{G} = \langle \mathcal{C}, W \rangle$ 中，都存在一个位置型策略，则称该目标在轮流制博弈中位置型子博弈几乎必胜（简称 *PSAWT*）。

定理 5（证明见附录E.2）考虑一个*PSAWT PI* 目标 $W \subseteq \mathsf{K}^\omega$ ，以及一个具有中性色的并发博弈 $\mathcal{G}$ ，其目标为 W。假设存在一个具有有限选择的子博弈最优策略。那么，存在一个位置型策略。

证明（概要）有限选择策略s_A 在每个状态下只从有限个GF策略中选择。因此，我们的思路是将前一小节的博弈 $\mathcal{G}_u$ 改造为一个博弈$\mathcal{G}'_u$ ，将其转化为一个（有限的）轮流制博弈。在每个状态下，玩家A 先选择她的GF策略。她只能从有限个策略中选择：在状态 q，她只能使用GF中的最优策略 $S_q^{\mathsf{s}_\mathsf{A}}$（回忆定义*20*）。我们在这个新博弈场中考虑目标 W，其中玩家B 的状态被赋予中性色。在 $\mathcal{G}$中存在一个具有有限选择的子博弈最优策略，这保证了 $\mathcal{G}'_u$ 中的所有状态都具有正值。于是，我们可以像定理*2*那样得出结论：通过拼接这些（轮流制）博弈中的子博弈几乎必然获胜策略 $\mathcal{G}'_u$ （根据假设，这些策略可以选为位置型），即可得到一个子博弈最优策略。
□

作为应用，可以看出，奇偶、均值收益和广义 B¨uchi 目标具有中性色，并且属于 PSAWT（[12, 18, 8]）。因此，对于这些目标，如果存在一种具有有限选择的最优策略，那么就存在一种位置型最优策略。

推论 2（证明见附录 E.3） 考虑一个具有奇偶（分别为均值收益、广义 *B ¨uchi*）目标的并发博弈 $\mathcal{G}$ 。假设在 $\mathcal{G}$中存在一种具有有限选择的最优策略。那么，就存在一种位置型最优策略。

参考文献

[1] Roderick Bloem、Krishnendu Chatterjee 和 Barbara Jobstmann。《模型检测手册》，第 Graph games and reactive synthesis 章，第 921–962 页。Springer，2018 年。

[2] Benjamin Bordais、Patricia Bouyer 和 St´ephane Le Roux。并发图博弈中的局部确定性到全局确定性。载于 Mikolaj Bojanczyk 和 Chandra Chekuri 编，《第41届 IARCS 软件技术与理论计算机科学基础年会，FSTTCS 2021，2021年12月15日至17日，线上会议》，LIPIcs 第213卷，第 41:1–41:14 页。Schloss Dagstuhl - Leibniz-Zentrum f¨ur I nformatik，2021年。[3] Benjamin Bordais、Patricia Bouyer 和 St´ephane Le Roux。并发可达性博弈中的最优策略。*CoRR*，abs/2110.14724，2021年。[4] Benjamin Bordais、Patricia Bouyer 和 St´ephane Le Roux。并发可达性博弈中的最优策略。载于 Florin Manea 和 Alex Simpson 编，《第30届 EACSL 计算机科学逻辑年会，CSL 2022，2022年2月14日至19日，G》¨*ottingen*，德国（线上会议），LIPIcs 第216卷，第 7:1–7:17 页。Schloss Dagstuhl - Leibniz-Zentrum f¨ur Informatik，2022年。[5] Benjamin Bordais、Patricia Bouyer 和 St´ephane Le Roux。在并发 b¨üchi 和 co-b¨üchi 博弈中实现（近乎）最优游玩。*CoRR*，abs/2203.06966，2022年。[6] Patricia Bouyer、St´ephane Le Roux、Youssouf Oualhadj、Mickael Randour 和 Pierre Vandenhove。可以使用与竞技场无关的有限记忆实现最优游玩的博弈。载于第*31*届国际并发理论会议，*CONCUR 2020*，*2020*年*9*月*1*日至*4*日，维也纳，奥地利（线上会议），第 24:1–24:22 页，2020年。[7] Krishnendu Chatterjee。具有尾部目标的并发博弈。理论计算机科学，388(1-3)：181–198，2007年。[8] Krishnendu Chatterjee、Luca de Alfaro 和 Thomas A. Henzinger。以记忆换取随机性。载于第*1*届系统定量评估国际会议（*QEST 2004*），*2004*年*9*月*27*日至*30*日，恩斯赫德，荷兰，第 206–217 页。IEEE计算机学会，2004年。[9] Krishnendu Chatterjee、Luca de Alfaro 和 Thomas A. Henzinger。定量并发奇偶博弈的复杂性。载于第十七届 *ACM-SIAM* 离散算法研讨会论文集，*SODA 2006*，*2006*年*1*月*22*日至*26*日，美国佛罗里达州迈阿密，第 678–687 页。ACM出版社，2006年。[10] Krishnendu Chatterjee、Laurent Doyen、Hugo Gimbert 和 Thomas A. Henzinger。免费的随机性。信息与计算，245：3–16，2015年。[11] Krishnendu Chatterjee 和 Rasmus Ibsen-Jensen。并发平均收益博弈的定性分析。信息与计算，242：2–24，2015年。[12]Krishnendu Chatterjee、Marcin Jurdzinski 和 Thomas A. Henzinger。定量随机奇偶博弈。载于 J. Ian Munro 编，《第十五届 ACM-SIAM 离散算法研讨会论文集，SODA 2004，2004年1月11日至14日，美国路易斯安那州新奥尔良》，第 121–130 页。SIAM，2004年。[13] Luca de Alfaro 和 Thomas A. Henzinger。并发 ω-正则博弈。载于第*15*届 *IEEE* 计算机科学逻辑研讨会，*2000*年*6*月*26*日至*29*日，美国加利福尼亚州圣巴巴拉，第 141–154 页。IEEE计算机学会，2000年。[14] Hugh Everett。递归博弈。数学研究年刊 –《博弈论贡献》，3：67–78，1957年。

[15] Hugo Gimbert 和 Florian Horn。《求解简单随机尾博弈》。载于 Moses Charikar（编），第二十一届 *ACM-SIAM* 离散算法年度研讨会论文集，*SODA 2010*，奥斯汀，得克萨斯州，美国，*2010*年*1*月*17*日至*19*日，第 847–862 页。SIAM，2010 年。[16] Stefan Kiefer、Richard Mayr、Mahsa Shirmohammadi、Patrick Totzke 和 Dominik Wojtczak。《如何在无限马尔可夫决策过程中进行游戏》（特邀报告）。载于 Artur Czumaj、Anuj Dawar 和 Emanuela Merelli（编），第*47*届国际自动机、语言与编程研讨会，*ICALP 2020*，*2020*年*7*月*8*日至*11*日，萨尔布 ¨吕肯，德国（线上会议），第 168 卷，LIPIcs，第 3:1–3:18 页。Schloss Dagstuhl - Leibniz 信息学中心，2020 年。[17] Marta Kwiatkowska、Gethin Norman、Dave Parker 和 Gabriel Santos。并发随机系统的自动验证。《系统设计中的形式化方法》，58:188–250，2021 年。[18] Thomas M Liggett 和 Steven A Lippman。具有完全信息和时间平均收益的随机博弈。《*SIAM*评论》，11(4):604–607，1969 年。[19] Donald A. Martin。Blackwell 博弈的确定性。《符号逻辑杂志》，63(4):1565–1581，1998 年。[20] Lloyd S Shapley 和 RN Snow。离散博弈的基本解。《博弈论贡献》，1(24):27–27，1950 年。[21] Wolfgang Thomas。无限博弈与验证。载于第*14*届国际计算机辅助验证会议论文集（*CAV'02*），第 2404 卷，《计算机科学讲义》，第 58–64 页。Springer，2002 年。特邀教程。[22] John von Neumann 和 Oskar Morgenstern。《博弈论与经济行为》。普林斯顿大学出版社，普林斯顿，1944年。[23] Wieslaw Zielonka。完全信息随机奇偶博弈。载于 Igor Walukiewicz（编），《软件科学与计算结构基础》，第七届国际会议，*FOSSACS 2004*，作为欧洲软件理论与实践联合会议 *ETAPS 2004* 的一部分举行，巴塞罗那，西班牙，*2004*年*3*月*29*日至*4*月*2*日，论文集，第 2987 卷，《计算机科学讲义》，第 499–513 页。Springer，2004 年。

A 其他预备知识

对于集合 Q 以及 Q的有限序列的子集 $S \subseteq Q^+$，我们用 $S^\omega \subseteq Q^\omega$表示 Q 的无限序列中有无穷多个前缀属于 S的集合：$S^\omega := \{\pi \in Q^\omega \mid \forall i \in \mathbb{N},\ \exists j \geq i,\ \pi_{\leq j} \in S\}$。

我们还定义（可数）马尔可夫链的概念。马尔可夫链是一个二元组 $(Q,\ \mathbb{P})$，其中 Q 是状态集合，$\mathbb{P}\colon Q \times Q \to [0,1]$ 是概率函数，并且对于所有状态 $q \in Q$，都有 $\sum_{q' \in Q} \mathbb{P}(q,\ q') = 1$。

B 第 5 节的证明

我们回顾一个 [3]中的命题（具体来说，是 [3]中的命题42），本节将用到它。

命题 3 考虑一个 *PI* 并发博弈 $\mathcal{G} = \langle \mathcal{C}, W \rangle$、一个状态 $q \in Q$，以及标准形 $\mathcal{F}_q$ 中该博弈双方玩家的策略 σ_{A}、$\sigma_{\mathsf{B}} \in \mathcal{D}(A) \times \mathcal{D}(B)$。我们有如下关系：

$$\sum_{q' \in Q} \mathbb{P}^{q,q'}(\sigma_{\mathsf{A}}, \sigma_{\mathsf{B}}) \cdot \chi_{\mathcal{G}}(q') = \mathsf{out}_{\mathcal{F}_q}(\sigma_{\mathsf{A}}, \sigma_{\mathsf{B}})$$

B.1 引理1的证明

下面给出命题。

命题 4 考虑一个 *PI* 并发博弈 $\mathcal{G}$，以及从状态 $q \in Q$出发的玩家 A 最优策略 s_{A}。那么，对于所有玩家 B 的动作 $b \in B_q$，有 $\mathsf{out}^{\mathcal{F}_q}(\mathsf{s}_{\mathsf{A}}(q)\ b) \geq \chi_{\mathcal{G}}(q)$。

证明与 [5]位置策略情形下引理17的证明相同。

证明 反设结论不成立，即存在一个玩家 B 的动作$b \in B_q$，使得对于某个 $\varepsilon > 0$，有 $out^{\mathcal{F}_q}(s_{\mathsf{A}}(q)\ b) \leq \chi_{\mathcal{G}}(q) - \varepsilon$。下面说明策略 s_{A}从 q出发并非最优。考虑一个玩家 B 策略 s'_{B}，使得对于所有 $q' \in Q$，该策略在状态 q' 的值至多为 $\chi_{\mathcal{G}}(q') + \varepsilon/2$：$\chi_{\mathcal{G}}[s_{\mathsf{B}}](q') \leq \chi_{\mathcal{G}}(q') + \varepsilon/2$。接着，我们定义玩家 B 策略 s_{B} 如下：$s_{\mathsf{B}}(q) := b$，并且对于所有 $q' \in Q$，有 $s_{\mathsf{B}}(q \cdot q') := s_{\mathsf{B}}(q')$。于是，由于 W 是前缀无关的，且根据命题 *3*，我们得到：

$$\begin{aligned}
\mathbb{P}_q^{\mathsf{s}_{\mathsf{A}},\mathsf{s}_{\mathsf{B}}}[W] &= \sum_{q' \in Q} \mathbb{P}^{\mathsf{s}_{\mathsf{A}},\mathsf{s}_{\mathsf{B}}}(q)[q'] \cdot \mathbb{P}_{q'}^{\mathsf{s}_{\mathsf{A}}^{q \cdot q'},\mathsf{s}_{\mathsf{B}}^{q \cdot q'}}[W] \\
&\leq \sum_{q' \in Q} \mathbb{P}^{\mathsf{s}_{\mathsf{A}},\mathsf{s}_{\mathsf{B}}}(q)[q'] \cdot (\chi_{\mathcal{G}}(q') + \varepsilon/2) \\
&= \sum_{q' \in Q} \mathbb{P}^{q,q'}(\mathsf{s}_{\mathsf{A}}(\rho), \mathsf{s}_{\mathsf{B}}(\rho)) \cdot \chi_{\mathcal{G}}(q') + \varepsilon/2 \\
&= \mathsf{out}_{\mathcal{F}_q}(\mathsf{s}_{\mathsf{A}}(q), b) + \varepsilon/2 \\
&\leq \chi_{\mathcal{G}}[q] - \varepsilon + \varepsilon/2 = \chi_{\mathcal{G}}[q] - \varepsilon/2
\end{aligned}$$

也就是说，策略 s^ρ A 从 q 出发并非最优。□

引理的证明1随即便可直接得出。

证明 反设s_{A} 不是局部最优策略。也就是说，存在某个$\rho = \rho' \cdot q \in Q^+$ 以及玩家B 的动作 $b \in B_q$，使得$\mathsf{out}^{\mathcal{F}_q}(\mathsf{s}_{\mathsf{A}}(\rho)\ b) \leq \chi_{\mathcal{G}}(q) - \varepsilon$对某个$\varepsilon > 0$成立。由命题*4* 可知，剩余策略$\mathsf{s}^\rho$ A 从 ρ 出发并非最优策略，因此得出矛盾。□

B.2 引理2的证明

首先，我们陈述一个关于局部最优策略的直接命题。

命题 5 在 *PI* 并发博弈 $\mathcal{G}$ 中，对于玩家 A 的局部最优策略 $\mathsf{s_A}$ 以及玩家 B 的确定性策略，对于所有有限路径 $\rho = \rho' \cdot q \in Q^+$，都有 $\chi_{\mathcal{G}}(q) \leq \sum_{u \in V_{\mathcal{G}}} u \cdot \mathbb{P}^{\mathsf{s_A},\ \mathsf{s}_\rho\ ^{\mathsf{B}}}[Q_u]$。

证明 令 $b := \mathsf{s_B}(\rho) \in B_q$。我们有：

$$\begin{aligned}\chi_{\mathcal{G}}(q) &\leq \sum_{a \in A} \mathsf{s_A}(\rho)(a) \cdot \nu_{\mathcal{G}} \circ \delta(q,a,b) = \sum_{a \in A} \mathsf{s_A}(\rho)(a) \cdot \Big(\sum_{q' \in Q} \mathsf{dist} \circ \delta(q,a,b)(q') \cdot \chi_{\mathcal{G}}(q')\Big) \\ &= \sum_{u \in V_{\mathcal{G}}} \sum_{a \in A} \mathsf{s_A}(\rho)(a) \cdot \Big(\sum_{q' \in Q_u} \mathsf{dist} \circ \delta(q,a,b)(q') \cdot \chi_{\mathcal{G}}(q')\Big) \\ &= \sum_{u \in V_{\mathcal{G}}} u \cdot \left(\sum_{q' \in Q_u} \sum_{a \in A} \mathsf{s_A}(\rho)(a) \cdot \mathsf{dist} \circ \delta(q,a,b) \right) = \sum_{u \in V_{\mathcal{G}}} u \cdot \mathbb{P}_\rho^{\mathsf{s_A},\mathsf{s_B}}[Q_u]\end{aligned}$$

□

现在可以开始证明引理 2。

证明 我们用 v 表示赋值 $\chi_{\mathcal{G}}$。我们将赋值 v 扩展到有限路径：$v^+ : Q^+ \to [0,1]$，使得对于所有 $\rho \cdot q \in Q^+$，都有 $v^+(\rho \cdot q) := v(q)$。如果所有状态的值都是 *0* 或 *1*，则该引理显然成立。现在假设有些状态的值介于 *0* 和 *1* 之间。我们用 $0 = u_0 < u_1 < \ldots < u_n = 1$ 表示 $V_{\mathcal{G}}$ 中的状态值，即 $|V_{\mathcal{G}}| = n+1$。对于所有 $i \leq n+1$，我们用 $Q_{\leq i}$ 表示集合 $Q_{\leq i} := v^{-1}[\{u_0, \ldots, u_i\}]$。我们通过对 $k \leq n$ 进行归纳来证明以下性质 $\mathcal{P}(k)$：

$$\mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}\Big(\bigcup_{0 \leq i \leq k} Q^* \cdot (Q_{u_i})^\omega \cap Q^* \cdot (Q_{\leq k})^\omega \Big) = \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(Q^* \cdot (Q_{\leq k})^\omega)$$

这对 $k = 0$显然成立。现在考虑某个 $k \leq n$，并假设 $\mathcal{P}(k)$ 成立。我们有：

$$\begin{aligned}\mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(Q^* \cdot (Q_{\leq k+1})^\omega) &= \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(Q^* \cdot (Q_{k+1})^\omega \cap Q^* \cdot (Q_{\leq k+1})^\omega) \\ &\quad + \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(Q^* \cdot (Q_{\leq k})^\omega \cap Q^* \cdot (Q_{\leq k+1})^\omega) \\ &\quad + \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}((Q^* \cdot Q_{\leq k})^\omega \cap (Q^* \cdot Q_{k+1})^\omega \cap Q^* \cdot (Q_{\leq k+1})^\omega)\end{aligned}$$

我们来证明术语 $\mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}((Q^* \cdot Q_{\leq k})^\omega \cap (Q^* \cdot Q_{k+1})^\omega \cap Q^* \cdot (Q_{\leq k+1})^\omega) = 0$。若 $k = n$，则此结论成立，因为此时 $u_{k+1} = 1$，而局部最优意味着所有以正概率被访问的状态的值都为 *1*。现在假设 $k < n$ 且 $u_{k+1} < 1$。对于所有 $l \in \mathbb{N}$，记事件 $E^l_{\leq k}$ 为 $E^l_{\leq k} := Q^l \cdot (Q_{k+1} \cap Q \cdot Q_{\leq k})$，记事件 $E^l_{\geq k+2}$ 为 $E^l_{\geq k+2} := Q^l \cdot (Q_{k+1} \cap Q \cdot Q_{\geq k+2})$。再记 $E^l := E^l_{\leq k} \cup E^l_{\geq k+2}$。我们还记 $E^\infty_{\leq k} := \cap_{d \in \mathbb{N}} \cup_{l \geq d} E^l_{\leq k}$、$E^\infty_{\geq k+2} := \cap_{d \in \mathbb{N}} \cup_{l \geq d} E^l_{\geq k+2}$ 和 $E^\infty := \cap_{d \in \mathbb{N}} \cup_{l \geq d} E^l$。若 $\mathbb{P}^{\mathsf{s_A}}$，则 $s^{\mathsf{B}}(E^\infty) = 0$，因为 $(\mathbb{P}^{\mathsf{s_A}})^{\mathsf{B}}(Q^* \cdot Q_{\leq k})^\omega \cap (Q^* \cdot Q_{k+1})^\omega \cap Q^* \cdot (Q_{\leq k+1})^\omega = 0$。现在假设 $\mathbb{P}^{\mathsf{s_A}}$，则 $s^{\mathsf{B}}(E^\infty) > 0$。

考虑某个 $\rho = \rho' \cdot q \in Q^+ \cdot Q_{k+1}$。由于$\mathsf{s_A}$ 是局部最优的，根据命题5，我们有：

$$\begin{aligned} u_{k+1} \leq \sum_{u \in V_{\mathcal{G}}} u \cdot \mathbb{P}_\rho^{\mathsf{s_A},\mathsf{s_B}}[Q_u] &= u_{k+1} \cdot \mathbb{P}_\rho^{\mathsf{s_A},\mathsf{s_B}}[Q_{u_{k+1}}] + \sum_{u \leq u_k} u \cdot \mathbb{P}_\rho^{\mathsf{s_A},\mathsf{s_B}}[Q_u] + \sum_{u \geq u_{k+2}} u \cdot \mathbb{P}_\rho^{\mathsf{s_A},\mathsf{s_B}}[Q_u] \\ &\leq u_{k+1} \cdot \mathbb{P}_\rho^{\mathsf{s_A},\mathsf{s_B}}[Q_{u_{k+1}}] + \sum_{u \leq u_k} u_k \cdot \mathbb{P}_\rho^{\mathsf{s_A},\mathsf{s_B}}[Q_u] + \sum_{u \geq u_{k+2}} \mathbb{P}_\rho^{\mathsf{s_A},\mathsf{s_B}}[Q_u] \\ &= u_{k+1} \cdot \mathbb{P}_\rho^{\mathsf{s_A},\mathsf{s_B}}[Q_{u_{k+1}}] + u_k \cdot \mathbb{P}_\rho^{\mathsf{s_A},\mathsf{s_B}}[Q_{\leq k}] + \mathbb{P}_\rho^{\mathsf{s_A},\mathsf{s_B}}[Q_{\geq k+2}] \end{aligned}$$

记 $\mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}_\rho[Q_{\leq k}]$ 为 p_k，并记 $\mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}_\rho[Q_{\geq k+2}]$ 为 p_{k+2}，我们得到：

$$\begin{aligned} & u_k \cdot p_k + p_{k+2} \geq u_{k+1} \cdot (p_k + p_{k+2}) \\ \Leftrightarrow\ & p_{k+2} \cdot (1 - u_{k+1}) \geq p_k \cdot (u_{k+1} - u_k) \\ \Leftrightarrow\ & p_{k+2} \geq p_k \cdot \frac{u_{k+1} - u_k}{1 - u_{k+1}} = p_{k+2} \cdot x \end{aligned}$$

对于 $x := \frac{u_{k+1}-u_k}{1-u_{k+1}} \geq 0$。因此，令 $c := \frac{x}{1+x} \in [0, 1[$，我们有 $\frac{p_{k+2}}{p_k + p_{k+2}} \geq c$。也就是说，对于 $l = |\rho|$，我们有：

$$\mathbb{P}_\rho^{\mathsf{s_A},\mathsf{s_B}}(\neg E^l_{\geq k+2} \cap E^l) \leq (1-c) \cdot \mathbb{P}_\rho^{\mathsf{s_A},\mathsf{s_B}}(E^l)$$

接下来，考虑概率 $\mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(\neg E^\infty_{\geq k+2} \mid E^\infty)$（回想一下，$\mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(E^\infty) > 0$）。这等于：

$$\lim_{d \to \infty} \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(\bigcap_{l \geq d} (\neg E^l_{\geq k+2}) \mid E^\infty) \leq \lim_{d \to \infty} (\lim_{t \to \infty} (1-c)^t) = 0$$

此外：

$$(Q^* \cdot Q_{\leq k})^\omega \cap (Q^* \cdot Q_{k+1})^\omega \cap Q^* \cdot (Q_{\leq k+1})^\omega \subseteq (Q^\omega \setminus E^\infty_{\geq k+2}) \cap E^\infty$$

由此可得：

$$\mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}((Q^* \cdot Q_{\leq k})^\omega \cap (Q^* \cdot Q_{k+1})^\omega \cap Q^* \cdot (Q_{\leq k+1})^\omega) \leq \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(\neg E^\infty_{\geq k+2} \cap E^\infty) = \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(\neg E^\infty_{\geq k+2} \mid E^\infty) \cdot \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(E^\infty) = 0$$

总而言之：

$$\begin{aligned} \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(Q^* \cdot (Q_{\leq k+1})^\omega) &= \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(Q^* \cdot (Q_{k+1})^\omega \cap Q^* \cdot (Q_{\leq k+1})^\omega) + \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(Q^* \cdot (Q_{\leq k})^\omega \cap Q^* \cdot (Q_{\leq k+1})^\omega) \\ &= \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(Q^* \cdot (Q_{k+1})^\omega) + \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(Q^* \cdot (Q_{\leq k})^\omega) \\ &= \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(Q^* \cdot (Q_{k+1})^\omega) + \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(\bigcup_{0 \leq i \leq k} Q^* \cdot (Q_{u_i})^\omega \cap Q^* \cdot (Q_{\leq k})^\omega) \\ &= \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(\bigcup_{0 \leq i \leq k+1} Q^* \cdot Q^\omega_{u_i} \cap Q^* \cdot (Q_{\leq k+1})^\omega) \end{aligned}$$

因此，$\mathcal{P}(k+1)$ 成立。事实上，对于所有 $k \leq n$ 都成立。因此，该引理恰好对应于$\mathcal{P}(n)$，因为 $Q_{\leq n} = Q$。 □

B.3 引理的证明3

事实上，我们证明了下面更一般的引理。

引理 7（证明 B.3）设有一个 *PI* 并发博弈 $\mathcal{G}$、一名玩家 A 局部最优策略s_A 以及一名玩家 B 确定性策略 s_B。对于有限序列 $\rho = \rho' \cdot q \in Q^+$:

$$\chi_{\mathcal{G}}(q) \leq \sum_{u \in V_{\mathcal{G}}} u \cdot \mathbb{P}_{\rho}^{\mathsf{s}_\mathsf{A}^{\rho},\mathsf{s}_\mathsf{B}^{\rho}}[Q^* \cdot (Q_u)^\omega]$$

证明 我们考虑局部最优策略 s_A、确定性玩家 B 策略 s_B 以及有限路径 $\rho \in Q^+$。考虑无限马尔可夫链 $\mathcal{M} = (\rho \cdot Q^+, \mathbb{P})$，其转移概率为$\mathbb{P}$: $\rho \cdot Q^+ \times \rho \cdot Q^+ \to [0,1]$ 其中，对所有 $\pi \in Q^+$，有 $\mathbb{P}(\rho \cdot \pi, \rho \cdot \pi \cdot q) := \mathbb{P}^{\mathsf{s}_\mathsf{A},\mathsf{s}_\mathsf{B}}(\rho \cdot \pi, q)$。所有其他概率均为 *0*。概率测度 $\mathbb{P}$ 扩展到从 ρ开始的有限路径、柱集以及任意博雷尔集（特别是 W）。

我们还按以下方式归纳地定义函数 f: $\rho \cdot Q^+ \to \mathbb{N}$: $f(\rho) := 0$并且对所有 $\pi \cdot q \cdot q' \in Q^+$，令 $f(\rho \cdot \pi \cdot q \cdot q') := f(\rho \cdot \pi \cdot q)$ 若 $\chi_{\mathcal{G}}(q) = \chi_{\mathcal{G}}(q')$ 且 $f(\rho \cdot q \cdot q') := f(\rho \cdot q) + 1$，否则取其他值。该函数用于计数值变化的次数。

令 $\varepsilon > 0$。根据引理 *2*，我们有

$$\lim_{n \to \infty} \mathbb{P}(Q^+ \cdot f^{-1}[\{n\}]) = 0$$

考虑满足 $\mathbb{P}(Q^+ \cdot f^{-1}[N+1]) \leq \varepsilon/2$ 的某个 $N \in \mathbb{N}$。对于所有 $0 \leq k \leq N$ 和所有值 $u \in V_{\mathcal{G}}$，我们用 $(Q_u, k) \subseteq \rho \cdot Q^+$ 表示以 Q_u 结尾且关于函数 f 的值为 k 的有限路径集合: (Q_u, k): $= \{\rho \cdot \pi \cdot q \in Q^+ \mid q \in Q_u,\ f(\rho \cdot \pi \cdot q) = k\}$。对于所有 $n \in \mathbb{N}$，我们用 $(Q_u, k)_n$: $= \rho \cdot Q^n \cap (Q_u, k)$ 表示。

对于所有 $0 \leq k \leq N$，我们用 $n_k \in \mathbb{N}$ 表示满足以下条件的一个指标:

$$\mathbb{P}(Q^+ \cdot f^{-1}[k]) \leq \mathbb{P}(Q^{\leq n_k} \cdot f^{-1}[k]) + \frac{\varepsilon}{2 \cdot (N+1) \cdot N_{V_{\mathcal{G}}}}$$

其中 $N_{V_{\mathcal{G}}} := |V_{\mathcal{G}}|$（这存在，因为 $\mathbb{P}(Q^+ \cdot S) = \lim_{n \to \infty} \mathbb{P}(Q^{\leq n} \cdot S)$）。特别地，这意味着 $\mathbb{P}(Q^{>n_k} \cdot f^{-1}[k]) = \mathbb{P}(Q^+ \cdot f^{-1}[k]) - \mathbb{P}(Q^{\leq n_k} \cdot f^{-1}[k]) \leq \varepsilon/(2 \cdot (N+1) \cdot N_V)$。现在，令$n := \max_{0 \leq k \leq N} n_k$。我们来证明，对于所有 $0 \leq k \leq N$ 和 $u \in V_{\mathcal{G}}$，有:

$$\mathbb{P}[(Q_u, k)_n] = \sum_{\pi \in (Q_u,k)_n} \mathbb{P}(\pi) \leq \mathbb{P}(Q^+ \cdot (Q_u, k)^\omega) + \frac{\varepsilon}{2 \cdot (N+1) \cdot N_V} \tag{1}$$

事实上，由于 $n \geq n_k$，我们有:

$$\begin{aligned}
\mathbb{P}(Q^+ \cdot (Q_u, k)^\omega) &\geq \mathbb{P}(Q^{\leq n} \cdot (Q_u, k)^\omega) \\
&= \sum_{\pi \in (Q_u,k)_n} \mathbb{P}(\pi) \cdot \mathbb{P}_\pi((Q_u, k)^\omega) \\
&= \sum_{\pi \in (Q_u,k)_n} \mathbb{P}(\pi) \cdot (1 - \mathbb{P}_\pi(Q^+ \cdot f^{-1}[k+1])) \\
&= \sum_{\pi \in (Q_u,k)_n} \mathbb{P}(\pi) - \sum_{\pi \in (Q_u,k)_n} \mathbb{P}(\pi) \cdot \mathbb{P}_\pi(Q^+ \cdot f^{-1}[k+1])) \\
&\geq \sum_{\pi \in (Q_u,k)_n} \mathbb{P}(\pi) - \mathbb{P}_{>n}(Q^+ \cdot f^{-1}[k+1])) \\
&\geq \sum_{\pi \in (Q_u,k)_n} \mathbb{P}(\pi) - \frac{\varepsilon}{2 \cdot (N+1) \cdot N_V}
\end{aligned}$$

我们得到公式（*1*）。下文中，我们将赋值 $\chi_{\mathcal{G}}$ 记为 $v : Q \to [0,1]$。下面用归纳法证明关于 i 的以下性质 $\mathcal{P}(i)$：“$\sum_{\pi\in\rho\cdot Q^i}\mathbb{P}(\pi)\cdot v(\pi) \geq v(\rho)$”，其中$v(\pi) \in [0,1]$ 指的是 π的最后一个状态对应的 $v(q)$。性质 $\mathcal{P}(0)$ 显然成立。现在假设该性质对某个 $i \in \mathbb{N}$成立。根据命题 *3*，且由于 s_{A} 是局部最优的，我们有：

$$\begin{aligned}
\sum_{\pi\in\rho\cdot Q^{i+1}}\mathbb{P}(\pi)\cdot v(\pi) &= \sum_{\pi=\rho\cdot q_1\cdots q_i\cdot q_{i+1}\in\rho\cdot Q^{i+1}}\mathbb{P}(\pi)\cdot v(q_{i+1})\\
&= \sum_{\pi'=\rho\cdot q_0\cdots q_i\in\rho\cdot Q^i}\sum_{q_{i+1}\in Q}\mathbb{P}(\pi'\cdot q_{i+1})\cdot v(q_{i+1})\\
&= \sum_{\pi'=\rho\cdot q_1\cdots q_i\in\rho\cdot Q^i}\sum_{q_{i+1}\in Q}\mathbb{P}(\pi')\cdot\mathbb{P}(\pi',\pi'\cdot q_{i+1})\cdot v(q_{i+1})\\
&= \sum_{\pi'=\rho\cdot q_0\cdots q_i\in\rho\cdot Q^i}\mathbb{P}(\pi')\cdot\sum_{q_{i+1}\in Q}\mathbb{P}(\pi',\pi'\cdot q_{i+1})\cdot v(q_{i+1})\\
&= \sum_{\pi'=\rho\cdot q_1\cdots q_i\in\rho\cdot Q^i}\mathbb{P}(\pi')\cdot\sum_{q_{i+1}\in Q}\mathbb{P}^{\mathsf{s}_{\mathsf{A}},\mathsf{s}_{\mathsf{B}}}(\pi')[q_{i+1}]\cdot v(q_{i+1})\\
&= \sum_{\pi'=\rho\cdot q_1\cdots q_i\in\rho\cdot Q^i}\mathbb{P}(\pi')\cdot\sum_{q_{i+1}\in Q}\mathbb{P}^{q_i,q_{i+1}}(\mathsf{s}_{\mathsf{A}}(\pi'),\mathsf{s}_{\mathsf{B}}(\pi'))\cdot v(q_{i+1})\\
&= \sum_{\pi'=\rho\cdot q_1\cdots q_i\in\rho\cdot Q^i}\mathbb{P}(\pi')\cdot\mathsf{out}_{\mathcal{F}_{q_i}}(\mathsf{s}_{\mathsf{A}}(\pi'),\mathsf{s}_{\mathsf{B}}(\pi'))\\
&\geq \sum_{\pi'=\rho\cdot q_1\cdots q_i\in\rho\cdot Q^i}\mathbb{P}(\pi')\cdot v(q_i)\\
&\geq v(\rho)
\end{aligned}$$

因此，对于所有 $i \in \mathbb{N}$，性质 $\mathcal{P}(i)$ 均得到保证，特别是 $\mathcal{P}(n)$ 成立。由此可得：

$$\begin{aligned}
\sum_{u\in V_{\mathcal{G}}} u\cdot\mathbb{P}(Q^+\cdot(Q_u)^\omega) &\geq \sum_{u\in V_{\mathcal{G}}} u\cdot\mathbb{P}(\bigcup_{k=0}^{N} Q^+\cdot(Q_u,k)^\omega) && \textit{since } \bigcup_{k=0}^{N} Q^+\cdot(Q_u,k)^\omega \subseteq Q^+\cdot Q_u^\omega\\
&\geq \sum_{k=0}^{N}\sum_{u\in V_{\mathcal{G}}} u\cdot\mathbb{P}(Q^+\cdot(Q_u,k)^\omega) && \textit{since } (Q_u,k)^\omega\cap(Q_u,j)^\omega=\emptyset \textit{ if } k\neq j\\
&\geq \sum_{k=0}^{N}\sum_{u\in V_{\mathcal{G}}} u\cdot(\sum_{\pi\in(Q_u,k)_n}\mathbb{P}(\pi)-\frac{\varepsilon}{2\cdot(N+1)\cdot N_{V_{\mathcal{G}}}}) && \textit{by Equation (1)}\\
&\geq \sum_{k=0}^{N}\sum_{u\in V}\sum_{\pi\in(Q_u,k)_n}\mathbb{P}(\pi)\cdot v(\pi)-\frac{\varepsilon}{2} && \textit{since } u=v(\pi),\forall\pi\in(Q_u,k)_n\\
&= \sum_{k=0}^{N}\sum_{\pi\in\rho\cdot Q^n\cap f^{-1}[k]}\mathbb{P}(\pi)\cdot v(\pi)-\frac{\varepsilon}{2} && \textit{since } \rho\cdot Q^n\cap f^{-1}[k]=\uplus_{u\in V_{\mathcal{G}}}(Q_u,k)_n\\
&\geq \sum_{\pi\in\rho\cdot Q^n}\mathbb{P}(\pi)\cdot v(\pi)-\varepsilon && \textit{since } \mathbb{P}(Q^+\cdot f^{-1}[N+1])\leq\frac{\varepsilon}{2}\\
&\geq v(\rho)-\varepsilon && \textit{by } \mathcal{P}(n)
\end{aligned}$$

由于这对所有 $\varepsilon > 0$ 都成立，我们得到：$\sum_{u\in V_{\mathcal{G}}} u\cdot\mathbb{P}(Q^+\cdot(Q_u)^\omega) \geq v(\rho) = \chi_{\mathcal{G}}(q)$（对于 q ρ 的最后一个状态）。 □

引理 3 因此是一个直接推论。

B.4 马尔可夫链上的 Levy 的 0-1 定律

下面给出 Levy 的 0-1 定律在具有前缀无关目标的无限马尔可夫链情境下的一个推广。该结果源自 [16]中的定理 5：

定理 6（前缀无关目标的 Levy 的 0-1 定律）考虑一个定义在状态集合 Q 上、概率为 $\mathbb{P}$ 的可数马尔可夫链 $\mathcal{M}$。对于所有前缀无关目标 $W \subseteq Q^\omega$，集合 W 和 $\{\rho \in Q^\omega \mid \lim_{n\to\infty} \mathbb{P}_{\rho_n}(W) = 1\}$ 除一个零测集外相等。对于 $Q^\omega \setminus W$ 也成立：集合 $Q^\omega \setminus W$ 和 $\{\rho \in Q^\omega \mid \lim_{n\to\infty} \mathbb{P}_{\rho_n}(W) = 0\}$ 除一个零测集外相等。

引理 4 因此是该定理的直接推论。注意，引理 4 也可由 [7] 中的引理 2 得出。

B.5 定理 1 的证明

证明：根据引理 *1*，第一个性质是必要的。下面证明第二个性质也是必要的。考虑一个玩家 A 策略 s_A，使得存在一条有限路径 $\rho \in Q^+$、一个确定性 B 玩家策略 s_B 和一个正值 $u \in V_{\mathcal{G}} \setminus \{0\}$，满足 $\mathbb{P}^{s^\rho_A, s^\rho_B}[W \cap Q^* \cdot (Q_u)^\omega] < \mathbb{P}^{s^\rho_A, s^\rho_B}[Q^* \cdot (Q_u)^\omega]$。考虑由两个策略 s^ρ_A 和 s^ρ_B 诱导、初始状态为 ρ 的可数马尔可夫链。我们将该马尔可夫链中的概率函数记为 $\mathbb{P}$。考虑前缀无关目标 $W_u := W \cup (Q^* \cdot (Q \setminus Q_u))^\omega$。我们有：

$$\begin{aligned}
\mathbb{P}[W_u] &= \mathbb{P}[W_u \cap Q^* \cdot (Q_u)^\omega] + \mathbb{P}[W_u \cap (Q^* \cdot (Q \setminus Q_u))^\omega] \\
&= \mathbb{P}[W \cap Q^* \cdot (Q_u)^\omega] + \mathbb{P}[(Q^* \cdot (Q \setminus Q_u))^\omega] \\
&< \mathbb{P}[Q^* \cdot (Q_u)^\omega] + \mathbb{P}[(Q^* \cdot (Q \setminus Q_u))^\omega] \\
&= 1
\end{aligned}$$

现在我们来证明，存在一条有限路径，其最后一个状态属于 Q_u，并且从该状态出发，W_u 的概率小于 $u/2$。根据引理*4*，马尔可夫链中存在状态 $\pi \in Q^*$（对应于一条有限路径），使得 $\mathbb{P}_\pi[W_u] < u/2$。如果 $\pi \in Q^* \cdot Q_u$，那么我们就证明了所需结论。现在反设结论不成立，并且对所有路径$\pi' \in Q^* \cdot Q_u$，都有 $\mathbb{P}_{\pi\cdot\pi'}[W_u] \geq u/2$。于是：

$$\begin{aligned}
\mathbb{P}_\pi[W_u] &= \mathbb{P}_\pi[W_u \cap (Q \setminus Q_u)^* \cdot Q_u \cdot Q^\omega] + \mathbb{P}_\pi[W_u \cap (Q \setminus Q_u)^\omega] \\
&= \sum_{\pi' \in (Q\setminus Q_u)^* \cdot Q_u} \mathbb{P}_\pi(\pi') \cdot \mathbb{P}_{\pi\cdot\pi'}[W_u] + \mathbb{P}_\pi[(Q \setminus Q_u)^\omega] \\
&\geq \sum_{\pi' \in (Q\setminus Q_u)^* \cdot Q_u} \mathbb{P}_\pi(\pi') \cdot \frac{u}{2} + \mathbb{P}_\pi[(Q \setminus Q_u)^\omega] \\
&= \mathbb{P}_\pi[(Q \setminus Q_u)^* \cdot Q_u \cdot Q^\omega] \cdot \frac{u}{2} + \mathbb{P}_\pi[(Q \setminus Q_u)^\omega] \\
&\geq \frac{u}{2}
\end{aligned}$$

这是一个矛盾，因为 $\mathbb{P}_\pi[W_u] < u/2$。因此，存在一条路径 $\pi \in Q^* \cdot Q_u$ 使得$\mathbb{P}_\pi[W_u] < u/2$。由于 $W \subseteq W_u$，可得 $\mathbb{P}_\pi[W] \leq \mathbb{P}_\pi[W_u] < u/2$。也就是说，剩余策略 $s^{\rho\cdot\pi}_A$ 从 $\rho \cdot \pi$ 的最后一个状态出发并非最优，换言之，玩家 A 的策略 s_A 并非子博弈最优。

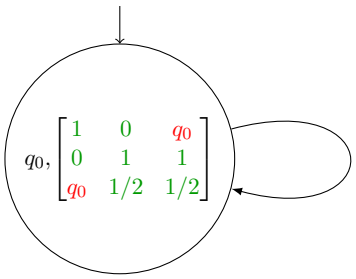


图 10：一个可达性博弈。

现在我们来证明这些条件是充分的。设 $\rho \in Q^+$。我们来证明剩余策略 $\mathsf{s}^\rho_{\mathsf{A}}$ 从 ρ出发是最优的。考虑一个玩家 B 的确定性策略 s_{B}。根据引理2:

$$\begin{aligned}
\mathbb{P}_\rho^{\mathsf{s}^\rho_{\mathsf{A}},\mathsf{s}^\rho_{\mathsf{B}}}[W] &= \mathbb{P}_\rho^{\mathsf{s}^\rho_{\mathsf{A}},\mathsf{s}^\rho_{\mathsf{B}}}[W \cap (\bigcup_{u\in V_{\mathcal{G}}} Q^+ \cdot (Q_u)^\omega)] = \sum_{u\in V_{\mathcal{G}}} \mathbb{P}_\rho^{\mathsf{s}^\rho_{\mathsf{A}},\mathsf{s}^\rho_{\mathsf{B}}}[W \cap Q^+ \cdot (Q_u)^\omega] \\
&\geq \sum_{u\in V_{\mathcal{G}}\setminus\{0\}} u \cdot \mathbb{P}_\rho^{\mathsf{s}^\rho_{\mathsf{A}},\mathsf{s}^\rho_{\mathsf{B}}}[W \cap Q^+ \cdot (Q_u)^\omega] \\
&\geq \sum_{u\in V_{\mathcal{G}}\setminus\{0\}} u \cdot \mathbb{P}_\rho^{\mathsf{s}^\rho_{\mathsf{A}},\mathsf{s}^\rho_{\mathsf{B}}}[Q^+ \cdot (Q_u)^\omega] + 0 \cdot \mathbb{P}_\rho^{\mathsf{s}^\rho_{\mathsf{A}},\mathsf{s}^\rho_{\mathsf{B}}}[Q^+ \cdot (Q_0)^\omega] \\
&= \sum_{u\in V_{\mathcal{G}}} u \cdot \mathbb{P}_\rho^{\mathsf{s}^\rho_{\mathsf{A}},\mathsf{s}^\rho_{\mathsf{B}}}[Q^+ \cdot (Q_u)^\omega] \geq \chi_{\mathcal{G}}(q)
\end{aligned}$$

最后一个不等式由引理 7得出。因此，玩家 A 的剩余策略$\mathsf{s}^\rho_{\mathsf{A}}$ 从 ρ 出发是最优的。□

C 第6节的证明

C.1 （正）最优但局部最优且非子博弈最优的策略

考虑图 10中的博弈：这是一个可达性博弈，也就是说，如果它在 q_0上无限循环，则值为 0。状态 q_0 的值为 $1/2$，玩家 A 采用位置策略，以概率 $1/2$ 选择顶端两行即可达到该值；玩家 B 采用位置策略，以概率 $1/2$选择两列也可达到该值。

然而，记 a_1，a_2 和 a_3 为状态 q_0 下玩家 A 可选的三个动作（从上到下），考虑以下玩家 A 策略 s_{A}，满足 $\mathsf{s}_{\mathsf{A}}(q_0)(a_1) = \mathsf{s}_{\mathsf{A}}(q_0)(a_2) := 1/2$且 $\mathsf{s}_{\mathsf{A}}(q_0^{n+1})(a_3) := 1$ 对所有 $n \geq 1$均成立。那么，该策略是局部最优的，也是最优的。事实上，如果博弈在 q_0上至少循环一次，那么在 q_0上循环和达到结果 1 的概率相同。因此，值的平均值至少为 $1/2$，这正是状态 q_0的值。然而，该策略并非子博弈最优，因为博弈在 q_0上循环一次后，玩家 B 可以通过始终以概率 1 选择左列来确保值为 0。

类似地，考虑一个玩家 A 策略 s_{A}，使得 $\mathsf{s}_{\mathsf{A}}(q_0)(a_1) = \mathsf{s}_{\mathsf{A}}(q_0)(a_2) := 1/4$，$\mathsf{s}_{\mathsf{A}}(q_0)(a_3) :=1/2$ 且对于所有 $n \geq 1$，$\mathsf{s}_{\mathsf{A}}(q_0^{n+1})(a_3) := 1$ 。那么，该策略是局部最优且正概率最优的，但并非最优，因为该策略的值为 $1/4$；玩家 B 可以通过以概率 1 无限期选择左列来确保这一点。

C.2 新博弈的形式化定义

C.2.1 所有状态的值均为 1 的博弈

命题 6 考虑一个 *PI* 目标 $\emptyset \subsetneq W \subseteq \mathsf{K}^\omega$。那么，存在一个并发博弈$\mathcal{G}_W = \langle \mathcal{C}_W, W \rangle$，其中所有状态的值均为 1。

证明 考虑并发竞技场 $\mathcal{C}_W = \langle Q, (A_q)_{q \in Q}, (B_q)_{q \in Q}, \mathsf{D}, \delta, \text{距离函数}, \mathsf{K}, \text{颜色函数} \rangle$，其中玩家 A 可以单独选择在颜色属于 K 的状态中所见状态的序列，即：

- $Q :=$ K，对于所有 $q \in Q$，均有 $A_q :=$ K 且 $B_q := \{b\}$；
- $\mathsf{D} := \{d_q \mid q \in Q\}$，并且对于所有 $q \in Q$，均有 距离函数$(d_q)(q) = 1$；
- 对于所有 $q \in Q$ 和 $q' \in A_q =$ K，均有 $\delta(q, q', b) := d_{q'}$；
- 对于所有 $q \in Q$，我们都有 颜色函数$(q) := q$。

事实上，这个博弈是轮流制的，因为只有玩家 A 的动作会影响状态序列。她可以准确决定所观察到的状态序列——以及颜色序列。因此，她可以严格遵循一条路径 $\rho \in W$（首个颜色除外，因为它无关紧要，因为 W 是前缀无关的），所以状态的值为 *1*。 □

C.2.2 博弈 $\mathcal{G}_u$ 的定义

定义 23 对于一个正值 $u \in V_{\mathcal{G}} \setminus \{0\}$，我们定义博弈 $\mathcal{G}_u = \langle \mathcal{C}_u, W \rangle$，其包含$\mathcal{C}_u = \langle Q'_u, (A'_q)_{q \in Q'_u}, (B'_q)_{q \in Q'_u}, \mathsf{D}\ \delta', \text{距离函数}', \mathsf{K}, \text{颜色函数} \rangle$，其中我们将博弈场 $\mathcal{C}_W$称为汇点状态，并且有若干条转移指向该状态：

- $Q := Q_u \cup \mathcal{C}_W$；
- 对于所有 $q \in Q_u$，都有 $A'_q := A_{\mathcal{F}^{\text{nfq}}}$ 以及 $B'_q := B_q$；
- $\mathsf{D}' := \{d_{q,\sigma_q,b} \mid q \in Q_u,\ \sigma_q \in A_q, b \in B\}$；
- 对于所有 $q \in Q_u$ 和 $\sigma_q \in A_q = S_{\mathcal{F}^{\text{nfq}}}$，都有：$\delta'(q, \sigma_q, b) := d_{q,\ \sigma_q,\ b}$；
- 对于所有 $q \in Q_u$，$\sigma_q = \sum_{a \in A} \sigma_q(a) \cdot \mathbb{1}_{\{a\}} \in A'_q$，$b \in B'_q$ 和 $q' \in Q$，都有：

$$\text{dist}'(d_{q,\sigma_q,b})(q') := \begin{cases} \sum_{a \in A} \sigma_q(a) \cdot \text{dist}(\delta(q,a,b))(q') & \text{if } q' \in Q_u \\ \sum_{a \in A} \sigma_q(a) \cdot \text{dist}(\delta(q,a,b))[Q \setminus Q_u] & \text{if } q' = \mathcal{C}_W \end{cases}$$

C.3 引理5的证明

事实上，我们证明了一个更强的结果。

引理 8 考虑某个 $u \in V_{\mathcal{G}} \setminus \{0\}$ 以及博弈 $\mathcal{G}_u$。再考虑目标（该目标不是前缀无关的）$W_u := W \cup Q^* \cdot (Q \setminus Q_u) \cdot Q^\omega \subseteq Q^\omega$，并假设在$\mathcal{G}$ 中存在一个局部最优策略，该策略从Q_u中的所有状态出发，针对目标 W_u 都是正最优的。那么，$\mathcal{G}_u$ 中所有状态的值均为正：$\chi_{\mathcal{G}_u}(q) > 0$，对于所有属于 $\mathcal{G}_u$ 的 q。

这个引理的证明相当直接，但篇幅较长。这是由博弈$\mathcal{G}_u$ 的形式化定义所致。

证明 首先注意，对于目标 W_u ， $\mathcal{G}_u$ 中所有状态的值都至少为u ，因为 $W \subseteq W_u$。现在考虑一个局部最优策略 s_{A} ，它从 Q_u 中的所有状态出发，对于目标 W_u 都是正最优的；令 $r = \min_{q \in Q_u} \chi_{\langle \mathcal{C}, W_u \rangle}[s_{\mathsf{A}}](q) > 0$表示策略 s_{A} 在 Q_u 中针对目标 Q_u所得值的最小值。假设状态 $q \in Q_u$ 在博弈 $\mathcal{G}_u$ 中（针对目标 W）的值为 0，并考虑一个玩家 B 策略 s_{B} ，使其在博弈 $\mathcal{G}_u$ 中从状态 q出发的值至多为 $r/2 > 0$ ，即$\chi_{\mathcal{G}_u}[s_{\mathsf{B}}](q) \leq r/2$ ，针对 W。注意，玩家 B 策略 s_{B} 既是 $\mathcal{C}$ 中的策略，也是 $\mathcal{C}_u$中的策略。对于所有 $\rho = \rho' \cdot q \in (Q_u)^+$，我们可以将 $s_{\mathsf{A}}(\rho)$写成 A'_q中各元素的凸组合：$s_{\mathsf{A}}(\rho) = \sum_{\sigma \in A'_q} \mu_\sigma(\rho) \cdot \sigma$。现在定义博弈 $\mathcal{G}_u$ 中的玩家 A 策略 s'_{A}，使其模拟策略 s_{A}；也就是说，对于所有 $\rho = \rho' \cdot q \in (Q_u)^+$ 和 $\sigma \in A'_q$，我们设定：$s'_{\mathsf{A}}(\rho)(\sigma) := \mu_\sigma(\rho)$。当博弈进入 $\mathcal{C}_W$时，策略 s'_{A} 就切换为确定性策略。因此，由于目标 W是前缀无关的，对于所有 $\rho \in (Q_u)^* \cdot \mathcal{C}_W$，都有 $\chi_{\mathcal{G}_u}(s'_{\mathsf{A}})(\rho) = 1$。我们用 $\mathbb{P}^{s_{\mathsf{A}}, s_{\mathsf{B}}}$ 表示策略 s_{A}、s_{B} 在 $\mathcal{G}$ 中从状态 q 出发所诱导的概率测度，并用 $\mathbb{P}^{s'_{\mathsf{A}}, s_{\mathsf{B}}, u}$ 表示 s'_{A}、s_{B} 在 $\mathcal{G}_u$ 中从状态 q出发所诱导的概率测度。

现在让我们通过对 n 归纳法来证明性质 $\mathcal{P}(n)$：对于所有 $\rho \in (Q_u)^n$，都有 $\mathbb{P}^{s_{\mathsf{A}}, s_{\mathsf{B}}}(\rho) = \mathbb{P}^{s'_{\mathsf{A}}, s_{\mathsf{B}}, u}(\rho)$，且 $\mathbb{P}^{s_{\mathsf{A}}, s_{\mathsf{B}}}(Q_u^{\leq n} \cdot (Q \setminus Q_u)) = \mathbb{P}^{s'_{\mathsf{A}}, s_{\mathsf{B}}, u}(Q_u^{\leq n} \cdot \mathcal{C}_W)$。这对 $n = 1$显然成立。现在假设这对某个 $n - 1 \geq 0$成立。考虑某个 $\rho = \rho' \cdot q = \rho'' \cdot q' \cdot q \in (Q_u)^n \cdot Q$。我们有：

$$
\begin{aligned}
\mathbb{P}^{s_{\mathsf{A}}, s_{\mathsf{B}}}(\rho) &= \mathbb{P}^{s_{\mathsf{A}}, s_{\mathsf{B}}}(\rho') \cdot \mathbb{P}^{s_{\mathsf{A}}, s_{\mathsf{B}}}_{\rho'}(q) \\
&= \mathbb{P}^{s_{\mathsf{A}}, s_{\mathsf{B}}, u}(\rho') \cdot \left(\sum_{a \in A} \sum_{b \in B} s_{\mathsf{A}}(\rho')(a) \cdot s_{\mathsf{B}}(\rho')(b) \cdot \mathsf{dist}(\delta(q', a, b))(q) \right) \\
&= \mathbb{P}^{s_{\mathsf{A}}, s_{\mathsf{B}}, u}(\rho') \cdot \sum_{b \in B} s_{\mathsf{B}}(\rho')(b) \cdot \left(\sum_{a \in A} s_{\mathsf{A}}(\rho')(a) \cdot \mathsf{dist}(\delta(q', a, b))(q) \right) \\
&= \mathbb{P}^{s_{\mathsf{A}}, s_{\mathsf{B}}, u}(\rho') \cdot \sum_{b \in B} s_{\mathsf{B}}(\rho')(b) \cdot \left(\sum_{a \in A} \sum_{\sigma \in A'_q} \mu_\sigma(\rho') \cdot \sigma(a) \cdot \mathsf{dist}(\delta(q', a, b))(q) \right) \\
&= \mathbb{P}^{s_{\mathsf{A}}, s_{\mathsf{B}}, u}(\rho') \cdot \sum_{b \in B} s_{\mathsf{B}}(\rho')(b) \cdot \left(\sum_{\sigma \in A'_q} \mu_\sigma(\rho') \cdot \sum_{a \in A} \sigma(a) \cdot \mathsf{dist}(\delta(q', a, b))(q) \right)
\end{aligned}
$$

若 $q \in Q_u$，则可得：

$$
\begin{aligned}
\mathbb{P}^{s_{\mathsf{A}}, s_{\mathsf{B}}}(\rho) &= \mathbb{P}^{s_{\mathsf{A}}, s_{\mathsf{B}}, u}(\rho') \cdot \sum_{b \in B} s_{\mathsf{B}}(\rho')(b) \cdot \sum_{\sigma \in A'_q} \mu_\sigma(\rho') \cdot \mathsf{dist}'(d_{q'}, \sigma_q, b)(q) \\
&= \mathbb{P}^{s_{\mathsf{A}}, s_{\mathsf{B}}, u}(\rho') \cdot \sum_{b \in B} s_{\mathsf{B}}(\rho')(b) \cdot \sum_{\sigma \in A'_q} s'_{\mathsf{A}}(\rho')(\sigma) \cdot \mathsf{dist}'(d_{q'}, \sigma, b)(q) \\
&= \mathbb{P}^{s_{\mathsf{A}}, s_{\mathsf{B}}, u}(\rho') \cdot \sum_{\sigma \in A'_q} \sum_{b \in B} s'_{\mathsf{A}}(\rho')(\sigma) \cdot s_{\mathsf{B}}(\rho')(b) \cdot \mathsf{dist}'(\delta'(q', \sigma, b))(q) \\
&= \mathbb{P}^{s_{\mathsf{A}}, s_{\mathsf{B}}, u}(\rho') \cdot \mathbb{P}^{s'_{\mathsf{A}}, s_{\mathsf{B}}, u}_{\rho'}(q) \\
&= \mathbb{P}^{s'_{\mathsf{A}}, s_{\mathsf{B}}, u}(\rho)
\end{aligned}
$$

此外，我们还有：

$$\begin{aligned}
\mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}[Q_u^n \cdot (Q \setminus Q_u)] &= \sum_{\rho' \in (Q_u)^n} \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(\rho') \cdot \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}_{\rho'}[Q \setminus Q_u] \\
&= \sum_{\rho' \in (Q_u)^n} \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(\rho') \cdot \sum_{q \in Q \setminus Q_u} \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}_{\rho'}(q) \\
&= \sum_{\rho' \in (Q_u)^n} \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(\rho') \cdot \sum_{q \in Q \setminus Q_u} \sum_{b \in B} \mathsf{s_B}(\rho')(b) \cdot \left(\sum_{\sigma \in A_q} \mu_\sigma(\rho') \cdot \sum_{a \in A} \sigma(a) \cdot \mathsf{dist}(\delta(q', a, b))(q) \right) \\
&= \sum_{\rho' \in (Q_u)^n} \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(\rho') \cdot \sum_{b \in B} \mathsf{s_B}(\rho')(b) \cdot \left(\sum_{\sigma \in A_q} \mu_\sigma(\rho') \cdot \sum_{a \in A} \sigma(a) \cdot \sum_{q \in Q \setminus Q_u} \mathsf{dist}(\delta(q', a, b))(q) \right) \\
&= \sum_{\rho' \in (Q_u)^n} \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(\rho') \cdot \sum_{b \in B} \mathsf{s_B}(\rho')(b) \cdot \left(\sum_{\sigma \in A_q} \mu_\sigma(\rho') \cdot \sum_{a \in A} \sigma(a) \cdot \mathsf{dist}(\delta(q', a, b))[Q \setminus Q_u] \right) \\
&= \sum_{\rho' \in (Q_u)^n} \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(\rho') \cdot \sum_{b \in B} \mathsf{s_B}(\rho')(b) \cdot \left(\sum_{\sigma \in A_q} \mu_\sigma(\rho') \cdot \mathsf{dist}'(d_{q',\sigma,b})[\mathcal{C}_W] \right) \\
&= \sum_{\rho' \in (Q_u)^n} \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(\rho') \cdot \sum_{\sigma \in A'_q} \sum_{b \in B} \mathsf{s}'_{\mathsf{A}}(\rho')(\sigma) \cdot \mathsf{s_B}(\rho')(b) \cdot \mathsf{dist}'(\delta(q', \sigma, b))[\mathcal{C}_W] \\
&= \sum_{\rho' \in (Q_u)^n} \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}(\rho') \cdot \mathbb{P}^{\mathsf{s}'_{\mathsf{A}},\mathsf{s_B},u}_{\rho'}[\mathcal{C}_W] \\
&= \mathbb{P}^{\mathsf{s}'_{\mathsf{A}},\mathsf{s_B},u}[Q_u^n \cdot \mathcal{C}_W]
\end{aligned}$$

由此可得：

$$\begin{aligned}
\mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}[(Q_u)^{\leq n} \cdot (Q \setminus Q_u)] &= \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}[(Q_u)^{\leq n-1} \cdot (Q \setminus Q_u)] + \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}[(Q_u)^n \cdot (Q \setminus Q_u)] \\
&= \mathbb{P}^{\mathsf{s}'_{\mathsf{A}},\mathsf{s_B},u}[(Q_u)^{\leq n-1} \mathcal{C}_W] + \mathbb{P}^{\mathsf{s}'_{\mathsf{A}},\mathsf{s_B},u}[(Q_u)^n \cdot \mathcal{C}_W] \\
&= \mathbb{P}^{\mathsf{s}'_{\mathsf{A}},\mathsf{s_B},u}[(Q_u)^{\leq n} \cdot \mathcal{C}_W]
\end{aligned}$$

因此， $\mathcal{P}(n)$ 成立。事实上，这一性质对所有 $n \in \mathbb{N}$ 都成立。由此可得：

- $\mathbb{P}^{\mathsf{s_A},\ \mathsf{s_B}}[(Q_u)^* \cdot (Q \setminus Q_u)] = \mathbb{P}_{\mathsf{s'_A},\mathsf{s_B}\ u}[(Q_u)^* \cdot \mathcal{C}_W]$ ；
- $\mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}[W \cap (Q_u)^\omega] = \mathbb{P}^{\mathsf{s'_A},\mathsf{s_B}\ u}[W \cap (Q_u)^\omega]$.

也就是说：

$$\begin{aligned}
\mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}[W_u] &= \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}[W_u \cap (Q_u)^\omega] + \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}[W_u \cap (Q_u)^* \cdot (Q \setminus Q_u)] \\
&= \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}[W \cap (Q_u)^\omega] + \mathbb{P}^{\mathsf{s_A},\mathsf{s_B}}[(Q_u)^* \cdot (Q \setminus Q_u)] \\
&= \mathbb{P}^{\mathsf{s}'_{\mathsf{A}},\mathsf{s_B},u}[W \cap (Q_u)^\omega] + \mathbb{P}^{\mathsf{s}'_{\mathsf{A}},\mathsf{s_B},u}[(Q_u)^* \cdot \mathcal{C}_W] \\
&= \mathbb{P}^{\mathsf{s}'_{\mathsf{A}},\mathsf{s_B},u}[W \cap (Q_u)^\omega] + \mathbb{P}^{\mathsf{s}'_{\mathsf{A}},\mathsf{s_B},u}[W \cap (Q_u)^* \cdot \mathcal{C}_W] \\
&= \mathbb{P}^{\mathsf{s}'_{\mathsf{A}},\mathsf{s_B},u}[W] \leq r/2 < r
\end{aligned}$$

根据玩家B 策略$\mathsf{s_B}$的选择，不等式成立。该结果与 r的定义矛盾：相对于策略$\mathsf{s_A}$和目标 W_u，Q_u 中各状态的值的最小值。事实上，不存在这样一个状态 $q \in Q_u$，使其在 $\mathcal{G}_u$ 中的值为 0。引理得证。 □

C.4 引理的证明6

证明 我们将策略$\mathsf{s_A}[(\mathsf{s_A^u})_{u\in V_\mathcal{G}\setminus\{0\}}]$记为$\mathsf{s_A}$，并应用定理*1*来证明它是子博弈最优的。根据博弈$\mathcal{G}_u\mathsf{s_A}$在$Q_u$中所有正$u\in V_\mathcal{G}\setminus\{0\}$状态处均为局部最优，这是由博弈$\mathcal{G}_u$的定义得出的。此外，由于所有状态$q\in Q_0$的值均为*0*，$Q_0$中的标准形博弈的值也为*0*。因此，对于所有状态$q\in Q_0$，标准形博弈$\mathcal{F}_q$中的任意策略在标准形博弈$\mathcal{F}_q$中都是最优的。也就是说，策略$\mathsf{s_A}$是局部最优的。

现在我们证明它保证了定理*1*的第二个性质。令$u\in V_\mathcal{G}\setminus\{0\}$。考虑确定性玩家*B*策略$s_\mathsf{B}$。考虑有限路径$\rho\in Q^+$和$\pi=\pi'\cdot q\in Q^*\cdot Q_u$。假设$\mathbb{P}^{\mathsf{s_A},s_{\rho\cdot\pi}\,\mathsf{B}}((Q_u)^\omega)>0$。那么，由于策略$s_\mathsf{A}$此时的行为与$s^u_\mathsf{A}$相同，并且这对$s^u_\mathsf{A}$也成立（因为它在$\mathcal{G}_u$中是子博弈几乎必然获胜的，且目标$W$是前缀无关的），我们有$\mathbb{P}^{\mathsf{s_A},s_{\rho\cdot\pi}\,\mathsf{B}}(W\mid(Q_u)^\omega)=1$。也就是说，$\mathbb{P}^{\mathsf{s_A},s_{\rho\cdot\pi}\,\mathsf{B}}(W\cap(Q_u)^\omega)=\mathbb{P}^{\mathsf{s_A},s_{\rho\cdot\pi}\,\mathsf{B}}((Q_u)^\omega)$。由于这对所有$\pi\in Q^*\cdot Q_u$都成立，因此可得$\mathbb{P}^{\mathsf{s_A},s_\rho\,\mathsf{B}}(W\cap Q^*\cdot(Q_u)^\omega)=\mathbb{P}^{\mathsf{s_A},s_\rho\,\mathsf{B}}(Q^*\cdot(Q_u)^\omega)$。

我们可以将应用定理*1*来得到该定理。 □

C.5 扩展到有限记忆

我们首先回顾什么是记忆骨架（例如，参见[6]），并说明它如何实现有限记忆策略。对于颜色集合K和状态集合Q，K上的记忆骨架是一个三元组$\mathcal{M}=\langle M,m_{\text{init}},\mu\rangle$，其中$M$是一个非空集合，称为记忆；$m_{\text{init}}\in M$是记忆的初始状态；$\mu\colon M\times\mathsf{K}\to M$是更新函数。注意，更新函数$\mu$可以归纳地扩展为函数$\mu^*\colon M\times\mathsf{K}^*\to M$，方式如下：对于所有$m\in M$，$\mu^*(m,\epsilon):=m$，以及所有$\rho\cdot k\in\mathsf{K}^+$，$\mu^*(m,\rho\cdot k):=\mu(\mu^*(m,\rho)\ k)$。给定每个$q\in Q$对应的动作集合$A_q$，带记忆的动作映射$M$是一个函数$\lambda\colon M\times Q\to\mathcal{D}(\cup_{q\in Q}A_q)$，使得对于所有$q\in Q$和$m\in M$，都有$\lambda(m,q)\in\mathcal{D}(A_q)$。事实上，记忆骨架和动作映射共同实现一个策略。

定义 24（有限记忆策略的实现）考虑一个定义在颜色集合K上的并发竞技场$\mathcal{C}$。定义在K上的记忆骨架$\mathcal{M}=\langle M,m_{\text{init}},\mu\rangle$和动作映射$\lambda\colon M\times Q\to\mathcal{D}(\cup_{q\in Q}A_q)$实现了策略$\mathsf{s_A}\colon Q^+\to\mathcal{D}(\cup_{q\in A_q}A_q)$，其定义如下：对所有$\rho=\rho'\cdot q\in Q^+$，$\mathsf{s_A}(\rho):=\lambda(\mu^*(m_{\text{init}},\text{颜色函数}(\rho))\ q)\in\mathcal{D}(A_q)$。

若存在一个记忆骨架$\mathcal{M}=\langle M,m_{\text{init}},\mu\rangle$（其中$M$有限）以及一个实现策略$\mathsf{s_A}$的动作映射$\lambda$，则称策略$\mathsf{s_A}$具有有限记忆。如果$M$是单元素集合，则可得到位置策略的定义。

我们可以将几乎必然可获胜目标的位置型定义推广到具有固定记忆骨架时可获胜的目标。

定义 25（$\mathcal{M}$-几乎必然可获胜目标）考虑颜色集合K上的博雷尔前缀无关目标$W\subseteq\mathsf{K}^\omega$以及一个记忆骨架$\mathcal{M}$。若对于所有颜色集合$\mathsf{K}$上的有限博弈$\mathcal{G}=\langle\mathcal{C},W\rangle$，只要存在子博弈几乎必胜策略，就存在一个可用$\mathcal{M}$实现的策略，则称其为$\mathcal{M}$-几乎必然可获胜（简称$\mathcal{M}$-*ASW*）。

现在我们调整定理 2 以用于有限记忆的迁移。

定理 7 考虑一个非空有限颜色集K和一个*PI*目标$\emptyset\subsetneq W\subseteq\mathsf{K}^\omega$。考虑一个目标为$W$的*PI*并发博弈$\mathcal{G}$。对于某个记忆骨架$\mathcal{M}$，如果目标$W$是$\mathcal{M}$-*ASW*，并且$\mathcal{G}$中存在子博弈最优策略，那么就存在一个可由$\mathcal{M}$实现的子博弈最优策略。

我们来定义给定记忆骨架 $\mathcal{M}$的相关记忆状态集合，也就是从 m_{init} 出发，经过任意有限颜色序列后可能到达的记忆状态子集。即:

定义 26（相关记忆状态）考虑颜色集合 K上的记忆骨架 $\mathcal{M} = \langle M, m_{\text{init}}, \mu\rangle$。相关记忆状态 $M_r \subseteq M$ 的集合定义为 $M_r := \{m \in M \mid \exists \rho \in \mathsf{K}^*,\ m = \mu^*(m_{\text{init}}, \rho)\}$。对于所有此类相关状态 $m \in M_r$，我们将记忆骨架 $\mathcal{M}^m := \langle M, m, \mu\rangle$记为 $\mathcal{M}^m$ 。

于是，我们有如下命题:

命题 7考虑颜色集合 K、K 上的记忆骨架 $\mathcal{M} = \langle M, m_{\text{init}}, \mu\rangle$ 以及相应的相关记忆状态集合 M_r。另考虑一个 $\mathcal{M}$-*ASW* 目标 $W \subseteq K^\omega$。那么，对于颜色集合 K 上的所有有限博弈 $\mathcal{G} = \langle \mathcal{C}, W\rangle$，只要存在子博弈几乎必然获胜策略，就存在一个动作映射 $\lambda:\ M \times Q \to \mathcal{D}(\cup_{q\in Q} A_q)$ ，使得对于所有相关状态 $m \in M_r$，由 $\mathcal{M}^m$ 和 λ 实现的策略在 $\mathcal{G}$ 中都是子博弈几乎必然获胜策略。

证明 考虑颜色集合 K，一个定义在 K 上的记忆骨架 $\mathcal{M} = \langle M, m_{\text{init}}, \mu\rangle$ 以及相关记忆状态集合 M_r。再考虑 $\mathcal{M}$-*ASW* 目标 $W \subseteq K^\omega$。另考虑一个定义在颜色集合 K 上的有限博弈 $\mathcal{G} = \langle \mathcal{C}, W\rangle$，其中存在一个子博弈几乎必胜策略。我们构造一个博弈 $\mathcal{G}^{\mathsf{B}}$ ，在其中玩家 B 先进行有限回合，以"选择"玩家 A 策略的记忆状态，然后博弈按 $\mathcal{G}$中的方式继续。更具体地，对于 $n := |M_r|$，并记 Q 为 $\mathcal{G}$中的状态集合，我们考虑有限博弈场 $\mathcal{C}^{\mathsf{B}} = \langle Q^{\mathsf{B}}, (A_q)_{q\in Q^{\mathsf{B}}}、(B_q)_{q\in Q^{\mathsf{B}}}、\mathsf{D}\ \delta、\mathsf{dist}、\mathsf{K}、\mathsf{col}\rangle$，其定义如下:

- $Q_{\mathsf{B}} := \mathsf{K} \times \{0, \ldots, n\}$ 仅玩家B 会行动的状态集合;
- 对于所有 $q = (k, i) \in Q_{\mathsf{B}}$，我们有 $A'_q := \{a\}$ 对于任意动作 a 以及 $B_q := \mathsf{Cont}_q \uplus \mathsf{Exit}_q$，其中 $\mathsf{Cont}_q := \{b_{k,i+1} \mid k \in \mathsf{K}\}$ 若 $i < n$，而$\mathsf{Cont}_q := \emptyset$ 则为其他情况。此外，$\mathsf{Exit}_q := \{b_{q'} \mid q' \in Q\}$;
- $\mathsf{D} := \{d_{k,i+1} \mid k \in \mathsf{K}\ \ 0 \le i \le n-1\} \uplus \{d_{q'} \mid q' \in Q\}$;
- 对于所有 $q = (k, i) \in Q_{\mathsf{B}}$ 和 $k \in \mathsf{K}$，我们有: $\delta(q, a, b_{k,\ i+1}) := d_{k,\ i+1}$ 并且对于所有 $q' \in Q$，我们有 $\delta(q, a, b_{q'}) := d_{q'}$;
- 对于所有 $k \in \mathsf{K}$ 和 $0 \le i \le n-1$，我们有 距离函数$(d_{k,i+1})((k, i+1)) := 1$ 并且对于所有 $q \in Q$距离函数$(d_q)(q) := 1$;
- 对于所有 $k \in \mathsf{K}$ 和 $0 \le i \le n$，我们有 颜色函数$((k, i)) := k$。

博弈场 $\mathcal{C}^{\mathsf{B}}$ 接入博弈场 $\mathcal{C}$ 之前（从而得到博弈场 $\mathcal{C}^{\mathsf{B}} \cdot \mathcal{C}$），最终通过自然状态 d_q 进入该博弈场 $\mathcal{C}$，进入时为 $q \in Q$ （这些状态必须在某个时刻被选中：对于所有 $k \in K$，它们是状态 (k, n) 的唯一后继）。我们声称，在博弈 $\mathcal{G}^{\mathsf{B}} := \langle \mathcal{C}^{\mathsf{B}} \cdot \mathcal{C}, W\rangle$ 中仍存在一个子博弈几乎必胜策略。这是因为在博弈 $\mathcal{G}$ 中情况如此，目标 W 是前缀无关的，而且博弈只能在 $\mathcal{C}^{\mathsf{B}}$ 中停留有限个回合。因此，由于目标 W 是 $\mathcal{M}$-*ASW*，存在一个动作映射 $\lambda: M \times Q^{\mathsf{B}} \cup Q \to \mathcal{D}(\cup_{q\in Q^{\mathsf{B}}\cup Q} A_q)$，使得由 M 和 λ 实现的策略 s_{A} 在 $\mathcal{G}^B$ 中是子博弈几乎必胜的。特别是，玩家 B 可以在博弈场 $\mathcal{C}^{\mathsf{B}}$ 中采取行动，使得当她离开该博弈场以"开始博弈 $\mathcal{G}$"时，玩家 A 策略 s_{A} 的记忆状态可以是任意相关记忆状态 $m \in M_r$。（注意，玩家 B 确实可以到达任意相关记忆状态，因为记忆骨架 $\langle M_r, m_{\text{init}}, \mu\rangle$

可以看作一个大小为 $|M_r| = n$ 的强连通图。）也就是说，对于所有相关记忆状态 $m \in M_r$，只要存在一个由颜色构成的有限序列 $\rho \in \mathsf{K}^k$，对应于 $k \leq n+1$，使得 $m = \mu^*(m_{\text{init}}, \rho)$，由 $\mathcal{M}^m$ 和 λ 实现的策略实际上就是剩余策略 $\mathsf{s}^{(\rho_0,0)\cdots(\rho_k,k)}_{\mathsf{A}}$，而且它也是子博弈最优的。

□

现在我们可以开始证明定理7。

证明 设 K 是一个非空有限颜色集，且 $\emptyset \subsetneq W \subseteq K^\omega$ 是一个 *PI* 目标。考虑一个目标为 W 的并发博弈 $\mathcal{G}$。假设目标 W 对某个记忆骨架 $\mathcal{M}$ 是 $\mathcal{M}$-*ASW*，并且 $\mathcal{G}$ 中存在一个子博弈最优策略。我们采用与定理 *2* 证明中相同的构造。不过，对于所有 $u \in V_{\mathcal{G}} \setminus \{0\}$，由于博弈 $\mathcal{G}_u$ 中存在一个子博弈几乎必胜策略，根据命题 *7*，存在一个动作映射 $\lambda_u : M \times Q_u \to \mathcal{D}(\cup_{q \in Q_u} A_q)$，使得对于所有相关状态 $m \in M_r$，由 $\mathcal{M}^m$ 和 λ 实现的策略在 $\mathcal{G}_u$ 中都是子博弈几乎必胜的。接着，我们将这些策略拼接成动作映射 $\lambda : M \times Q \to \mathcal{D}(\cup_{q \in Q} A_q)$，使得对于所有 $\rho = \rho' \cdot q \in Q^+$，$\lambda(m, q) := \lambda_{\chi_{\mathcal{G}}(q)}(m, q)$，也就是说，所选的动作映射取决于值区域。接下来我们声称，由记忆骨架 $\mathcal{M}$ 和动作映射 λ 实现的策略 s_{A} 是子博弈最优的。与引理 *6* 的证明一样，该策略是局部最优的。此外，同样按照引理 *6* 的证明，我们来证明它满足定理 *1* 的第二个性质。令 $u \in V_{\mathcal{G}} \setminus \{0\}$。考虑一个确定性玩家 *B* 策略 s_{B}。考虑一条有限路径 $\rho \in Q^+$ 和 $\pi = \pi' \cdot q \in Q^* \cdot Q_u$。假设 $\mathbb{P}^{s_{\mathsf{A}}, s_{\rho \cdot \pi}{}^{\mathsf{B}}}((Q_u)^\omega) > 0$。那么，我们有 $\mathbb{P}^{s_{\mathsf{A}}, s_{\rho \cdot \pi}{}^{\mathsf{B}}}(W \mid (Q_u)^\omega) = 1$，因为策略 s_{A} 此后表现得像 s^u_{A}（由 $\mathcal{M}$ 和 λ_u 实现），而这对 s^u_{A} 成立（因为它在 $\mathcal{G}_u$ 中是子博弈几乎必胜的，不受起始相关记忆状态的影响，而且目标 W 是前缀无关的）。□

C.6 从 [15]获取原始结果

我们首先证明，在回合制设定中，正向最优策略总是存在局部最优的策略（针对略作修改的目标）。

命题 8 考虑一个轮流制博弈 $\mathcal{G}$（即状态集合可以划分为玩家 A 状态 q，其中 B_q 是单例集，以及玩家 B 状态 q，其中 A_q 是单例集）。考虑赋值 v： $Q \to [0,1]$，该赋值给出了关于行动策略从所有状态出发的博弈值。考虑某个值 $u \in V_{\mathcal{G}} \setminus \{0\}$（关于赋值 v），并考虑目标（首次在引理 *8* 中定义） $W_u := W \cup Q^* \cdot (Q \setminus Q_u) \cdot Q^\omega \subseteq Q^\omega$（其中 Q_u 是关于赋值 v 定义的）。那么，存在一个局部最优行动策略，使其从 Q_u 中的所有状态出发的值均为正。

证明：考虑一个轮流制博弈 $\mathcal{G}$，并令 $Q = Q_{\mathsf{A}} \uplus Q_{\mathsf{B}}$，其中对于所有 $q \in Q_{\mathsf{A}}$（分别为 Q_{B}），B_q（分别为 A_q）都是单元素集合。考虑某个值 $u \in V_{\mathcal{G}} \setminus \{0\}$，并反设存在一个状态 $q \in Q_u$，从该状态出发，不存在其值为正的局部最优行动策略 q。令 $D_{<u} \subseteq D$ 为值小于 u 的自然状态集合，并令 $x < u$ 为这些自然状态中的最高值。考虑博弈 $\mathcal{G}' = \langle \mathcal{C}', W_u \rangle$，其博弈场 $\mathcal{C}'$ 是博弈场 $\mathcal{C}$ 的一个限制，其中移除了通向 $D_{<u}$ 中自然状态的玩家 *A* 动作（即玩家 *A* 在 Q_u 中的状态只能采用局部最优策略）。根据假设，在该博弈 $\mathcal{G}'$ 中存在一个值为 *0* 的状态 $q \in Q_u$。令 s_{B} 为该博弈 $\mathcal{G}'$ 中一个从 q 出发、值为 $(u-x)/4 > 0$ 的玩家 *B* 策略。现在考虑一个玩家 *A* 确定性策略 s_{A}（回想一下，在轮流制博弈中，确定性策略与任意策略能取得相同的值，例如参见 *[10]* 中定理 *1* 的说明）。记 *NLO*（即非局部最优）为以下集合：

在这些有限路径之后，玩家 A 的确定性策略 s_{A} 不再是局部最优的、且以 Q_u 结尾的有限路径集合：$NLO := \{\rho = \rho' \cdot q' \in Q^* \cdot Q_u \mid \delta(q, s_{\mathsf{A}}(\rho)\ b) \in D_{<u}\}$，因为 b 中只有一个玩家 B 动作。现在我们可以如下定义玩家 B 策略 s'_{B}：只要对局仍处于 Q_u 中，且尚未到达 NLO 中的有限路径，它就模仿策略 s_{B}；一旦到达这样的路径，s'_{B} 就切换到值为 $(x+u)/2$ 的策略（根据 NLO 的定义，这是可行的）。现在考虑一个玩家 A 策略：在到达 NLO 中的有限路径 $\cup Q^* \cdot (Q \setminus Q_u)$ 之前，它完全按照 s_{A} 行动；若到达该路径，则切换为任意局部最优策略（因此它是博弈 $\mathcal{G}'$ 中的一个策略）。采用这些策略后，s_{A} 与 s'_{A} 在路径 $(Q_u \setminus NLO)^*$ 上一致。策略 s_{B} 和 s'_{B} 在这些路径上也一致。考察策略 s'_{A} 和 s_{B} 的情形——回想一下，对于目标 W_u，s_{B} 对阵局部最优策略时的值至多为 $u - x/4$：

$$\begin{aligned}
\mathbb{P}_q^{s'_{\mathsf{A}},s_{\mathsf{B}}}[W_u] &= \mathbb{P}_q^{s_{\mathsf{A}},s_{\mathsf{B}}}[W_u \cap (Q_u \setminus \mathsf{NLO})^\omega] + \mathbb{P}_q^{s'_{\mathsf{A}},s_{\mathsf{B}}}[W_u \cap (Q_u \setminus \mathsf{NLO})^* \cdot (Q \setminus Q_u)] + \mathbb{P}_q^{s'_{\mathsf{A}},s_{\mathsf{B}}}[W_u \cap Q_u^* \cdot \mathsf{NLO}] \\
&= \mathbb{P}_q^{s'_{\mathsf{A}},s_{\mathsf{B}}}[W \cap (Q_u \setminus \mathsf{NLO})^\omega] + \mathbb{P}_q^{s'_{\mathsf{A}},s_{\mathsf{B}}}[(Q_u \setminus \mathsf{NLO})^* \cdot (Q \setminus Q_u)] + \mathbb{P}_q^{s'_{\mathsf{A}},s_{\mathsf{B}}}[W_u \cap Q_u^* \cdot \mathsf{NLO}] \\
&= \mathbb{P}_q^{s_{\mathsf{A}},s'_{\mathsf{B}}}[W \cap (Q_u \setminus \mathsf{NLO})^\omega] + \mathbb{P}_q^{s_{\mathsf{A}},s'_{\mathsf{B}}}[(Q_u \setminus \mathsf{NLO})^* \cdot (Q \setminus Q_u)] + \mathbb{P}_q^{s'_{\mathsf{A}},s_{\mathsf{B}}}[W_u \cap Q_u^* \cdot \mathsf{NLO}] \\
&\leq \frac{u-x}{4}
\end{aligned}$$

由此可得 $\mathbb{P}^{s_{\mathsf{A}},\,s'_{\mathsf{B}}}_q[W \cap (Q_u \setminus \mathsf{NLO})^\omega] + \mathbb{P}^{s_{\mathsf{A}},\,s'_{\mathsf{B}}}_q[(Q_u \setminus \mathsf{NLO})^* \cdot (Q \setminus Q_u)] \leq \frac{u-x}{4}$。此外，根据策略 s'_{B} 的选择，我们有 $\mathbb{P}^{s_{\mathsf{A}},\,s'_{\mathsf{B}}}_q[W \cap Q_u^* \cdot \mathsf{NLO}] \leq \frac{u+x}{2}$。因此得到：

$$\begin{aligned}
\mathbb{P}_q^{s_{\mathsf{A}},s'_{\mathsf{B}}}[W] &= \mathbb{P}_q^{s_{\mathsf{A}},s_{\mathsf{B}}}[W \cap (Q_u \setminus \mathsf{NLO})^\omega] + \mathbb{P}_q^{s_{\mathsf{A}},s'_{\mathsf{B}}}[W \cap (Q_u \setminus \mathsf{NLO})^* \cdot (Q \setminus Q_u)] + \mathbb{P}_q^{s_{\mathsf{A}},s'_{\mathsf{B}}}[W \cap Q_u^* \cdot \mathsf{NLO}] \\
&\leq \mathbb{P}_q^{s_{\mathsf{A}},s_{\mathsf{B}}}[W \cap (Q_u \setminus \mathsf{NLO})^\omega] + \mathbb{P}_q^{s_{\mathsf{A}},s'_{\mathsf{B}}}[(Q_u \setminus \mathsf{NLO})^* \cdot (Q \setminus Q_u)] + \frac{u+x}{2} \\
&\leq \frac{u-x}{4} + \frac{u+x}{2} \\
&= u - \frac{u-x}{4} < u
\end{aligned}$$

因此，所有玩家 A 的确定性策略的值都至多为 $u - \frac{u-x}{4} < u$。这与状态 q 的值为 u 这一事实相矛盾。 □

现在我们可以推导出，轮流制博弈中始终存在子博弈最优策略。事实上，引理 8 已证明：如果对于所有值 $u \in V_{\mathcal{G}} \setminus \{0\}$，存在从 Q_u 中所有状态出发、相对于目标 W_u（我们将其称为假设 H_u）的值为正的局部最优策略，那么博弈 $\mathcal{G}_u$ 中的所有状态（定理 2的证明中）均为正。因此，通过对定理 2的证明进行直接调整，可以证明该假设 H_u 蕴含着子博弈最优策略的存在性。由此可知，轮流制博弈中始终存在子博弈最优策略。

此外请注意，在这种情况下存在子博弈最优确定性策略。这是因为定理 3 在轮流制博弈中的表述保证了子博弈几乎必胜确定性策略的存在。这一点已在 [15] 中证明。我们对证明的改编（见附录 D）也可以证明这一点，只需注意到，在轮流制博弈中，ε-最优策略可以从确定性策略中找到。

D 定理3的证明

在开始证明定理 3之前，我们需要一些额外的记号和一个非常有用的定理。

额外记号 设 Q 为一个有限集。对于所有 $\rho \in Q^*$，我们用 $\mathsf{Cyl}(\rho) \subseteq Q$表示集合 $\mathsf{Cyl}(\rho) := \{\rho \cdot \pi \mid \pi \in Q^*\}$。于是，对于 Q^* 的子集 S，我们用 $\mathsf{Suff}(S) \subseteq Q^*$表示 S 中元素的后缀集合，即 $\mathsf{Suff}(S) := \cup_{\rho \in S}\mathsf{Cyl}(\rho)$。如果集合 $S \in Q^*$ 满足后缀封闭 条件，即 $S = \mathsf{Suff}(S)$，我们就称其为后缀封闭 集合。此外，对于 Q^+ 的所有后缀封闭集合 S，存在唯一的集合 $\mathsf{Pref}(S) \subseteq S$，使得 S 可以写成如下不相交并：$S = \uplus_{\rho \in \mathsf{Pref}(S)}\mathsf{Cyl}(\rho)$（集合 $\mathsf{Pref}(S)$ 可定义为 $\mathsf{Pref}(S) := \{\rho \in S \mid \rho$ 没有属于 $S\}$ 的前缀）。最后，对于所有有限路径 $\rho = q_0 \cdots q_n \in Q^+$，我们用 $\mathsf{head}(\rho) \in Q$ 表示 ρ 的最后一个状态，即 $\mathsf{head}(\rho) := q_n$。

考虑一个 PI 并发博弈 $\mathcal{G} = \langle \mathcal{C}, W \rangle$，以及玩家 A 策略 s_A。现在，对 $u \in [0,1]$，我们定义一个作用于策略的算子 $\mathsf{ValInf}_u : \mathsf{S}^\mathsf{A}_\mathcal{C} \times \mathcal{P}(Q^+) \to \mathcal{P}(Q^+)$：给定策略 $\mathsf{s} \in \mathsf{S}^\mathsf{A}_\mathcal{C}$ 以及一个我们关注的前缀封闭有限路径集合 $S \in \mathcal{P}(Q^+)$，该算子将其映射为有限路径集合 ρ，其中路径相对于剩余策略 s^ρ 的值小于 $1\ 2$。也就是说，$\mathsf{ValInf}_u(\mathsf{s}\ S) := \{\rho \in S \mid \chi_\mathcal{G}[\mathsf{s}^\rho](\rho) < u\}$。

证明 对所有状态 $q \in Q$，记 $s_q : Q^+ \to \mathcal{D}(A) \in S^\mathsf{A}_\mathcal{C}$ 为一个玩家 A 策略，其从状态 q 出发的值至少为 $c \cdot (1 - c/4)$（即 $\chi_\mathcal{G}[s_q](q) \geq c \cdot (1 - c/4)$）。现在，令 $q_0 \in Q$ 为一个状态，并展示一个从 q_0出发的几乎必胜策略。我们归纳地构造一个策略序列 $(s_n)_{n \in \mathbb{N}} \in (S^\mathsf{A}_\mathcal{C})^\mathbb{N}$ 以及一个有限路径集合序列 $(S_n)_{n \in \mathbb{N}} \in (\mathcal{P}(Q^+))^\mathbb{N}$，使得对所有 $n \in \mathbb{N}$，以下性质均成立：

(a) S_n 是后缀封闭的；

(b) $S_n \subseteq S_{n-1}$ 和 $S_n \cap \mathsf{Pref}(S_{n-1}) = \emptyset$;

(c) 对所有 $k < n$，s_n 在 $Q^+ \setminus S_{k+1}$上与 s_k 一致；

(d) 对所有 $\rho \in \mathsf{Pref}(S_n)$，剩余策略 s_n 在 ρ 中的值至少为 $c \cdot (1 - c/4)$，即 $\chi_\mathcal{G}[\mathsf{s}^\rho_n](\rho) \geq c \cdot (1 - c/4)$

初始时，我们设定 $s_0 := s_{q_0}$ 和 $S_0 := q_0 \cdot Q^+$。注意，这确实是后缀封闭的，因此满足性质 *(a)*；并且由于 $\mathsf{Pref}(S_0) = \{q_0\}$，性质 *(d)* 也得到保证。接着，假设对于某个 $n \geq 1$，对所有 $k \leq n-1$，$s_k \in S^\mathsf{A}_\mathcal{C}$和 $S_k \subseteq Q^+$ 均已定义，且性质 *(a)* –*(d)* 成立。令我们定义 s_n 和 S_n。若 $S_{n-1} = \emptyset$，则 $S_n := S_{n-1}$ 且 $s_n := s_{n-1}$（并且性质 $(a) - (d)$ 得到保证）。

现在假设 $S_{n-1} \neq \emptyset$。我们先定义 $S_n \subseteq Q^*$，并证明它满足性质（a）–（b）。令 $V_{n-1} = \mathit{ValInf}^{c \cdot \frac{1}{2}}(s_{n-1}, S_{n-1}) \subseteq S_{n-1}$为 S_{n-1} 的有限路径集合，其中 s_{n-1} 的剩余策略的值小于 $c \cdot \frac{1}{2}$。然后，我们将 $S_n \subseteq Q^+$ 定义为 V_{n-1} 的后缀集合：$S_n := \mathit{Suff}(V_{n-1})$。特别地，注意到 $\mathit{Pref}(S_n) = \mathit{Pref}(V_{n-1})$。显然，$S_n$是后缀封闭的，因此性质（$a$）得以保证。

现在考虑性质（b）。我们有 $S_n = \mathit{Suff}(V_{n-1}) \subseteq \mathit{Suff}(S_{n-1}) = S_{n-1}$，因为 $V_{n-1} \subseteq S_{n-1}$，且 S_{n-1} 是后缀封闭的。此外，考虑一条路径 $\rho \in \mathit{Pref}(S_{n-1}) \neq \emptyset$。反设 $\rho \in S_n = \mathit{Suff}(V_{n-1})$。那么存在某个 $\rho' \in V_{n-1} \subseteq S_{n-1}$，使得 $\rho \in \mathit{Cyl}(\rho')$。由于 $\rho' \in S_{n-1}$，因此存在某个 $\rho'' \in \mathit{Pref}(S_{n-1})$，使得 $\rho' \in \mathit{Cyl}(\rho'')$，从而 $\rho \in \mathit{Cyl}(\rho'')$。我们得到 $\rho \in \mathit{Cyl}(\rho) \cap \mathit{Cyl}(\rho'')$，其中 $\rho, \rho'' \in \mathit{Pref}(S_{n-1})$。根据定义，这意味着 $\rho = \rho''$ 且 $\rho' \in \mathit{Cyl}(\rho)$。也就是说，$\rho = \rho' \in V_{n-1}$。然而，根据索引 $n-1$ 处的性质（d），我们有 $\chi_\mathcal{G}[s^\rho_{n-1}](\mathit{head}(\rho)) \geq c \cdot (1 - c/4) > c/2$，因此 $\rho \notin V_{n-1}$。矛盾由此得证。事实上，$\rho \notin S_n$，且性质（b）得以保证。

现在定义策略 $\mathsf{s}_n : Q^+ \to \mathcal{D}(A)$。我们令：

- $\mathsf{s}_n|_{Q^+ \setminus S_n} := \mathsf{s}_{n-1}$;
- 对所有 $\rho \in \mathsf{Pref}(V_{n-1}) = \mathsf{Pref}(S_n)$，均有 $\mathsf{s}^\rho_n := \mathsf{s}_{\mathsf{head}(\rho)}$。

下面证明该定义确保性质 (c) –(d)。令 $k \leq n-1$。那么，s_k 与 s_{n-1} 在 $Q^+ \setminus S_{k+1} \subseteq Q^+ \setminus S_n$上一致。因此，$s_k$ 也与 s_n 在 $Q^+ \setminus S_{k+1}$上一致。此外，s_{n-1} 与 s_n 在 $Q^+ \setminus S_n$上一致。因此，性质 (c) 成立。最后，由于 W 是一个前缀无关目标，对于所有 $\rho \in \mathit{Pref}(S_n)$ 和 $q := \mathit{head}(\rho) \in Q$，均有 $\chi_{\mathcal{G}}[s_n^\rho](\mathit{head}(\rho)) = \chi_{\mathcal{G}}[s_q](q) \geq c \cdot (1 - c/4)$。也就是说，性质 (d) 成立。

至此，我们完成了序列 $(s_n)_{n\in\mathbb{N}} \in (S_{\mathcal{C}}^{\mathsf{A}})^{\mathbb{N}}$ 和 $(S_n)_{n\in\mathbb{N}} \in (\mathcal{P}(Q^+))^{\mathbb{N}}$的归纳定义。接着，我们可以将策略 s_{A} 定义为策略 s_n的极限。也就是说，如果在某个时刻 $S_n = \emptyset$，并且对所有 $k \geq 0$都有 $s_n = s_{n+k}$，那么我们令 $s_{\mathsf{A}} := s_n$。否则，请注意，根据性质 (b)，我们有 $\cap_{n\in\mathbb{N}} S_n = \emptyset$。事实上，考虑一条有限路径 $\rho \in \cap_{n\in\mathbb{N}} S_n \subseteq Q^*$。对所有$n \in \mathbb{N}$，都存在唯一一条有限路径 $\rho_n \in \mathit{Pref}(S_n)$，使得 $\rho \in \mathit{Cyl}(\rho_n)$。此外，对所有$k < l \in \mathbb{N}$，我们有 $\rho_k \neq \rho_l$；否则就会有 $\mathit{Pref}(S_k) \ni \rho_k \in \mathit{Pref}(S_l) \subseteq S_l \subseteq S_{k+1}$（因为 $l \geq k+1$），但这并不成立。因此，ρ 有无限多个不同的有限前缀，这是不可能的。由此可知，对所有 $\rho \in Q^+$，恰有一个索引 $n \in \mathbb{N}$ 使得$\rho \in S_n \setminus S_{n-1}$。然后，我们按以下方式定义策略 s_{A}：

$$\forall n \in \mathbb{N},\ \mathsf{s}_{\mathsf{A}}|_{S_n \setminus S_{n-1}} := \mathsf{s}_n$$

具有性质 (c)时，该定义保证：

$$\forall n \in \mathbb{N},\ \mathsf{s}_{\mathsf{A}}|_{Q^+ \setminus S_{n+1}} = \mathsf{s}_n \tag{2}$$

请注意，这在第一种情况下也成立，其中 $S_n = \emptyset$ 对某个 $n \in \mathbb{N}$成立。

我们声称，该玩家 A 策略几乎必然成立。考虑一个玩家 B 策略s_{B}。定义并发博弈 $\mathcal{G}'$，它是通过展开状态空间，从 $\mathcal{G}$ 得到的，因此状态空间变为可数的 $q_0 \cdot Q^*$（策略集合保持不变）。下面证明，从所有状态出发，采用策略 $\mathcal{G}'$ s_{A} 和s_{B} 时，博弈值至少为 $\mathcal{G}$（这对应于$c \cdot \frac{1}{2} > 0$中的有限路径）。由引理4 可知，这意味着从所有状态出发，博弈值实际上为1。

对于所有集合 $S \subseteq q_0 \cdot Q^*$，我们用 $\Diamond S$ 表示指定集合 S 被到达这一事件。

设 $n \in \mathbb{N}$和 $\rho \in S_n \setminus S_{n+1}$。首先，考虑一个玩家 B 策略 $\tilde{}\mathsf{s}_{\mathsf{B}}$，它在 $Q^+ \setminus S_{n+1}$上与 s_{B}一致，并且对于所有路径 $\rho' \in \mathit{Pref}(S_{n+1}) \cap \mathit{Cyl}(\rho)$，都有 $\mathbb{P}^{\rho'}_{s_n^{\rho'}, \tilde{s}_{\mathsf{B}}^{\rho'}}[W] \leq c \cdot \frac{1}{2}$。请注意，这是可行的，因为 $\rho' \in \mathit{Pref}(S_{n+1}) = \mathit{Pref}(V_n) \subseteq V_n = \mathit{ValInf}^{c\cdot\frac{1}{2}}(s_n, S_n)$，这意味着 $\chi_{\mathcal{G}}[s_n^\rho](\rho') < c \cdot \frac{1}{2}$。根据此定义，若 $\mathbb{P}^{\rho}_{s_n^{\rho}, \tilde{}s^{\rho}_{\mathsf{B}}}[\Diamond S_{n+1}] > 0$，则有：

$$\begin{aligned}
\mathbb{P}^{\rho}_{s_n^{\rho}, \tilde{s}_{\mathsf{B}}^{\rho}}[W \mid \Diamond S_{n+1}] &= \frac{\sum_{\pi \in \mathsf{Cyl}(\rho) \cap \mathsf{Pref}(S_{n+1})} \mathbb{P}^{\pi}_{s_n^{\pi}, \tilde{s}_{\mathsf{B}}^{\pi}}[W] \cdot \mathbb{P}^{\rho}_{s_n^{\rho}, \tilde{s}_{\mathsf{B}}^{\rho}}[\pi]}{\mathbb{P}^{\rho}_{s_n^{\rho}, \tilde{s}_{\mathsf{B}}^{\rho}}[\Diamond S_{n+1}]} \\
&\leq c \cdot \frac{1}{2} \frac{\sum_{\pi \in \mathsf{Cyl}(\rho) \cap \mathsf{Pref}(S_{n+1})} \mathbb{P}^{\rho}_{s_n^{\rho}, \tilde{s}_{\mathsf{B}}^{\rho}}[\pi]}{\mathbb{P}^{\rho}_{s_n^{\rho}, \tilde{s}_{\mathsf{B}}^{\rho}}[\Diamond S_{n+1}]} \\
&= c \cdot \frac{1}{2} \cdot \frac{\mathbb{P}^{\rho}_{s_n^{\rho}, \tilde{s}_{\mathsf{B}}^{\rho}}[\Diamond S_{n+1}]}{\mathbb{P}^{\rho}_{s_n^{\rho}, \tilde{s}_{\mathsf{B}}^{\rho}}[\Diamond S_{n+1}]} = c \cdot \frac{1}{2}
\end{aligned}$$

此外，由于 $\tilde{}\mathsf{s}_{\mathsf{B}}$ 和 s_{B} 在 $Q^+ \setminus S_{n+1}$ 上一致：

$$\begin{aligned}
\mathbb{P}^{\rho}_{s_n^{\rho}, \tilde{s}_{\mathsf{B}}^{\rho}}[W] &= \mathbb{P}^{\rho}_{s_n^{\rho}, \tilde{s}_{\mathsf{B}}^{\rho}}[W \mid \Diamond S_{n+1}] \cdot \mathbb{P}^{\rho}_{s_n^{\rho}, \tilde{s}_{\mathsf{B}}^{\rho}}[\Diamond S_{n+1}] + \mathbb{P}^{\rho}_{s_n^{\rho}, \tilde{s}_{\mathsf{B}}^{\rho}}[W \mid \neg\Diamond S_{n+1}] \cdot \mathbb{P}^{q}_{s_n^{\rho}, \tilde{s}_{\mathsf{B}}^{\rho}}[\neg\Diamond S_{n+1}] \\
&\leq c \cdot \frac{1}{2} \cdot \mathbb{P}^{\rho}_{s_n^{\rho}, s_{\mathsf{B}}^{\rho}}[\Diamond S_{n+1}] + \mathbb{P}^{\rho}_{s_n^{\rho}, s_{\mathsf{B}}^{\rho}}[W \mid \neg\Diamond S_{n+1}] \cdot \mathbb{P}^{\rho}_{s_n^{\rho}, s_{\mathsf{B}}^{\rho}}[\neg\Diamond S_{n+1}] \\
&= c \cdot \frac{1}{2} \cdot p_1 + p_2 \cdot (1 - p_1)
\end{aligned}$$

对于 $p_1 := \mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_n,\ \mathsf{s}^{\rho}_{\mathsf{B}}}[\Diamond S_{n+1}]$ 和 $p_2 := \mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_n,\ \mathsf{s}^{\rho}_{\mathsf{B}}}[W \mid \neg\Diamond S_{n+1}]$。我们得到：

$$\mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_n,\tilde{\mathsf{s}}^{\rho}_{\mathsf{B}}}[W] \leq \frac{c}{2} \cdot p_1 + p_2 \cdot (1 - p_1) \tag{3}$$

该不等式对所有 $\rho \in S_n \setminus S_{n+1}$ 和 $n \in \mathbb{N}$都成立。

现在，考虑 $\rho \in \mathsf{Pref}(S_n)$ 的情形，我们来证明 $\mathbb{P}^{\rho}_{\mathsf{s}^{\rho}\mathsf{A},\mathsf{s}^{\rho}\mathsf{B}}[W] \geq c \cdot \frac{1}{2}$。事实上，由性质 (d) 可得：$\chi_{\mathcal{G}}[\mathsf{s}^{\rho}_n](\rho) \geq c \cdot (1 - c/4)$。这意味着 $\mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_n,\tilde{\mathsf{s}}^{\rho}\mathsf{B}}[W] \geq c \cdot (1 - c/4)$。因此：

$$c \cdot (1 - \frac{c}{4}) \leq \frac{c}{2} \cdot p_1 + p_2 \cdot (1 - p_1)$$

因此，$p_2 > \frac{c}{2}$（因为 $1/2 < 1 - c/4$），并且：

$$p_1 \leq \frac{p_2 - c \cdot (1 - \frac{c}{4})}{p_2 - \frac{c}{2}}$$

也就是说：

$$p_2 \cdot (1 - p_1) \geq p_2 \cdot \frac{p_2 - \frac{c}{2} - p_2 + c \cdot (1 - \frac{c}{4})}{p_2 - \frac{c}{2}} = p_2 \cdot \frac{c \cdot (\frac{1}{2} - \frac{c}{4})}{p_2 - \frac{c}{2}} = \frac{c}{2} \cdot \frac{p_2 - p_2 \cdot \frac{c}{2}}{p_2 - \frac{c}{2}} \geq \frac{c}{2}$$

现在，我们可以考虑在给定策略 s_{A} 和 $\tilde{\mathsf{s}}_{\mathsf{B}}$ 的情况下，满足 W 的概率 $\mathbb{P}^{\rho}_{\mathsf{s}^{\rho}\mathsf{A},\mathsf{s}^{\rho}\mathsf{B}}[W]$。注意，$\mathsf{s}_{\mathsf{A}}$ 与 s_n 在 $S_n \setminus S_{n+1}$ 上一致，尤其是在 $\mathsf{Cyl}(\rho) \setminus S_{n+1}$ 上。因此，我们有：

$$\begin{aligned}
\mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{\mathsf{A}},\mathsf{s}^{\rho}_{\mathsf{B}}}[W] &= \mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{\mathsf{A}},\mathsf{s}^{\rho}_{\mathsf{B}}}[W \mid \Diamond S_{n+1}] \cdot \mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{\mathsf{A}},\mathsf{s}^{\rho}_{\mathsf{B}}}[\Diamond S_{n+1}] + \mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{\mathsf{A}},\mathsf{s}^{\rho}_{\mathsf{B}}}[W \mid \neg\Diamond S_{n+1}] \cdot \mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{\mathsf{A}},\mathsf{s}^{\rho}_{\mathsf{B}}}[\neg\Diamond S_{n+1}] \\
&\geq \mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{\mathsf{A}},\mathsf{s}^{\rho}_{\mathsf{B}}}[W \mid \neg\Diamond S_{n+1}] \cdot \mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{\mathsf{A}},\mathsf{s}^{\rho}_{\mathsf{B}}}[\neg\Diamond S_{n+1}] \\
&= \mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{n},\mathsf{s}^{\rho}_{\mathsf{B}}}[W \mid \neg\Diamond S_{n+1}] \cdot \mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{n},\mathsf{s}^{\rho}_{\mathsf{B}}}[\neg\Diamond S_{n+1}] \\
&= p_2 \cdot (1 - p_1) \geq \frac{c}{2}
\end{aligned}$$

这对所有 $\rho \in$前缀(S_n) 以及所有 $n \in \mathbb{N}$ 都成立。

现在考虑任意一个 $\rho \in S_n \setminus S_{n+1}$。在这种情况下，由于 $\rho \notin \mathsf{ValInf}^{c,\frac{1}{2}}(\mathsf{s}n, S_n) = V_n \subseteq S_{n+1}$，我们有 $\chi_{\mathcal{G}}[\mathsf{s}^{\rho}_n](\rho) \geq \frac{c}{2}$且 $\mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_n,\tilde{\mathsf{s}}^{\rho}\mathsf{B}}[W] \geq \frac{c}{2}$。因此，根据公式 *3* 可得：

$$\frac{c}{2} \leq \mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_n,\tilde{\mathsf{s}}^{\rho}_{\mathsf{B}}}[W] \leq \frac{c}{2} \cdot p_1 + p_2 \cdot (1 - p_1)$$

因此，假设 $1 - p_1 = \mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_n,\mathsf{s}^{\rho}\mathsf{B}}[\neg\Diamond S_{n+1}] > 0$，我们可以得出 $p_2 = \mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_n,\mathsf{s}^{\rho}\mathsf{B}}[W \mid \neg\Diamond S_{n+1}] \geq \frac{c}{2}$。此外，我们已证明，对于所有 $\rho' \in \mathsf{Pref}(S_{n+1})$，都有 $\mathbb{P}^{\rho'}_{\mathsf{s}^{\rho'}\mathsf{A},\mathsf{s}^{\rho'}\mathsf{B}}[W] \geq \frac{c}{2}$。因此，假设 $\mathbb{P}^{\rho}_{\mathsf{s}^{\rho}\mathsf{A},\mathsf{s}^{\rho}\mathsf{B}}[\Diamond S_{n+1}] \neq 0$，则有：

$$\begin{aligned}
\mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{\mathsf{A}},\mathsf{s}^{\rho}_{\mathsf{B}}}[W \mid \Diamond S_{n+1}] &= \frac{1}{\mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{\mathsf{A}},\mathsf{s}^{\rho}_{\mathsf{B}}}[\Diamond S_{n+1}]} \cdot \sum_{\rho' \in \mathsf{Pref}(S_{n+1})} \mathbb{P}^{\rho'}_{\mathsf{s}^{\rho'}_{\mathsf{A}},\mathsf{s}^{\rho'}_{\mathsf{B}}}[W] \cdot \mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{\mathsf{A}},\mathsf{s}^{\rho}_{\mathsf{B}}}[\rho'] \\
&\geq \frac{1}{\mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{\mathsf{A}},\mathsf{s}^{\rho}_{\mathsf{B}}}[\Diamond S_{n+1}]} \cdot \sum_{\rho' \in \mathsf{Pref}(S_{n+1})} \frac{c}{2} \cdot \mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{\mathsf{A}},\mathsf{s}^{\rho}_{\mathsf{B}}}[\rho'] \\
&= \frac{c}{2} \cdot \frac{1}{\mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{\mathsf{A}},\mathsf{s}^{\rho}_{\mathsf{B}}}[\Diamond S_{n+1}]} \cdot \sum_{\rho' \in \mathsf{Pref}(S_{n+1})} \cdot \mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{\mathsf{A}},\mathsf{s}^{\rho}_{\mathsf{B}}}[\rho'] \\
&= \frac{c}{2} \cdot \frac{1}{\mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{\mathsf{A}},\mathsf{s}^{\rho}_{\mathsf{B}}}[\Diamond S_{n+1}]} \cdot \mathbb{P}^{\rho}_{\mathsf{s}^{\rho}_{\mathsf{A}},\mathsf{s}^{\rho}_{\mathsf{B}}}[\Diamond S_{n+1}] \\
&= \frac{c}{2}
\end{aligned}$$

接着，假设 $0 < \mathbb{P}^{\rho}_{s^{\rho}_A, s^{\rho}_B}[\Diamond S_{n+1}] < 1$，则有：

$$\begin{aligned}
\mathbb{P}^{\rho}_{s^{\rho}_A, s^{\rho}_B}[W] &= \mathbb{P}^{\rho}_{s^{\rho}_A, s^{\rho}_B}[W \mid \Diamond S_{n+1}] \cdot \mathbb{P}^{\rho}_{s^{\rho}_A, s^{\rho}_B}[\Diamond S_{n+1}] + \mathbb{P}^{\rho}_{s^{\rho}_A, s^{\rho}_B}[W \mid \neg\Diamond S_{n+1}] \cdot \mathbb{P}^{\rho}_{s^{\rho}_A, s^{\rho}_B}[\neg\Diamond S_{n+1}] \\
&\geq \frac{c}{2} \cdot \mathbb{P}^{\rho}_{s^{\rho}_A, s^{\rho}_B}[\Diamond S_{n+1}] + \mathbb{P}^{\rho}_{s^{\rho}_A, s^{\rho}_B}[W \mid \neg\Diamond S_{n+1}] \cdot \mathbb{P}^{\rho}_{s^{\rho}_A, s^{\rho}_B}[\neg\Diamond S_{n+1}] \\
&= \frac{c}{2} \cdot \mathbb{P}^{\rho}_{s^{\rho}_n, s^{\rho}_B}[\Diamond S_{n+1}] + \mathbb{P}^{\rho}_{s^{\rho}_n, s^{\rho}_B}[W \mid \neg\Diamond S_{n+1}] \cdot \mathbb{P}^{\rho}_{s^{\rho}_n, s^{\rho}_B}[\neg\Diamond S_{n+1}] \\
&= \frac{c}{2} \cdot \mathbb{P}^{\rho}_{s^{\rho}_n, s^{\rho}_B}[\Diamond S_{n+1}] + p_2 \cdot \mathbb{P}^{\rho}_{s^{\rho}_n, s^{\rho}_B}[\neg\Diamond S_{n+1}] \\
&\geq \frac{c}{2} \cdot \mathbb{P}^{\rho}_{s^{\rho}_n, s^{\rho}_B}[\Diamond S_{n+1}] + \frac{c}{2} \cdot \mathbb{P}^{\rho}_{s^{\rho}_n, s^{\rho}_B}[\neg\Diamond S_{n+1}] \\
&= \frac{c}{2}
\end{aligned}$$

注意，如果概率 $\mathbb{P}^{\rho}_{s^{\rho}_A, s^{\rho}_B}[\Diamond S_{n+1}] = \mathbb{P}^{\rho}_{s^{\rho}_n, s^{\rho}_B}[\Diamond S_{n+1}]$ 等于 *0* 或 *1*，这一点也成立。

总而言之，对于所有 $\rho \in Q^+$，我们都有 $\chi_{\mathcal{G}'}(\rho) = \mathbb{P}^{\rho}_{s^{\rho}A, s^{\rho}B}[W] \geq \frac{1}{2}$。因此，$\inf_{q \in Q} \chi_{\mathcal{G}'}(q) \geq \frac{1}{2}$，这意味着，根据引理 *4*，对于所有 $\rho \in q_0 \cdot Q^*$，我们都有 $1 = \chi_{\mathcal{G}'}(q) = \mathbb{P}^{\mathcal{G}}_{\rho, s^{\rho}A, s^{\rho}B}[W]$。特别地，$\mathbb{P}^{\mathcal{G}}_{s_A, s_B}[W] = 1$。由于这对所有玩家 *B* 策略 s_B都成立，因此策略 s_A 几乎必然。接着，我们可以从所有状态 $q \in Q$ 出发进行同样的论证，从而得到一个几乎必然的策略。此外，我们已经证明，该策略从所有有限路径出发的值至少为$c/2 > 0$。事实上，这意味着该策略是子博弈几乎必然的。 □

E 有限选择策略

E.1 定理4的证明

我们将4定理推广到更一般的策略，即正有界策略。

定义 27（正有界策略）玩家A 策略s_A 称为正有界（简称 *p.b.*），如果存在常数 $c > 0$，使得对所有 $\rho \cdot q \in Q^+$、所有 $a \in A_q$，均有：$s_A(\rho \cdot q)(q) > 0 \Rightarrow s_A(\rho \cdot q)(q) \geq c$。

我们给出下面的定理。

定理8 考虑一个 *PI* 并发博弈 $\mathcal{G}$。假设存在一个正有界最优策略。那么，存在一个正有界子博弈最优策略。

证明 记s^{pb} 为一个最优且正有界的玩家A 策略。我们归纳地构造一个玩家A 子博弈最优策略s_A。其定义如下：对于所有有限路径$\rho = \rho' \cdot q \in Q^+$，令$s_A(\rho)$ 为：

$$s_A(\rho) := \begin{cases} s_{pb}(\rho) & \text{if } s_{pb} \text{ is optimal from } \rho, \text{ i.e. } \chi_{\mathcal{G}}(s^{\rho}_{pb})[q] = \chi_{\mathcal{G}}[q] \\ s_{pb}(q) & \text{otherwise} \end{cases}$$

由于策略 s^{pb} 是正有界的，因此策略s_A 也是正有界的。下面应用定理 *1*来证明它是子博弈最优的。

令 $\rho = \rho' \cdot q \in Q^+$。在所有情况下，策略 s_A 都与策略 s^{pb} （在ρ 或 q处）一致，而后者从 q出发是最优的，因此，根据命题 *4*，对于玩家 B 的任意动作$b \in B_q$，都有：$\text{out}^{\mathcal{F}_q}(s_A(q)\ b) \geq \chi_{\mathcal{G}}(q)$。由此可知，策略 s_A 是局部最优的。

现在我们来证明它能确保第二个性质。令 $\rho \in Q^+$，并用s^{ρ} 表示剩余策略$s^{\rho}{}_A$。考虑一个玩家B 的确定性策略s_B 以及某个值 $u \in V_{\mathcal{G}} \setminus \{0\}$。我们引入两种记号：

- 我们用$\mathsf{Exit}_u \subseteq Q^+$表示以 Q_u 结尾且以正概率离开该值区域的有限路径集合：$\mathsf{Exit}_u := \{\pi \cdot Q^* \cdot Q_u \mid \mathbb{P}^{s^\rho}_{s_\pi} [Q \setminus Q_u] > 0\}$,sB。
- 我们也将偏离 $\subseteq Q^+$ 记作策略spb A 非最优的有限路径集合：偏离 $:= \{\pi = \pi' \cdot q \in Q^+ \mid \chi_{\mathcal{G}}[s^{\rho \cdot \pi}$ pb $](q) < \chi_{\mathcal{G}}(q)\}$。

下面证明以下事实：

(a). $\mathbb{P}^{s_\rho, s_\rho}$ B$[Q^* \cdot (Q_u)^\omega \cap (Q^* \cdot$ 退出$_u)^\omega] = 0$; ***(b).*** $\mathbb{P}^{s_\rho, s^B}[Q^* \cdot (Q_u)^\omega \cap Q^* \cdot (Q \setminus$ 退出$_u)^\omega] \leq \mathbb{P}^{s_\rho, s^B}[Q^* \cdot (Q_u)^\omega \cap Q^* \cdot (Q \setminus$ 偏离$)^\omega]$; ***(c).*** $\mathbb{P}^{s_\rho, s_\rho}$ B$[Q^* \cdot (Q_u \setminus$ 偏离$)^\omega] = \mathbb{P}^{s_\rho,}$ $_{s_\rho}$ B$[W \cap Q^* \cdot (Q_u \setminus$ 偏离$)^\omega]$。

如果我们假设这些事实都成立，那么可以得到：

$$
\begin{aligned}
\mathbb{P}^{s_\rho, s_B}_\rho [Q^* \cdot (Q_u)^\omega] &= \mathbb{P}^{s_\rho, s_B}_\rho [Q^* \cdot (Q_u)^\omega \cap Q^* \cdot (Q \setminus \mathsf{Exit}_u)^\omega] && \textit{by fact (a)} \\
&\leq \mathbb{P}^{s_\rho, s_B}_\rho [Q^* \cdot (Q_u)^\omega \cap Q^* \cdot (Q \setminus \mathsf{Deviate})^\omega] && \textit{by fact (b)} \\
&= \mathbb{P}^{s_\rho, s_B}_\rho [Q^* \cdot (Q_u \setminus \mathsf{Deviate})^\omega] \\
&= \mathbb{P}^{s_\rho, s_B}_\rho [W \cap Q^* \cdot (Q_u \setminus \mathsf{Deviate})^\omega] && \textit{by fact (c)} \\
&\leq \mathbb{P}^{s_\rho, s_B}_\rho [W \cap Q^* \cdot (Q_u)^\omega] \\
&\leq \mathbb{P}^{s_\rho, s_B}_\rho [Q^* \cdot (Q_u)^\omega]
\end{aligned}
$$

事实上，所有这些不等式都是等式。因此我们可以应用定理*1* 得出结论。现在我们逐一证明这些事实。

(a). 考虑某个 $\pi = \pi' \cdot q \in \mathit{Exit}_u$。我们有 $\mathbb{P}^{s^\pi_\rho, s_{\rho \cdot \pi}}$ B$[Q \setminus Q_u] > 0$。令 $b := s_B(\pi)$（回忆一下，s_B是一个确定性策略），并令 $A_{Q \setminus Q_u} := \{a \in A_q \mid$距离函数 $\circ\, \delta(q, a, b)[Q \setminus Q_u] > 0\}$。于是，$s^\pi_\rho[A_{Q \setminus Q_u}] > 0$，因此对某个固定的 $c > 0$ 有 $s^\pi_\rho[A_{Q \setminus Q_u}] \geq c$（因为 s_ρ 是 *p.b.*）。此外，令$x := \mathit{min}_{d \in \mathsf{D}}\, \mathit{min}^{q \in \mathsf{Supp}(\mathsf{dist}(d))}\, \mathit{dist(d)(q)} > 0$。由此可得 $\mathbb{P}^{s^\pi_\rho, s_{\rho \cdot \pi}}$ B$[Q \setminus Q_u] \geq c \cdot x$。事实上，这对所有 $\pi \in \mathit{Exit}_u$ 都成立。因此，对所有 $\pi \in Q^*$，都有 $\mathbb{P}^{s^\pi_\rho, s_{\rho \cdot \pi}}$ B$[(Q_u)^\omega \mid (Q^* \cdot \mathit{Exit}_u)^\omega] \leq \mathit{lim}_{n \to \infty}(1 - c \cdot x)^n = 0$。由此可得 $\mathbb{P}^{s_\rho, s_\rho}$ B$[Q^* \cdot (Q_u)^\omega \cap (Q^* \cdot \mathit{Exit}_u)^\omega] = 0$。

(b). 下面证明 $\mathbb{P}^{s_\rho, s^B}[Q^* \cdot (Q_u)^\omega \cap Q^* \cdot (Q \setminus$ 退出$_u)^\omega \cap (Q^* \cdot$ 偏离$)^\omega] = 0$。令 $\theta \in Q^* \cdot (Q_u)^\omega \cap Q^* \cdot (Q \setminus$退出$_u)^\omega$。令 $n \in \mathbb{N}$ 为满足 $\theta_{\geq n} \in (Q_u \setminus$退出$_u)^\omega$的一个指标。假设其存在，考虑满足 *Deviate* 的最小指标 $i \geq n+1$ $\theta_i \in$ $\chi_{\mathcal{G}}[$ $]\, \theta_i <$。也就是说，$\chi_{\mathcal{G}}\theta_i$ $\chi_{\mathcal{G}}[\,^{\rho \cdot \theta_{\leq i-1}}]\, \theta_{i-1} = \chi_{\mathcal{G}}\, \theta_{i-1}$ $s^{\rho \cdot \theta_{\leq i}}$ *pb ( ) ( ) and spb ( ) ( )*。通过对命题 *4* 作直接的调整，若 $\mathbb{P}^{s_\rho, s_{\theta_{\leq i-1}}}$ B$[\theta_i] > 0$，对于 $b := s_B(\rho \cdot \theta_{\leq i-1})$（回想一下，$s_B$ 是确定性的），我们有 $\mathit{out}^{\mathcal{F}_{\theta_{i-1}}}(s_\rho(\theta_{\leq i-1})\ b) > \chi_{\mathcal{G}}(\theta_{i-1}) = u$。因此，在 $\theta_{\leq i-1}$ 处，以非零概率到达一个值不同于 u的状态，即 $\mathbb{P}^{s_\rho, s_{\theta_{\leq i-1}}}$ B$[Q \setminus Q_u] > 0$。也就是说，$\theta_{i-1} \in$退出$_u$。这是一条路径——其发生的概率为正——它不访问 Exit_u，也不访问 *Deviate*。因此，几乎必然地，一条仅有限次访问 Exit_u 的路径也仅有限次访问 *Deviate*。由此可得 $\mathbb{P}^{s_\rho, s^B}[Q^* \cdot (Q_u)^\omega \cap Q^* \cdot (Q \setminus$退出$_u)^\omega \cap (Q^* \cdot$偏离$)^\omega] = 0$。也就是说：$\mathbb{P}^{s_\rho, s^B}[Q^* \cdot (Q_u)^\omega \cap Q^* \cdot (Q \setminus$ 退出$_u)^\omega] = \mathbb{P}^{s_\rho, s^B}[Q^* \cdot (Q_u)^\omega \cap Q^* \cdot (Q \setminus$ 退出$_u)^\omega \cap Q^* \cdot (Q \setminus$ 偏离$)^\omega] \leq \mathbb{P}^{s_\rho, s^B}[Q^* \cdot (Q_u)^\omega \cap Q^* \cdot (Q \setminus$偏离$)^\omega]$。

(c). 我们采用与证明定理 *1* 必要条件相似的方法，但在得出结论时还需克服一个额外的困难。实际上，反设 $\mathbb{P}^{s_\rho,}$ $_{s_\rho}$ B$[Q^* \cdot (Q_u \setminus$ 偏离$)^\omega] > \mathbb{P}^{s_\rho,}$ $_{s_\rho}$ B$[W \cap Q^* \cdot (Q_u \setminus$ 偏离$)^\omega]$。考虑由两个策略 s_ρ 和 s_B 诱导的、初始状态为 ρ的可数马尔可夫链。我们将该马尔可夫链中的概率

函数记为 $W_u := W \cup (Q^* \cdot (Q \setminus Q_u \cup \text{偏离}))^\omega$。我们有：

$$\begin{aligned}\mathbb{P}[W_u] &= \mathbb{P}[W_u \cap Q^* \cdot (Q_u \setminus \mathsf{Deviate})^\omega] + \mathbb{P}[W_u \cap (Q^* \cdot (Q \setminus Q_u \cup \mathsf{Deviate})^\omega] \\ &= \mathbb{P}[W \cap Q^* \cdot (Q_u \setminus \mathsf{Deviate})^\omega] + \mathbb{P}[(Q^* \cdot (Q \setminus Q_u \cup \mathsf{Deviate}))^\omega] \\ &< \mathbb{P}[Q^* \cdot (Q_u \setminus \mathsf{Deviate})^\omega] + \mathbb{P}[(Q^* \cdot (Q \setminus Q_u \cup \mathsf{Deviate}))^\omega] \\ &= 1\end{aligned}$$

现在我们来证明，存在一条有限路径，其最后一个状态属于 $Q_u \setminus$偏离，并且从该路径出发，W_u 的概率小于 $u/2$。根据引理 4，由于 W_u 与前缀无关，因此马尔可夫链中存在一个状态 $\pi \in Q^*$（它对应一条有限路径），使得$\mathbb{P}_\pi[W_u] < u/2$。

现在，反设存在矛盾：对于所有 $\pi' \in \pi \cdot Q^* \cdot$（$Q_u \setminus$偏离），要么有 $\mathbb{P}_{\pi'}[W_u] \geq \frac{u}{2}$，要么有 $\mathbb{P}_{\pi'}[Q^* \cdot$（$Q \setminus Q_u \cup$ 偏离）$] \succ \frac{u}{2}$。记 $Q_1 := \{\pi' \in \pi \cdot Q^* \cdot$（$Q_u \setminus$ 偏离）$_- \mid \mathbb{P}_{\pi'}[W_u] \geq \frac{u}{2}\}$，并记 $Q_2 :$ $= \pi \cdot Q^* \cdot$（$Q_u \setminus$ 偏离）$\setminus Q_1$。根据定义，$_-\pi \cdot Q^* \cdot$（$Q_u \setminus$偏离）$= Q_1 \uplus Q_2$；根据假设，对于所有 $\pi' \in Q_2$，都有$\mathbb{P}_{\pi'}[Q^* \cdot$（$Q \setminus Q_u \cup$偏离）$] > u/2$。由此可得，$\mathbb{P}_\pi[Q^* \cdot$（$Q_u \setminus$偏离）$^\omega \cap$（$Q^* \cdot Q_2$）$^\omega] = 0$。也就是说，$\mathbb{P}_\pi[Q^* \cdot$（$Q_u \setminus$ 偏离）$^\omega] = \mathbb{P}_\pi[Q^* \cdot$（$Q_1 \uplus Q_2$）$^\omega] = \mathbb{P}_\pi[Q^* \cdot (Q_1)^\omega]$。此外，根据 Q_1 的定义（因为在某个时刻，只会观察到属于某个集合且其概率至少为 $u/2$ 的状态，其中该集合的概率为 W_u）以及定理 6，有$\mathbb{P}_\pi[W_u \cap Q^* \cdot$（$Q_1$）$^\omega] = \mathbb{P}_\pi[Q^* \cdot$（$Q_1$）$^\omega]$。于是：

$$\begin{aligned}\mathbb{P}_\pi[W_u] &= \mathbb{P}_\pi[W_u \cap Q^* \cdot (Q_u \setminus \mathsf{Deviate})^\omega] + \mathbb{P}_\pi[W_u \cap (Q^* \cdot (Q \setminus Q_u \cup \mathsf{Deviate}))^\omega] \\ &= \mathbb{P}_\pi[W_u \cap Q^* \cdot (Q_1)^\omega] + \mathbb{P}_\pi[(Q^* \cdot (Q \setminus Q_u \cup \mathsf{Deviate}))^\omega] \\ &= \mathbb{P}_\pi[Q^* \cdot (Q_1)^\omega] + \mathbb{P}_\pi[(Q^* \cdot (Q \setminus Q_u \cup \mathsf{Deviate}))^\omega] \\ &= \mathbb{P}_\pi[Q^* \cdot (Q_u \setminus \mathsf{Deviate})^\omega] + \mathbb{P}_\pi[(Q^* \cdot (Q \setminus Q_u \cup \mathsf{Deviate}))^\omega] \\ &= 1\end{aligned}$$

这与事实 $\mathbb{P}_\pi[W_u] \leq u/2$ 矛盾。事实上，存在某个 $\pi' \in \pi \cdot Q^* \cdot (Q_u \setminus$偏离$)$使得我们有 $\mathbb{P}_{\pi'}[W_u] < \frac{u}{2}$和 $\mathbb{P}_{\pi'}[Q^* \cdot (Q \setminus Q_u \cup$偏离$)] \leq u/2$。令 $\pi' = \pi'' \cdot q$。由于 $\pi' \notin$偏离，我们有 $\chi_{\mathcal{G}}[\mathsf{s}^{\rho \cdot \pi'}{}_{\mathsf{pb}}](q) = \chi_{\mathcal{G}}(q) = u$。因此：

$$\begin{aligned}u \leq \mathbb{P}_{\rho \cdot \pi'}^{\mathsf{s}_{\mathsf{pb}}^{\rho \cdot \pi'}, \mathsf{s}_{\mathsf{B}}}[W] &\leq \mathbb{P}_{\rho \cdot \pi'}^{\mathsf{s}_{\mathsf{pb}}^{\rho \cdot \pi'}, \mathsf{s}_{\mathsf{B}}}[W \mid (Q_u \setminus \mathsf{Deviate})^\omega] \cdot \mathbb{P}_{\rho \cdot \pi'}^{\mathsf{s}_{\mathsf{pb}}^{\rho \cdot \pi'}, \mathsf{s}_{\mathsf{B}}}[(Q_u \setminus \mathsf{Deviate})^\omega] \\ &\quad + \mathbb{P}_{\rho \cdot \pi'}^{\mathsf{s}_{\mathsf{pb}}^{\rho \cdot \pi'}, \mathsf{s}_{\mathsf{B}}}[Q^* \cdot (Q \setminus Q_u \cup \mathsf{Deviate})] \\ &\leq \mathbb{P}_{\rho \cdot \pi'}^{\mathsf{s}_{\mathsf{pb}}^{\rho \cdot \pi'}, \mathsf{s}_{\mathsf{B}}}[W \mid (Q_u \setminus \mathsf{Deviate})^\omega] \cdot (1 - \frac{u}{2}) + \frac{u}{2}\end{aligned}$$

注意，我们确实可以将这些概率与先前的概率（用 $\mathbb{P}$ 表示）联系起来，采用策略 s^ρ，而不是 $\mathsf{s}^{\rho \cdot \pi'}{}_{\mathsf{pb}}$，因为这两个策略在偏离之外一致。我们得到：

$$\mathbb{P}_{\rho \cdot \pi'}^{\mathsf{s}_{\mathsf{pb}}^{\rho \cdot \pi'}, \mathsf{s}_{\mathsf{B}}}[W \mid (Q_u \setminus \mathsf{Deviate})^\omega] = \mathbb{P}_{\pi'}[W \mid (Q_u \setminus \mathsf{Deviate})^\omega] \geq \frac{u}{2-u}$$

因此我们可以得出：

$$\begin{aligned}\mathbb{P}_{\pi'}[W_u] &\geq \mathbb{P}_{\pi'}[W_u \cap (Q_u \setminus \mathsf{Deviate})^\omega] \\ &= \mathbb{P}_{\pi'}[W \mid (Q_u \setminus \mathsf{Deviate})^\omega] \cdot \mathbb{P}_{\pi'}[(Q_u \setminus \mathsf{Deviate})^\omega] \\ &\geq \frac{u}{2-u} \cdot (1 - \frac{u}{2}) \\ &= \frac{u}{2}\end{aligned}$$

这与事实 $\mathbb{P}_{\pi'}[W_u] < \frac{u}{2}$ 相矛盾。事实上，我们的假设 $\mathbb{P}^{\mathsf{s}_\rho,\ \mathsf{s}_\rho\ \mathsf{B}}[Q^* \cdot (Q_u \setminus \text{偏离})^\omega] > \mathbb{P}^{\mathsf{s}_\rho,\ \mathsf{s}_\rho\ \mathsf{B}}[W \cap Q^* \cdot (Q_u \setminus \text{偏离})^\omega]$ 并不成立。也就是说，我们有$\mathbb{P}^{\mathsf{s}_\rho,\ \mathsf{s}_\rho\ \mathsf{B}}[Q^* \cdot (Q_u \setminus \text{偏离})^\omega] = \mathbb{P}^{\mathsf{s}_\rho,\ \mathsf{s}_\rho\ \mathsf{B}}[W \cap Q^* \cdot (Q_u \setminus \text{偏离})^\omega]$。

□

定理 4 的证明可以类似地完成，此外需要注意的是，如果s_{fc} 是一个有限选择策略，那么本证明中定义的策略 s_{A} 也同样是有限选择策略。

E.2 定理 5的证明

首先注意到，W_{proj}是前缀无关的。事实上，对于 $\rho \in W_{\mathsf{proj}}$。我们有 $\rho = k^n \cdot \pi_0 \cdot k \cdot \pi_1 \cdots$对于 $n \in \{0,1\}$和 $\pi \in W$。于是，对于所有 $i \in \mathbb{N}$，存在某个$n' \in \{0,1\}$使得 $\rho_{\geq i} = k^{n'} \cdot \pi_j \cdot k \cdot \pi_{j+1} \cdots$以及 $j \in \mathbb{N}$。然后，由于 W是前缀无关的，因此 $\pi_{\geq j} \in W$。

现在，对于一个有限选择策略 s_{A} 和一个值 $u \in V_{\mathcal{G}} \setminus \{0\}$，我们定义博弈 $\mathcal{G}_u^{\mathsf{tb}}$：

定义 28 对于一个正值 $u \in V_{\mathcal{G}} \setminus \{0\}$，我们定义博弈 $\mathcal{G}_u^{\mathsf{tb}} = \langle \mathcal{C}_u, W_{\mathsf{proj}} \rangle$，其包含 $\mathcal{C}_u^{\mathsf{tb}} = \langle Q_u^{\mathsf{A}} \uplus Q_u^{\mathsf{B}} \uplus \mathcal{C}_W, (A'_q)_{q \in Q'_u}, (B'_q)_{q \in Q'_u}, \mathsf{D}\ \delta', \text{距离函数}', \mathsf{K}, \text{颜色函数}' \rangle$，其中我们将博弈场 $\mathcal{C}_W$称为汇点状态，并且有若干条转移指向该状态：

- $Q_{\mathsf{A}u} := Q_u$ 玩家 A 的状态集合；
- 对于所有 $q \in Q_{\mathsf{A}u}$，我们有 $A'_q := S_q$ 以及 $B'_q := \{b_{\mathsf{id}}\}$，其中 b 是一个新的动作 id；
- $Q_{\mathsf{B}u} := \{(q, \sigma) \mid q \in Q_u,\ \sigma \in A'_q\}$ 玩家 B 的状态集合；
- 对于所有 $q \in Q_{\mathsf{B}u}$，我们有 $A'_q := \{a_{\mathsf{id}}\}$，对于一个新的动作 a_{id} 和 $B'_q := B_q$；
- $\mathsf{D}' := \{d_{q,\sigma} \mid q \in Q_u,\ \sigma \in A'_q\} \uplus \{d_{q,\sigma,b} \mid q \in Q_u,\ \sigma_q \in A'_q, b \in B\}$;
- 对于所有 $q \in Q_u$和 $\sigma \in A'_q$，我们有：$\delta'(q, \sigma, b_{\mathsf{id}}) := d_{q,\sigma_q}$；
- 对于所有 $q \in Q_u$、$\sigma \in A'_q$和 $b \in B_q$，我们有：$\delta'((q, \sigma)\, a_{\mathsf{id}}\, b) := d_{q,\sigma,b}$；
- 对于所有 $q \in Q_u$、$\sigma \in A'_q$，我们有 距离函数$'(d_{q,\sigma})((q, \sigma)) := 1$。此外，对于所有 $b \in B'_q$ 和 $q' \in Q$，我们有：

$$\text{dist}'(d_{q,\sigma_q,b})(q') := \begin{cases} \sum_{a \in A} \sigma_q(a) \cdot \text{dist}(\delta(q,a,b))(q') & \text{if } q' \in Q_u \\ \sum_{a \in A} \sigma_q(a) \cdot \text{dist}(\delta(q,a,b))[Q \setminus Q_u] & \text{if } q' = \mathcal{C}_W \end{cases}$$

- 最后，对于所有 $q \in Q_u$，都有 颜色函数$'(q) :=$ 颜色函数(q)；并且对于所有 $\sigma \in A'_q$：颜色函数$'(q, \sigma) := k$ 对于某个任意颜色 $k \in \mathsf{K}$。

现在开始证明定理 5。

证明 与引理 *5* 的证明类似，我们可以证明，如果存在一个具有有限选择的子博弈最优策略，那么对于所有 $u \in V_{\mathcal{G}} \setminus \{0\}$，博弈 $\mathcal{G}_u^{\mathsf{tb}}$中的所有状态都具有正值。事实上，如果玩家 *A* 在轮流制博弈 $\mathcal{G}_u^{\mathsf{tb}}$ 中采用具有有限选择的子博弈最优策略，我们得到的马尔可夫决策过程（除去颜色为 k 的中间状态）与在并发博弈 $\mathcal{G}$ （限制于 Q_u）中玩家 *A* 采用相同的具有有限选择的子博弈最优策略时得到的马尔可夫决策过程相同。此外，所有博弈 $\mathcal{G}_u^{\mathsf{tb}}$都是有限的轮流制博弈。因此，我们可以应用定理 *2* 的证明：根据定理 *3*，在所有博弈 $\mathcal{G}_u$ 中，对于 $u \in V_{\mathcal{G}} \setminus \{0\}$，都存在子博弈几乎必胜策略。随后，我们将这些策略粘合为一个策略，从而得到子博弈最优策略；这由引理 *6* 给出。同样，如果对于所有 $u \in V_{\mathcal{G}} \setminus \{0\}$，所有策略 s_{A}^u 都是位置型的，那么粘合策略 $s_{\mathsf{A}}[(s_{\mathsf{A}}^u)_{u \in V_{\mathcal{G}} \setminus \{0\}}]$ 也具有位置型。

□

E.3 推论 2 的证明

我们来正式定义这些目标，并说明它们具有中性色，且属于 PSAWT。

首先考虑奇偶性目标。其形式化定义见定义 7。在 K 中，最小颜色（对于奇偶性目标而言是一个整数）显然是中性色（因为我们考虑的是无限次出现的颜色中的最大值）。此外，奇偶性目标是PSAWT，例如文献 [12, 23]中就证明了这一点。

现在来定义均值收益目标。

定义 29（均值收益）令 $\mathsf{K} := Q \cap [0,1]$ 且 $m \in \mathsf{K}$。相对于 m 的均值收益目标$W_{\mathsf{MP}(\mathsf{m})}$ 为 $W_{\mathsf{MP}(\mathsf{m})} := \{\rho \in \mathsf{K}^\omega \mid \mathrm{limsup}_{n\to\infty} \frac{1}{n+1}\sum_{i=0}^{n} \rho_i \geq m\}$。——

给定某个 $m \in Q \cap [0,1]$ 的均值收益目标 $W_{\mathsf{MP}(\mathsf{m})}$，可以看出，$m$ 实际上是 $W_{\mathsf{MP}(\mathsf{m})}$ 的中性色。此外，正如 [18] 中所证明的，它也是 PSAWT。

最后，考虑广义 B¨uchi 目标。

定义 30（广义 B¨uchi）设 $\mathsf{K} \subseteq \mathbb{N}$ 是一个有限整数子集。广义 B¨uchi 目标是 K 上 B ¨*uchi* 目标的交集。

通过添加一个在任何相交的 B¨uchi 目标中都未出现的新颜色，该目标便具有一个中性色。此外，正如 [8] 中所证明的，它也是 PSAWT，因为它属于向上闭合 Muller 目标的子类。

推论 2 的证明是直接的。

证明 根据定理 4，存在一个 $\mathcal{G}$ 中的子博弈最优策略，且它具有有限选择。由于所考虑的目标具有中性色，并且是*PSAWT*，根据定理 5，可知存在一个位置子博弈最优策略。 □