arXiv:2305.00474v1 [cs.SI] 30 Apr 2023

变化环境中的学习、多样性与适应：弱连接的作用*

Daron Acemoglu † Asuman Ozdaglar‡ Sarath Pattathil §

摘要

适应动态条件需要一定程度的多样性。如果所有智能体都采取最佳当前行动，那么它们要更慢才能学习到底层状态已经发生变化、行为应当随之适应。智能体之间沟通迅速时，多样性更难维持，因为它们往往会迅速发现并采取最佳行动。我们通过一个社会网络上的（贝叶斯）学习模型来探讨这些问题。智能体能够迅速从与自己通过强连接相连的其他智能体那里学习，也可能有动机与其协调。然而，我们证明，当底层环境变化足够迅速时，任何仅由强连接构成的网络，长期表现都只比随机选择略好。相较之下，同时结合强连接与弱连接的网络——弱连接仅缓慢传递信息——能够实现高得多的长期平均收益。最优社会网络会将很大比例的智能体纳入一个强连通分量，同时保留足够数量规模较小的社群；这些社群做出多样化的选择，并通过弱连接与该分量沟通。

关键词：适应、贝叶斯学习、变化环境、多样性、网络、强连接、弱连接。

JEL 分类：D83,D85.

*我们衷心感谢 Hewlett基金会、MIT Quest 和 Mathworks工程奖学金提供的资金支持。感谢 Ben Golub 和 Alex Wolitzky 提出的建议。

†麻省理工学院经济系。电子邮箱：daron@mit.edu。

‡麻省理工学院电气工程与计算机科学系。电子邮箱：asuman@mit.edu。§麻省理工学院电气工程与计算机科学系。电子邮箱：sarathp@mit.edu。

1 引言

经济动荡、气候变化、新型疫情以及复苏的民族主义和其他极端主义意识形态所带来的威胁日益加剧，重新激发了人们对理解社会如何增强韧性、应对挑战的兴趣（Keck 和 Sakdapolrak (2013)；Brunnermeier (2022年)。生态学和生物学领域的大量研究始于Fisher (1958)颇具影响力的研究，这些研究指出，多样性对于在变化环境中保持韧性至关重要。缺乏多样性的物种或许很适应特定环境，但面对重大变化时便难以适应。本文探讨社会背景下多样性与适应之间的关系。我们关注这一问题的一个方面：学习并适应变化环境。

1.1 主要论点

我们设想一组彼此互动并在两种行动之间进行选择的智能体，其中一种行动带来的收益更高。局部互动和知识流动会促使智能体选择与其邻居相同的行动。然而，当所有或大多数智能体都选择同一种行动时，即便该行动目前收益更高，学习动态改进的替代方案也会变得更加困难，并可能降低长期收益。

形式上，我们研究一个简单博弈的贝叶斯-纳什均衡，其中每个智能体的效用由物质收益和网络收益构成：前者取决于其行动是否与底层状态相符，后者取决于其行动与关系密切的邻居所采取的行动有多相符。关键在于，底层状态会随时间按照马尔可夫链变化，因此智能体必须适应不断演变的环境。我们假设智能体通过贝叶斯更新形成对底层状态的信念，但在基准模型中，它们只最大化当前效用，因此没有理由进行试验。（随后，我们将这些结果扩展到智能体最大化其贴现效用的情形。）

我们模型中的适应需求，形成了经典“开发—探索”权衡的网络版本：一些智能体应在多大程度上偏离当前最优行动，并进行实验，以判断底层环境是否已经发生变化？然而，与标准的实验问题不同，网络结构在这里至关重要。如果智能体彼此紧密相连，它们往往会采取相同的行动，既因为局部信息流，也因为局部交互。但如果所有智能体今天都采取最优行动，就会产生适应问题：当环境发生变化、需要适应（转而采取如今收益更高的行动）时，这种改变要么不会发生，要么发生得极其缓慢。另一方面，如果社会网络包含多个彼此断开的连通分量，且其中一些分量采取着不同的行动，那么整个社会将迅速发现环境变化；但由于社会图彼此断开，这一信息无法从一个局部社群传递到社会的其他部分。

本文的主要论点是，Granovetter 关于“弱连接”的观点（Granovetter（1977），即弱连接虽然不像紧密连接那样频繁互动，却可以偶尔充当信息传递渠道，为这一问题提供了一个有力的解决方案。在此观点的基础上，

我们将弱连接建模为间歇性地传递有关行为和收益的信息，但不会促使局部行动趋于一致。我们证明，由若干彼此强烈隔离但通过弱连接相互联系的集群组成的社会，能够快速适应不断变化的环境，同时确保大多数智能体在大部分时间采取高收益行动。

我们还从最大化长期平均收益的角度刻画了最佳社会网络。结果发现，一种星形网络能够实现最高的长期平均收益。这种星形网络由大量紧密连接的智能体组成，它们聚集在中央的星形节点中；此外，还有数量充足的小型社群，作为该星形节点弱连接的叶节点。叶节点负责实验，并确保整个社会能在底层环境发生变化时迅速学会适应。中央的星形社群既能利用局部交互带来的收益，也能受益于信息优势，而这些信息能从叶节点迅速获取。

我们通过证明没有弱连接时的长期平均收益仅略高于随机行动，确认了弱连接对这一结果至关重要，因为大多数智能体会固守曾经有效、但后来已不再有效的行动。在这类网络中，对变化环境的适应只能依靠缓慢的突变或错误；在我们的假设下，这类转换的发生速度远低于环境变化的速度。这一结果表明，正是弱连接的存在以及恰当的强连接拓扑结构，确保了社会能够实现接近最高可能水平的收益。

我们的基础分析针对的是智能体选择使当前收益最大化的行动这一情形。我们还表明，只要折扣率不太高，我们的结果也适用于智能体最大化折扣收益的情形。具体而言，我们推导出了该折扣率的一个上界：当折扣率低于此上界时，我们的所有主要结果仍然成立。尤其是，在不存在弱连接的情况下，长期平均收益与随机行动大致相当；而星形网络则能实现高得多的长期平均收益。这个上界取决于局部互动的强度和网络的最大度数。

1.2 更广泛的背景

我们认为，我们的研究结果不仅与博弈论和经济学文献相关，也与关于多样性和适应的更广泛文献相关。多样性这一主题在生物学中至关重要，但社会系统还面临一个关键问题：如何激励智能体采取能够保留多样性的行动。

多样性带来的适应益处，得到了对若干不同物种的研究支持。例如，Agha 等人(2018)通过实验表明，蓝细菌在遗传上同质时，更容易受到真菌寄生物的侵害。事实上，在保持同质的宿主群体中，寄生物能够迅速传播；而遗传多样性较高的宿主群体则能更有效地抵御寄生物，因为其中含有较不易受特定寄生物侵害的基因，并且这些基因会在入侵发生后更快增殖。在Fischer (2004)的研究中也观察到了蜜蜂多样性在应对温度波动方面的类似益处。每只蜜蜂抱团取暖和扇风的温度阈值都与

一种遗传关联性状有关。在这一性状上缺乏遗传多样性的蜂巢，其内部温度会出现异常大的波动；而具有遗传多样性的蜂巢，其内部温度则稳定得多。因此，蜜蜂的遗传多样性有助于维持相对稳定的温度，最终改善蜂巢的健康状况。

即使在生物系统中，维持多样性也是一项重大挑战。关于有性生殖益处的理论中，最广为接受的一种恰恰认为，有性生殖通过混合双亲的等位基因，确保个体和群体内部拥有足够的多样性（例如参见Weismann 等(1904）；Barton 和 Charlesworth (1998)；Burt (2000））。因此，与无性生殖相比，有性生殖能够通过适应变化环境而获得更高的适合度。对酵母的实验为这一假说提供了证据。具体而言，Goddard 等(2005)对一个菌株进行基因改造，培育出两种酵母菌株；二者完全相同，唯一区别在于繁殖方式。实验确认，有性生殖的菌株比无性生殖的菌株更能适应严酷的环境。

人们也曾假设，多样性在社会环境中具有类似的适应优势，并且有时已得到证实。Granovetter(1977、1983、2017)认为，新型且更优越的技术会通过弱连接在科技集群中迅速传播，例如硅谷。弱连接既可能由不同公司的员工或经理之间的沟通形成，也可能直接由员工在公司间流动而形成（关于硅谷的这一模式，参见Saxenian(1996)；关于城市环境中沟通这一方面的更一般性论述，参见Jacobs(2016）。[1]其他研究强调了能够连接社群不同部分之间“结构洞”的智能体的重要性（Burt (1992）。这一视角也重新诠释了 Robert Putnam 对多样化组织（例如保龄球馆、体育俱乐部和地方宗教组织）重要性日渐下降的担忧（Putnam 等(2000）。这些组织能够提供一种弱连接，跨越结构洞，在原本鲜少互动的不同社会群体之间传递信息。我们的讨论还强调，这种情况尤其重要的一点是，它不应引发强连接往往会造成的强烈同质化倾向。

1.3 经济学与博弈论文献

在经济学和博弈论文献中，本文与若干不同领域的研究相关。首先是一小部分关于适应与多样性的文献。Gross (1996研究了物种内部存在大量非典型变异的原因，并将其与适应联系起来。与本文更为密切相关的是Santos etal. (2008)，他们分析了多样性在公共物品博弈中的作用，并指出多样性能够促进合作。经济学领域的许多研究通常认为，现代社会中的多样性会导致冲突（例如，参见La Ferrara and Mele(2006)的综述），不过也有少数论文，例如Montalvo and Reynal-Querol (2021)，记录了多样性带来的各种

[1]Rajkumar 等人（2022年）利用 Linkdin 数据证实，弱连接仍然是求职的核心因素。他们强调，弱关系强度与工作机会传递之间存在“倒U形关系：关系较弱会增加工作机会传递，但这种作用只在一定范围内成立；超过这一范围后，关系弱化带来的边际收益递减。”这与我们命题1的结果一致。

经济收益。

经济学中有几篇论文研究了社交网络中的学习动态。我们的研究与其中聚焦于贝叶斯模型的分支最为直接相关，例如 Gale 和 Kariv（2003）、Banerjee 和 Fudenberg（2004）、Smith 和 Sørensen（2008）、Callander 和 H¨orner（2009），以及 Acemoglu 等人（2011）。此外，还有几篇论文探讨了社交网络中的非贝叶斯学习，其中尤以 Bala 和 Goyal（1998、2001）、DeMarzo 等人（2003）以及 Golub 和 Jackson（2010）的研究最为著名。这些论文都没有考虑对变化环境的适应问题，不过，如何在强连接带来的从众倾向与为智能体采取不同的行动提供足够激励之间取得平衡，这一问题在 Smith 和 Sørensen（2008）及 Acemoglu 等人（2011）中有所涉及。与我们的研究关系更密切的是近期几篇探讨相关问题中学习速度的论文。例如，Acemoglu 等人（2022年）刻画了贝叶斯代理人在观察不同的在线过往评论样本时的学习速度；关于文献中学习速度研究结果的讨论，请参阅他们的论文。

与我们的研究关系更为密切的，是几篇探讨底层状态变化时学习问题的论文。Moscarini 等人(1998年指出，除非底层状态“足够持久”，否则单一行动上不会出现（贝叶斯）级联。Frongillo 等人 (2011)考察了各种非贝叶斯学习规则，并表明，尽管环境不断变化，这些规则在完全图上仍会收敛到稳态分布。Dasaratha 等人 (2018研究了一种学习模型，其中个体从他人和自己的私人信号中学习。他们表明，私人信号越多样，学习效果越好；这一逻辑与我们的主要结果相关。最后，L´evy 等人 (2022年)的研究也与本文密切相关。他们指出，在智能体对称的情况下，即使底层状态发生变化，所有参与者也会迅速收敛到相同的（共识）行动。据我们所知，这些论文以及其他研究都未曾考察一般网络和状态变化条件下的贝叶斯学习，也未曾刻画哪些类型的网络能带来更好的学习表现，或对弱连接的重要性进行建模和观察。

我们的模型结构也与关于进化或学习动态和均衡选择的文献相关。在这一文献中，开创性研究Kandori 等人(1993) 考察了一个进化模型：有限数量的智能体随机匹配并进行双人协调博弈，同时受到噪声或突变的影响。他们表明，噪声的存在会缩小长期“均衡”（稳定配置）的范围；特别是在一个 2×2 博弈中，进化动态会促使系统趋向帕累托占优的纳什均衡。在相关研究中，Young(1993) 刻画了一个大型有限群体博弈中的随机稳定均衡，该博弈存在随机匹配和噪声。与Kandori 等人(1993) 的研究一样，噪声起到了均衡选择机制的作用。Ellison (1993) 指出，Kandori 等人 (1993) 和 Young (1993) 中的均衡选择过程非常缓慢，并提出局部匹配而非随机匹配能够显著加快收敛。我们的研究与这类文献有几项重要区别。首先，据我们所知，这类文献并未研究对变化环境的适应问题或多样性的作用。相反，这类文献关注的是多重均衡博弈中的均衡选择。其次，我们关注的是底层状态变化的博弈中的贝叶斯-纳什均衡，而非进化规则或经验法则式行为。

最后，我们使用的一些数学方法也常见于一般信念动态的相关文献。Holley 和 Liggett (1975) 例如研究了所谓的“投票者模型”，该模型与Kandori 等人(1993) 和 Young (1993) 基于随机匹配的进化动态相似（影响在随机匹配的双方之间传递）。相比之下，我们的模型所产生的随机动态更接近Kanoria 等人 (2011) 和 Yildiz 等人 (2010) 研究的“多数派动力学”。

1.4 论文其余部分

论文其余部分的安排如下。在第 2 节中，我们介绍模型并定义贝叶斯-纳什均衡。在第 3 节中，我们刻画均衡，并提供一种分析一般网络的方法。在第 4节中，我们比较强连接和弱连接网络，并分析能够带来最高福利的网络。在第 5节中，我们将这些结果推广到前瞻型智能体，最后在第 6节讨论研究结果。

2 模型

本节介绍基本环境，描述由强连接和弱连接构成的网络、收益与平均福利，并定义贝叶斯-纳什均衡。

2.1 网络

我们考虑一组智能体 $V = \{1, 2, \cdots, n\}$，它们由无向图 G 中的节点表示。连接分为两类：强连接和弱连接。我们用对称矩阵 $S^G \in \{0,1\}^{n\times n}$ 表示强连接，约定

$$S_{ij}^G = \begin{cases} 1 & \text{if agents } i \text{ and } j \text{ have a strong link between them} \\ 0 & \text{otherwise} \end{cases}$$

智能体 i 的邻域是相对于强连接定义的，即 $N^G(i) = \{j \in V : \ S_{ij} = 1\}$。网络的最大度记为 $d_{\max}^G = \max_{i\in V} |N^G(i)|$。

另一方面，弱连接由对称矩阵 $W^G \in \{0,1\}^{n\times n}$描述，其中同样

$$W_{ij}^G = \begin{cases} 1 & \text{if agents } i \text{ and } j \text{ have a weak link between them} \\ 0 & \text{otherwise} \end{cases}$$

我们还分别用 $\mathcal{E}_s^G$和 $\mathcal{E}_w^G$表示强连接和弱连接的集合。只要不会引起混淆，我们就省略上标 G。

2.2 行动与奖励

时间是连续的，并延伸至无穷远。在每个时刻 $t \in [0, \infty)$，智能体 $i \in V$ 选择行动$a_i(t) \in \mathcal{A} = \{0, 1\}$。智能体由此获得的收益由两部分组成：

1. 物质收益 $R_a(t)$，它仅取决于智能体所采取的行动 a 以及底层自然状态（因此具有随机性）；

2. 网络收益，它取决于智能体邻域内的行动，具体将在第 2.4节介绍。

之所以需要适应，是因为底层状态，以及两种行动 $R_0(t)$ 和 $R_1(t)$ 所带来的物质收益，会随时间变化。我们假设这些变化按照速率为 λ 的泊松时钟到达，并将这些变化发生的（随机）时刻记为 $\{T_k\}_{k=0}^{\infty}$，称之为收益冲击（并令 $T_0 = 0$）。不失一般性，我们假设在时刻 T_k 泊松时钟触发后，奖励于 T_k^+ 发生变化，即紧接在 T_k 之后。这意味着，对于所有 k，某一行动在区间 $(T_k, T_{k+1}]$ 内的奖励保持不变。我们还通过假设两种行动 $R_0(t)$ 和 $R_1(t)$ 之间的收益差恒定，并将其归一化为 1，来简化分析。不过，至关重要的是，哪种行动的收益更高会随着泊松时钟的触发而自然改变。此外，我们将 $A(t)$ 定义为时刻 t 奖励较高的行动，并将 $\{A_k\}_{k=1}^{\infty}$ 记为区间 $(T_{k-1}, T_k]$ 内奖励较高的行动。

概括这一奖励结构，我们可以写出，对于所有（随机）时间点 $\{T_k\}_{k=0}^{\infty}$，有

$$R_0(T_k^+) - R_1(T_k^+) = \begin{cases} +1 & \text{w.p. } 1/2 \\ -1 & \text{w.p. } 1/2 \end{cases}$$

对于所有 $t \in (T_k,\ T_{k+1}]$，$R_0(t) - R_1(t) = R_0(T_k^+) - R_1(T_k^+)$。

另请注意，$\lambda = 0$ 的情况构成了物质收益恒定且已知的特殊情形。我们假设所有智能体在时间 $t = 0$ 均初始化为执行动作 0。

2.3 信息结构

接下来我们描述信息结构，该结构取决于强连接和弱连接的性质。

强连接： 在所有时间 t，每个智能体 i 都将获得其强连接邻居在集合 $\mathcal{N}(i)$ 中的行动历史及相关收益的完整信息。

弱连接： 与强连接相比，弱连接传递信息的速度较慢。我们通过假设弱连接起初处于“休眠”状态，并以随机方式激活来对此进行建模。具体而言，存在一个速率为 γ 的泊松时钟，每当时钟滴答一次，就会激活一条休眠的弱连接。此外，一旦弱连接被激活，它就会传递信息，随后进入“非活跃”状态，直到另一个独立的泊松时钟（此次速率为 ϕ）使其重新进入“休眠”状态。我们将在下文解释这种两阶段激活的原因。首先，我们说明如何从所有弱连接中选出被激活的弱连接。

令 $\mathcal{W}(t) = \{(i, j) \in \mathcal{E}_w^G: a_i(t) \neq a_j(t), (i, j)$ 处于休眠状态$\}$。换言之，这是那些处于休眠状态、且涉及在时刻 t 采取不同行动的两个相连智能体的弱连接集合。这是与信息传递相关的弱连接集合，因为采取相同行动的智能体之间没有相关信息需要传递。我们假设，一旦相关的泊松时钟滴答，就会从 $\mathcal{W}(t)$ 中均匀随机地选取一条连接。此后，该连接变为激活状态，信息便通过这条连接传递；也就是说，如果完全激活的连接是 (i, j)，则个体 i 的当前行动和收益会传递给 j，反过来，j 的信息也会被 i 观察到。信息传递完成后，该连接进入非激活状态，并保持这一状态，直到速率为 ϕ 的泊松时钟滴答，此后它再次变为休眠状态。

此处有几点值得说明。首先，弱连接上的信息传递速度较慢，与强连接上极快的信息传递形成对比。强连接体现的是频繁互动，例如家庭成员、同事或联系紧密的智能体之间的互动；弱连接则通过闲聊或随机观察偶尔传递信息。从我们的数学表述来看，弱连接只有在从非活跃状态转为休眠状态，并等待到激活后，才会传递信息。正如我们将看到的，这种缓慢的信息传递对我们的结果起着关键作用。其次，已激活的弱连接存在于行动不同的智能体之间，这与弱连接因闲聊或信息交换而激活这一观点相符。在我们的设定中采用这一假设，主要是为了简化分析：如果不作此假设，部分已激活的弱连接可能不会传递相关信息。尽管这不会影响我们的总体结果，但已激活的连接若不传递有用信息，会使我们在证明中使用的耦合论证更加困难。第三，两阶段激活对于确保信息传递足够缓慢十分重要。具体而言，如果不存在非活跃状态，同一条弱连接可能会被多次选中（因为弱连接是从行动不同的智能体之间选出的），而其他弱连接则始终不会被激活。通过两阶段激活，我们可以确保一条弱连接在传递相关信息后会转为非活跃状态，并在由速率为 ϕ的泊松时钟决定的某个“退避”时段内无法传递任何信息。此后，它会进入休眠状态，成为信息传输通道的候选者。在这一表述中，$\phi \to \infty$ 对应于弱连接永远不会处于非活跃状态的情形，而 $\phi \to 0$对应于弱连接被激活以传递信息后便进入非活跃状态、此后永远不会再传递信息的情形。

2.4 总体收益与信念

智能体最大化其静态的当前收益（直到第 5 节，我们将在该节引入前瞻性行为）。如上所述，智能体 i 在采取行动 a_i 于时间 t时的每期总效用为

$$\mathcal{U}_i^{a_i}(t) = R_{a_i}(t) + \tau f_i(a, t),$$

其中，$R_a(t)$ 是该智能体的物质收益，如上所述；$\tau f_i(a,\ t)$ 是其网络收益，其中 $a = [a_1, a_2, \cdots \cdots, a_N]$ 表示该群体的完整行动组合（不过，起作用的是智能体 i 的邻居所采取的行动）。具体而言，我们将这一网络收益定义为智能体 i 的邻居在时间 t 采取行动 a_i 的人数。也就是说，

$$f_i(a,t) = \sum_{j \in N^G(i)} \mathbb{I}_{a_j(t)=a_i},$$

其中，$\mathbb{I}_{a_j(t)=a_i}$ 是指示函数，用于表示智能体 i 的邻居 j 在时间 t是否采取与该智能体相同的行动 a_i 。直观地说，该项刻画了与联系紧密的智能体进行协调所带来的收益。参数 $\tau \geq 0$ 表示该局部网络收益的重要性。[2]

虽然网络收益是确定性的（给定其他智能体的行动组合），但物质收益具有随机性，并取决于上文所述的底层状态。因此，智能体的最优回应将取决于其信念，下面对此进行描述。

令 $\mu_i(t)$表示智能体 $i \in V$ 在时间 t认为行动1具有更高奖励的信念，即，$R_1(t) - R_0(t) = +1$。更正式地说，

$$\mu_i(t) = \mathbb{E}_{i,t}[\mathbb{I}_{R_1(t)>R_0(t)}],$$

其中，$\mathbb{E}_{i,t}$ 表示根据智能体 i 在时间 t的信息集所得到的期望，并且$\mathbb{I}_{R_1(t)>R_0(t)}$是 $R_1(t) > R_0(t)$ 的指示函数。我们假设，对于所有智能体 $i \in V$，均有$\mu_i(0) = 1/2$，即在时间 $t = 0$ ，智能体对于哪个行动具有更高的物质收益没有任何信息。

T智能体最大化当前收益这一假设意味着

$$a_i(t) = \underset{a \in \{0,1\}}{\operatorname{argmax}}\ \mathbb{E}_{i,t}[\ \mathcal{U}_i^a(t)\ ]. \tag{1}$$

最后，和Kandori 等人（1993）以及Young （1993）一样，我们引入个体颤抖。我们假设，另一个速率为 $\epsilon > 0$ 的泊松时钟会引发行为变化。具体而言，每当时钟滴答一次，就会随机均匀地选取一名智能体，而她最终采取的行动与原本打算采取的行动相反。我们将这一现象称为 ϵ-颤抖。在全文中，我们假设 ϵ 很小，事实上远小于底层状态变化的速率（λ）以及弱连接传递信息的速率（γ）。

2.5 贝叶斯-纳什均衡

我们关注这个博弈的贝叶斯-纳什均衡。贝叶斯-纳什均衡（BNE）的定义采用标准形式。

[2]本节中的所有结果在 $\tau = 0$时仍然成立，因此不存在网络收益；但这种局部收益互动在第5节分析的前瞻情形中变得十分重要。

定义 1（贝叶斯-纳什均衡）若行动组合 $a = [a_1, a_2, \cdots \cdots, a_N]$ 满足$a_i \in \{0, 1\}$，则它是纯策略贝叶斯纳什均衡，当且仅当对每个 i，a_i 在其他智能体的行动组合给定时，使方程(1)中的预期收益最大 a_{-i}；其中，(1)中的期望$\mathbb{E}_{i,\ t}$是根据贝叶斯法则计算的。

2.6 平均福利

我们通过考察不同社会网络的长期平均收益（在贝叶斯纳什均衡中）来评估其适应成效。这一指标很有吸引力，因为只有能够迅速响应变化环境的社会，才能实现较高的长期平均收益。[3] 形式上，在时间点 T_k，由 n 个智能体组成的社会的平均收益为

$$\mathcal{S}_k^G = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{a_i(T_k)=A_k},$$

其中，$\{T_k\}_{k=0}^{\infty}$ 是收益冲击发生的时点，而指示函数$\mathbb{I}_{a_i\ (T_k)\ =A_k}$ 在 $a_i(T_k) = A_k$ 时取值为 1，否则取值为 0。我们以图 G所表示的社会网络为条件。长期平均福利定义如下：

$$\mathcal{S}^G = \lim_{k\to\infty} \mathcal{S}_k^G. \tag{2}$$

有几点值得注意。首先，我们只关注收益冲击发生的时点，因为在两次收益冲击之间，节点处于（可能的）暂态状态，并试图通过强连接和弱连接进行学习，以判断哪种行动具有更高的物质收益。其次，我们也可以将长期平均收益等价地定义为所有时期收益的平均值。不过，这种定义取决于初始行动，尽管随着取极限，初始行动的权重会趋于零。当前的定义使表述更加简洁，且不失一般性。

3 均衡刻画

本节将刻画贝叶斯纳什均衡（BNE），并给出任意 BNE 下平均福利的表达式。我们的刻画过程如下。首先，我们证明贝叶斯信念具有单调性，即在下一次信息到达之前，信念动态的变化方向不会逆转，并且在信息到达时会跳至正确的概率。基于这一刻画，我们证明智能体只会在信息到达时刻改变其行动。结合强连接的结构，我们进一步表明，在任意 BNE 下，除信息到达时刻外，强连接组件始终采取相同的行动。本节最后一小节将基于这一均衡结构刻画平均福利，并使用一个专门构造的嵌入式马尔可夫链，该链定义在社会网络中智能体的行动组合之上。

[3]如果我们转而关注　　贴现平均收益，这会降低未来未能适应变化的重要性　　s.

3.1 信念动态

下一个定义将引入信息到达时刻（集合）。直观而言，这些是智能体 i 接收到"新信息"的时刻。这可能是因为与该智能体相邻的弱连接被激活、强连接邻居改变了行为，或者该智能体自身发生了一次 ϵ-颤动。形式化地说：

定义 2（新信息的最后时刻）实例 t是智能体 i 的信息到达时刻，若在时刻 t发生以下任一情况：

- 与智能体 i 相邻的弱连接被激活。
- 对于某个 $j \in N(i)$，有 $a_i(t) \neq a_j(t)$。
- 智能体 i 发生了一次 ϵ-颤动。

智能体 i 的信息到达时刻定义如下

$$\mathcal{T}_i = \{t : \ t \ is \ a \ time \ of \ information \ arrival \ for \ agent \ i \ \},$$

以及在 t之前最后一次信息到达的时间是

$$T_i(t) = \sup\{t_i : t_i \in \mathcal{T}_i, \ and \ t_i \leq t\}.$$

提醒读者，鉴于迄今为止所规定的信息结构，每次信息到达都会完全揭示哪项行动具有更高的（物质）奖励。因此，在信息到达时，智能体 i相信行动 1 是较优行动的概率只能是 0 或 1。

我们还注意到，智能体 i 在时刻 t的信息集，记作 $\mathcal{I}_i(t)$，可由时刻 t之前最后一次信息到达 $T_i(t)$以及智能体此时观察到的行动组合完全概括。回顾一下， $\mu_i(t)$表示智能体 i在时刻 t 相信行动 1 的物质收益高于行动 0 的概率，因此 $\mu_i(t) = \mathbb{P}(A(t) = 1 \mid \mathcal{I}_i(t))$（还要回顾一下， $A(t)$表示在时刻 t具有较高物质收益的行动，该行动对所有智能体均相同）。

借助这一记号，我们现在可以确立贝叶斯更新的一项关键性质，从而能够刻画贝叶斯纳什均衡。

引理1 时刻 t、 μ_i的贝叶斯信念(t)满足以下单调性：

$$\mu_i(T_i(t)) = 1 \implies \mu_i(t) > \frac{1}{2}$$
$$\mu_i(T_i(t)) = 0 \implies \mu_i(t) < \frac{1}{2}.$$

引理 1 指出，一旦智能体意识到物质收益较高的行动（这发生在信息到达之后），她就会一直认为该行动更有可能是高回报行动，直到下一次信息到达

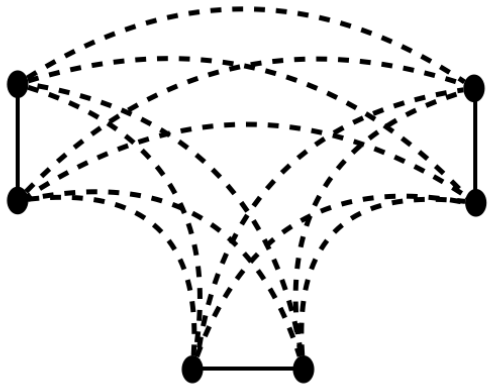

图 1：满足例 1 中假设的网络

。因此，一旦智能体在时间 $T_i(t)$ 认为，例如，行动 1 优于行动 0，她就会继续认为行动 1 优于行动 0 ($\mu_i > 1/2$)，直到收到新信息。

尽管引理1 证明了贝叶斯信念的单调性，但并未完整刻画信念动态。一般而言，这种刻画很难实现，不过在某些特殊情形下可以做到，正如下一个例子所示。此例纯粹用于说明；在本文其余部分，我们只使用引理1中的单调性结论。

例 1 现在，我们给出模型的一个特殊情形。在该情形下，网络具有足够的对称性，因此可以明确刻画贝叶斯更新；当然，这也验证了引理1。在此例中，我们首先对图结构施加一些限制。具体而言，我们假设图 G 满足

$$S_{ij}^G = 1, S_{ik}^G = 1 \implies S_{jk}^G = 1 \qquad \forall\ i, j, k \in V$$

and

$$W_{ij}^G = 1 \qquad \forall\ i, j \in V \text{ such that } S_{ij}^G = 0.$$

我们刚刚定义了一类网络：任意两个未通过强连接相连的智能体之间都存在弱连接，且智能体组成由强连接构成的团簇之岛；也就是说，每个智能体都属于一个由强连通智能体组成的团。这样的团簇有多个，形成“岛”。图1 展示了这类网络的示意图。接下来，我们进一步限定弱连接结构，假设每当泊松时钟（速率为 γ）跳动时，网络中的所有弱连接都会被激活。在这些假设下，随着我们取

极限 $\epsilon \to 0$，我们可以如下刻画精确的信念更新

$$\mu_i(t) = (1 - e^{-\lambda(t-T_i(t))})\frac{1}{2} + e^{-\lambda t}\mu_i(T_i(t)).$$

也很容易看出，该方程中的信念满足引理 1中的单调性。

引理 1中的信念单调性立即得出我们的下一个结果，该结果表明，智能体只会在信息到达时改变自己的行动。

引理 2 对于任意智能体 $i \in V$ 以及所有 t，均有：

$$a_i(t^-) \neq a_i(t) \implies t \in \mathcal{T}_i.$$

因此，如果任意智能体在时刻 t改变了自己的行动，那么 t 必定是信息到达的时刻。引理 2的一个直接但重要的推论是，所有智能体都会维持当前行动，直到发生以下两种事件之一：弱连接激活或 ϵ-颤动。[4]

有了这些结果，我们现在可以刻画整个网络的贝叶斯纳什均衡行动组合。对于该定理，我们定义 $\mathfrak{s}_{ij} = 1$：如果存在一条强连通路径连接智能体 i和 j（即存在一条由智能体 k_1、……、k_K组成的路径，连接 i和 j，且满足 $S_{ik_1} = S_{k_1k_2} = \ldots S_{k_Kj} = 1$）。如果网络中对所有 i，$j \in V$都成立，我们也称该网络是强连通的。最后，如果所有智能体的邻居数量相同（因此 $d_{\max} = d_{\min}$），则称该图是正则图。

定理 1 贝叶斯纳什均衡总是存在。设 $a(t) = [a_1(t)\ a_2(t)\ \cdots\ \cdots, a_N(t)]$ 是时间 t 的一个贝叶斯纳什均衡行动组合。那么：

- 如果 $\tau \leq 1/d_{\text{最大值}}$，所有由强连通路径连接的智能体都会采取相同行动。也就是说，对于所有 $i, j \in V$ 以及所有时期 t，

$$\mathfrak{s}_{ij} = 1 \implies a_i(t) = a_j(t).$$

- 如果 $\tau > 1/d_{\text{最小值}}$，所有智能体在所有时期 t 都继续采取其初始设定的相同行动，即行动 0。

  尤其是，如果 G 也满足正则条件（$d_{\max} = d_{\min}$），那么对于所有 $i, j \in V$ 以及所有时期 t，都有 $a_i(t) = a_j(t)$。

定理 1 大大简化了对任意贝叶斯纳什均衡的刻画。具体而言，如果局部收益交互度（由参数 τ 衡量）不太大，那么所有强连通智能体以及所有通过强连通路径相连的智能体，始终会采取

[4]智能体也可以从其一位强连接邻居那里获得新信息，但该邻居要改变她的行动，还需要发生弱连接激活或 ϵ-颤动。

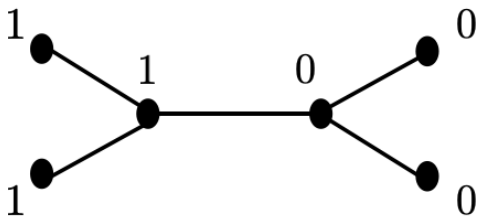

图 2：当 $\tau = 2/5$ 时的一种异质贝叶斯纳什均衡。

相同行动。值得注意的是，即使在 $\tau = 0$ 的情况下，这一点依然成立，因为强连接能够完美传递有关底层状态的信息，从而有力地促使所有智能体采取相同行动。鉴于这些信息，强连通分量中的所有智能体对于哪种行动能带来更高的物质收益，都有相同的信念。因此，当 $\tau = 0$ 时，他们都会采取相同行动。当 τ 不太大时，同样的结论也成立。在这种情况下，还存在一种额外的作用力，即希望与自身局部邻域中的行为保持一致。这通常会进一步促使强连通分量中的所有智能体采取相同行动。尽管如此，接下来的例子表明，当参数 τ 大于 $1/d_{\max}$ 时，与邻居保持一致的愿望可能导致强连通分量的不同部分采取不同的行动。然而，定理的第二部分指出，即便如此，只要阈值 τ 足够高，仍可实现协调。但如此高的阈值也有弊端：即使某个节点知道行动1的物质收益更高，她仍会继续采取动作 0，因为她的所有邻居都在采取动作 0，而且与邻居保持一致所带来的效用高于采取物质收益更高的行动所带来的效用。

例2图2 展示了一个网络，在这个网络中，强连通的智能体之间可以维持不同的行动。图中展示的网络具有阈值 $\tau = 2/5$ ，在强连通分量内存在不同智能体采取不同行动的贝叶斯纳什均衡。直观地说，尽管该网络是强连通的，但它包含两个不同的部分，而且每个部分内的局部行动对收益的影响都大于另一半中的行动。这足以维持一种均衡：左侧采取行动1，右侧采取动作 0。这个例子表明，只要阈值足够高，即使在强连通网络中也能保持多样性。

3.2 平均福利

本小节将对贝叶斯纳什均衡下的平均福利进行一般性刻画。该刻画基于对网络中智能体的行动组合定义一个嵌入式马尔可夫链。

请注意，我们使用“嵌入式”一词，是因为我们考察的是离散时间的马尔可夫链，尽管其底层学习过程是在连续时间内进行的。具体而言，转变

仅在发生收益冲击时才会在此马尔可夫链中发生，这足以让我们跟踪长期平均收益（根据方程（2））。[5]

定义 3（激活马尔可夫链）激活马尔可夫链（*AMC*）是一种嵌入式马尔可夫链，其中 i^{th} 转移发生在时间 T_i。

AMC 的状态：该马尔可夫链的状态表示为（P、 R、 B），其中：

- $P \in \{0,1\}^n$ 表示智能体所采取的贝叶斯纳什均衡行动组合。
- $R \in \{0,1\}$ 表示奖励较高的行动，即，如果我们处于时期 k，则$R = A_k$。
- $B \subseteq \mathcal{E}_w^G$ 表示在一个时期结束时处于休眠状态的弱连接集合。

AMC 的转移概率：该链的转移概率定义如下：

$$\mathbb{P}^G((P_l, R_l, B_l)|(P_m, R_m, B_m)) = \frac{1}{2} \times \mathbb{P}^G(P_l, B_l|P_m, R_l, B_m),$$

其中 $\mathbb{P}^G(P_l B_l|P_m, R_l, B_m)$ 表示在行动组合初始化为 P_m、B_m 中的弱连接处于休眠状态，且物质收益较高的行动为 R_l 的条件下，按照 P_l 采取行动且 B_l 中的弱连接处于休眠状态的概率。

定义 3 中的 AMC 概括了收益冲击发生时网络中智能体的行为。例如，假设收益冲击发生时我们处于状态 m，该状态由元组 (P_m, R_m, B_m) 给出。根据定义，冲击到来之前，智能体按照 P_m采取行动，且物质收益较高的行动是 R_m。此外，B_m 中的弱连接处于休眠状态，意味着它们可能被激活。这也意味着 $\mathcal{E}_w^G \setminus B_m$ 中的弱连接处于非活跃状态。冲击发生后， R_l 成为物质收益较高的行动。此后，弱连接激活和 ϵ-颤动可能导致智能体的行动组合发生变化。使用嵌入式马尔可夫链的一大便利之处在于，我们无需追踪这些行动组合的中间变化。相反，只需关注所有这些变化发生后的行动组合，即下一次收益冲击发生时正在采取的行动组合，记作 P_l。此外，由于一个时期内传递的信息取决于哪些弱连接处于休眠状态，我们也在状态 B_l中记录这些信息。这也解释了为什么转移概率中会有一个 $1/2$：下一次收益冲击发生时，两种行动中的每一种成为物质收益较高行动的概率均为 $1/2$。

总之，AMC 概括了收益冲击发生时行动组合之间转移的信息。这一点特别有用，因为根据我们在公式（ $\mathcal{S}^G$ 2）中对长期平均收益的定义，只需了解收益冲击发生时的收益。

[5]嵌入马尔可夫链用于排队论。在排队论中，作业到达和作业离开发生在连续时间内，但有时根据作业到达和离开时间构建的离散时间表示更有用（例如，Wolff （1989））。

下一个定理利用了这一特性，并通过 AMC 的平稳分布刻画长期平均收益。

定理 2 对于任意（弱）连通图 G，定义 3中的 ***AMC*** 平稳分布（记为 η^G）存在。此外，长期平均福利可以表示为该平稳分布的函数：

$$\mathcal{S}^G = \sum_q \eta_q^G f(q),$$

其中，η_q^G 是状态 q 的平稳概率，而 $f(q)$ 表示状态 q中采取高回报行动的智能体比例，定义如下：

$$\begin{aligned} f(q) &= f(P_q, R_q, B_q) \\ &= \frac{1}{n} \sum_{v \in V} \mathbb{1}_{P_q(v)=R_q}. \end{aligned}$$

定理 2是本文的主要结果之一，它对长期平均福利给出了紧确刻画。在本文余下部分，我们将运用这一刻画，确定在变化环境中哪些社会结构能够实现较高程度的适应和福利。正如我们将看到的，计算许多图（包括我们将在主要结果定理 3和 4中研究的图）的 AMC 平稳分布相对简单，这使得上述分析得以进行。

这里我们将介绍一些附加符号，这些符号将在本文余下部分中使用。我们将所有节点都采取物质奖励较高的行动的状态定义为一致状态，并用 C表示所有一致状态的集合。类似地，将并非所有智能体都采取相同行动的状态定义为多样状态——即至少一个节点采取动作 0，且至少一个节点采取行动1。我们用 D表示多样状态的集合。我们将转移到一致状态的条件概率定义为：

$$p^G = \sum_{s \in D} \mathbb{P}^G(C|s)\eta_s^G.$$

由于 C 是所有节点采取相同行动且该行动奖励较高时的所有可能状态的集合，因此我们有 $\mathbb{P}^G(C|s) = \sum_B \mathbb{P}^G((1,1,B)|s) + \mathbb{P}^G((0,0,B)|s)$。

4 适应与变化

在本节中，我们研究哪些网络结构更能适应变化环境，也就是能够带来较高的长期平均福利。在下一小节中，我们首先介绍另一项主要结果：对于任何不存在弱连接的网络，长期平均福利都非常低，实际上只比选择随机行动略高。接下来的结果

表明，岛屿网络——其中智能体在岛内（或组成部分内）强连通，而岛屿之间则通过弱连接相连——有可能实现更高的福利。最后，我们全面刻画了最有利于实现长期适应的网络结构；结果发现，这类结构呈星形，中间有一个大型强连通组成部分，弱连接的叶节点则向该星形组成部分提供信息。

4.1 没有弱连接时的低福利

下一个定理是我们的主要结果之一，它表明，没有弱连接时，社会的福利会很低，因为社会无法适应底层状态的变化。

定理 3（没有弱连接就无法快速学习）考虑一个没有弱连接的图 G 。假设 $\tau \leq 1/d_{\max}$。则：

$$\mathcal{S}^G \leq \frac{1}{2} + \frac{\epsilon}{2(\lambda+\epsilon)}.$$

此外，当 $\tau \geq 1/d_{\min}$时，我们有 $\mathcal{S}^G = \frac{1}{2}$。 –

定理3表明，在没有弱连接的网络中，长期平均福利较低，且其上界为 $1/2+\mathcal{O}(\epsilon)$。回想一下，我们关注的是满足以下条件的经济体：$\epsilon$ 非常小（因此颤动远比收益冲击罕见）。具体而言，当 $\epsilon \to 0$时，长期平均福利与这样一种环境并无不同：没有智能体掌握有关底层状态的信息，且所有参与者都随机选择行动。此外，如果阈值 τ 足够高，则没有节点会改变其行动，因此这类网络的平均福利恰好为 $1/2$。

虽然这一结果乍看之下似乎有悖常理，但实际上相当直观。设想一个社会网络，其中的智能体在某个时刻学到了底层状态，并据此协调一致地采取高回报行动。如果没有弱连接，也没有 ϵ-颤动，它们就会继续采取这一行动；但随着时间推移，底层状态会发生变化，长期来看，它仅以 $1/2$ 的概率与初始状态（因而也与初始行动）一致。在这种情形下，长期平均福利恰好为 $1/2$。没有弱连接但存在 ϵ-颤动的社会网络，表现会略好于这一假设情形，因为颤动会不时揭示底层状态，使所有收到这一信息的强连接智能体能够转而采取高回报行动。但当 ϵ 较小时，这种适应过程非常缓慢，对长期平均福利的影响也很小，定理 3对此作了形式化证明。

一个直接的推论是，正如我们在引言中所指出的，弱连接对于在变化环境中实现快速学习和适应至关重要。然而，下一小节将说明，用弱连接替代强连接并不足够。最后两个小节将进一步完整刻画：在这一设定下，通过弱连接相连的岛屿网络如何实现更高的福利，以及哪些类型的网络能够实现最高福利。



图3：两个简单网络

4.2 弱连接是否必然改善福利?

在本小节中，我们比较图中所示的两个简单网络3 以初步了解弱连接所发挥的作用。

图 3 所示的两个网络各有两个智能体。在第一个网络中，两个智能体通过一条强连接相连；在第二个网络中，两个智能体通过一条弱连接相连。我们在附录中证明，事实上，网络 2（有弱连接）的长期平均福利低于网络 1（有强连接）。原因在于，用弱连接替代强连接会减慢信息传递，却无法缓解定理3所描述的学习缓慢问题。相反，长期适应要求弱连接成为信息的额外通道，而非强连接的替代品。

4.3 岛屿网络中的适应

在本小节中，我们考虑通过弱连接相连的岛屿网络。尽管很难准确刻画这些网络的福利，但下一个命题给出了这些网络平均福利的上界。对于该命题，请回忆 p^G，其定义见第3.2节，对应于转移到一致状态的条件概率（也就是说，所有智能体都采取奖励较高的行动的状态）。

命题1 考虑一个岛屿网络 G，其中包含 k 个岛，每个岛有 m_c 个节点，且满足 $m_1 \geq m_2 \geq \cdots \geq m_k$。此外，这些岛通过弱连接相连。于是，平均福利的上界如下：

$$\mathcal{S}^G \leq \underbrace{\left(\frac{2p^G + \frac{\lambda}{\epsilon}p^G}{1 + 2p^G + 2\frac{\lambda}{\epsilon}p^G}\right)}_{(I)} + \underbrace{\left(\frac{1}{1 + 2p^G + 2\frac{\lambda}{\epsilon}p^G}\right)}_{(II)} \times \underbrace{\left(\frac{1}{2} + \frac{\sum_{i=1}^{\min\{\lceil \gamma/\lambda \rceil, k\}} m_i}{\sum_{i=1}^{k} m_i}\right)}_{(III)}$$

命题 1 给出了同时具有强连接和弱连接的岛屿网络平均福利的上界。尽管该上界并不紧，但它揭示了任何网络在变化环境中实现较高平均福利时面临的权衡。具体而言，命题 1 右侧对应于网络可能处于的两类状态对平均福利的贡献：网络可能处于一致状态，即所有节点都采取奖励较高的行动，其

奖励较高，而这由项 (I) 表示；或者网络可能处于多样状态，即有些节点采取一种行动，另一些节点采取另一种行动，此类状态的贡献由 (II)×(III) 表示。

从项 (I) 开始，我们可以看到，如果 p^G 很大，那么当 $\lambda >> \epsilon$ 时，一致状态下的福利将接近 $1/2$ （这正是我们关注的情形）。直观地说，这反映了一个问题：当向一致状态的转移发生得非常迅速时，针对底层环境变化的适应就会很少。因此，只有 p^G 的值相当小的网络才能实现较高的福利。

接下来考察其余各项： p^G 的值较小，便能确保项 (II) 也较大，但这还需要与较大的 (III) 相结合。这意味着 γ/λ 较大，或者对于 $d << k$， $\sum_{i=1}^{d} m_i \approx \sum_{i=1}^{k} m_i$。然而，$\gamma/\lambda$ 不可能很大，因为这意味着所有弱连接都能在一个时期内被激活，从而导致 p^G 非常大。因此，对于 $d << k$，我们必须有 $\sum_{i=1}^{d} m_i \approx \sum_{i=1}^{k} m_i$，这意味着网络中最大的连通分量必须包含大部分节点。因此，要实现较高的长期平均福利上界，岛屿网络必须满足这样的条件：其最大的连通分量包含大部分智能体。

总的来说，命题1 中的上界凸显了促成较高平均福利的一般性因素。具体而言，我们看到，要在变化环境中实现适应：

- 网络在大多数时候应当是不连通的，否则行动会产生过多的从众效应，导致底层环境变化时学习速度减慢。在岛屿网络中，这可以通过让岛屿集合高度断开来实现。这对应于要求 p^G 不能太大，这也包含了 γ/λ 不能太大的要求；

- 然而，不连通的组成部分之间仍应以合理的频率传递信息。在岛屿网络中，这可以通过设置以足够高频率激活的弱连接来实现。这对应于要求 γ/λ 不能太小；

- 即便处于不连通状态，我们仍应确保有相当一部分智能体采取正确行动。在岛屿网络中，这可以通过让每个岛都强连通，并让弱连接将相关信息传递到足够多的岛屿来实现。这对应于要求图中较大的岛屿包含大多数节点，即 $\sum_{i=1}^{\lceil \gamma/\lambda \rceil} m_i \approx \sum_{i=1}^{k} m_i$）。

4.4 适应性最强的网络

在上一小节中，我们看到，弱连接的岛屿网络能够实现远高于定理3中无任何弱连接的基准网络的长期平均福利。本小节将探讨其他网络能否表现得更好，并从对变化环境的适应这一角度刻画最优网络。我们将

看到，命题1所强调的相同原则也适用于回答这一问题。具体而言，我们将证明，一种兼顾一致状态下大多数智能体采取正确行动的需求，以及为信息传递保留一定多样性的必要性的网络结构，能够实现可行的最高收益。

为了预见哪些类型的网络会具有这些性质，我们将一个星形网络定义为具有 m 个连通分量和 n 个节点（$n > m$）的网络，其中一个连通分量包含 $n - m + 1$个强连通节点，其余 $m - 1$ 个连通分量的大小均为 1。此外，我们假设这些 $m - 1$ 个连通分量各自通过一条弱连接与大小为 $n - m + 1$的较大连通分量相连。星形网络示例见图4 。

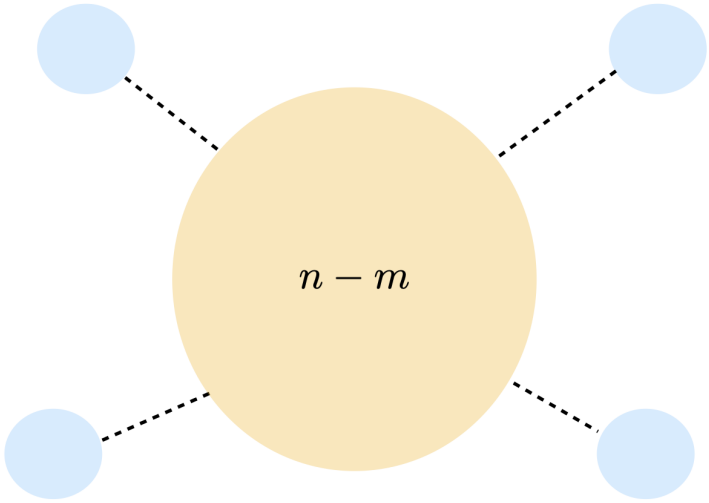

图4：一个星形网络，其中中心的 $n - m$ 个智能体彼此强连通，其余 m 个智能体通过弱连接与核心相连。

下一个定理证明，图4所示的星形网络，在所有网络中实现了最高的长期平均福利。

定理4 对于任意具有 G 个 n 节点的网络，存在一个具有相同节点数的星形网络（如图4所示），其长期平均福利高于 G，因为 $\epsilon \to 0$和 $\phi \to 0$ 或 $\phi \to \infty$。

此外，对于 $\phi << \lambda << \gamma$，当节点数 n 和弱连接数 m 趋于无穷大且 $m/n \to 0$ 时，星形网络的平均福利趋近于1 。

该定理得出了两个重要结果。首先，当扰动（由 ϵ 给出）较小，且弱连接从非活跃转为休眠的速度较快（ϕ较大）时，星形网络（如本文所定义）能够实现最高的长期平均福利。这些条件既有技术层面的意义，也有实质意义。就实质而言，该结果要求 ϵ–颤抖或错误不足以成为（外生）多样性的充分来源。就技术而言， $\epsilon \to 0$ 的极限使我们能够聚焦于这样一种情形：对变化环境的所有适应都来自智能体向那些

采取不同行动的人学习。假设 $\phi \to 0$ 或 $\phi \to \infty$ 成立，使我们能够聚焦于边界情形，从而得到更精确的刻画。

定理的第二部分表明，在我们施加的充分条件下，星形网络的平均福利趋近于 1，这是此情形下可实现的最高收益。这些条件要求节点数量相对于弱连接数量足够大（这凸显了与上一小节强调的相同作用机制；我们需要相当比例的智能体在“多样”状态下选择高回报行动）。出于技术原因，在这一结果中，我们还令节点数量和弱连接数量都趋于无穷大。最后，我们也考虑 ϕ 较小的情形（尽管不一定趋近于零）。这一条件仍能确保弱连接一旦被激活、随后又变为非活跃状态，就需要很长时间才能摆脱非激活状态。

现在我们来解释这一性质为何对我们的结果有用，以及弱连接的两阶段激活过程为何对我们的整体分析至关重要。首先请注意，如果 ϕ 不小，我们可能遇到这样的情形：初始网络中所有节点都采取错误行动，随后 $\epsilon-$ 颤动影响其中一个智能体，网络便转变为多样状态。由于弱连接激活发生在采取不同行动的智能体之间，它首先会选中受到 $\epsilon-$颤动影响的智能体，由该智能体传递相关信息。下一次发生激活时，再次选中受到 $\epsilon-$颤动影响的节点的概率很高；如此一来，网络可能很快又转变为一致状态。因此，引入带有退避期的两阶段激活过程，可以避免同一节点在短时间内连续被选中，并有助于维持一定程度的多样性。换言之，这一机制让弱连接在非激活状态下停留更长时间，从而使我们能够避免网络转变为多样状态后几乎立即又回到一致状态的情形。

该定理的证明依赖于定理2中给出的平均福利刻画。我们首先证明，对于任意具有 G 个 k 连通分量的岛屿网络，我们总能构造出一个网络，其中一个连通分量的规模为 $n-k+1$，其余 $k-1$ 个连通分量的规模均为 1。因此，这一结果意味着，只需考察具有这种特殊结构的网络。其次，我们证明，在所有具有这种结构的网络中，星形网络的平均福利最高。这一步的证明直观明了，并利用了这样一个事实：星形网络能够使规模最大的连通分量在中间采取相同行动，同时有足够多具有多样性的叶节点向这个中间连通分量传递信息。

5 前瞻性智能体的适应

到目前为止，我们关注的是最大化当前（即时）收益、完全不考虑未来收益的智能体。如果智能体足够有耐心，它们也可以通过实验来找出哪种行动是最优的。尽管这类实验问题

既重要又有趣，但超出了当前论文的讨论范围。[6] 尽管如此，探究前瞻性行为是否会消解我们所确定的主要经济力量仍有意义。下一个定理表明，答案是否定的；只要智能体赋予未来收益的权重不过高，我们的所有结果都可以推广。智能体的折现因子上限 β 取决于网络结构；正如下文所示，对于强连通图，该上限可以任意接近 1。

我们现在考虑的唯一差异是，与方程（1）中一样，每个智能体 $i \in V$ 不再选择行动以最大化当前收益，而是在每个时刻选择行动，以最大化其 β折现收益：

$$a_i(t_0) = \underset{a\in\{0,1\}}{\operatorname{argmax}}\ \mathbb{E}\left[\sum_{j=0}^{\infty} \beta^j \mathcal{U}_i^{a_t(t_j)}(t)\right]. \tag{3}$$

这里，时刻 t_i 按照速率为 1 的泊松时钟确定。

下一个定理表明，当折现因子 β 不太大时，该问题的解与我们迄今为止所关注的问题完全相同，因此智能体会选择使当前收益最大化的行动，并且只会在新信息到达时改变行动。

定理 5 如果折现因子 β 满足

$$\beta < \frac{\tau d_{\min}}{2 + \tau d_{\max}},$$

那么就不会进行实验，且方程（3）的解与

$$\underset{a\in\{0,1\}}{\operatorname{argmax}}\ \mathbb{E}_{i,t}[\ \mathcal{U}_i^a(t)].$$

因此，定理*1-4* 在智能体具有前瞻性时同样适用 *l.*

注意到，τ 会带来更高的 β上界。这很容易理解。当局部互动相对于个体试验的潜在收益而言，对收益的影响更大时，智能体更有动机遵循其（强连通的）邻居的行动。事实上，回到定理1的第二部分，我们可以看到，对于强连通网络或正则网络，我们可以将 τ 设得足够大，从而使 β 的上界足够接近 1。

然而，反过来也可以证明，存在一些网络与折现因子的组合，使得个体智能体愿意进行实验；但我们将把对这类情形的进一步探索留待未来研究。

[6]网络中的实验研究见尤其：Keller 等人(2005）、Bonatti 和 H¨orner(2011、2017）、Board 和 Meyer-ter Vehn(2022)

6 讨论与结论

多样性有助于生物和社会系统适应环境。在生物学中，种群多样性越高，就越有可能具备足够的遗传变异，从而形成有效策略，应对新物种入侵、食物短缺或新的气候条件。在社会系统中，多样性可以让人们更快地发现环境变化，并促进采取适当的应对措施。但在社交网络中，维持多样性更加困难。这是因为多样性本质上意味着探索或尝试不同的策略，而个体智能体往往有动机利用当前高奖励行动所带来的更高收益。然而，当社会网络能够快速传播信息时，多样性和充分的实验就极难维持，因为所有智能体都会了解到当前的最优行动，并倾向于转而采取这一行动。

本文将多样性与利用高收益行动之间的张力进行了形式化。在我们的模型中，一群贝叶斯代理人围绕底层状态形成信念，并选择行动，以最大化物质收益（来自正确匹配底层状态）与网络收益（与其强连接的其他智能体协调所带来的收益，例如家人、亲族、亲密朋友或同事）之和。我们的形式化分析刻画了这一博弈的贝叶斯-纳什均衡，并给出了明确公式，用于确定社会网络中的长期平均收益。

我们的主要结果之一表明，在仅由强连接构成的网络中，长期平均收益约等于所有人随机选择行动时的收益。原因在于，强连接使信息传递十分迅速，因此所有强连接组件都会收敛到相同的行动。一旦如此，得知底层状态已经改变的过程就会变得十分缓慢；从长远来看，最初的最优行动出错与正确的可能性相同，因此收益约等于随机选择时的收益。

我们的主要结果表明，将强连接与弱连接结合起来，可以获得高得多的收益。正如（Granovetter (1977)）所设想的那样，弱连接所涉及的互动比强连接更加不频繁、也不那么紧密；在我们的设定中，它们传递信息的速度也比强连接更慢（具有间歇性）。当弱连接与一种能够产生多个彼此不同的（不连通的）社群的强连接结构相结合时，就有足够的多样性空间。每个强连通分量仍将采取相同的行动，但不同的分量可以采取多样化的行动。当底层状态发生变化时，其中一个分量会迅速发现这一变化。随后，这一信息将通过弱连接传递给社会的其他部分。我们展示了这样一种网络如何实现这种平衡并因此获得远高于仅由强连接构成的网络的收益：该网络由彼此之间通过弱连接相连的强连通岛组成。能够实现最高长期平均福利的社会网络具有星形结构，即中央是一个大型强连接分量，通过弱连接从规模小得多的叶节点获取信息。

我们的主要结果适用于贝叶斯型主体，但他们只最大化当前收益，因此不存在个体实验激励。在论文的最后部分，我们证明，只要前瞻性贝叶斯型主体的贴现因子低于某个阈值，我们的结果同样适用；前提是他们不会受到个体实验的激励，即其贴现因子

低于该阈值。该阈值本身在很大程度上取决于局部互动。

本文提出了若干留待未来研究探讨的问题，最后我们简要讨论其中几个。

- 纳入更多前瞻性主体，允许其进行一定程度的实验，但不考虑多样化行动带来的全部社会效益，是一个重要的研究方向。这相当于将网络中的实验（例如脚注6所引述的研究）领域的新兴文献中的见解，与本文强调的内容结合起来；本文关注的是对变化环境的适应问题。

- 相关的是，我们的分析通过假设智能体具有相同偏好而得以简化。偏好异质性会带来更多多样性，将这一点与我们的整体框架结合起来，也是理论研究的另一个重要方向。

- 除了未来的理论研究之外，对社会系统中适应与多样性之间的联系进行实证评估也会很有意义。正如我们在引言中的讨论所指出的，生物学领域日益丰富的研究记录了多样性带来的适应益处。这些益处在社会环境中的大小和作用机制有何不同，是未来研究面临的一个重大问题。

- 最后，从更广泛的意义上说，我们的论文提供了一个例子，说明在社会环境中，信息交换和沟通的改善可能会带来不利后果——在这个例子中，这是因为它损害了多样性和适应能力。随着通信技术以惊人的速度不断进步，多样性、过度从众和适应方面是否会出现意想不到的后果，是一个重要且多层面的问题，值得认真研究。

参考文献

Acemoglu, D., M. A. Dahleh, I. Lobel, and A. Ozdaglar (2011). 社交网络中的贝叶斯学习。《经济研究评论》*78*(4), 1201–1236. Acemoglu, D., A. Makhdoumi, A. Malekian, and A. Ozdaglar (2022年). 从评论中学习：选择效应与学习速度。《计量经济学》*90*(6), 2857–2899. Agha, R., A. Gross, T. Rohrlack, and J. Wolinska (2018). 壶菌寄生虫对其蓝藻宿主的适应受到宿主种内多样性的阻碍。《微生物学前沿》*9*, 921. Aldous, D. and J. Fill (1995). 可逆马尔可夫链与图上的随机游走。Bala, V. and S. Goyal (1998). 从邻居中学习。《经济研究评论》*65*(3), 595–621. Bala, V. and S. Goyal (2001). 社会学习中的从众主义与多样性。《经济理论》*17*, 101–120. Banerjee, A. and D. Fudenberg (2004). 口碑学习。博弈与经济行为 *46*(1), 1–22. Barton, N. H. and B. Charlesworth (1998). 为什么要有性繁殖和重组？《科学》*281*(5385), 1986–1990. Board, S. and M. Meyer-ter Vehn (2022年). 网络中的实验。可见于 *SSRN 4271882* . Bonatti, A. and J. H¨orner (2011). 合作。《美国经济评论》*101*(2), 632–663. Bonatti, A. and J. H¨orner (2017). 在实验博弈中学习产生分歧。《经济理论杂志》*169*, 234–269. Brunnermeier, M. (2022年).《韧性社会：新冠疫情后的经济学》. 哈珀柯林斯。Burt, A. (2000). 观点：性、重组与选择效力——Weismann 的观点是否正确？《进化》*54*(2), 337–351. Burt, R. S. (1992). 结构洞. 哈佛大学出版社。 Callander, S. and J. H¨orner (2009). 少数人的智慧。《经济理论杂志》*144*(4), 1421–1439. Dasaratha, K., B. Golub, and N. Hak (2018). 从邻居那里了解变化状态。*arXiv* 预印本 *arXiv:1801.02042*.

DeMarzo, P. M., D. Vayanos 和 J. Zwiebel（2003）。说服偏差、社会影响与单维观点。《经济学季刊》*118*(3), 909–968。Ellison, G.（1993）。学习、局部交互与协调。《计量经济学》：计量经济学会期刊, 1047–1071。Fischer, S. H.（2004）。蜜蜂真酷。《科学新闻》, 1–2。Fisher, R. A.（1958）。自然选择的遗传理论。Frongillo, R. M., G. Schoenebeck 和 O. Tamuz（2011）。变化世界中的社会学习。载于互联网与网络经济学：第七届国际研讨会，*WINE 2011*，新加坡，*2011*年*12*月*11—14*日。会议论文集 *7*，第146–157页。施普林格。Gale, D. 和 S. Kariv（2003）。社交网络中的贝叶斯学习。《博弈与经济行为》*45*(2), 329–346。Goddard, M. R., H. C. J. Godfray 和 A. Burt（2005）。性别提高了自然选择在实验酵母种群中的效力。《自然》*434*(7033), 636–640。Golub, B. 和 M. O. Jackson（2010）。社交网络中的朴素学习与群体智慧。《美国经济学杂志：微观经济学》*2*(1), 112–149。Granovetter, M.（1983）。弱关系的力量：重新审视网络理论。《社会学理论》*1*, 201–233。Granovetter, M.（2017）。社会与经济。哈佛大学出版社。Granovetter, M. S.（1977）。弱关系的力量。载于社交网络，第347–367页。爱思唯尔。Gross, M. R.（1996）。替代性繁殖策略与战术：性别内的多样性。《生态学与进化趋势》*11*(2), 92–98。Holley, R. A. 和 T. M. Liggett（1975）。弱相互作用无限系统及投票者模型的遍历定理。《概率年刊》*3*(4), 643–663。Jacobs, J.（2016）。城市经济。Vintage。Kandori, M., G. J. Mailath 和 R. Rob（1993）。博弈中的学习、突变与长期均衡。《计量经济学》：计量经济学会期刊, 29–56。Kanoria, Y., A. Montanari 等（2011）。树上的多数派动力学与动态腔方法。《应用概率年刊》*21*(5), 1694–1748。Keck, M. 和 P. Sakdapolrak（2013）。什么是社会韧性？经验教训与未来方向。《地理学》, 5–19。

Keller, G., S. Rady 和 M. Cripps（2005）。指数型老虎机的策略性实验。《计量经济学》*73*(1), 39–68。La Ferrara, E. 和 A. Mele（2006）。种族隔离与公立学校支出。L´evy, R., M. Pkeski 和 N. Vieille（2022）。变化环境中的稳态社会学习。*arXiv* 预印本 *arXiv:2201.02122*。Montalvo, J. G. 和 M. Reynal-Querol（2021）。族群多样性与增长：重新审视相关证据。《经济学与统计学评论》*103*(3), 521–532。Moscarini, G., M. Ottaviani 和 L. Smith（1998）。变化世界中的社会学习。《经济理论》*11*, 657–665。Putnam, R. D. 等（2000）。《独自打保龄：美国社群的衰落与复兴》。西蒙与舒斯特。Rajkumar, K., G. Saint-Jacques, I. Bojinov, E. Brynjolfsson 和 S. Aral（2022）。对弱关系力量的因果检验。《科学》*377*(6612), 1304–1310。Santos, F. C., M. D. Santos 和 J. M. Pacheco（2008）。社会多样性促进公共物品博弈中合作的出现。《自然》*454*(7201), 213–216。Saxenian, A.（1996）。《区域优势：硅谷与*128*号公路的文化与竞争，作者新序》。哈佛大学出版社。Smith, L. 和 P. Sørensen（2008）。随机抽样下的理性社会学习。技术报告，工作论文。Weismann, A., J. A. Thomson 和 M. R. Thomson（1904）。《进化论》，第2卷。E. Arnold。Wolff, R. W.（1989）。《随机建模与排队论》。培生学院分部。Yildiz, M. E., R. Pagliari, A. Ozdaglar 和 A. Scaglione（2010）。随机网络中的投票模型。载于*2010*年信息论与应用研讨会（*ITA*），第1–7页。IEEE。Young, H. P.（1993）。惯例的演化。《计量经济学》：计量经济学会期刊, 57–84。

附录

我们在附录中证明正文所述的结果。首先给出以下引理，后续的若干证明将用到它：

引理 A1 考虑一个包含 n 个智能体的网络。假设所有智能体最初都选择行动0，并且 $r_1 > r_0$，因此行动1 具有更高的物质收益。那么，在收益冲击发生之前，图中至少有一个智能体了解到行动1的概率为：

$$q = \frac{\epsilon}{\lambda + \epsilon}.$$

证明。 证明非常直接。速率为 ϵ 的泊松时钟跳动时会发生随机冲击，而速率为 λ 的独立泊松时钟跳动时会发生收益冲击。

令 X 表示下一次随机切换前的随机时间，令 Y 表示下一次收益冲击前的随机时间。因此，若 $X < Y$，则至少有一个节点会在收益冲击发生之前了解到行动 1。然而，由于这些是相互独立的泊松时钟，我们有 $X \sim \exp(\epsilon)$， $Y \sim \exp(\lambda)$，且 X 与 Y 相互独立。因此，

$$q = \mathbb{P}(X < Y) = \frac{\epsilon}{\lambda + \epsilon}.$$

从而完成证明。 ■

引理 1

这里，我们证明引理的第一部分。注意，第二部分完全遵循我们刚才的论证，只需将信念重新定义为与动作 0 具有更高物质收益这一事件相关的概率。首先，注意在时间 $T_i(t)$，我们有 $\mu_i(T_i(t)) = 1$。这意味着在时间 $T_i(t)$，节点 i 以概率 1 知道动作 1 的物质收益高于动作 0。

现在，智能体 i 的信念更新可以分解为两部分：

- 由于了解环境动态而产生的更新（$\mu_i^e(t)$）：这体现了收益冲击导致的信念更新。

- 由于与邻居互动而产生的更新（$\mu_i^n(t)$）：如果某个相邻节点已经“了解到”另一个行动，这体现了节点i的信念更新。

由于获知环境而产生的信念更新过程如下：

$$\mu_i^e(t) = \left(1 - e^{-\lambda(t - T_i(t))}\right) \frac{1}{2} + e^{-\lambda(t - T_i(t))} \mu_i(T_i(t)).$$

这一点很容易说明如下：以 $e^{-\lambda(t - T_{i(t)})}$ 的概率，在时刻 $T_i(t)$ 与 t 之间没有发生收益冲击，且信念保持为 $\mu_i(T_i(t))$. 剩余概率下，

发生了一次或多次收益冲击，在这种情况下，两个行动成为具有更高奖励的行动的可能性相等。在此情况下，信念为 $1/2$。需要注意的关键性质是：$\mu_i^e(t) \geq 1/2$。

刻画精确信念更新 $\mu_i(t)$ 的主要难点，在于写出信念 $\mu_i^n(t)$ 的显式形式。请注意，由于智能体 i 不断与她的邻居互动，从时间 $T_i(t)$ 到 t 一直没有新信息这一事实意味着，可能有某个节点探索了另一种行动（通过一次 $\epsilon-$ 翻转或弱连接的激活），发现它并没有更好，于是继续采取当前奖励更高的行动。换句话说，节点 i 从时间 $T_i(t)$ 到 t 没有收到关于另一种行动的新信息这一事实，应当进一步强化这样一种判断：在时间 t，行动1也是奖励更高的行动。这是证明信念单调性所需的唯一性质。更正式地说，我们可以将时间 t 的信念写为：

$$\mu_i(t) = \int_{T_i(t)}^{t} [\mu_i^n(\tau)\mu_i(T_i(t)) + (1 - \mu_i^n(\tau))\mu_i^e(\tau)]\, d\tau.$$

证明的最后一步很容易得出，只需注意被积函数中的项始终大于 $1/2$（因为对于所有 $\tau \in (T_i(t)$ $t)$，$0 \leq \mu_i^n(\tau) \leq 1$ 且 $\mu_i^e(\tau) \geq 1/2$）。

引理2的证明

我们用反证法证明此引理。为简便起见，假设 $a_i(t^-) = 1$，即节点 i 在时间 t^-采取行动1。这意味着

$$\mathbb{E}_{i,t^-}[R_1(t-)] + \tau f_i(1, t^-) \geq \mathbb{E}_{i,t^-}[R_0(t-)] + \tau f_i(0, t^-). \tag{A1}$$

现在，由于 t 不是信息到达的时刻，信念 $\mu_i(t)$ 在时间 t连续。特别地，这意味着我们可以在方程（A1）中令 $t^- \to t$ 取极限，且不等式的符号保持不变。这意味着 $\mathbb{E}_{i,\ t}[R_1(t)] + \tau f_i($ $1,\ t) \geq \mathbb{E}_{i,\ t}[R_0(t)] + \tau f_i(0,\ t)$，这表明智能体 i 将在时间 t继续采取行动1。

因此，如果节点 i 在时间 t改变其行动，这必定意味着 t 是节点 i获得新信息的时刻。证明完毕。

定理1的证明

回想一下，智能体 i在时间 t选择行动 a，以最大化 $\mathbb{E}_{i,\ t}[R_a] + \tau f_i(a,\ t)$。假设智能体 i 采取动作 0，而她的一位邻居 j 采取行动1。在这种情况下，两位智能体都确切地知道哪种行动的物质收益更高。假设行动1的收益更高。这意味着智能体 i 采取的是物质收益较低的行动。

然而，如果智能体 i切换到行动1，她的效用将为 $1+\tau f_i(1,t) > 0+\tau f_i(0,t)$，因为$\tau \leq 1/d_{\min}$。因此，智能体 i 会选择行动1，而非0。这表明，智能体 i 及其所有邻居必须选择相同的行动。将同样的论证推广到所有与智能体

i相连的节点，我们就得到该定理的第一部分。

第二部分源于所有节点最初都选择相同行动这一事实。假设所有节点最初都选择动作 0。考虑节点 i。它的所有邻居都在选择动作 0。现在，假设节点 i 在时间 t 发生了一次 ϵ–翻转，并得知事实上行动1的物质收益更高。那么，我们有 $\mathbb{E}_{i,\ t}[R_1 - R_0] = 1$。然而，由于 $\tau > 1/d_{\min}$，我们有$\tau f_i(0,\ t) > 1$（因为所有邻居也都在选择动作 0）。因此，尽管节点 i知道行动1的物质收益更高，她仍继续选择动作 0，因为她所有的邻居都在选择动作 0。证明完毕。

定理 2 的证明

我们首先证明该马尔可夫链既不可约又非周期。

引理 A2 定义 *3* 中的激活马尔可夫链对于任意图 G 均不可约且非周期

证明。考虑时间 k 的状态，记为 (P_k, R_k, B_k)。我们有（注意，为方便起见，我们省略了 $\mathbb{P}(\cdot)$ 上的上标 G ）：

$$\begin{aligned}\mathbb{P}((P_{k+1}, R_{k+1}, B_{k+1}) &= (P_k, R_k, B_k)|(P_k, R_k, B_k)) \\ &= \frac{1}{2} \times \mathbb{P}(\text{No weak link activation or } \epsilon\text{−flip before shock}) \\ &> 0.\end{aligned}$$

因此，对于马尔可夫链中的任意状态，都有正概率停留在同一状态。这说明该马尔可夫链是非周期的。

令 $P_z = \{z\}^n$ 表示所有智能体都采取行动 z 时，$z \in \{0, 1\}$ 对应的行动向量。现在，考虑两个状态 (P_1, R_1, B_1) 和 (P_2, R_2, B_2)。我们证明这两个状态之间存在一条具有正概率的路径。很容易看出如下（此处 B 是弱连接的任意子集）：

$$\begin{aligned}\mathbb{P}((\bar{P}_{R_1}, R_1, B_1)|(P_1, R_1, B_1)) &> 0 \\ \mathbb{P}((\bar{P}_{1-R_1}, 1 - R_1, B_2)|(\bar{P}_{R_1}, R_1, B_1)) &> 0 \\ \mathbb{P}((\bar{P}_{1-R_2}, 1 - R_2, B_2)|(\bar{P}_{1-R_1}, 1 - R_1, B_2)) &> 0 \\ \mathbb{P}((P_2, R_2, B_2)|(\bar{P}_{1-R_2}, 1 - R_2, B_2)) &> 0\end{aligned}$$

第一个不等式可证明如下：

$$\begin{aligned}&\mathbb{P}((\bar{P}_{1-R_2}, 1 - R_2, B_1)|(P_1, R_1, B_1)) \\ &\geq \frac{1}{2} \times \mathbb{P}(\text{Every agent learns the right action by random flipping before shock}) \\ &> 0.\end{aligned}$$

同样的论证可用于证明下一个不等式。因此，在这条马尔可夫链中，从任意状态转移到任何其他状态的概率都为正。这表明定义3中的马尔可夫链是不可约的，从而完成了证明。

■

现在，根据引理 A2，我们知道该马尔可夫链既是非周期的，也是不可约的。因此，由于它的状态数有限，它存在唯一的平稳分布（例如参见 Aldous 和 Fill (1995)）。因此，平稳分布 η_i^G 定义良好。现在，根据 Aldous 和 Fill (1995)，我们知道，马尔可夫链的极限行为可以用其遍历行为来刻画，因此我们有

$$\lim_{k\to\infty} \mathcal{S}_k^G = \mathbb{E}_{\eta^G}[f],$$

由此完成定理的证明。

定理 3 的证明

在一般图的情形下，可能的均衡只有两种：所有智能体都采取奖励较高的行动，或者所有智能体都采取奖励较低的行动（根据定理 1，因为 $\tau \leq d_{\max}$）。我们给出采取正确行动的智能体比例的一个上界。我们用以下两状态马尔可夫链来近似第 3.2 节中描述的马尔可夫链。

- G - 所有智能体都采取奖励较高的行动。
- B - 所有智能体都采取奖励较低的行动。

该两状态马尔可夫链的转移概率如下：

$$\mathbb{P}(\mathtt{G}|\mathtt{B}) = \frac{1}{2}(1+q), \qquad \mathbb{P}(\mathtt{B}|\mathtt{G}) = \frac{1}{2}(1-q).$$

这里 q 表示某个智能体通过随机翻转获知更优行动的概率，推导过程见引理A1。

利用这些转移概率，可得（这里 η_i 表示状态i下的平稳分布）：

$$\eta_{\mathtt{B}} = \frac{1}{2}(1-q) \qquad \eta_{\mathtt{G}} = \frac{1}{2}(1+q).$$

根据定理 2，可得仅包含强连接的任意连通图的平均福利为：

$$\mathcal{S}^G = \frac{1}{2}(1+q).$$

现在，将引理 q 中的数值代入A1，即可得到最终结果 t.

图 3 分析

为便于阐述，我们考察 $\phi \to \infty$时的极限行为，此时弱连接始终处于激活或休眠状态。根据定理3，我们知道网络 1 的平均福利

网络 1 的平均福利为:

$$\mathcal{S}^{G_1} = \frac{1}{2}\left(1 + \frac{\epsilon}{\epsilon + \lambda}\right).$$

接下来，我们计算网络 2 的平均福利。共有三种可能的状态：(i)G （两个节点都采取高回报动作）(ii) B （两个节点都采取低回报动作）以及 (iii) M （恰好一个节点采取高回报动作）。考虑转移到状态 B的转移概率:

$$\mathbb{P}(\mathtt{B}|\mathtt{G}) = \frac{1}{2} \times (1-q), \qquad \mathbb{P}(\mathtt{B}|\mathtt{B}) = \frac{1}{2} \times (1-q), \qquad \mathbb{P}(\mathtt{B}|\mathtt{M}) = 0.$$

这表明， G 和 B 的稳态概率之间的关系为:

$$\eta_B^{G_2} = \eta_G^{G_2} \times \frac{1-q}{1+q}.$$

现在，第二个网络的平均福利为:

$$\begin{aligned}\mathcal{S}^{G_2} &= \eta_G^{G_2} + \frac{1}{2}\eta_M^{G_2} \\ &= \eta_G^{G_2} + \frac{1}{2}(1 - \eta_G^{G_2} - \eta_B^{G_2}).\end{aligned}$$

由此可见， G_2 的福利低于 G_1。假设对于某个 $t > 0$， $\eta_G^{G_2} = \eta_G^{G_1} - t$ ，我们有:

$$\mathcal{S}^{G_1} - \mathcal{S}^{G_2} = t - \frac{t}{2} \times \left(1 + \frac{1-q}{1+q}\right) > 0.$$

证明至此完成。

命题的证明1

该证明通过精心设计的马尔可夫链代入定理 2 得出。考虑具有以下状态的马尔可夫链:

- G - 所有智能体都选择奖励较高的行动。
- B - 所有智能体都选择奖励较低的行动。
- M - 至少有一个节点选择动作 0，且至少有一个节点选择行动1。

这些状态之间的转移概率由下式给出（我们忽略涉及更高次幂 ϵ的项。此外，我们使用引理A1 代入 q的值）：

$$\mathbb{P}(\mathtt{M}|\mathtt{G}) = \frac{1}{2} \times \frac{\epsilon}{\epsilon+\lambda}, \qquad \mathbb{P}(\mathtt{B}|\mathtt{G}) = \frac{1}{2} \times \frac{\lambda}{\epsilon+\lambda}$$
$$\mathbb{P}(\mathtt{M}|\mathtt{B}) = \frac{1}{2} \times \frac{\epsilon}{\epsilon+\lambda}, \qquad \mathbb{P}(\mathtt{B}|\mathtt{B}) = \frac{1}{2} \times \frac{\lambda}{\epsilon+\lambda}$$
$$\mathbb{P}(\mathtt{M}|\mathtt{M}) = 1 - p^G, \qquad \mathbb{P}(\mathtt{B}|\mathtt{M}) = 0.$$

利用这些转移概率，我们得到平稳分布

$$\eta_M^G = \frac{1}{1 + \frac{2\lambda p^G}{\epsilon} + 2p^G}, \qquad \eta_G^G = \frac{\frac{\lambda p^G}{\epsilon} + 2p^G}{1 + \frac{2\lambda p^G}{\epsilon} + 2p^G}.$$

现在，为了推导平均福利的上界，我们只需找出多样状态下平均福利的上界（记作 $\mathcal{S}_M^G$）。

首先请注意，如果没有弱连接激活，平均福利就只是 $\approx 1/2$（因为我们对收益冲击取了平均值）。然而，平均而言，每个时期会发生 γ/λ 次弱连接激活。现在，为了推导上界，我们假设这些弱连接激活会通知岛屿网络中最大的连通分量。因此，我们可以给出多样状态下平均福利的上界：

$$\mathcal{S}_M^G \leq \frac{1}{2} + \frac{\sum_{i=1}^{\min\{\lceil \gamma/\lambda \rceil, k\}} m_i}{\sum_{i=1}^{k} m_i}.$$

因此，利用图的平均福利可以写成 $\mathcal{S}^G = \eta_G + \eta_M \times \mathcal{S}_M^G$ 这一事实，我们便完成了证明。

定理 4 的证明

我们首先 证明定理中关于星形网络是最优的部分。

注意，$\phi \to 0$ 的情形很容易得到。这种情形对应于这样一种情况：一旦弱连接传输信息，该连接就会变为非活跃状态，且再也无法使用。由于中间组成部分拥有最多节点的星形网络，对应于最多节点能够获取通过弱连接流动的信息的情形，因此它从弱连接中学习的概率最高，进而具有最高的福利。接下来我们关注 $\epsilon \to 0$且 $\phi \to \infty$的情形。具体而言，这意味着所有弱连接始终处于休眠或活跃状态；此外，由于 ϵ 微不足道，我们只需考虑其从一致状态转移时的影响（因为 ϵ 冲击是脱离这些状态的唯一方式）。

我们首先定义一些量，以便阐述我们的结果。请注意，这些结果基于定理2，为此我们需要刻画马尔可夫链定义3。首先，我们定义：

$$\mathcal{H}_{n,\ m} = \{\text{包含 } n \text{ 个节点和 } m \text{ 个连通分量的岛屿网络}\}.$$

每个图$\mathcal{H}_{n,\ m}$都由 $m+1$个强连通智能体岛组成，这些岛通过弱连接相连。定义

$$d_k^G = \sum_{\text{states } i \text{ where exactly k components play the right action}} \eta_i^G.$$

换言之， d_k^G 表示恰有 k 个连通分量采取物质收益较高的行动时的稳态概率。现在，对于类$\mathcal{H}_{n,\ m}$中的所有图，我们有关于 d_k^G 的以下性质。

引理 A3 对于所有图$G \in \mathcal{H}_{n,\ m}$，我们有：

$$d_k^G = d_k,$$

即，在类$\mathcal{H}_{n,\ m}$中的所有图里，采取正确行动的连通分量数量的分布相同。

证明。 考虑定义 3 中一般马尔可夫链的以下表示，其中网络 G:

- 状态： s_k，其中 $0 \leq k \leq m+1$。

这里，状态 s_k 表示所有恰有 k 个区块采取正确行动的行动组合。注意，对于图 G， d_k^G 是该马尔可夫链中状态 s_k 的平稳概率。这是一个嵌入式马尔可夫链，系统状态在每次发生收益冲击时被观测。

现在，对于图 G ，令 $\mathbb{P}_{s_k}^G(\cdot)$ 表示图 G 中从状态 s_k 转移到任意状态 s_j 的转移概率。我们将证明，对于所有图 $G \in \mathcal{H}_{n,\ m}$，该转移概率都相同。原因如下。

- 从状态 s_0 或 s_{m+1} 转移：这种情况最初由一次随机翻转引起，其概率为 ϵ。这对所有图 G 都是相同的。
- 从其他任意状态转移：这是由弱连接的激活引起的。由于每次弱连接激活都会增加一个采取正确行动的区块，因此决定转变的唯一因素是弱连接激活的次数（因为我们考虑的是极限 $\epsilon \to 0$ 和 $\phi \to \infty$）。由于 $\mathcal{H}_{n,\ m}$ 中的所有图都具有相同数量的连通分量，而且每条弱连接都会使采取正确行动的区块数量恰好增加一个，因此，对于所有图 $G \in \mathcal{H}_{n,\ m}$，该转移概率也相同。

因此，由于触发状态间转移的事件对所有图都相同G 且具有相同的概率，转变 $\mathbb{P}_{s_k}^G(\cdot)$ 对所有图也都相同 G。最后，

由于所有图的转移概率都相同 G，最终的平稳分布也将相同，从而完成引理的证明。

■

现在请注意，在 k 个区块采取正确行动的约束下，存在多种可能的定向。然而，对于 $k = 0$ 和 $k = m+1$，定向恰好只有一种：所有智能体都采取劣行动，或所有智能体都采取优行动，分别对应这两种情况。令η_G^G 和η_B^G 分别表示图 G处于全优状态和全劣状态的稳态概率。我们有以下推论。

推论 A1 对于所有图$G \in \mathcal{H}_{n,\ m}$，我们有：

$$\eta_G^G = \eta_G, \qquad \eta_B^G = \eta_B.$$

接下来，我们定义：

$$\mathcal{I}_{n,\ m} = \{\text{具有 } n \text{ 个节点和 } m \text{ 个连通分量的岛屿网络，其中一个连通分量有} n-m+1 \text{ 个节点，其余连通分量均只有1个节点}\}.$$

首先，我们证明，对于$\mathcal{H}_{n,\ m}$中的任意岛屿网络，总能找到另一个属于$\mathcal{I}_{n,\ m}$的网络，其平均福利更高。以下引理将证明这一点：

引理 A4 对于任意图$G \in \mathcal{H}_{n,\ m}$，存在一个图$G' \in \mathcal{I}_{n,\ m}$ 使得 $S^G \leq S^{G'}$。

证明。注意，一个包含 m 个连通分量、每个分量包含 k_i 个节点（i =1，2，$\cdots m$）的图，其平均福利可写为：

$$\begin{aligned}
\mathcal{S}^G &= \sum_s \eta_s^G f(s) \\
&= \sum_s \eta_s^G \left[\sum_{i=1}^{k} \mathbb{1}(\text{component } i \text{ plays the action with higher reward in state } s) \right] \\
&= \sum_{i=1}^{k} \zeta_i^G m_i.
\end{aligned}$$

这里，ζ_i 表示连通分量 i 采取物质收益较高的行动的稳态概率。令 k^* =argmax ζ_i^G，即 k^* 是 G 中在稳态下采取较优行动的概率最高的连通分量。

现在，考虑另一个具有相同弱连接结构的网络，但所有节点都属于连通分量 k^*，而其他所有连通分量都只有一个节点。将这个网络记为 G'。注意，$G' \in \mathcal{I}_{n,\ m}$。此外，注意到 G' 将具有与图 d_k 相同的分布G，特别是，$\sum \zeta_i^G = \sum \zeta_i^{G'}$（引理 A3）。另外，由于我们向连通分量 k^* 中添加了更多节点，因此有 $\zeta_{k^*}^{G'} \geq \zeta_{k^*}^{G}$。这清楚地表明

$$
\begin{aligned}
\mathcal{S}^G &= \sum_{i=1}^{k} \zeta_i^G m_i \\
&= \zeta_{k^*}^G m_{k^*} + \sum_{i \neq k^*} \zeta_i^G m_i \\
&\leq \zeta_{k^*}^G (n-m+1) + \sum_{i \neq k^*} \zeta_i^G \\
&\leq \zeta_{k^*}^{G'} (n-m+1) + \sum_{i \neq k^*} \zeta_i^{G'} \\
&= \mathcal{S}^{G'}.
\end{aligned}
$$

由此完成证明。 ■

因此，引理 A4 告诉我们，只需将注意力限制在$\mathcal{I}_{n,\ m}$中的图。我们将包含 $n-m+1$ 个节点的组成部分称为图 G 的核心。

我们将注意力转向马尔可夫链的另一种表示。考虑图中所示的链

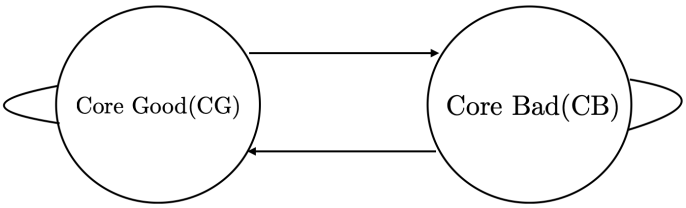

图 5：两个待比较的网络

这两个状态如下：

- 核心良好（CG）——这些表示核心连通分量中的 $n-m+1$ 个智能体采取正确行动的状态。

- 核心不良（CB）——这些表示核心连通分量中的 $n-m+1$ 个智能体采取错误行动的状态。

令 $\eta_{\mathtt{CG}}^G$ 和 $\eta_{\mathtt{CB}}^G$ 表示该链的平稳分布。注意，这些平稳分布必须满足：

$$
\begin{aligned}
\eta_{\mathtt{CG}}^G &= \sum_{\text{states } i \text{ where core is good}} \eta_i^G \\
\eta_{\mathtt{CB}}^G &= \sum_{\text{states } i \text{ where core is bad}} \eta_i^G.
\end{aligned}
$$

我们有 f 以下重要引理刻画了星形网络的行为：

引理 A5 我们有：

$$\eta_{\texttt{CG}}^{G_{star}} \geq \eta_{\texttt{CG}}^{G}, \qquad \forall\, G \in \mathcal{I}_{n,m}.$$

证明。 首先，根据推论 A1，我们知道，对于所有图 $G \in \mathcal{H}_{n,\ m}$，所有区块都采取正确行动（或所有区块都采取错误行动）的状态具有相同的概率。

接下来注意，对于核心处于不良状态的任意多样状态，即 CB 中的任意多样状态，转移到 CG 中某个状态的概率对于 G_{star} 都相同，即

$$\mathbb{P}^{star}(\texttt{CG}|s) = p_1 > 0, \qquad \forall\ s \in \texttt{CB}.$$

这是因为，在具有多样状态的星形网络中，任何一次弱连接激活都会让核心获知正确行动，随后新状态就会处于 CG。其余的弱连接激活无关紧要。其他导致状态进入 CG的事件，例如收益冲击或ϵ翻转，则保持不变，不受 CB中状态 s 的影响。

接下来，对于任意其他图 $G \in \mathcal{H}_{n,\ m}$，我们有

$$\mathbb{P}^{G}(\texttt{CG}|s) \leq p_1, \qquad \forall\ s \in \texttt{CB}.$$

注意，这是因为从多样状态出发，至少需要一次弱连接激活才能让核心获知正确行动。也可能需要多次弱连接激活（取决于 G 的结构以及状态 s）。此外，其他导致状态进入 CG 的事件，例如收益冲击或 ϵ 翻转，则保持不变，不受 CB 中的状态 s 或图 G 的影响。

这两个观察结果导出了图中马尔可夫链转移概率的以下不等式5。

$$\mathbb{P}^{star}(\texttt{CG}|\texttt{CB}) \geq \mathbb{P}^{G}(\texttt{CG}|\texttt{CB}).$$

完全采用相同的论证，我们得到

$$\mathbb{P}^{star}(\texttt{CG}|\texttt{CG}) \geq \mathbb{P}^{G}(\texttt{CG}|\texttt{CG}).$$

Th这些关于转移概率的不等式给出了我们所需的结果。 ■

现在，我们将这些结果汇总起来，得到最终定理。由引理 2，可得

图 G 的平均福利为:

$$\begin{aligned}\sum_i \eta_i^G f(i) &= \frac{1}{n}\sum_{k=1}^{m}\sum_{i_k}\eta_{i_k}^G\left[(n-m+1)\mathbb{1}(\text{core good})+k-\mathbb{1}(\text{core good})\right]\\ &= \frac{1}{n}\sum_{k=1}^{m}\left[(n-m-1)\sum_{i_k}\eta_{i_k}^G\mathbb{1}(\text{core good})+k\sum_{i_k}\eta_{i_k}^G\right]\\ &= \frac{1}{n}\sum_{k=1}^{m}\left[(n-m)\sum_{i_k}\eta_{i_k}^G\mathbb{1}(\text{core good})+kd_k\right]\\ &= \frac{1}{n}\left[(n-m)\sum_{k=1}^{m}\sum_{i_k}\eta_{i_k}^G\mathbb{1}(\text{core good})+\sum_{k=1}^{m}kd_k\right]\\ &= \frac{1}{n}\left[(n-m)\eta_{\mathtt{CG}}^G+\sum_{k=1}^{m}kd_k\right].\end{aligned}$$

当 $\eta_{\mathtt{CG}}^G$ 达到最大值时，该值也达到最大。现在，根据引理 A5，我们知道星形网络使这一概率最大，因此定理第一部分的证明得证。

定理的第二部分通过尽可能提高核心连通分量采取物质收益较高行动的概率来推导。为此，注意，若有 $\gamma = \mathcal{O}(m^{1/2}\lambda)$，并假设 $\phi = \mathcal{O}(m^{-1/4}\lambda)$，则在一个时期内弱连接被激活的概率为 $\to 1$ as $m \to \infty$。这意味着随着 m 增长至 ∞，在每个时期，核心连通分量采取较优行动的概率都趋近于 1。极限情况下的平均福利可以用核心连通分量的平均福利作为下界，该平均福利为 $(n-m+1)/m$，并趋近于 1，因为 $m/n \to 0$。定理第二部分的证明得证。

定理证明 5

当智能体 i 的泊松时钟（速率为 1）响起时，我们称该智能体变为“活跃”。当智能体 i 在时间 t变为活跃时，设其信念为 $\mu_i(t) = 1/2$。这意味着它的所有邻居都在采取相同的行动。为方便起见，假设该行动为动作 0。在这种情况下，每个时间步的奖励将为 $R_0 + \tau f_i(0) > R_0 + \tau d_{\min}$，其中 d_{min} 是图的最小度数。

如果智能体 i 决定探索并采取行动1，奖励将为 R_1。不过，探索之后，智能体 i 必定会知道哪个行动更好。

由于奖励之间的差值始终为 1，我们假设较高奖励的行动获得奖励 1，较低奖励为 0。

据此，探索后的总期望奖励可由以下上界限定：

$$\begin{aligned}
&\mathbb{E}[R_1] + \mathbb{E}\left[\sum_{j=1}^{\infty} \beta^j \left(R_{a(i,t_j)} - R_{1-a(i,t_j)} + \tau(f_i(a(i,t_j)) - f_i(1-a(i,t_j)))\right)\right] \\
&\leq \frac{1}{2} + \sum_{j=1}^{\infty} \beta^j \mathbb{E}\left[R_{a(i,t_j)} - R_{1-a(i,t_j)} + \tau(f_i(a(i,t_j)) - f_i(1-a(i,t_j)))\right] \\
&\leq \frac{1}{2} + \sum_{j=1}^{\infty} \beta^j \left(1 + \tau d_{\max}\right) \\
&= \frac{1}{2} + \frac{\beta}{1-\beta}(1 + \tau d_{\max}).
\end{aligned}$$

另一方面，如果智能体不进行探索，则期望总和的下界为：

$$\begin{aligned}
&\mathbb{E}[R_1] + \mathbb{E}[\sum_{j=1}^{\infty} \beta^j \left(R_{a(i,t_j)} - R_{1-a(i,t_j)} + \tau(f_i(a(i,t_j)) - f_i(1-a(i,t_j)))\right)] \\
&\geq \frac{1}{2} + \tau d_{\min}.
\end{aligned}$$

现在，如果

$$\frac{1}{2} + \tau d_{\min} > \frac{1}{2} + \frac{\beta}{1-\beta}(1 + \tau d_{\max}),$$

那么智能体就没有偏离的动机。

化简这个不等式，得到：

$$\tau d_{\min} > \beta(2 + \tau d_{\max}),$$

由此可得条件：如果：

$$\beta < \frac{\tau d_{\min}}{2 + \tau d_{\max}},$$

那么智能体就没有偏离的动机。