

集群系统建模综述

集群系统模型一般包含两部分：拓扑模型和节点动力学模型。拓扑模型决定了智能体间的连接关系，通常用图表示，如：无向图、有向图、加权图等。节点动力学模型用于描述每个智能体自身的运动规则，通常采用一阶、二阶积分器模型、一般系统模型等方法。图 1 为哈尔滨工业大学孙延超博士论文中多智能体分布式协同控制问题分类。

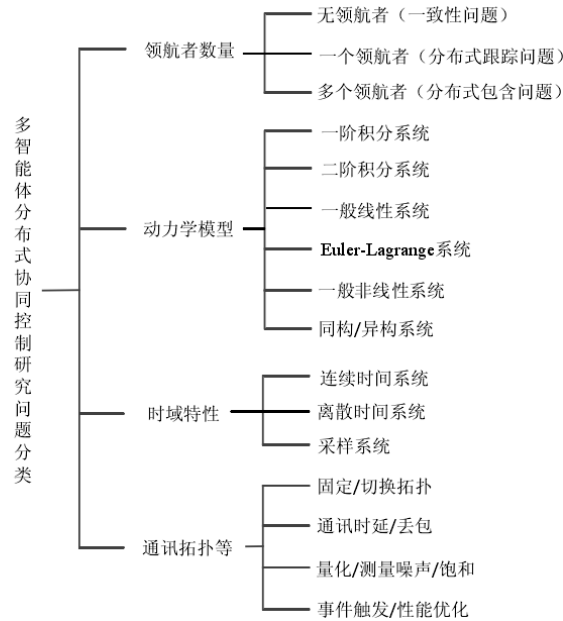


图 1 多智能体分布式协同控制问题分类

最早的集群系统模型是由 Reynolds[1]于 1987 年提出的 Boids 模型，该模型用计算机来模拟群体行为，并给出了智能集群系统满足的三个规则：

- 1) 速度匹配：agent 尽量与邻居 agents 速度和方向的平均值保持一致；
- 2) 聚集：agent 尽量向邻居 agents 的平均位置运动；
- 3) 避免碰撞：相邻 agents 之间避免发生碰撞。

2001 年，Reynolds 将 Boids 模型所有资料公布于网站¹上，包括程序、示例、相关论文等。在 Boids 模型的基础上，1995 年 Vicsek 等人[2]将其简化，提出了一种离散多智能体模型----Vicsek 模型，用于模拟大量粒子涌现出一致性的现象，其实质是 Boids 模型中速度匹配的动力学表示。Vicsek 模型刻画了多个粒子构成的自治系统的同步运动。在这个模型中粒子遵循如下规则：

- 1) 系统中运动的粒子具有常速率 v ；
- 2) 粒子存在一个影响半径 r ，即系统中的任意一对粒子，只有这对粒子之间的直线距离小于 r 时，他们才存在相互的影响；
- 3) 粒子每一时刻的运动方向跟上一时刻它的影响半径范围内的其他所有粒子的平均运动方向相同。

同时 Vicsek 模型引入了数学化描述，如公式(1.1)所示：

$$\begin{aligned} \mathbf{x}_i(t+1) &= \mathbf{x}_i(t) + \mathbf{v}_i(t)\Delta t \\ \theta_i(t+1) &= \langle \theta_i(t) \rangle_r + \Delta_i(t) \end{aligned} \quad (1.1)$$

其中： $\mathbf{x}_i(t)$ 为 agent i 在 t 时刻的位置， $\mathbf{v}_i(t)$ 为 agent i 在 t 时刻的速度， $\theta_i(t)$

¹ <http://www.red3d.com/cwr/boids/>

为 agent i 在 t 时刻的航向, $\langle \theta_i(t) \rangle_r$ 为 agent i 及其周围 agents 的航向的平均, $\Delta_i(t)$ 为扰动项。利用 φ 评判是否同步的标注, 如公式(1.2):

$$\varphi = \frac{1}{Nv_0} \left| \sum_{i=1}^N v_i \right| \quad (1.2)$$

φ 可以描述为归一化的平均速度, 当速度一致时, $\varphi = 1$ 。完全杂乱无章时, $\varphi = 0$ 。在[2]中也给出了结果, 如图 2 所示:

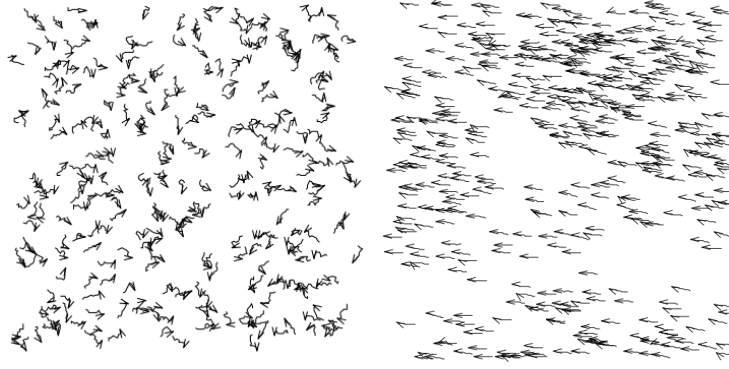


图 2 左: $\varphi = 0$ 右: $\varphi = 1$

2003 年, Jadbabaie 等人[3]在无噪声的假设条件下对 Vicsek 模型进行了简化, 用矩阵论和图论(无向图)给出了 Vicsek 模型的收敛性的理论证明, 指出: 只要满足联通, 粒子的运动方向就能达到一致。与 Vicsek 描述类似, Jadbabaie 引入了图论, 将无领导的模型描述为:

$$\theta_i(t+1) = \frac{1}{1+k_i(t)} (\theta_i(t) + \sum_{j \in N_i(t)} \theta_j(t)) \quad (1.3)$$

其中 $k_i(t)$ 为 agent i 周围的 agents 数目, $\sum_{j \in N_i(t)} \theta_j(t)$ 描述为周围 agents 航向之和, $N_i(t)$ 表示 t 时刻 agent i 的拓扑连接矩阵。如果 agents 间距离较远, 系统容易发生聚堆、簇分离的情况, 因此 Jadbabaie 等人又引入了虚拟领航员, 作为导航项 $b_i(t)$, 将公式(1.3)进行修改, 得到如公式(1.4)的系统模型:

$$\theta_i(t+1) = \frac{1}{1+k_i(t)+b_i(t)} (\theta_i(t) + \sum_{j \in N_i(t)} \theta_j(t) + b_i(t)\theta_0) \quad (1.4)$$

当需要领航员的情况下, $b_i(t)=1$, 否则 $b_i(t)=0$ 。集群系统模型自此从群体动力学模型时代进入了网络化系统与图论描述时代。

Olfati-Saber 等人[4]–[7]在 Jadbabaie 工作的基础上研究了系统的网络拓扑结构与系统收敛性之间的关系, 他指出如果系统的网络拓扑结构是强连通的有向图, 则对于任意初始状态, 系统的状态是渐近收敛的, 且对于强连通有效拓扑结构下的多智能体系统, 平均一致收敛的充要条件是它的信息交换图是平衡图。

Olfati-Saber 将系统描述为 $G=(\mathcal{V}, \mathcal{E}, \mathcal{A})$, 其中 \mathcal{V} 是节点, \mathcal{E} 为边, \mathcal{A} 为邻接矩阵, 它的元素均为非负。agent 邻居的集合为:

$$N_J := \bigcup_{v_i \in J} N_i = \left\{ v_j \in \mathcal{V} : v_i \in J, (v_i, v_j) \in \mathcal{E} \right\}, J \subseteq \mathcal{V}$$

动力学模型建立为: $\dot{x}_i = f(x_i, u_i), i \in \mathcal{I}$

定义有向拉普拉斯矩阵为: $\mathcal{L}(\mathcal{G}) = L = \Delta - \mathcal{A}$, 其中 Δ 为入度矩阵。

对于切换拓扑的模型定义为:

$$\Gamma_n = \{G = (\mathcal{V}, \mathcal{E}, \mathcal{A}) : \text{rank}(\mathcal{L}(G)) = n-1, \mathbf{1}^T \mathcal{L}(G) = \mathbf{0}\} \quad (1.5)$$

$$\dot{x}(t) = -\mathcal{L}(G_k)x(t) \quad k = s(t), G_k \in \Gamma_n$$

第一个式子是对强连通网络的描述, 第二个式子是对系统的描述。其中 $s(t): \mathbb{R}_{\geq 0} \rightarrow \mathcal{I}_{\Gamma_n}$ 是切换信号。

具有时延的 agent i 模型为:

$$\dot{x}_i(t) = \sum_{v_j \in N_i} a_{ij} [x_j(t - \tau_{ij}) - x_i(t - \tau_{ij})] \quad (1.6)$$

其中, $\tau_{ij} > 0$ 是 agent i 与 agent j 之间的时延。整个系统模型的传递函数为:

$$G(s) = (sI + L(s))^{-1} \quad (1.7)$$

其中 $L(s) = e^{-\tau s} L(G)$

在 Olfati-Saber 的基础上, 美国加州大学河滨分校任伟教授等人[8]固定拓扑结构的假设下, 将强连通拓扑条件弱化为只要在一段时间内网络拓扑子图的联合图包含一条有向生成树, 则系统可实现一致性; 在动态变化的交互拓扑下, 如果有向交互图的并集在系统演化过程中有足够频繁的生成树, 则也可以实现信息渐近一致。其离散系统建模(1.8)与连续系统建模(1.9)为:

$$\xi[k+1] = D[k]\xi[k] \quad (1.8)$$

$$\dot{\xi}(t) = C(t)\xi(t) \quad (1.9)$$

其中 $\xi = [\xi_1, \dots, \xi_n]^T, D = [d_{ij}], d_{ij} = \alpha_{ij}[k]G_{ij}[k] / \sum_{j=1}^n \alpha_{ij}[k]G_{ij}[k]$,

$\alpha_{ij}[k]$ 为权重, 在 i 和 j 有信息交互时:

$$G_{ij}[k]=1; \quad C = [c_{ij}], (i, j) \in \mathcal{I}, \quad c_{ii} = -\sum_{j \neq i} (\sigma_{ij}(t)G_{ij}(t)) \quad c_{ij} = \sigma_{ij}(t)G_{ij}(t), j \neq i$$

除此之外, Cucker 和 Smale[9]在 Vicsek 和 Jadbabaie 工作的基础上提出了一个非常有意义的集群模型(CS 模型), 在模型中所有 agent 具有惯性, 且整个系统完全驱动。在 CS 模型中, 每个 agent 会对其速度进行自我调节, 即通过自己在时刻 t 的速度跟其他 agent 在时刻 t 的速度差的加权平均值来调节自己下一时刻的速度。之前的模型需要在无限时间序列上的一个假设, CS 模型的收敛结果只依赖于初始状态条件和影响参数。在 CS 模型中, 也做了一些理想化处理:

1). 所有 agent 之间都有相互影响;

2). agent 受其它 agent 影响的强弱跟它们之间绝对距离和速度差有关。

连续 CS 模型: 考虑一个具 N 个 agent 的多智能体复杂系统, 对于 agent i,

它在 t 时刻的位移记为 $\mathbf{x}_i(t)$ 的, 速度记为 $\mathbf{v}_i(t)$ 的, 则 $(\mathbf{x}_i(t), \mathbf{v}_i(t))$ 满足:

$$\begin{aligned} \frac{d\mathbf{x}_i(t)}{dt} &= \mathbf{v}_i(t) \\ \frac{d\mathbf{v}_i(t)}{dt} &= \alpha \sum_{j \in N_j \neq i} a_{ij} (\|\mathbf{x}_j(t) - \mathbf{x}_i(t)\|) (\mathbf{v}_j(t) - \mathbf{v}_i(t)) \end{aligned} \quad (1.10)$$

$$\text{这里 } a_{ij}(t) = \frac{1}{N} \phi(\|\mathbf{x}_j(t) - \mathbf{x}_i(t)\|), \phi(r) = \frac{1}{(1 + |\mathbf{r}|^2)^\beta}, \alpha, \beta > 0$$

2011 年美国马里兰大学终身杰出教授 Tadmor 等人[10]改进了 CS 模型, 在 CS 模型基础上建立一个新的多智能体复杂系统集群模型(MT)。在这个模型中不仅考虑了 agent 的数量还考虑了 agent 在空间中的几何关系。由于相对距离的引进使原 CS 模型中的对称性质遭到破坏, Tadmor 通过引进一个全新的分析方法对复杂系统的集群性质进行了开创性的研究。

连续 MT 模型: 考虑一个具 N 个 agent 的复杂系统, 对于 agent i , 它在 t 时刻的位移记为 $\mathbf{x}_i(t)$: 速度记为 $\mathbf{v}_i(t)$ 的, 则 $(\mathbf{x}_i(t), \mathbf{v}_i(t))$ 满足:

$$\begin{aligned} \frac{d\mathbf{x}_i(t)}{dt} &= \mathbf{v}_i(t) \\ \frac{d\mathbf{v}_i(t)}{dt} &= \alpha \sum_{j \in N_j \neq i} b_{ij} (\|\mathbf{x}_j(t) - \mathbf{x}_i(t)\|) (\mathbf{v}_j(t) - \mathbf{v}_i(t)) \end{aligned} \quad (1.11)$$

$$\text{其中 } b_{ij}(t) = \frac{\phi(\|\mathbf{x}_j(t) - \mathbf{x}_i(t)\|)}{\sum_{k \in \mathbb{N}} \phi(\|\mathbf{x}_k(t) - \mathbf{x}_i(t)\|)}, \phi(r) = \frac{1}{(1 + |\mathbf{r}|^2)^\beta}, \alpha, \beta > 0$$

以 CS 模型中 agent 之间的影响机制为基础, 美国的 Shen 教授[11]提出了一个特殊的具领导机制和分等级机制的集群模型(HL)。HL 模型中每个成员都属于和它对应的一个等级, 对它们从高到低依次排序。成员们遵循这样一个影响机制一低等级的成员只能被高等级的成员影响和领导。

HL 模型: 考虑由 $(k+1)$ 个 agent 组成的具等级制度的复杂系统, 对 agent 进行标记 $[0 \ 1 \ \cdots \ k]$ 每个等级对应相应的 agent, 满足:

- 1) 当 $j < i$ 时, $a_{ij}(t) \neq 0$;
- 2) 若 agent i 的领导组成的集合是 $L = \{a_{ij}(t) > 0\}$, 那么对于任意的 $i > 0$ 我们都有 $L \neq \emptyset$

$a_{ij}(t)$ 表示 agent i 被 agent j 领导。这样的模型称为 HL 模型。

哈工大 Li 等人[12]在 Shen 的基础上改进了 HL 模型, 建立了一个更为一般的单领航机制的多智能复杂系统集群模型(RL)。在这个模型中存在一个全局领导者, 它不受其它 agent 的影响, 但是直接或者间接地影响着其它所有 agent。这个模型更好地揭示了全体的合作信息交换的优点。

2013 年哈工大 Dong 等人[13]在 Shen 的基础上研究了具自由意志的离散 H-L 模型, 给出了一个自由意志函数控制的条件以确保系统的集群性。在上述基础上, 2016 年湖南大学李乐博士在其博士论文中提出了 HR 模型, 考虑一个有 N 个 agent 的自适应复杂系统, 假设系统中存在 K 个等级(整数 $K > 1$), 等级 i 记

为“ R_i ”，并且这个等级存在 N_i 个 agent。对于等级 R_i 中的 agent i ，它在 t 时刻的位移记为 $x_i(t)$ ，并且速度记为 $v_i(t)$ ， $(x_i(t), v_i(t))$ 满足：

$$\begin{aligned} \frac{dx_i(t)}{dt} &= v_i(t) \\ \frac{dv_i(t)}{dt} &= \alpha \sum_{j \in \mathbf{R}_{i-1} + \mathbf{R}_i, j \neq i} b_{ij} (\|x_j(t) - x_i(t)\|) (v_j(t) - v_i(t)), i \notin \mathbf{R}_1 \end{aligned} \quad (1.12)$$

[14]–[16]提出了一种经典的非线性动力学模型：Kuramoto 模型，该模型由 n 个耦合振荡器组成，其动力学方程[16]为(1.13)：

$$\begin{aligned} \dot{\theta} &= \omega - \frac{K}{N} B \sin(B^T \theta) \\ r^2 &= \frac{N^2 - 2e + 2\mathbf{1}_e^T \cos(B^T \theta)}{N^2} \end{aligned} \quad (1.13)$$

其中 B 为具有 N 个节点、 e 个边的有向图的 incidence 矩阵， j 传输到 i 的时候 $B_{ij}=1$ ，反之 $B_{ij}=-1$ ，其他情况时 $B_{ij}=0$ 。 θ, ω 分别表示振荡器的相位和固有频率， K 表示耦合的强度。

考虑时延的模型[15]为(1.14)：

$$\dot{\theta}_i = \omega_i + \frac{K}{n_i} \sum_{j=1}^N A_{ij} \sin(\theta_j(t - \tau_{ij}) - \theta_i(t)) \quad (1.14)$$

文献[17]中构建多智能体系统的领导编队的线性多智能体系统如下所示：

$$\begin{aligned} \dot{v} &= Ev + Hw \\ \dot{x}_i &= A_i x_i + B_i u_i + D_i v \\ e_i &= C_i x_i + F_i v, \quad i = 1, 2, \dots, N \end{aligned} \quad (1.15)$$

其中 $x_i \in \mathbb{R}^{n_i}, u_i \in \mathbb{R}, e_i \in \mathbb{R}$ 是 follower i 的状态、控制信号、跟随误差； $v \in \mathbb{R}^q$

是 leader 的状态，由第一个式子构建。 $D_i v$ 是扰动项， $F_i v$ 为参考项。

Brunovsky 型是一种具有代表性的标准高阶非线性集群系统[9]，这类的 agent 模型通过一个高阶积分器耦合未知非线性动力学以及未知扰动来表示：

$$\begin{aligned} \dot{x}_{ij} &= x_{i+1,j} \\ \dot{x}_{nj} &= f_j(x_j) + u_j + \zeta_j \end{aligned} \quad (1.16)$$

其中 $i=1, \dots, n$ ， $x_{ij} \in \mathbf{R}$ 是智能体 j 的第 i 阶状态； $x_j = [x_{1j} \ \dots \ x_{nj}]^T$ 是智能体 j 的状态向量；未知函数 $f_j(\bullet): \mathbf{R}^n \rightarrow \mathbf{R}$ 在 \mathbf{R}^n 上局部 Lipschitz，且 $f_j(0)=0$ ， $\zeta_j \in \mathbf{R}$ 是未知的但是有界的外部扰动。

文献[18]中考虑 N 个跟随者， M 个领航者（与其他 agent 不连接）的情况，

系统的 Laplacian 矩阵 L 写为:

$$L = \begin{bmatrix} 0_{m \times m} & 0_{m \times n} \\ L_{n \times m} & L_{n \times n} \end{bmatrix} \quad (1.17)$$

考虑时延的跟随者的状态方程可以写为:

$$\dot{X}_F(t) = -(L_{n \times n} \otimes I_p)X_F(t - \tau) - (L_{n \times m} \otimes I_p)X_L(t - \tau)$$

其中 p 为状态量 x 的维数。

清华大学王龙等人[19]在 2011 年提出异质集群系统模型, 将一阶、二阶模型混合, 建立系统模型, 如下:

$$\begin{cases} \dot{x}_i(t) = v_i(t), & i \in \mathcal{I}_m \\ \dot{v}_i(t) = u_i(t), & i \in \mathcal{I}_m \\ \dot{x}_i(t) = u_i(t), & i \in \{m+1, \dots, n\} \end{cases} \quad (1.18)$$

$x_i \in \mathbb{R}, v_i \in \mathbb{R}, u_i \in \mathbb{R}$ 是 agent i 的位置、速度、控制信号。

在[19]基础上, Kim 等人[20]则在考虑环境因素影响时, agent 之间的通信连接可能中断或重连的情况, 他们把这种情况描述为一个伯努利概率序列的数学模型, 根据通信中断前的一步信息来设计控制协议, 然后利用线性矩阵不等式工具, 解决了一阶、二阶异质集群系统在离散时间情况下的均方一致性控制问题。

Liu 则在王龙等人[19]基础上, 将异质系统扩展到更为复杂的情况: 即异质系统由线性一阶、线性二阶、非线性 Euler-Lagrange 三类智能体组成, 并在非线性 agent 的参数已知和非已知的情况下解决了异质多 agent 的编队和聚集问题。

$$\begin{aligned} \dot{x}_i(t) &= v_i(t), & i \in \mathcal{I}_l \\ \dot{v}_i(t) &= u_i(t), & i \in \mathcal{I}_l \\ \dot{x}_i(t) &= u_i(t), & i \in \{l+1, \dots, m\} \\ \dot{x}_i(t) &= v_i(t), & i \in \{m+1, \dots, n\} \\ M_i(x_i)\dot{v}_i + C_i(x_i, v_i)v_i &= u_i(t), & i \in \{m+1, \dots, n\} \end{aligned} \quad (1.19)$$

$x_i \in \mathbb{R}, v_i \in \mathbb{R}, u_i \in \mathbb{R}$ 是 agent i 的位置、速度、控制信号, $M_i(x_i) \in R$ 为一般惯性矩阵, $C_i(x_i, \dot{x}_i) \in R$ 为科氏力和离心力矩阵, E-L 方程需满足以下 4 条假设条件:

$$\begin{aligned} 0 &< \lambda_m \{M_i(x_i)\} I \leq M_i(x_i) \leq \lambda_M \{M_i(x_i)\} I < \infty \\ M_i(x_i)\ddot{x}_i + C_i(x_i, \dot{x}_i)\dot{x}_i &= Y_i(x_i, \dot{x}_i, \ddot{x}_i)\theta_i \\ r^T (M_i(x_i) - 2C_i(x_i, \dot{x}_i))r &= 0 \\ \|C_i(x_i, \dot{x}_i)\| &\leq k_c \|\dot{x}_i\| \end{aligned} \quad (1.20)$$

事件驱动建模与上述情况相似, 不再赘述。事件驱动集群系统控制典型文献 [21], [22]。

在强化学习领域, 建模主要围绕 MDP 和博弈论进行:

马尔可夫决策过程(MDP)可以由一个元组 (S, A, c, P, ρ) 来表示, 其中

S : 表示状态空间;

A : 动作空间;

$c(s, a) \in [0, \infty)$: 代价函数;

$P(s'|s,a)$: 状态转移概率;

$\rho(s)$: 初始状态概率分布。

部分可观的马尔可夫决策过程(POMDP)是 MDP 的更一般性描述。一般来说, POMDP 中智能体 i 可以描述为 $POMDP_i = \langle S, A_i, T_i, \odot_i, O_i, R_i \rangle$, 其中:

S : 环境的有限状态集;

A_i : 智能体 i 的行为集;

T_i : 智能体 i 在一种状态下采取行为 a , 到达某一状态的概率的集合;

\odot_i : 智能体 i 的观测(observation);

O_i : 智能体 i 的观测函数(observation function):定义了给定动作的观察概率;

R_i : 智能体 i 的奖励函数, 代表 i 的偏好。

由于代理是针对越来越复杂的环境构建的, 一组决策者需要以分散的方式做出选择时, 可以将问题建模为分散的部分可观察的马尔可夫决策过程 (Decentralized POMDP, Dec-POMDP) [23]。虽然 Dec-POMDP 模型为不确定性下的协同顺序决策提供了一个丰富的框架, 但该模型的计算复杂度是一个重要的研究挑战。它是部分可观测马尔可夫决策过程(POMDP)框架的扩展。一般来说, Dec-POMDP 中智能体 i 可以描述为 $POMDP_i = \langle I, S, A_i, T_i, R_i, \odot_i, O_i, h \rangle$, 其中:

$I = \{1, \dots, n\}$: n 个 agents 的集合

S : 环境的有限状态集;

A_i : agent i 的行为集, 结连行为的集合: $A = \times_i A_i$;

T_i : agent i 在一种状态下采取行为 a , 到达某一状态的概率的集合;

\odot_i : agent i 的观测(observation), 结连观测的集合 $\odot = \times_i \odot_i$

O_i : 智能体 i 的观测函数(observation function):定义了给定动作的观察概率;

R_i : 智能体 i 的奖励函数, 代表 i 的偏好, $R(s, a)$ 表示状态 s 下采取动作 a 的立即回报。

h : 问题的维度, 始终为一个正整数

Multi-agent Markov decision process(MMDP)[23]:相较于 Dec-POMDP, 多智能体 MDP 的描述与之相似, 即为: $MMDP_i = \langle I, S, A_i, T_i, R_i, h \rangle$ 。

与博弈论结合, Hu 等人[24]提出了 Stochastic game 模型:

$I = \{1, \dots, n\}$: n 个 agents 的集合

S : 环境的有限状态集;

对于每个 agent i , 有限行为集 A_i , 回报函数 $S \times A \rightarrow R_i$, 其中 $A = A_1 \times A_2 \times \dots \times A_n$

T_i : agent i 在某状态下采取行为 a , 到达另一状态的概率 $S \times A \times S \rightarrow [0 \ 1]$

与上述的模型不同, 一般随机博弈为了寻找一个 Nash 均衡点, 其满足对于所有的 $\pi^i \in \Pi^i$:

$$v^i(s, \pi_*^1, \dots, \pi_*^n) \geq v^i(s, \pi_*^1, \dots, \pi_*^{i-1}, \pi^i, \pi_*^{i+1}, \dots, \pi_*^n)$$

与图论结合 Zhang 等人[25]提出了 Networked Multi-Agent MDP 模型，其同样可以描述为一个元组： $(\mathcal{S}, \{\mathcal{A}^i\}_{i \in \mathcal{N}}, P, \{R^i\}_{i \in \mathcal{N}}, \{\mathcal{G}_t\}_{t \geq 0})$

其中： $\{\mathcal{G}_t = (\mathcal{N}, \mathcal{E}_t)\}_{t \geq 0}$ 是一个带有时延的 N 个 agent 的通信网络， \mathcal{S} 为全局状态空间， $\mathcal{A} = \prod_{i=1}^N \mathcal{A}^i$ 为所有 agent 的结连行为空间， $R^i: \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{R}$ 为局部回报函数， $P: \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow [0,1]$ 表示状态转移概率。Zhang 等人在文中假设状态和结连行为空间全局可观，回报函数仅局部可观。

参考文献

- [1] R. C.W, “Flocks, Herds, and Schools: A Distributed Behavioral Model,” *Tech. Coloproctol.*, vol. 21, no. 2, pp. 25–34, 1987.
- [2] T. Vicsek, A. Czirak, E. Ben-Jacob, I. Cohen, and O. Shochet, “Novel type of phase transition in a system of self-driven particles,” *Phys. Rev. Lett.*, vol. 75, no. 6, pp. 1226–1229, 1995.
- [3] A. Jadbabaie, J. Lin, and A. S. Morse, “Coordination of Groups of Mobile Autonomous Agents Using Nearest Neighbor Rules,” *IEEE Trans. Automat. Contr.*, vol. 48, no. 6, pp. 988–1001, 2003.
- [4] R. Olfati-Saber, J. A. Fax, and R. M. Murray, “Consensus and Cooperation in Networked Multi-Agent Systems,” *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, 2007.
- [5] R. Olfati-Saber, “Flocking for multi-agent dynamic systems: Algorithms and theory,” *IEEE Trans. Automat. Contr.*, vol. 51, no. 3, pp. 401–420, 2006.
- [6] L. Moreau, “Stability of Multiagent Systems With Time-Dependent Communication Links,” *IEEE Trans. Automat. Contr.*, vol. 50, no. 2, pp. 169–182, 2005.
- [7] R. Olfati-Saber and R. M. Murray, “Consensus problems in networks of agents with switching topology and time-delays,” *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1520–1533, 2004.
- [8] W. Ren and R. W. Beard, “Consensus seeking in multiagent systems under dynamically changing interaction topologies,” *IEEE Trans. Automat. Contr.*, vol. 50, no. 5, pp. 655–661, 2005.
- [9] H. Zhang and F. L. Lewis, “Adaptive cooperative tracking control of higher-order nonlinear systems with unknown dynamics,” *Automatica*, vol. 48, no. 7, pp. 1432–1439, 2012.
- [10] S. Motsch and E. Tadmor, “A New Model for Self-organized Dynamics and Its Flocking Behavior,” *J. Stat. Phys.*, vol. 144, no. 5, pp. 923–947, 2011.
- [11] J. Shen, “Cucker-smale flocking under hierarchical leadership,” *SIAM J. Appl. Math.*, vol. 68, no. 3, pp. 694–719, 2007.
- [12] Z. LI and XIAOPING XUE, “CUCKER–SMALE FLOCKING UNDER ROOTED LEADERSHIP WITH FIXED AND SWITCHING TOPOLOGIES,” *SIAM J. Appl. Math.*, vol. 70, no. 8, pp. 3156–3174, 2010.
- [13] Jiu-Gang Dong, “Flocking under hierarchical leadership with a free-will leader,” *Int. J. Robust Nonlinear Control*, vol. 18, no. October 2014, pp. 557–569, 2012.
- [14] R. Sepulchre, D. A. Paley, and N. E. Leonard, “Stabilization of planar collective motion: All-to-all communication,” *IEEE Trans. Automat. Contr.*, vol. 52, no. 5, pp. 811–824, 2007.
- [15] A. Papachristodoulou and A. Jadbabaie, “Synchronization in oscillator networks: Switching topologies and non-homogeneous delays,” *Proc. 44th IEEE Conf. Decis. Control. Eur. Control Conf. CDC-ECC ’05*, vol. 2005,

pp. 5692–5697, 2005.

[16] A. Jadbabaie, N. Motee, and M. Barahona, “On the stability of the Kuramoto model of coupled nonlinear oscillators,” *Proc. Am. Control Conf.*, vol. 5, pp. 4296–4301, 2004.

[17] W. Gao, Z. P. Jiang, F. L. Lewis, and Y. Wang, “Leader-to-Formation Stability of Multiagent Systems: An Adaptive Optimal Control Approach,” *IEEE Trans. Automat. Contr.*, vol. 63, no. 10, pp. 3581–3587, 2018.

[18] Y. Wang and Y. Song, “Leader-following control of high-order multi-agent systems under directed graphs: Pre-specified finite time approach,” *Automatica*, vol. 87, pp. 113–120, 2018.

[19] Y. Zheng, Y. Zhu, and L. Wang, “Consensus of heterogeneous multi-agent systems,” *IET Control Theory Appl.*, vol. 5, no. 16, pp. 1881–1888, 2011.

[20] J. M. Kim, J. B. Park, and Y. H. Choi, “Leaderless and leader-following consensus for heterogeneous multi-agent systems with random link failures,” *IET Control Theory Appl.*, vol. 8, no. 1, pp. 51–60, 2014.

[21] D. Yang, W. Ren, X. Liu, and W. Chen, “Decentralized event-triggered consensus for linear multi-agent systems under general directed graphs,” *Automatica*, vol. 69, pp. 242–249, 2016.

[22] C. Nowzari, E. Garcia, and J. Cortés, “Event-triggered communication and control of networked systems for multi-agent consensus,” *Automatica*, vol. 105, pp. 1–27, 2019.

[23] Frans A. Oliehoek & Christopher Amato, *A Concise Introduction to Decentralized POMDPs*. 2015.

[24] J. Hu and M. P. Wellman, “Nash Q-learning for general-sum stochastic games,” *J. Mach. Learn. Res.*, vol. 4, no. 6, pp. 1039–1069, 2004.

[25] K. Zhang, Z. Yang, H. Liu, T. Zhang, and T. Başar, “Fully Decentralized Multi-Agent Reinforcement Learning with Networked Agents,” pp. 1 – 45, 2018.