## 多智能体系统概述

**Agent概念：**Agent是处在某个环境中，能灵活自主地行动以实现其目标的个体

**智能体基本属性**：自治性：对行为及自我状态进行控制。反应性：可感知环境并根据环境作出反应。目标主动性：基于既定目标主动发起行为。 社会交互性：通过群体合作与协调实现设计目标。

**特征**：智能性：根据推理、学习等技术分析解释各种信息和知识的能力。代理能力：Agent能感知外界发生的消息，并根据自己所具有的知识、动作作出反应。

**架构体系**：基于推理的架构；基于反应的架构：包容体系架构 、马尔可夫决策过程；分层架构：水平分层架构、垂直分层架构；信念-愿望-意图架构。

**分类**：Agent 的分类可以根据体系结构、属性或功能的不同来进行分类。

#### 多智能体系统

**背景**：Agent具有自主性、交互性、反应性、主动性、学习性和移动性等智能特性，但单个的Agent对问题的解决能力有限，很难完成动态分布、网络和异构情况下的大型、复杂问题。Agent的研究最终要融入多Agent系统，研究单个Agent的最终目的是将它放入MAS中，解决大型、复杂问题。

**定义**：一组具有感知，相互通讯，计算，控制能力的智能体组成的系统 ；具有感知周围环境的能力；具有与其他智能体交互的能力 ；信息处理，计算能力； 需求响应，决策能力。

**分布式**(感知:局部性,交互:局部性,计算:异步性)->**开放性**(性质:异质,目标:异质,可加入和退出)->**动态性**(状态,交互结构,交互策略)

**交互结构**： 纯分布式结构、 联盟结构、等级层次结构。

**交互结构特征**: 小世界特性、无标度特性、网络层次特性。

## 多智能体系统模型及框架

#### 智能体与多智能体系统的基本模型与形式化定义

**Agent的抽象模型**：Agent从环境获取信息输入（环境感知、信息交互）;Agent经过决策作出动作作用于环境.

**与环境的交互:** Agent依赖于：Agent与其环境的先验知识与环境的交互历史(当前环境的观察值、先前经验和观察值，或是从学习中获得到的其他数据)；试图达到的目标或整个世界状态的偏好；能够执行的最原始动作，即能力。

**单Agent基本形式化定义**：

**多Agent系统的基本形式化定义**：

#### 多智能体系统问题求解基本框架

**A picture containing text, screenshot, font, number

Description automatically generated**

**知识**是可以用来解决本领域问题的领域信息。**表示方案**是Agent中使用的知识形式。**知识库**则是Agent中存储的所有知识的表达。

**确定问题**：确定问题解的构成，对问题的描述进行提炼。

**表示问题**：对问题、环境、系统进行建模。包括符号和符号系统；抽象和抽象层次：Agent可以用物理符号系统对世界进行建模、世界模型是世界中真实存在的详情和世界的动态行为的表达、所有的模型都是抽象出来的、抽象的层次是指按照抽象程度排列的一个偏序、低层抽象比高层抽象拥有更多的细节；层次：知识层是将Agent所知道的、所相信的以及它的目标抽象出来、符号层是对Agent的推理进行描述。

**推理与行为（计算并输出）：设计时刻推理**：是在设计Agent时进行的推理。它由Agent 的设计者而不是 Agent本身来完成。**离线计算**：在观察世界、需要行动之前，Agent能做的计算。Agent离线获取背景知识和数据，并将其汇编成可用的形式，称之为知识库。背景知识可以在设计时或离线状态时给出。**在线计算**：是在观察环境及在环境中进行某些行为时由Agent完成的计算。在线状态获取的信息称为观察。Agent必须使用知识库和观察来决定要做什么。**构建Agent的两大一般策略：**简化环境并为这些环境建立复杂的推理系统、在复杂环境中建立简单的agent

**定义解（解的类型）：最优解**：在所有的可能解中最符合求解质量度量的最好的解。这个度量用序数度量和基数度量来描述。**可满足解**：满足条件的解。**近似最优解**：质量度量接近理论上存在的最优解。**可能解**：一个有希望的解

#### 智能体设计的复杂性维度

**模块性维度**：**扁平结构**：Agent会将所有需要完成的工作按照某一顺序完成 。**模块化结构**：Agent会将任务分解为若干子任务，再考虑子任务中的细节 **分层结构**：Agent将用分层的方式解决这些子任务，直至将问题分解为一些简单的问题

**表示方案维度（如何描述世界）**：主要通过以下几个方面进行推理，**状态**：内部（信念）状态、环境状态。**特征**：用状态特征或状态真假的命题进行推理，状态可以用一系列特征描述，其中每个特征在每个状态都有一个值。**个体与关系**：特征可以依赖于关系与个体。单个个体上的关系是一种属性，个体之间的每一种可能关系上都存在特征。Agent有时需要对无限个体的集合进行处理，例如数的集合。此时无法用状态或特征描述，只能在关系层面上进行推理，例如用函数关系推理。

**规划期维度**：用来说明Agent规划的向前时间的程度。当Agent决定做出某个动作时，能够观察到未来的远近，被称为规划期。规划期是Agent认为它的动作结果所能影响的向前程度。 Agent做规划时所考虑的时间点被称为阶段。根据规划期对**Agent分类**：非规划Agent（做出动作时，不考虑未来影响）、有限期规划（遵循有限的时间步）、不确定期规划 （能够向前探索但不确定步数）、无限期规划（ 一直向前进行规划的Agent）。

**不确定性维度**：**感知不确定性**（能否感知环境中的状态），分为完全可观察（Agent直接观察世界状态）和部分可观察（Agent不能直接观察状态或相同的观察结果导致很多状态）。**效用不确定性**（做出动作的效果不确定）分为确定性的（动作所导致的状态由动作及之前的状态决定）和随机的（对于结果状态，只能给出概率分布)。

**偏好维度**：Agent会为自身获取更优的结果，做出某一动作优于另一动作选择的唯一原因是其偏好动作会导致更理想的结果，取值是**目标**和**复杂偏好**。**目标：**在某一最终状态下要达到的完成目标，或是在所有已访问过的状态中必须背保持的目标。**复杂偏好：**在不同时期权衡各种期望的结果。序数偏好表示只注重偏好的排序。基数偏好涉及有关值的大小

**Agent数量维度**：**单Agent推理**和**多Agent推理**（是否要考虑其他Agent的目标和偏好，其他Agent的动作会不会对自身造成影响）。

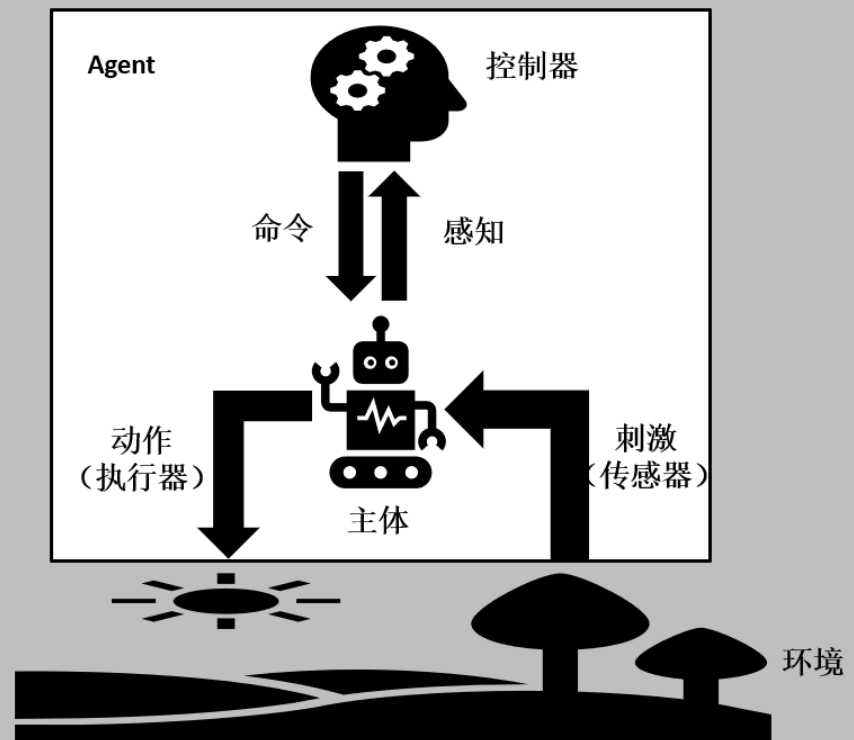
**学习维度**：Agent需要一些先验知识或其他资源来帮助它进行决策，分为：**已有的知识**和**学到的知识**（从数据或经验获取）。**学习**一般指找到与数据相符的最好模型。

**计算限制维度**：计算限制维度中Agent性质：**完全理性**（Agent可以推理出最佳方案，不需要考虑有限的计算资源）。**有限理性** （在给定的计算资源限制上，Agent决定所能找到的最佳方案）。

**多维交互**：**表示方案维度与模块性维度交互**： 一些层次上的简单模块可以用一系列有限状态集进行推理，而其他抽象层次可能要对个体及关系进行推理。**规划期维度与模块性维度交互**：在高层次时，小狗跑过来并得到治疗可以得到即时奖励。当决定将它的爪子放在哪里时，需要很长时间来得到奖励。**效用不确定性与模块性维度交互**：在分层结构的某个层次上，某个动作可能是确定性的，但在另外的层次则可能是随机的。**偏好模型与不确定性交互**： Agent必须权衡满足由一定概率的主要目标，还是满足具有更高概率的策略目标。

**模块性**中可以使用**多Agent**。设计单个Agent的一种方式是构造多个拥有共同目标的交互Agent，这样能够使较高层次的Agent智能地执行动作。**学习**可以通过**特征**来进行描述，决定哪个特征值能够最好地预测其他特征的值。学习也可以通过个体及关系来进行。维度中**模块性**和**有限理性**能使推理更加**有效**。虽然形式化体系变得复杂，但却能够通过将系统**分解**成更小的组件并提供所需的近似值，使Agent能够在**有限时间**和**有限内存**中**及时**执行动作，有利于构建复杂系统。

#### Agent组成结构和构建



**控制器的构建**：**离散时间和稠密时间**（设T为时间点集合，若任意两个时间点间仅存在有限个时间点，则称T是离散的。但如果T中任意两个时间点间都存在另一个时间点，则认为T是稠密的）；**感知轨迹**（假定P为在T时间内所有可能感知对象的集合，从T到P的函数为感知轨迹函数，描述了每一时间点所观察到的事物）；**命令轨迹**（假定C是所有命令的集合，一个为从T到C的函数为命令轨迹函数，其代表在每个时间点的命令）

**信念状态和因果转换**：**因果转换**指控制器必须保存Agent的**信念状态**（Agent所能记住的所有以前时间的信息的总和）并决定每个时刻发出何种命令(指令函数)。当它做这些时，它应该获得的信息需包括自身信念状态和当前感知，离散时间下的信念转换函数，其中， S 是信念状态集， P 是可能认知的集合; 表示状态 S 是在信念状态S 之后观察到P得出的信念状态。指令函数为，其中， C是可能指令的集合；C表示当前的信念状态为 S 观察到 P 时控制器需要发出的指令 C 。

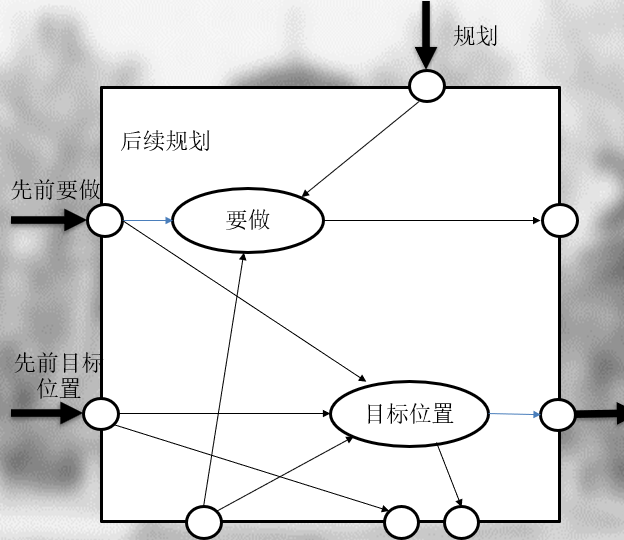
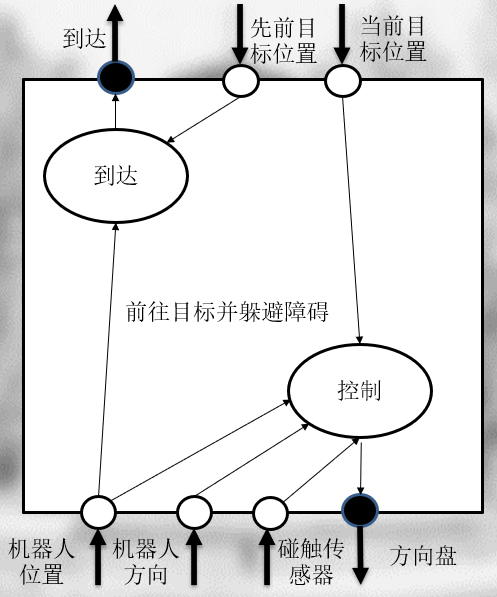
#### 分层控制

- 一个理想的分层Agent系统结构。箭头表示信息流。虚线说明某时刻的输出作为下一时刻的输入

A picture containing text, diagram, plan, line

Description automatically generated A picture containing text, screenshot, font, diagram

Description automatically generated



## 多智能体系统交互结构

**交互结构的表示方法**：用有向图<a,b>或无向图(a,b)表示，结点为agent，边为交互连接。

**交互结构的度量指标**：**度**；**平均（最短）路径长度**，用于评估智能体之间交互的平均耗费：

**局部聚类系数**：，其中 表示节点的邻居之间存在的实际边数，表示节点 的邻居个数。**平均聚类系数**：；**全局聚类系数**：

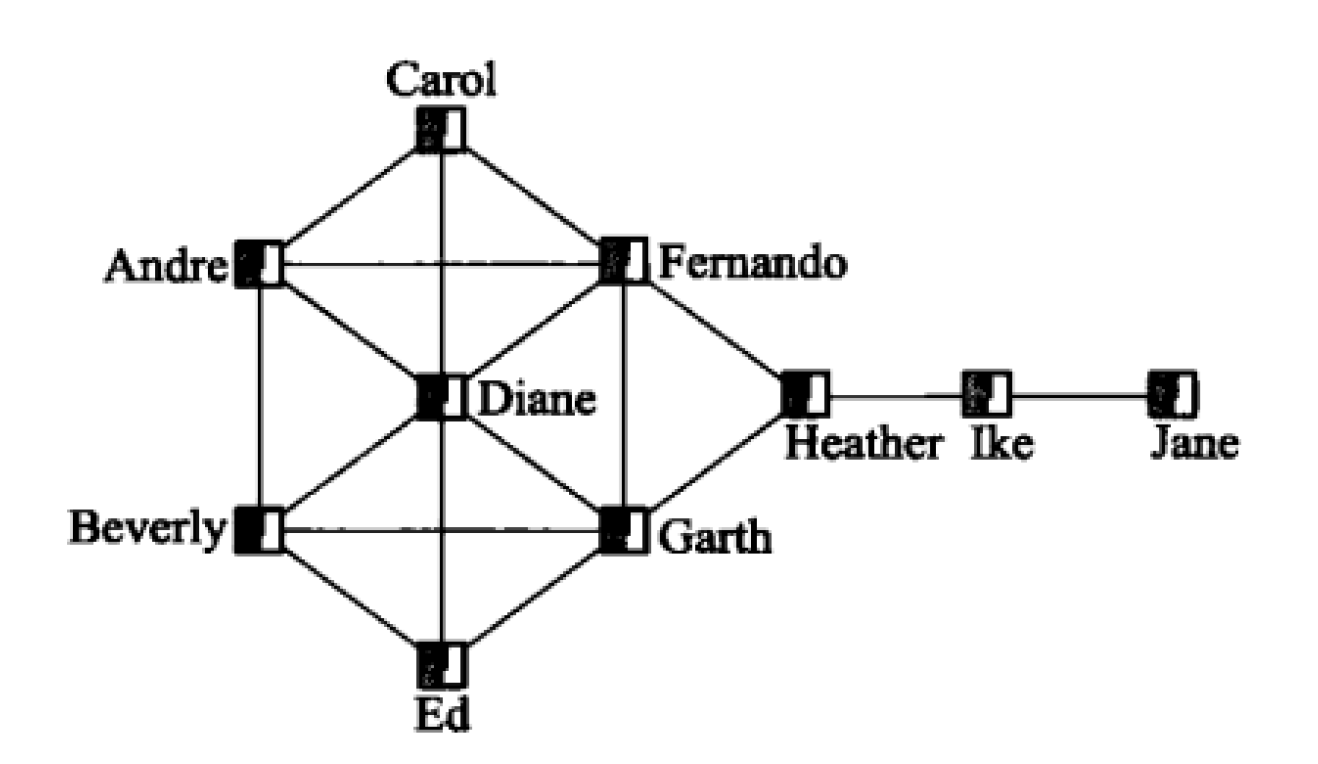
**交互结构的基本模型**：**全局耦合网络** （任意两个节点之间均有边相连），有聚类性、小世界性质，便于全局协调，, , ；**最近邻耦合网络**（任意节点均与其左右K/2个邻居节点相连，其中K为偶数），有聚类性、不具有小世界性质、难以全局协调。,

**星型网络**（存在一个中心节点，其余节点均只与该中心节点连接），具有小世界性质、便于全局协调、鲁棒性差，（非中心节点），，。**随机网络**（任意两点之间以一定概率p建立交互链接）、具有小世界性质、p很小时没有高聚类特性，（非中心节点）, ；**WS小世界模型**：建立规则图（最近邻耦合网络），随机化重连（以概率p随机地重连网络中的每条边，但无重边和自环）。；**NW小世界模型**：建立规则图（最近邻耦合网络），随机化加边（以概率p随机在任意一对节点间增加边，但无重边和自环）。

。

**无标度网络**：网络增长（首先建立一个具有 个节点的网络, 每次增加一个节点, 将连接到已有的 个节点上, ，优先连接（新节点与已存在节点建立连接的概率为），演化特性为富者更富 赢者通吃。**BA网络**：同无标度网络，但新节点与已存在节点建立连接的概率为。**鲁棒性**：进行一定数量的节点移除后，网络中大部分节点仍是连通。无标度网络对于随机故障策略具有良好鲁棒性，对于蓄意攻击策略具有高度脆弱性。

**交互结构的应用性指标**：**度中心性**（度越大越重要）。**介数中心性**（最短路径通过该点越多越重要）：，其中是从节点s到节点t的最短路径数量，是从节点s到节点t的最短路径中经过节点i的数量。**接近中心性**（越大损耗越小），其中是结点i到其他所有节点距离均值，越小意味着节点i更接近其他节点。**Kite网络**（度最大、介数最大、接近数最大的节点各不相同）





**交互连接强度**：用邻居重叠度定义， ，其中*、*分别为*i、j*不包含两端点的邻居集合。**交互连接的嵌入性**：两个节点共同邻居数量。**交互连接介数中心性**（最关键的边）：交互连接上所有节点对间最短路径流量之和。对于节点A和B，设想1个单位的流量从A到B，均分到它们之间所有的最短路径上：k条路径，则每条路径上分得。若一条边被m条路径共用，则在它上面流过。所有节点对都考虑后一条边的累积流量就是它的介数。**交互连接正负性**：给边分配正负，若交互结构中所有三角关系都是或，则为平衡的。

## 多智能体协作

**需要协作的情况：**缺少完成某一任务的信息；只具有完成部分任务的能力；通过多智能体的合作，可以提高解决问题的效率。

**多智能体协作方式**：**集中式协作方式**：存在一个智能体做决策并集中控制整个系统， 是一种规划与决策的自上而下式的层次控制结构。主要方法是集中式协商方法（投票协商）、运筹学方法、集中式学习方法。优点是协调性较好、可以得到全局最优解。缺点是实时性和动态性差、对环境变化的响应能力差。**分布式协作方式**：各智能体真有高度的自治能力，自行处理信息、 规划与诀策、执行指令，与其它智能体通讯进行协调，没有集中控制单元。主要方法有合同网模型、分布式黑板模型、分布式学习方法、博弈论。优点是有较好的容错能力和扩展能力、不需要集中控制单元。缺点是对智能体间通讯要求高、多边协作效率低，全局目标的实现难以保证。

**黑板模型**（黑板+知识源+监控机制）：为了解决分布在不同物理环境下多个实体协作完成任务的并行和分布计算模型。该模型能实现异构知识源的集成。分为中心式黑板和分布式黑板。

A screenshot of a computer screen

Description automatically generated with low confidence

**知识源**：描述某个独立领域问题的知识及其知识处理方法的知识库。在系统中具有多个知识源，每个知识源可用来完成某些特定的解题功能。知识源具有“条件-动作”的形式。条件描述了知识源应用求解的前提，动作描述了知识源的行为。当条件满足时，知识源被触发，其动作部分增加或修改黑板上的内容。**黑板**：是用来存储数据、传递信息和处理方法的动态数据库，是系统中的全局工作区，主要存放知识源所需要的信息和求解过程中的解状态数据，如初始数据、部分解、替换解、最终解等。分为中心式黑板（简单易行，易产生性能瓶颈）和分布式黑板（可靠性高，灵活性强）。**监控机制**：根据黑板上的问题求解状态和各知识源的求解技能依据某种控制策略，动态地选择和激活合适的知识源，使知识源能实时地响应黑板的变化。**模型特点**：黑板模型的各主体（即知识源）之间相互独立，主体之间不存在相互作用；黑板模型能够灵活表示信息；黑板模型使用共同的交互语言； 黑板模型具有独立的监控机制；黑板模型适合于在多重抽象级上描述与处理问题；黑板模型是一种机遇问题求解机制，适合于事先无法确定问题求解次序的复杂问题；黑板模型提供了集成现有软件的一种方法，适用于软件的开发。

**合同网模型**（管理者+工作者）：**管理者**：对每一待求解任务建立任务通知书，将任务通知书发送给有关的工作者主体，接收并评估来自工作者的投标，然后从投标中选择最合适的工作者，与之建立合同，最后监督任务的完成，并综合结果。**工作者**：接收相关任务通知书，评价自己的资格，对感兴趣的子任务返回任务投标。如果投标被接受，按合同执行分配给自己的任务，然后向管理者报告求解结果。**任务通知书**：不需要预先定义主体的角色，任何主体通过发布任务通知书而成为管理者，任何主体通过应答任务通知书而成为工作者。这一灵活性使任务能够被层次地分解分配，形成一个动态确定的树结构。**构成**：收件人（任务通知书可发送给一个或多个可能求解任务的主体）、资格说明（规定投标主体应满足的条件）、任务简述（对任务的简要描述，是工作者决策是否投标的依据）、 投标说明（向各主体说明投标必须提供的信息）、**终止时间**（接收投标的截止时间）。**特点**：合同网中任务的产生、任务的分配、管理者、合同者的产生均是动态的，因此其灵活性较好。智能体间订立合同需要预先订立协议，因而对于任务不明确的大规模、动态环境下的任务分配问题，管理者并不了解实时情况下有哪些智能体可以完成任务。当更有能力的主体在合同建立之前处于忙状态时，管理者只能选择一个有限能力的智能体，也即解可能不是最优的。合同网模型忽略了任务之间可能存在的各种关联，因此更适用于任务能够较容易地独立分解，即分解后各个的子任务之间不存在相互作用的问题。在协商过程中合同网的各智能体间以广播的方式互相通信，若智能体数量多则通信效率较低。**联盟形成模型**：自治的智能体会结成联盟解决单个智能体无法解决的问题，提高问题求解的收益。结成联盟具有有效性、稳定性、简单性、分布性和对称性。**产生联盟结构**（找能达成目标的智能体构成联盟）：分为非重叠联盟结构和重叠联盟结构。常用方法有基于动态规划的启发式算法、基于图论的启发式算法、智能算法与进化算法。**任务分配**（将任务分派给哪个智能体联盟）：一个智能体联盟可以完成多项任务； 同个智能体联盟完成不同的任务有不同的收益；不同智能体联盟完成相同的任务有不同的收益。优化目标是使得所有智能体的收益和最大。**效用分配**：需要满足集体理性原则（联盟效用要分配给联盟中的全部联盟成员）、个体理性原则（加入联盟中的智能体所获得的效用要大于智能体独立工作所获得的收益）。如果智能体对收益不满，那么其将退出联盟，破坏联盟形成的稳定性。

**强化学习模型**：解决了智能体如何通过与环境交互学习以最大化累积奖励值的问题。**解决智能体间学习过程交互问题的方法**：**智能体独立训练**：每个智能体独立学习，互不干涉，将其他智能体视为环境；通过智能体与环境的交互间接实现智能体之间的交互；由于存在多个智能体的环境极其不稳定，在部分可观察的环境中这种方法效果较差。**智能体共享参数**：智能体间共享参数可以充分利用所有样本训练出一个较好的策略，提高样本的利用效率；使用智能体共享参数的方式会交互没有必要的信息，很难实现智能体间的协作，策略收敛速度较慢；由于学习样本极易过期，使用过期的样本训练策略会导致策略越学越差，所以使用参数共享机制需要识别并处理过期的样本。

## 多智能体协作与协同-博弈

**博弈基础**：**基础要素**：**参与者**（至少两个）、**策略集**（一组如何执行具体行为的备选项，备选项称为参与者的可能策略）、**收益**: 每个行为策略均会使参与人得到一个收益, 此收益还会受到博弈中其他参与者策略选择的影响, 同一种策略，不同的参与人得到的收益可能不同。

**基本假设**：**已知的收益矩阵**（每个参与人对于收益矩阵均有充分的了解）、**理性的参与者**（参与者始终追求博弈中自身收益最大化）、**独立的决策**（参与者进行决策时不能进行协商）。**严格占优策略**：对于一个参与人来说，如果存在一个策略，无论对方如何决策，该策略都是最佳选择，称该策略对于此参与人为严格占优策略。参与人会选择严格占优策略。理性个体的理性选择可能会得到对于群体收益的次优结果。**（严格）最佳应对**：设 分别为参与人甲和乙选择的一个策略。当参与人乙采取策略，如果甲采用策略 得到的收益大于（等于）任何其他策略，则称甲的策略 是乙的策略 的（严格）最佳应对。对于某个策略 ，最多只能存在一个严格最佳应对，对于不同策略，最佳应对可能相同，也可能不同。**（严格）占优策略**：如果甲的策略 ，对于乙的每一个策略都是（严格）最佳应对，则称策略为甲的（严格）占优策略。仅有一个参与者具有严格占优策略时，该参与者会采取严格占优策略，另一方将采取此策略的最佳应对。**纳什均衡**：两个参与人，选择的策略组为；如果为的最佳应对，且同时为的最佳应对，则称策略组为纳什均衡。在纳什均衡状态，任意一方均没有动机去改变策略；如果两人同时改变策略可能会更好，但单方面改变不会得到额外好处；纳什均衡可能有多个。**多重均衡**：**协调博弈**：达到对多方都有利的结果）。某策略组合下双方收益相同为**对等协调博弈**，反之为**非对等**。通常可以预期参与者会选择收益更高的纳什均衡点。

**基本思路**：

A picture containing text, screenshot, font, line

Description automatically generatedA picture containing text, screenshot, font, line

Description automatically generated

**零和博弈**：参与人收益的和为0，一方获益造成另一方损失，不存在纯策略纳什均衡。**混合策略**：以一定的概率进行策略的选择。 **混合策略预期收益**：设A和B的混合策略分别是和，收益矩阵分别是和，那么收益分别为和。**纳什均衡存在性定理**：具有有限参与者和有限纯策略的博弈一定存在纳什均衡（包括混合策略均衡）。**求解混合纳什均衡**（在何种混合策略下存在纳什均衡）：设策略集，则A的采用策略的预期收益。利用和解出。

**博弈策略的迭代剔除**：**严格劣势策略**：对于参与人A的策略S，无论对方如何决策，其收益总小于另一个策略T，则称A的策略S严格劣势于T。**弱劣势策略**：对于参与人A的策略S，无论对方如何决策，其收益总小于等于另一个策略T，且至少存在一个策略组合(S, U)的收益小于(T, U)，则称A的策略S弱劣于为T。**策略迭代剔除**：迭代剔除弱劣势策略直至没有弱劣势策略。

#### 多人博弈

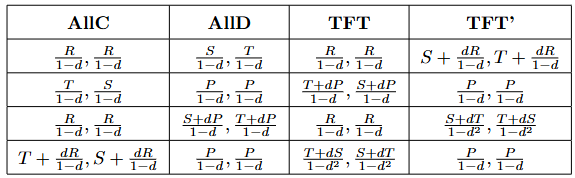
**公共物品博弈**：每个参与人有资源可以用于公共投资、每个参与人均可选择策略”投资的资源“或者策略”不投资“，全体投资得到的公共资源会乘以系数r(r>1)再平均分配给每个参与者，每个参与者的收益为。当时，不投资策略占优。

**重复博弈**：多次同结构博弈的重复**分为**有限次重复博弈和无限次重复博弈。

**重复囚徒博弈**：T > R > P > S

|  |  |  |
| --- | --- | --- |
|  | *C* (合作) | *D*(背叛) |
| *C* | R, R | S, T |
| *D* | T, S | P, P |

**经典策略**： AllC（总选择C策略）、AllD（总选择D策略）、Grim（在对手选择D策略之前选择C，当对手选择D后一直选择D）、TFT (Tit for Tat) （在第一轮选择C策略，随后选择对手上一轮采用的策略）、TFT‘（在第一轮选择D策略，随后选择对手上一轮采用的策略）。**总收益**：是Agent 在各轮阶段博弈中的收益，则Agent 在无限重复博弈中的收益是，其中为折损因子, 。 **总收益矩阵**（省略了第一列）：

**存在噪声的无限重复博弈**：TFTT ：初始采用C策略并维持；当对手连续两次采用D策略，则下一轮采用D策略；如果对手采用C策略，则下一轮采用C策略。GTFT ：发现对手采用D策略时，在下一轮仍有一定几率选用C策略。

## 多智能体自组织机制

**概念**：**组织**：系统内部按照一定规则形成的特定结构和功能，系统中的要素按照某种指令形成特定结构和功能的过程。组织化的必要条件是信息。**自组织**：系统的要素通过彼此的相干性，协同性或某种默契而形成的特定结构和过程。特点是局部化和分散化。

自组织的应用场景：多智能体系统、网格计算、传感器网络

**多智能体中的自组织**：使系统在执行期间无需明确命令即可更改其组织的机制或过程。**特征**：没有明确外部控制（系统的适应和改变只基于内部组件的决定，而不遵循任何明确的外部命令）、分散控制（自组织过程通过组件间的局部交互来实现，无需内部或外部的集中控制。此外，对全局信息的访问也受到交互位置的限制）、动态性和进化性（自组织系统可以随着环境的改变自适应动态进化）。**分类**：按照是否具有明确的控制中心分为**弱自组织系统（白蚁社会）**和**强自组织系统（细胞分化）**。

**多智能体系统交互结构自组织**：**根据交互结构分类**：**基于直接交互的自组织机制**（智能体之间的信息交换是通过直接交互进行的，专注于改变agent组织的结构。优点是简单高效，缺点是在智能体数量过大时收敛速度慢，不适用于大规模系统）、**基于间接交互的自组织机制**（智能体之间的信息交换不直接进行，而是通过某种中介进行）。

**基于直接交互的方法**：**基于度连接**（agent在自己的所有邻居中选择度最大并且与自己不相连的agent建立连接，同时在自己的邻居agent中选择一个度最小的agent取消连接。优点：拥有很多连接的个体一般能较快的学习到主流策略，所以与度最大的个体建立连接可以更快的完成学习）、**基于表现连接**（agent在自己的所有邻居中选择收益最高并且与自己不相连的agent建立连接，同时在自己的邻居agent中选择一个收益最低的agent取消连接。优点：通常使用好策略的agent收益会更高，与收益高的agent建立连接可以更好的学习到这种策略）、**随机全局连接**（agent在所有agent中随机选择一个与自己不相连的agent建立连接，同时在与自己相连的agent中随机选择一个agent取消连接）、**基于全局建议连接**（agent更加认可和自己做出相同选择的agent，因此agent选择一个当前时刻与自己做出相同策略并且与自己不相连的agent建立连接，同时在自己邻居agent中选择一个当前时刻与自己策略不同的agent取消连接）。

**基于间接交互的方法**：**基于中介节点的自组织机制**（agent通过中介来建立连接，主要分为执行agent和中介agent）：**执行agent**（负责执行任务、记录资源利用率，取消不必要的连接），**中介agent**（接受执行agent发来的调整连接请求，为执行agent）。**agent状态切换机制**（当前任务很少的执行agent可以切换为中介agent，为当前任务繁忙的超载agent承担起调整连接的工作）。优点：通过分工，提升系统效率，通过agent状态切换机制，使系统负载均衡。缺点：通过中介进行连接切换，有一定系统开销。

**多智能体系统能力资源自组织**：**基于强化（**初始时分配一定量资源，然后根据agent表现（负载率、收益、任务完成量等）调整资源。初始时刻为所有agent等量分配资源，一段时间后根据agent的负载率调整资源分配，负载率高的分配更多资源，负载率低的收回资源）、**基于贪心（**优先满足当前时刻表现最好的agent的资源请求。优点是简单、表现好；缺点是不够灵活）、**基于市场**（根据系统资源利用情况，动态调整资源价格，agent通过支付一定的价格来使用资源。优点是灵活、鲁棒性好）、 **基于拍卖**（招标-投标-中标，优点是可靠性高、灵活性高）***、***基于博弈论、基于学习、基于谈判、基于公平分配。

**结合自组织的多智能体协作：众包自组织**：无组织关系的非特定大众在主观意愿上互不相干，但却在相互默契的某种规则下解决众包任务，由此形成的具有类似组织效果的有序结构。**特点**：非强制性、无层级型、类组织性、进化性。**网格计算**：**网格**：一个集成的资源与计算环境，由多个位置不同的计算机资源组成。**网格计算**：基于网格的问题求解。**特点**：分布式、资源整合、均衡负载。**无人机自组织网络**：无人机群建立的动态网络，处于该网络中的无人机可以实现信息传递共享，以协作完成复杂任务。**特点**：拓扑变化频繁、节点密度不均匀、可用能量有限（基于电池容量）、任务场景和移动模型特殊。

**任务分配**：当一个Agent无法独立完成任务时，如何将合适的任务分配给合适的Agent以实现整体执行效果最优。分为**集中式任务分配** 和**分布式任务分配**。**集中式任务分配**：**典型方法**：整数规划方法、搜索算法（启发式、基于满意决策）、智能优化算法（群智能算法、遗传算法）。**特点**：实现简单，具备产生全局最优解的潜力，适用于小规模系统。**分布式任务分配**：**典型方法**：基于行为激励的方法、基于市场机制的方法（合同网协议、拍卖算法）、基于空闲链的方法、基于群智能的方法。**特点**：可并行计算、可以快速计算方案、应对动态环境效果好、适用于大规模系统。

**负载均衡**：到达系统的任务会分配给不同的Agent，当一个任务被分配到某个Agent上如果该Agent正在执行另一任务，此时就需要将到达任务放入任务等待队列中直到该Agent空闲。随着任务的不断到达，可能会出现一些拥有资源多的Agent上的任务等待队列很长，而另一些资源少的Agent的任务等待队列始终为空，此时系统就处于不均衡状态。**典型方法**：**静态负载均衡**：根据系统当前状态来调整各节点的负载，从而达到当前状态下系统的负载均衡。**特点**：可以很好的实现某个时刻的负载均衡、对未来的负载情况无能为力、动态任务环境下开销较大。**例子**：基于博弈论的负载均衡、基于任务信息的负载均衡。**动态负载均衡**：利用历史任务信息预测未来任务到达强度，根据预测出的未来任务到达强度进行当前状态的负载均衡。**特点**：可以缓解未来负载不均衡情况、减少系统计算负载均衡的开销 、适用于动态任务环境。**例子**：基于指数平滑方法预测负载、基于服务复制与转移的动态均衡。

## 多智能体群集扩散

**群集现象：**群体-个体组层，群集行为-大量个体行为所共同构成群集行为，群集现象是个体决策+系统规则构成的。

**群集策略演化**（分析社会群体中策略得以维持的规则与核心因素）：交互模式为两人博弈，收益矩阵对称。研究内容是**演化稳定策略**：当某个具有原始策略S的系统，产生了采用策略T的突变个体；如果经过演化，采用策略T的突变个体最终消失，那么称策略S为演化稳定策略。**系统演化特性**（自然选择）：具有相对更高收益的策略的个体数量会增长；具有相对更少收益的策略的个体数量会减少。 **从收益矩阵分析策略演化稳定性**：计算各策略预期收益，预期收益最大的策略为稳定演化策略。**从纳什均衡看策略演化的稳定性**：策略组为严格纳什均衡（），则策略为演化稳定策略；策略组不是纳什均衡（），则策略不是演化稳定策略（存在策略突变使得个体收益更高）；策略组为弱纳什均衡（），并且 ，则策略为演化稳定策略。

**亲属选择：**亲属选择指个体或群体仅对其同类或亲属表现出来的利他行为。根据个体或群体的社会生活经验，同类或亲属在其情感的认知判断中具有优先呗选择的权力。

**PD Game类博弈**：采用策略C的个体自身将耗费 c ，使对方获得收入 b。采用策略D的个体自身耗费为 0 ，对方获得收入为 0

|  |  |  |
| --- | --- | --- |
|  | C | D |
| C |  |  |
| D |  |  |

**合作（策略C）成为演化稳定策略的经典系统规则**：亲属选择、直接互惠、间接互惠。**亲属选择**：群体中的任意个体，与另一个随机选择的个体为亲属关系的概率为 r（亲属系数）。如果双方为亲属，随机选择的个体将采用与原个体相同的策略。假设群体中采用策略C的个体占群体的比例为，则采用策略D的个体占群里的比例为。采用策略C/D的个体收益分别为。得到收益矩阵

|  |  |  |
| --- | --- | --- |
|  | C（时） | D（时） |
| C |  |  |
| D |  |  |

C策略为演化稳定策略的条件是。**直接互惠**：策略集为{TFT,AllD}的群体中的随机配对个体，采用无限重复博弈模式进行交互。收益矩阵为

|  |  |  |
| --- | --- | --- |
|  | TFT | AllD |
| TFT |  |  |
| AllD |  |  |

TFT策略为演化稳定策略的条件是。**间接互惠**：策略C个体将与策略C个体合作，将有的概率与策略D个体合作。收益矩阵：

|  |  |  |
| --- | --- | --- |
|  | C | D |
| C |  |  |
| D |  |  |

C策略为演化稳定策略的条件是。

**群集运动 理解：**系统中的个体按照自身的属性赋予其行为规则，并结合个体间的相互通信，进行相应运动，最后随着时间的演化，整体系统形成某种规律性的场景。群集行为理论是多智能体系统研究中的基础核心。个体间的信息交互方式决定群集行为的类型，因此研究群集行为的本质是分析和设计相应的信息交互方式，即分布式协议。**研究意义：**促进群集理论在实际工程中的应用，为社会网络中出现的复杂现象提供更为合理的理论解释，进而提高生产率。

**群集运动场景：**多无人机系统、工业环境下多机器人系统、城市交通网络

**分类：**个体动力学方程角度（个体模型分为同质个体与异质个体）、个体相互关系角度（个体拓扑结构分为合作网络与合作竞争网络）

**同质个体研究：**一致性是核心，一致性是指多智能体系统中的个体在局部协作和相互通信下，调整更新自己的行为，最终使得每个个体均达到相同的状态。**问题关键：**设计仅利用有限邻居信息的分布式协议（事件驱动）。**群一致性：**多智能体系统中的个体根据某种规则被划分成若干个子群，同一个子群中的个体趋于同一状态，同时不同子群中的个体趋于不同的状态。**难点：**呈现出多个一致状态，个体不仅与群内部的邻居进行信息交换，还和不同群的个体也有信息交互。**解决方案：**分布式协议往往带有补偿特征，补偿个体异质带来的不利影响，从而实现整体系统一致性。

**异质个体研究：**异质多智能体系统是指系统中个体的自身属性不同，在数学模型上用不同的动力学方程描述。**难点：**需要设计带有异质特征的分布式协议。

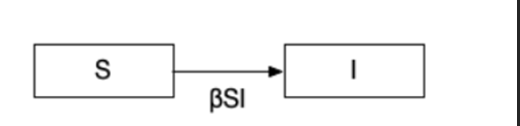
**合作竞争网络：**交互平衡图、子平衡图、非平衡图，对于交互平衡图和子平衡图，多智能体系统可实现群一致性，对于非平衡图，需施加额外的潜质控制器才可以。**展望：**个体异质普遍存在、个体间有竞争、群集行为存在定性表现

**Boids模型**：

**基本规则：**分离规则（计算范围内质心，远离质心）、对齐规则（计算范围内平均速度，保持一致）、聚合规则（计算范围内质心，指向质心）

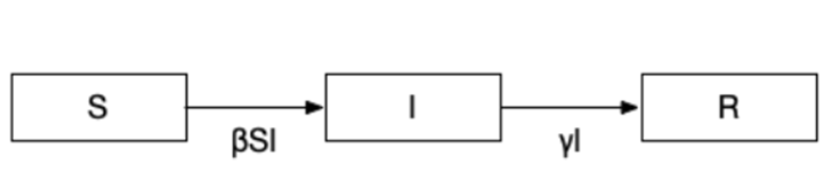
**群集信息传播**：**线性阈值模型**：网络表示为一个有向图 ，节点存在两种状态：激活、非激活；如果节点一旦被激活，则保持激活状态，直到信息传播结束。每个节点会拥有一个独立的激活阈值，激活阈值越大，表示此节点越难以被激活；反之则越容易被激活。当节点受到邻居的激活权重总和（边权和）大于等于其激活阈值时，该节点被激活。可以用来分析社交网络的传播影响。**独立级联模型**：设置类似线性阈值模型，但结点无权。每个节点对于其每一条入边具有独立的激活概率，激活概率越大，表示此节点越容易被该边激活。

当一个节点 i 在第 t 步被激活后，其只能在第 t+1 步去激活它的邻居节点；每个节点只有一次机会去激活邻居节点。重复执行，直至再没有新的节点被激活。**传染病模型**：**SI模型**：



分为易感节点 (S) 、传播节点 (I，是信息的源头 ，通过一定的概率 把信息传播给易感节点，且无法停止传播) 两类 。每个节点对于其每一条入边具有独立的感染概率。其中 , 表示初始状态下传播节点在总人群中的比例。

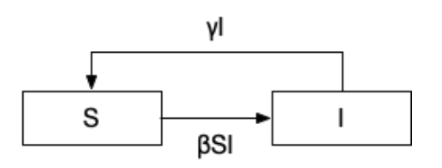
##### SIR模型



分为易感节点 (S) 、传播节点 (I)、免疫节点 (R)。I 通过比例系数 变为R , R不可再进行信息传播。

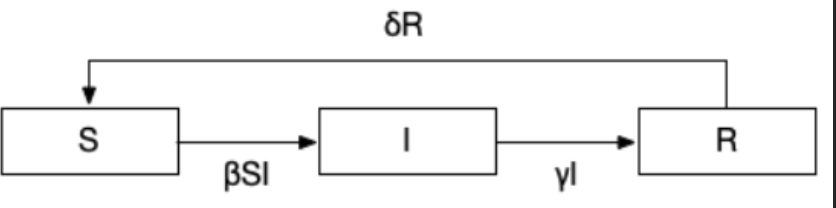
,

##### SIS模型



分为易感节点 (S) 、传播节点 (I)。I以一定的概率被治愈 ，变为S。

##### SIRS模型



分为易感节点 (S) 、传播节点 (I)、免疫节点 (R)。I 通过比例系数 变为R , R不可再进行信息传播。但在单位时间内将有暂时免疫节点再次变为S。

##### SEIR模型

A picture containing font, screenshot, line, number

Description automatically generated

分为易感节点 (S) 、传播节点 (I)、免疫节点 (R)和潜伏节点 (E)。S 在接触传播节点 I之后以一定概率 变为E， E在接触S 时，不可以将S变为I或E。接着E以 一定概率变为I。这时候的I可以将S变为E。I会以一定概率 变为R。 R具有永久的免疫力。