智能无人系统控制实践

第十章 集群协同与博弈对抗

作者：戴训华 中南大学

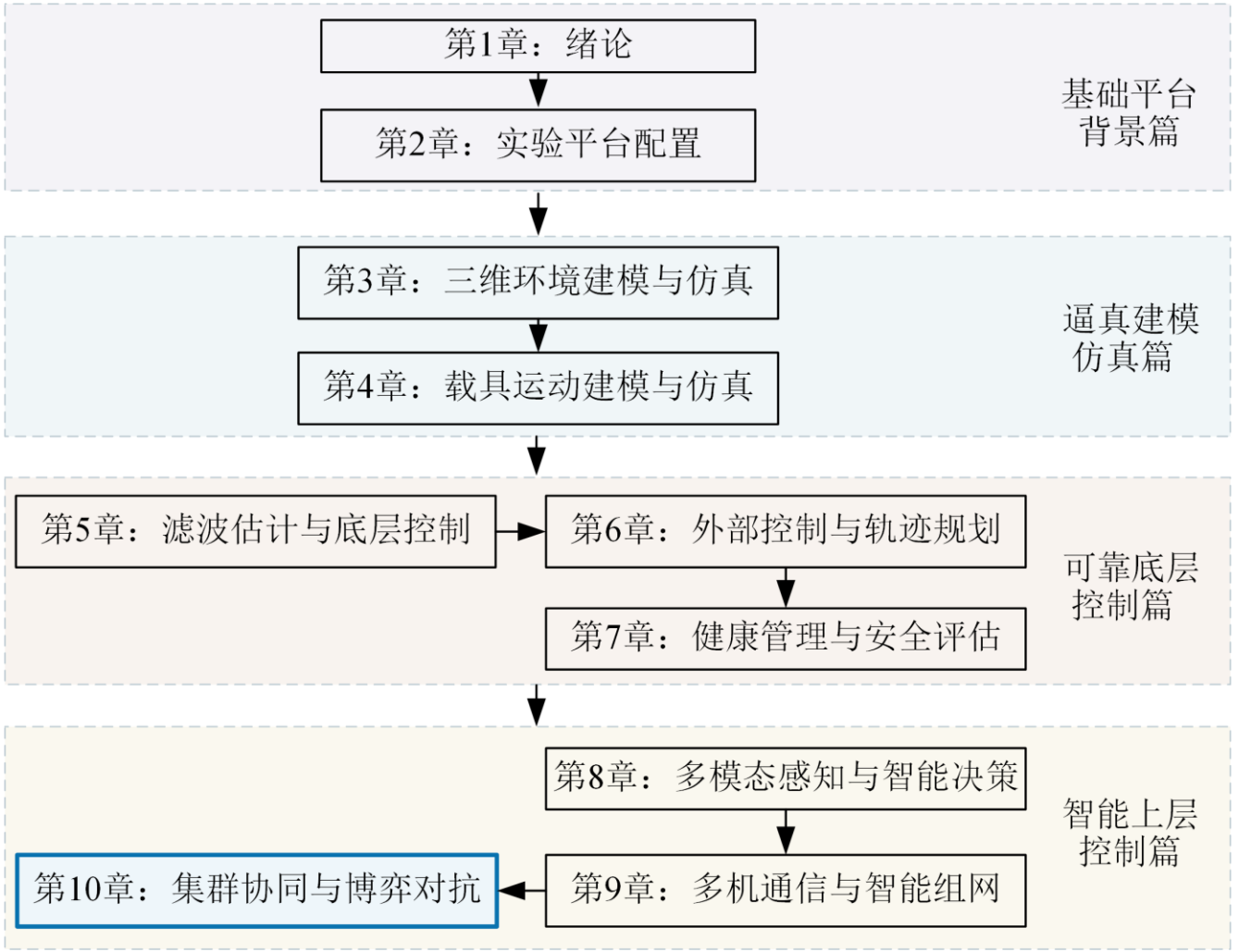
主页：https://faculty.csu.edu.cn/daixunhua

邮箱：dai.xh@csu.edu.cn

平台：https://rflysim.com

配套教材





教材：戴训华，全权，任雪峰. 智能无人系统控制实践[M]. 北京：电子工业出版社，2026.

本章学习目标

学完第十章能够完成什么

1. 解释集群的信息、决策与执行关系，选择控制架构
2. 计算一致性与编队指令，说明收敛条件和安全边界
3. 建立序贯决策模型，比较单智能体与多智能体学习
4. 连接RflySim接口，以误差、距离和时效评价实验

本章目录

集群协同基础与控制架构

第十章 集群协同与博弈对抗

10.1 集群协同基础与控制架构

10.2 一致性与编队控制

10.3 智能博弈与强化学习基础

10.4 单智能体强化学习经典算法

10.5 多智能体强化学习与规模扩展

10.6 集群关键技术与挑战

10.7 RflySim集群框架与接口

10.8 本章实验案例与综合验证

10.9 习题与课堂讨论

本节 3 个教学单元

10.1.1 集群协同的背景与任务

教材对应：第10.1.1节

10.1.2 集群控制架构

教材对应：第10.1.2.1节

10.1.3 集群控制任务分解

教材对应：第10.1.2.2节

10.1.1 集群协同的背景与任务：巡检任务

三架载具同向运动，还需要哪些协同？

共同任务：沿走廊巡检；几何要求：相邻间距4 m；安全要求：全程不碰撞。



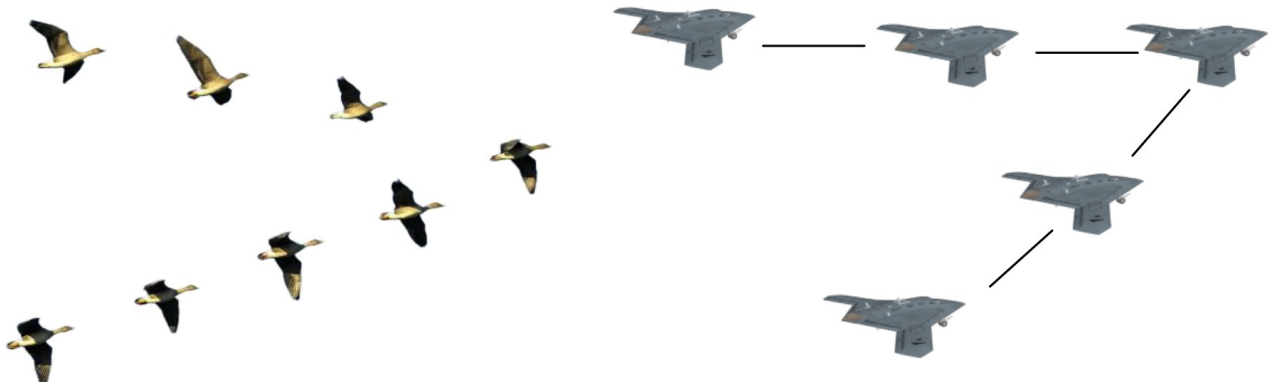
单机控制：各机根据自身测量跟踪速度或位置指令。

集群协同：根据共同任务与其他成员的信息，协调分工、运动和异常处置。

同飞只描述现象；协同还需要明确成员之间的约束与分工。

10.1.1 集群协同的背景与任务：基本概念

从生物集群到多智能体无人系统



(a) 鸟群编队　　(b) 无人机编队

图10.1 生物集群到无人集群

集群：多个个体通过交互形成整体行为的群体系统。

多智能体系统：由多个能感知、决策和行动，并相互合作或竞争的主体组成。

留迹机制：经环境留下的信息间接协调，不必直接发消息。

10.1.1 集群协同的背景与任务：核心特性

集群能力来自局部交互与系统组织

系统特性	巡检中的具体动作	需要检查的结果
自组织	邻居交换状态，形成指定相对关系	队形误差是否减小
自适应	风扰或障碍出现后调整路线与速度	扰动后是否恢复任务
容错与重构	一机退出，剩余成员重新分配区域	覆盖缺口和恢复时间
可扩展性	增加节点时限制通信与计算开销	时延、负载与控制误差

10.1.1 集群协同的背景与任务：发展脉络

研究主线从局部规则走向系统集成

阶段	主要问题	留下的基础
1990年代至2000年代初	群体行为怎样产生；局部规则怎样设计	概念、仿生算法与理论模型
2000年代中期至2010年代	怎样在平台上实现规模与异构协作	多机实验与应用验证
2010年代以来	学习、微型化、人机协同怎样集成	迁移、鲁棒性与韧性评价

共同难题持续存在：模型是否合适，信息是否及时，约束是否始终满足。

10.1.1 集群协同的背景与任务：协同方向

集群协同研究共同目标下怎样配合

研究方向	巡检问题	可观察量
分布式控制与估计	局部信息如何形成共同判断	状态差、估计误差
队形与运动协调	怎样保持间距并通过转弯	队形误差、最小距离
任务与行为协同	谁负责哪一区域，何时交接	覆盖率、冲突、耗时
通信与感知共享	怎样补充盲区与陈旧数据	数据龄期、可见范围
鲁棒与自适应	节点退出后怎样恢复协作	重构时间、剩余能力

10.1.1 集群协同的背景与任务：策略互动

博弈对抗把其他决策者纳入决策模型

方向	要回答的问题
博弈建模	参与者、信息、动作和收益怎样定义
对手建模与意图推理	对方可能采用什么行为，推断有多可靠
策略规划	资源有限时怎样选择行动与优先级
强化学习与策略演化	怎样从交互数据改善长期决策
感知到决策一体化	观测不完整时怎样依据反馈调整行动

10.1.1 集群协同的背景与任务：联合选择

同一个“前进”，为何会有不同后果？

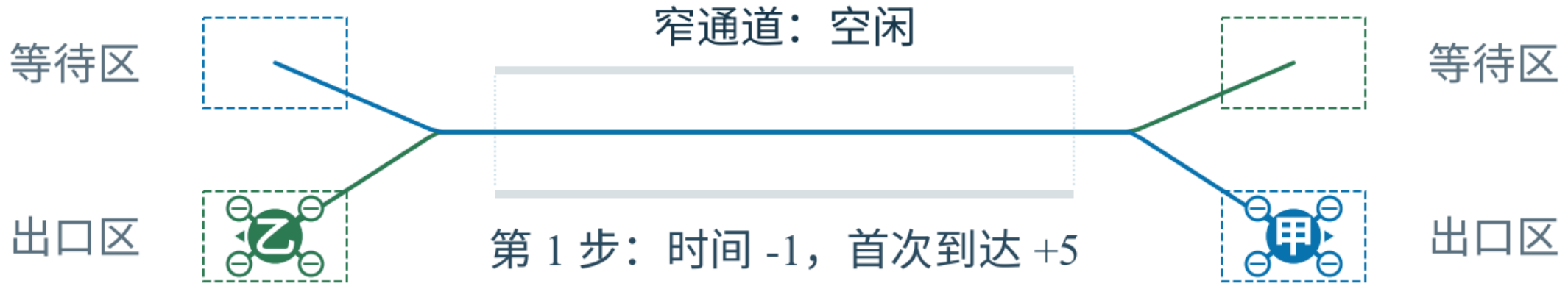


图：同一窄通道中，甲乙的联合选择决定能否安全通行

10.1.1 集群协同的背景与任务：交互经验

一次行动之后，哪些反馈会成为学习经验？

两步回顾：甲先等待 -1，再通行 +4



局部观测	实际动作	奖励	下一观测	是否结束
通道空，乙退出 甲未到达	通行 第 1 步	+4 -1 + 5	通道空，乙退出 甲已到达	是 true 甲已完成

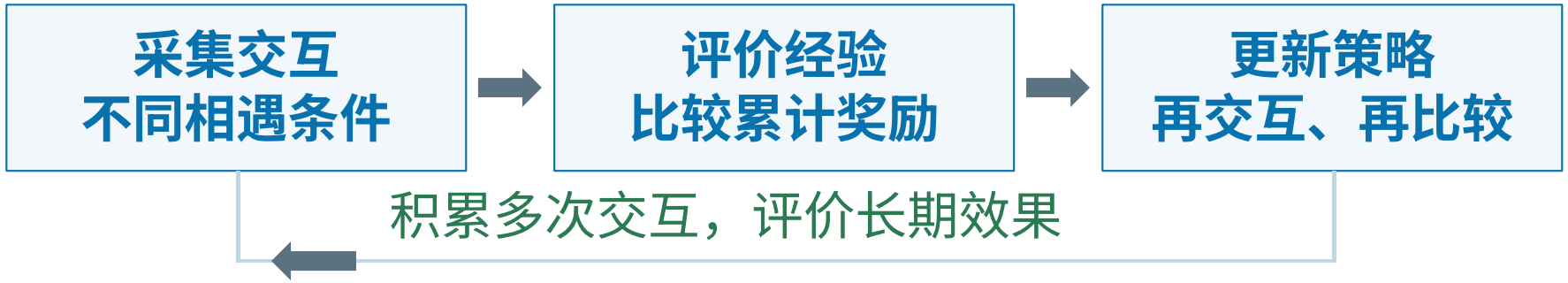
图：记录甲的观测、动作、奖励、下一观测及任务是否结束

10.1.1 集群协同的背景与任务：训练与执行

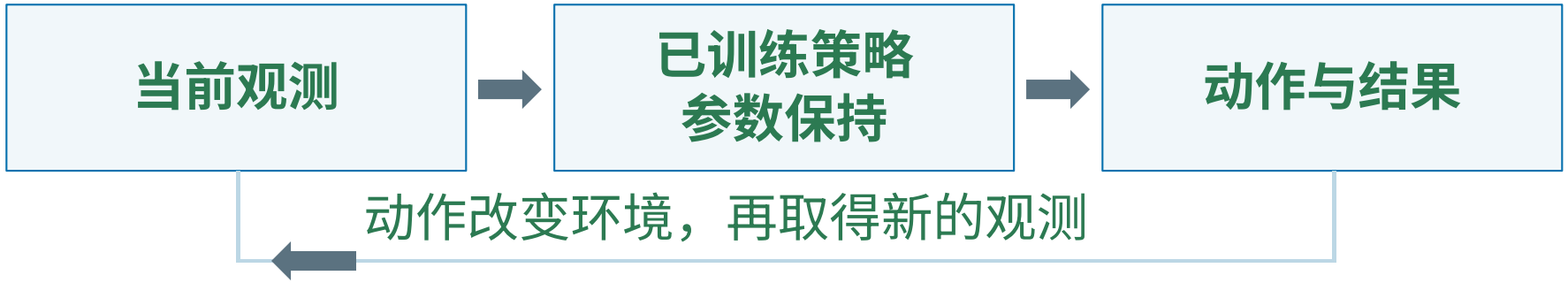
反馈修正当下动作，学习改进选择动作的策略

策略：依据当前观测（必要时结合历史）选择动作的规则。

训练阶段



固定策略执行



同一套策略参数，也能根据当前情境选择等待或通行。

10.1.1 集群协同的背景与任务：信息可得性

训练时看得更全，执行时仍要用拿得到的信息

训练：可以汇总多机资料

甲乙的状态与动作
两机奖励、下一状态及结束标志

联合评价，改进各自策略

全局资料可用于评价与更新；
各机策略仍按执行接口取输入。

执行：甲依据实际输入行动

本机观测：位置、通道占用
邻居消息：乙的位置、时刻与有效性

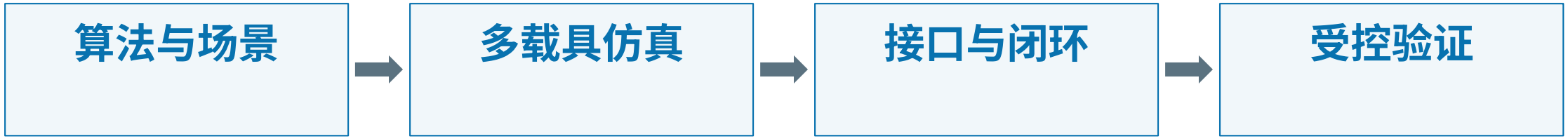
甲的策略 → 等待或通行

消息缺失或过期，要按设计处理；
不能用取不到的全局真值补空。

工具分工：环境产生交互，训练程序更新并保存策略，执行程序加载后使用。

10.1.1 集群协同的背景与任务：开发工具链

仿真让复杂任务具备可重复的检查条件



可重复：固定初值、任务和随机扰动，比较方案差异。

分层增加真实因素：运动模型 → 飞控跟踪 → 网络与感知误差 → 实物约束。

结果判据：同时记录任务完成、队形误差、安全距离和通信时效。

本章目录

集群协同基础与控制架构

第十章 集群协同与博弈对抗

10.1 集群协同基础与控制架构

10.2 一致性与编队控制

10.3 智能博弈与强化学习基础

10.4 单智能体强化学习经典算法

10.5 多智能体强化学习与规模扩展

10.6 集群关键技术与挑战

10.7 RflySim集群框架与接口

10.8 本章实验案例与综合验证

10.9 习题与课堂讨论

本节 3 个教学单元

10.1.1 集群协同的背景与任务

教材对应：第10.1.1节

10.1.2 集群控制架构

教材对应：第10.1.2.1节

10.1.3 集群控制任务分解

教材对应：第10.1.2.2节

10.1.2 集群控制架构：架构全景

三种架构按信息流与指令组织方式区分

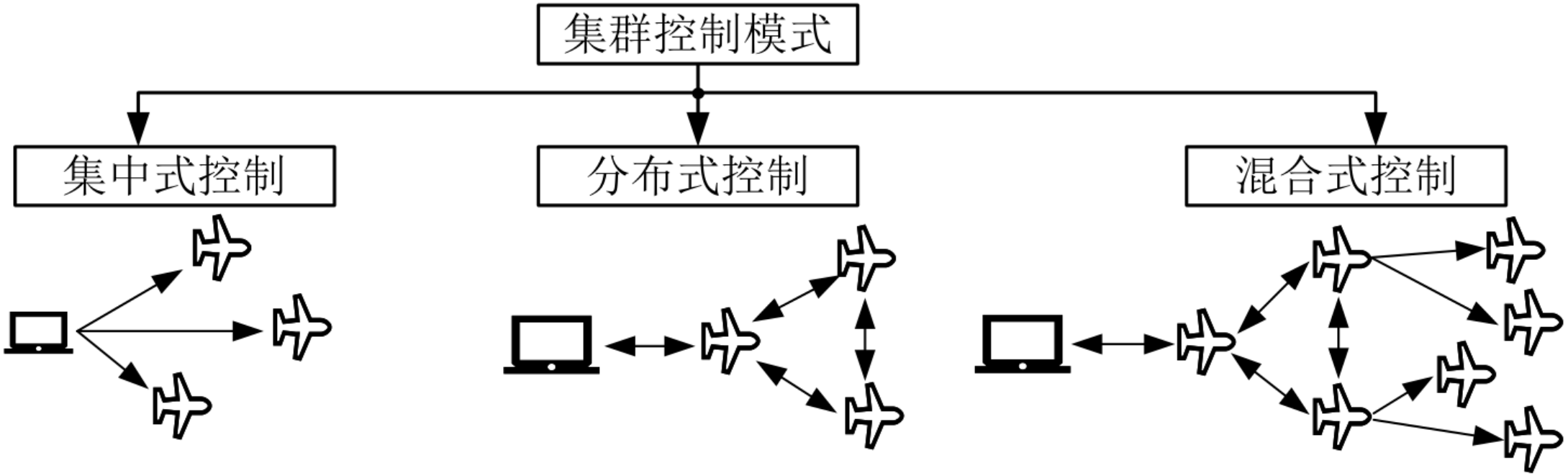


图10.2 集群系统控制架构

判断顺序：谁观测状态 → 谁计算协同指令 → 谁跟踪并执行指令。

10.1.2 集群控制架构：集中式信息流

集中式：各机观测，中心统一计算

甲、乙、丙的位置与速度　　共同目标与三机状态　　三份运动指令分别下发

各机传感与估计 → **中心协同计算** → **各机跟踪与执行**

中心已知：全局任务、已收到的各机状态及约束。

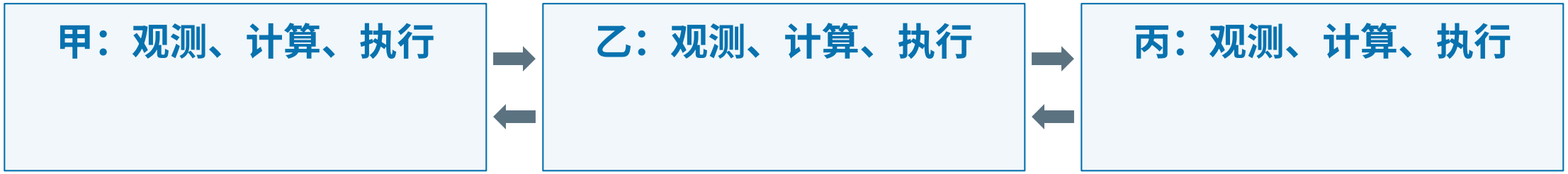
本地仍负责：状态估计、快速跟踪控制、执行器动作与健康反馈。

全局信息便于统一优化，也把通信与计算需求集中到中心。

10.1.2 集群控制架构：分布式信息流

分布式：每机利用允许获得的信息计算

共同目标可预先给定；邻居状态连同时间戳与有效标记参与本轮计算。



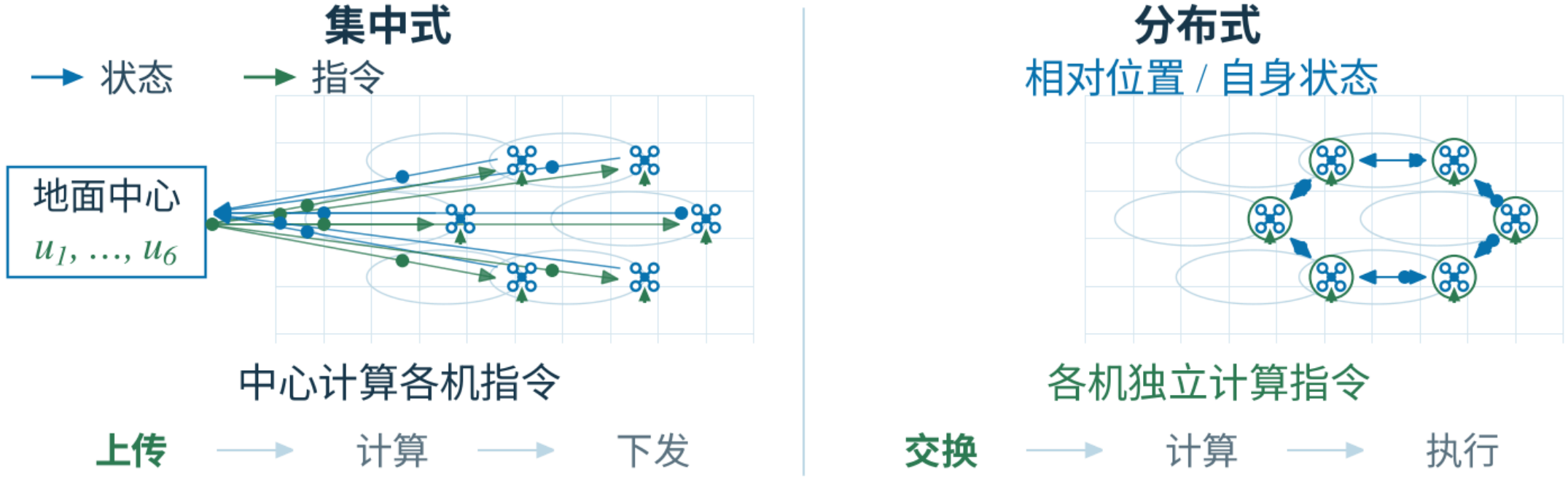
邻居：乙　　邻居：甲、丙　　邻居：乙

乙的消息过期，甲就不能把它当作当前状态；按设计预测、保持或降级，并重新检查控制条件。

10.1.2 集群控制架构：同任务对照

同一编队任务中的集中与分布式计算

共同参考轨迹已预先给定；两侧采用相同六机初始状态与队形目标。



图：同任务、不同计算组织；各机仍负责本地观测与执行。

10.1.2 集群控制架构：同任务比较

相同控制律可在不同计算位置产生相同指令

无量纲状态取0、4、8；每机按“邻居减自身”求和。

节点	本轮需要的数据	邻居差之和
甲	自身0，乙4	$4-0=4$
乙	自身4，甲0，丙8	$(0-4)+(8-4)=0$
丙	自身8，乙4	$4-8=-4$

集中计算：中心用全部状态生成三份输出。

分布计算：每个节点只生成自己的输出，持续交换邻域状态。

10.1.2 集群控制架构：混合式架构

混合式：全局分工与局部闭环分层协作

全局调度：区域、角色与约束

任务下发 → 子群甲乙

任务下发 → 成员丙

局部协同：邻域状态与队形调整

局部协同：邻域状态与队形调整

指令下发；状态向上反馈

指令下发；状态向上反馈

本地闭环：跟踪控制与执行

本地闭环：跟踪控制与执行

执行状态与异常逐层反馈，必要时重新分配任务。

10.1.2 集群控制架构：三个维度

控制架构、计算部署与模型精度分别描述什么

维度	要问的问题	可区分的例子
控制架构	协同指令依赖哪些信息、由谁计算	集中式、分布式、混合式
计算部署	程序运行在哪些计算设备上	单机多进程、多电脑分担
模型精度	每个载具的运动由什么模型产生	飞控动力学闭环、质点模型

判定架构先追踪控制所需信息，再查看程序运行位置。

10.1.2 集群控制架构：取舍与边界

架构选择要同时考虑信息、负载和失效

组织方式	主要收益	需要处理的约束
集中式	全局任务易统一安排	中心故障、汇总带宽与时延
分布式	局部通信与自治计算	连通性、收敛速度与信息不完整
混合式	全局分工与局部响应结合	层间接口、冲突仲裁与恢复策略
预置轨迹独立执行	按统一时序执行既定任务	偏差、突发障碍与在线重规划

本章目录

集群协同基础与控制架构

第十章 集群协同与博弈对抗

10.1 集群协同基础与控制架构

10.2 一致性与编队控制

10.3 智能博弈与强化学习基础

10.4 单智能体强化学习经典算法

10.5 多智能体强化学习与规模扩展

10.6 集群关键技术与挑战

10.7 RflySim集群框架与接口

10.8 本章实验案例与综合验证

10.9 习题与课堂讨论

本节 3 个教学单元

10.1.1 集群协同的背景与任务

教材对应：第10.1.1节

10.1.2 集群控制架构

教材对应：第10.1.2.1节

10.1.3 集群控制任务分解

教材对应：第10.1.2.2节

10.1.3 集群控制任务分解：任务到指令

从共同任务到各机动作需要四个层次

任务分配 → 轨迹规划 → 编队协调 → 跟踪执行

层次	输入	输出
任务分配	区域、能力、剩余资源	每机任务与优先级
轨迹规划	目标、障碍与运动约束	带时间的位置或速度参考
编队协调	参考与自身、邻居状态	修正后的各机运动指令
跟踪执行	运动指令与本机状态	控制量与实际运动反馈

10.1.3 集群控制任务分解：任务分工

轨迹规划与编队保持约束不同的对象

问题	约束对象	三机巡检中的判据
轨迹规划	每机随时间的参考运动	轨迹可行、区域可达、障碍不相交
编队保持	成员之间的相对几何与速度	相对位置误差小、速度协调
碰撞避免	全程机间距离与障碍距离	最小距离大于规定安全值



同时沿直线互换目标，终点间距合格，途中仍会相遇。

10.1.3 集群控制任务分解：简化计算

固定目标算例：把位置误差变为速度指令

教学假设：一维理想速度响应；无扰动、无时延；目标固定；增益0.5 s^{-1}。
本例目标间距2 m；每0.5 s重算速度并保持，向右为正。
i为节点号；p_i为位置，p_i^{des}为目标，v_i^*为速度指令，p_i^+为下一步位置。

$$v_i^* = 0.5(p_i^{\mathrm{des}} - p_i), \qquad p_i^+ = p_i + 0.5v_i^*$$

节点	当前位置 / m	目标 / m	速度 / (m/s)	0.5 s后 / m
甲	0	2	1	0.5
乙	4	4	0	4
丙	8	6	−1	7.5

10.1.3 集群控制任务分解：固定目标响应

完整结果：误差收敛与安全需要分别评价

$$p_i[k] = p_i^{\mathrm{des}} + (p_i[0] - p_i^{\mathrm{des}})(0.75)^k, \quad t_k = 0.5k\,\mathrm{s}$$

k为更新次数；$p_i[0]$为初始位置；t_k为经过时间。

时刻 / s	甲位置 / m	乙位置 / m	丙位置 / m	相邻距离 / m
0	0	4	8	4
2	1.367	4	6.633	2.633
4	1.800	4	6.200	2.200
10	1.994	4	6.006	2.006

本算例相邻距离由4 m趋近2 m，全程不小于2 m。

10.1.3 集群控制任务分解：验收判据

同一任务必须同时报告多个结果指标

评价对象	记录内容	能支持的判断
共同任务	覆盖率、漏检区域、完成时间	是否完成所分配的任务
运动与队形	位置跟踪误差、相对偏移误差	是否跟随参考并维持结构
全程安全	全时段最小机间距离、障碍距离	是否满足规定的间隔约束
信息与异常	数据龄期、掉线和恢复时间	结果是否依赖理想信息条件

10.1.3 集群控制任务分解：进入一致性

邻居误差怎样逐步影响整个集群



信息关系：本轮有效的邻域数据决定参与计算的差值项；先检查消息是否收到、是否过期。

控制关系：相对误差决定指令方向，状态改变后继续反馈。

队形关系：扣除各自期望偏移再协调，保留成员之间的空间间隔。

练习与讨论

讨论：怎样判断三机巡检方案是否协同

情境一：三台电脑仿真三机，一个控制程序读取全部状态并统一下发指令。判断架构并说明依据。

情境二：各机按预置轨迹同步飞行，没有交换邻居状态。哪些能力尚不能由此推出？

情境三：末态队形误差很小。请列出另外三类必须记录的结果。

情境四：策略参数不变，却按观测改变等待/通行动作。这还是闭环吗？
若训练时用了乙的全局真值，执行时应检查什么？

参考要点

参考答案：先看信息依赖，再看任务证据

情境一：集中式控制；全局状态集中参与协同指令计算。多电脑描述部署方式。

情境二：可确认预置任务同步执行；在线邻域协同、重构与避碰仍需单独证据。

情境三：任务覆盖与耗时、全时段最小距离、数据龄期及异常恢复；还要查看误差随时间的变化。

情境四：仍是闭环，观测在持续影响动作；但未进行在线参数更新。
训练与执行要约定相同的策略输入，额外全局资料只用于允许的评价与更新。

本节小结

本节小结：让共同任务落到可计算的闭环

概念：多智能体描述多个决策主体；协同强调共同目标下的相互配合。

架构：集中、分布与混合按信息依赖和指令组织判断。

任务分层：任务分配 → 轨迹规划 → 编队协调 → 本地跟踪。

易错点：部署数量、末态图案与算法名称不能代替收敛和安全证据。

选择控制架构之后，下一步是把邻域误差写成有条件的控制律。