本章目录

安全决策与整机改进

第七章 健康管理与安全评估

7.1 健康管理与安全评估基础
7.2 统一故障建模框架
7.3 虚实安全强化与故障数据方法
7.4 故障危害与安全评估
7.5 故障诊断与健康预测
7.6 安全决策与整机改进
7.7 故障注入接口与算法开发框架
7.8 例程实验与虚实验证
7.9 习题与课堂讨论

本节 3 个教学单元

7.6.1 处置与控制取舍
教材对应：第7.1.5.5节

7.6.2 分层决策与训练
教材对应：第7.1.5.5节

7.6.3 整机回归与改进
教材对应：第7.1.5.6节

7.6 安全决策与整机改进

7.6.1 处置与控制取舍：安全决策的输入

诊断结果要与任务和剩余能力一起使用

诊断与状态
故障、位置、电量
→
任务与约束
剩余路程、场地
→
可执行动作
等待、退出、着陆

输入	巡检电机效能下降案例	决策中解决的问题
故障证据	异常电机、效能变化、诊断可信度	是否改任务；是否切换控制
当前状态	高度、速度、定位、电量与链路	保持、移动或下降是否可行
任务场景	返航路径、备用着陆区、周围人员	哪种可行动作的后果更可接受

7.6.1 处置与控制取舍：PX4保护动作

警告、保持、返航与降落改变不同的行为

PX4 v1.12.3 多旋翼（版本差异见课件5.5「控制输出的版本边界」页）

动作	作用机制	典型前提与后果
警告	发送告警，当前任务模式继续运行	适用于允许继续飞行的告警级条件
保持	进入 Hold 模式，控制当前位置	依赖有效定位与足够控制能力
返航	进入 Return 模式，按返航配置移动	依赖可用导航与返航路径/目的地
降落	进入 Land 模式，执行下降与着陆	按可用估计执行；着陆区决定后果

7.6.1 处置与控制取舍：电量触发条件

同一电量条件按配置落到不同保护级别

教学阈值：15%、10%、5%；COM_LOW_BAT_ACT=3（临界返航、紧急降落）。

阈值参数	本例进入的电量区间	保护结果
BAT_LOW_THR	12%：低于15%，尚高于10%	警告；继续当前模式
BAT_CRIT_THR	8%：低于10%，尚高于5%	返航；导航无效则降落
BAT_EMERGEN_THR	4%：低于5%	紧急降落； 不等待保持延迟

百分数为剩余容量比例；v1.12.3 默认值为 15%、7%、5%，本例为教学设定。

7.6.1 处置与控制取舍：链路丢失时序

超时检测与动作延迟是两段不同的等待

v1.12.3示例：COM_RC_LOSS_T=0.5 s、COM_RCL_ACT_T=3 s，配置返航。

6.0 s
最后一次遥控输入 → **6.5 s**
确认丢失，进入保持 → **9.5 s**
延迟结束，请求返航

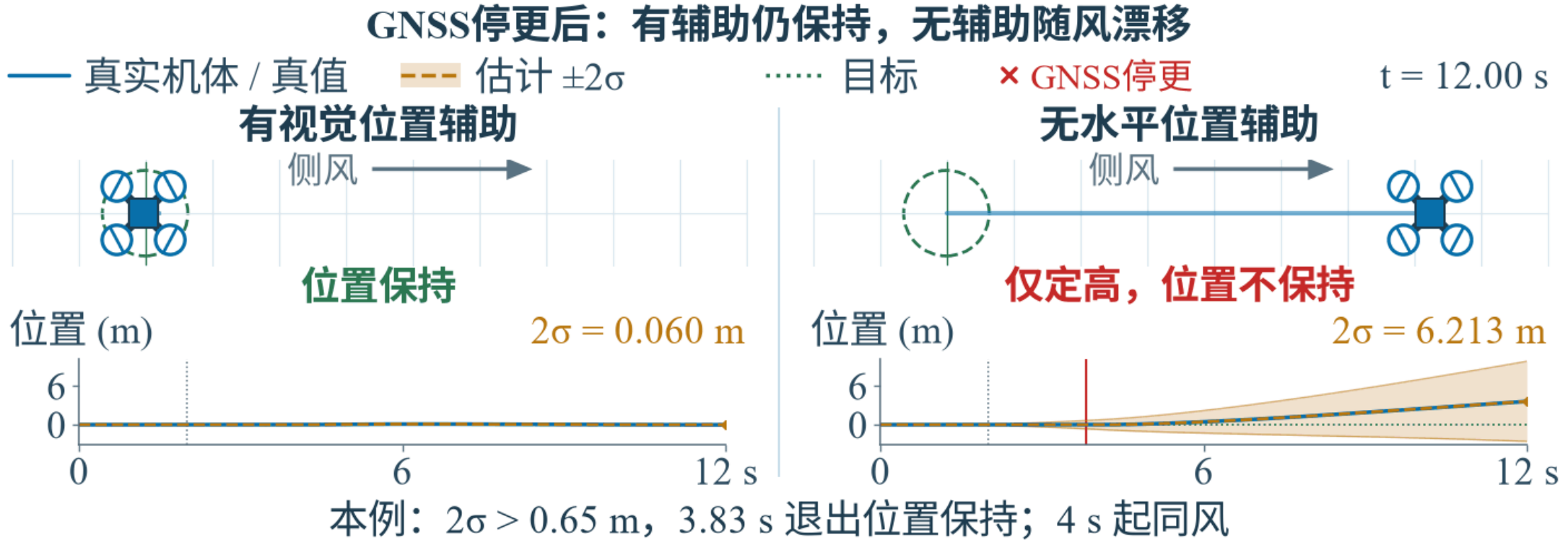
两段时间：0.5 s 用于确认输入丢失；3 s 原地保持，等待遥控信号恢复。

动作竞争：保持期间若同时触发紧急电量降落，执行更严重的降落动作。

PX4 v1.12保护条件：https://docs.px4.io/v1.12/en/config/safety.html

7.6.1 处置与控制取舍：剩余定位能力

GNSS失锁后，还能不能稳在原地？

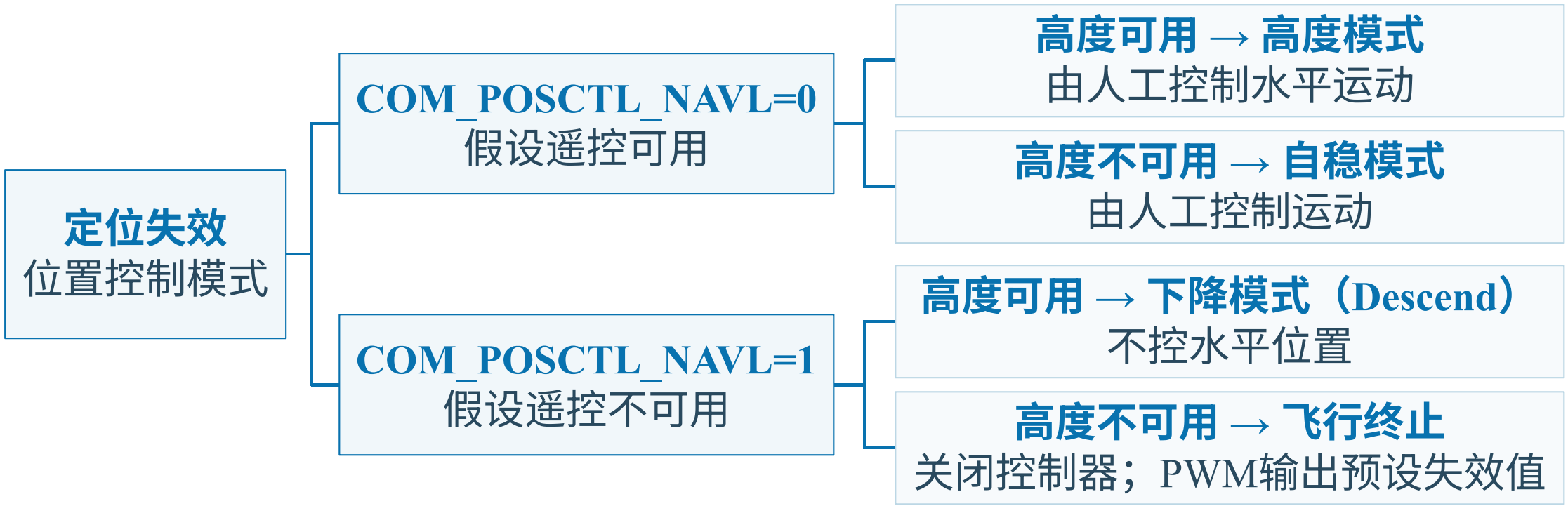


图：相同GNSS（卫星导航）中断和侧风下的水平运动（MuJoCo 仿真）
右侧假设遥控可用、高度有效、无人工修正；定位失效后退到定高。

7.6.1 处置与控制取舍：模式的能力前提

定位丢失后，保持和返航可能失去依据

v1.12.3位置控制模式中定位失效：先看COM_POSCTL_NAVL，再看高度估计。
树叶是固件请求的模式；切换后能否控住运动，还要另外检查。



7.6 安全决策与整机改进

7.6.1 处置与控制取舍：主流安全决策方法

规则组织条件，学习优化行为选择，混合方法分配执行权

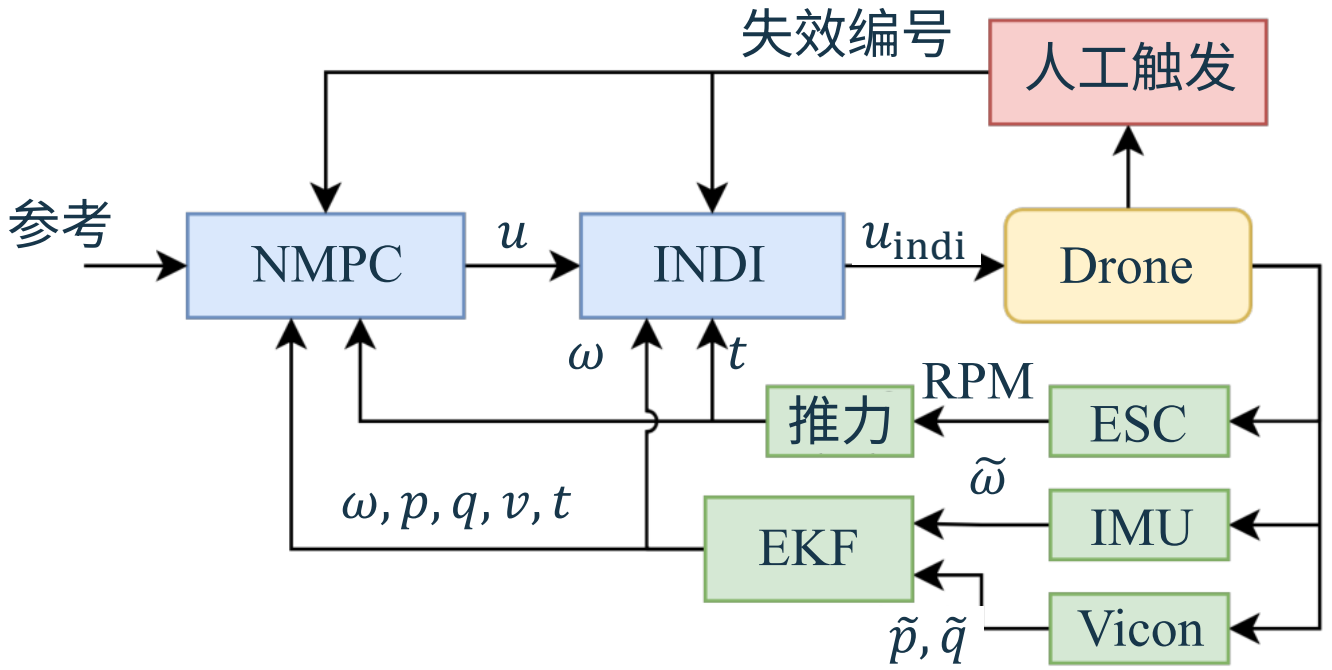
方法	工作方式	主要取舍
规则	状态机：条件触发模式转换； 行为树：按优先级检查可行分支	容易检查； 规则覆盖依赖设计
学习	从观测和任务状态选择行为； 用动作后的结果更新策略	适合复杂状态； 样本效率与可解释性差
混合	规则先限定可执行动作； 学习策略再优化性能或任务结果	约束更明确； 接口与回退要一致

本章主线：先把触发条件与执行前提写清，再优化不同场景中的行为选择。

7.6 安全决策与整机改进

7.6.1 处置与控制取舍：已知故障的容错控制

单旋翼失效后放弃偏航，优先保持推力方向与位置



NMPC 非线性模型预测控制
INDI 增量非线性动态逆

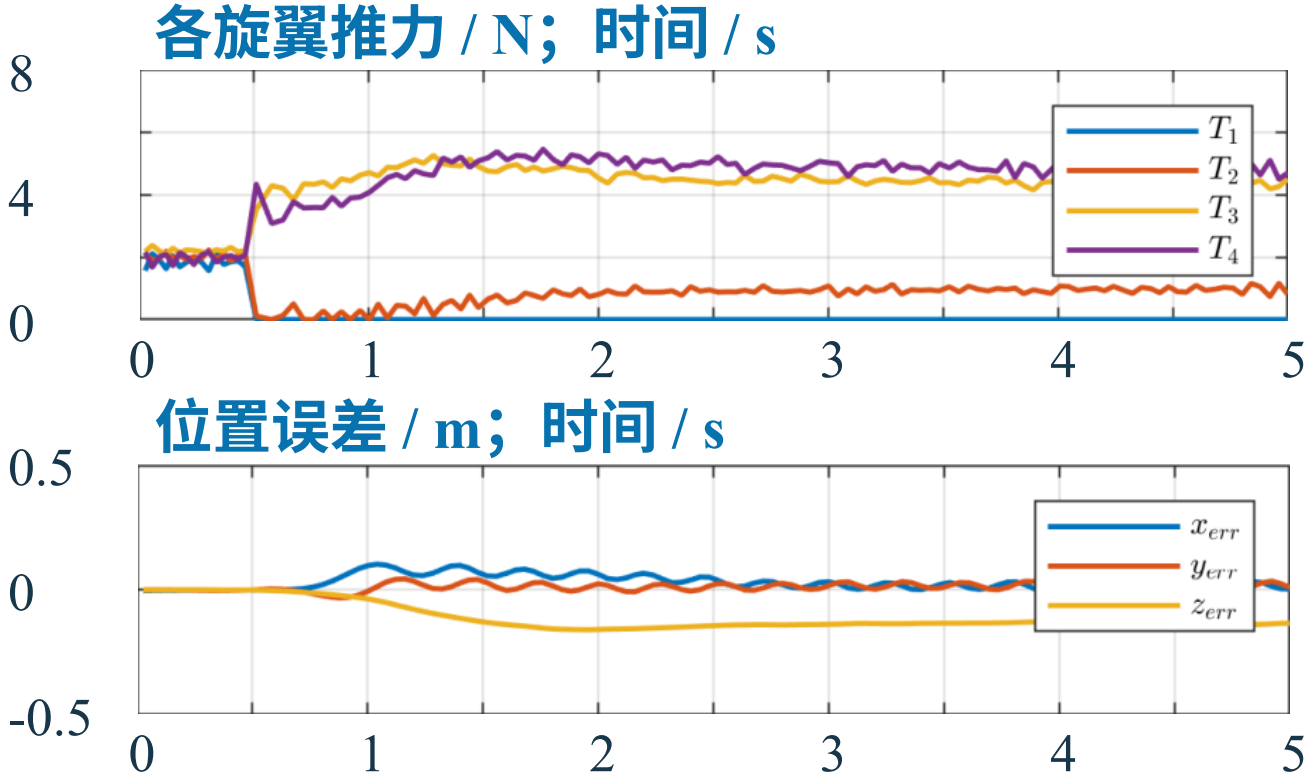
EKF 扩展卡尔曼滤波
IMU 惯性测量单元
VICON 动作捕捉系统
ESC 电子调速器

原因： 3个推力无法同时控制总推力与三轴力矩。

u, u_{indi}：NMPC及INDI的旋翼推力指令；t：由转速估计的实际推力（N）。
p, v：惯性系位置/速度（m、m/s）；q：姿态四元数；ω：机体系角速度（rad/s）。
Nan等，IEEE RA-L，2022，Fig.7；DOI：https://doi.org/10.1109/LRA.2022.3154033

7.6.1 处置与控制取舍：容错控制的试验结果

失效旋翼推力归零，位置仍受到控制



T_1 T_2 T_3 T_4

0.5 s 后第一旋翼推力归零，其他旋翼重新分配推力。

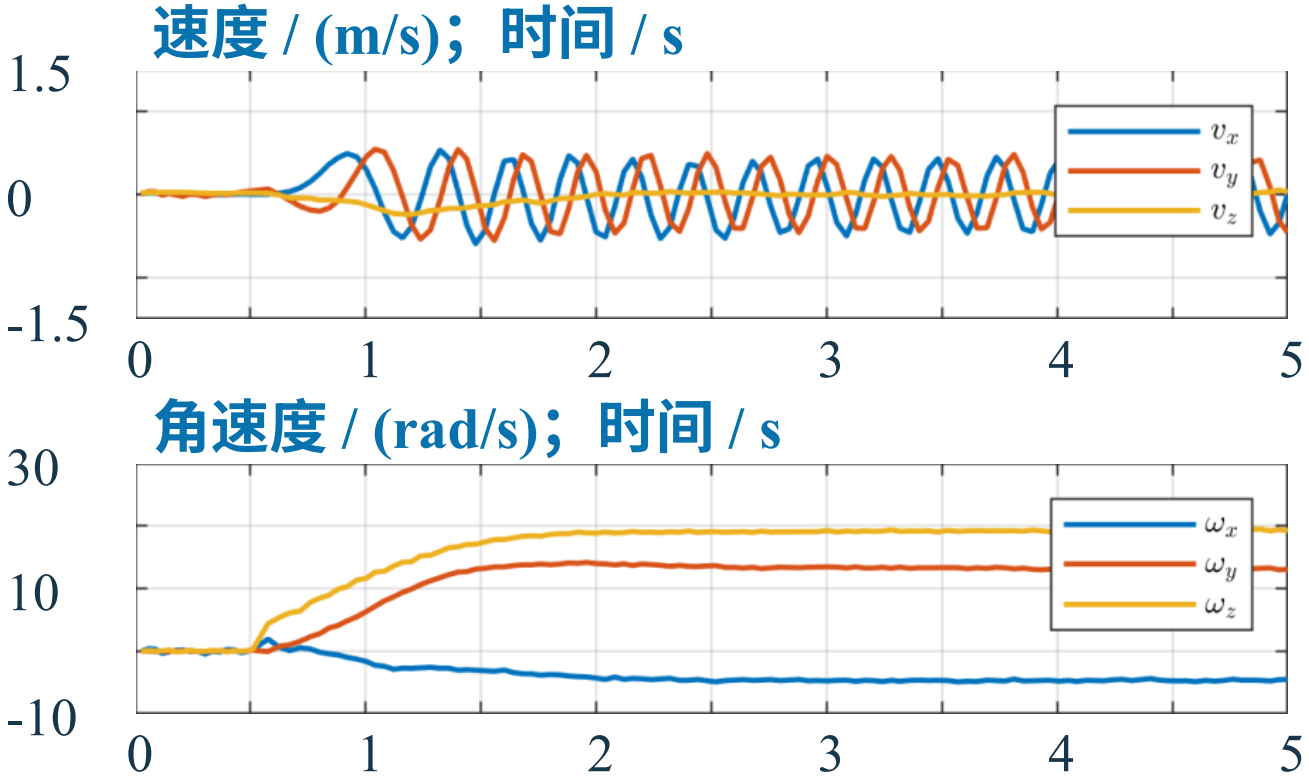
x_{err} y_{err} z_{err}

横向误差接近零；垂向误差仍有偏置。

图：同上论文Fig.8前两组曲线（旋翼推力与位置误差）

7.6.1 处置与控制取舍：剩余控制能力的边界

保住位置的代价是持续旋转与速度振荡



v_x v_y v_z

位置近零仍伴随速度分量的振荡。

ω_x ω_y ω_z

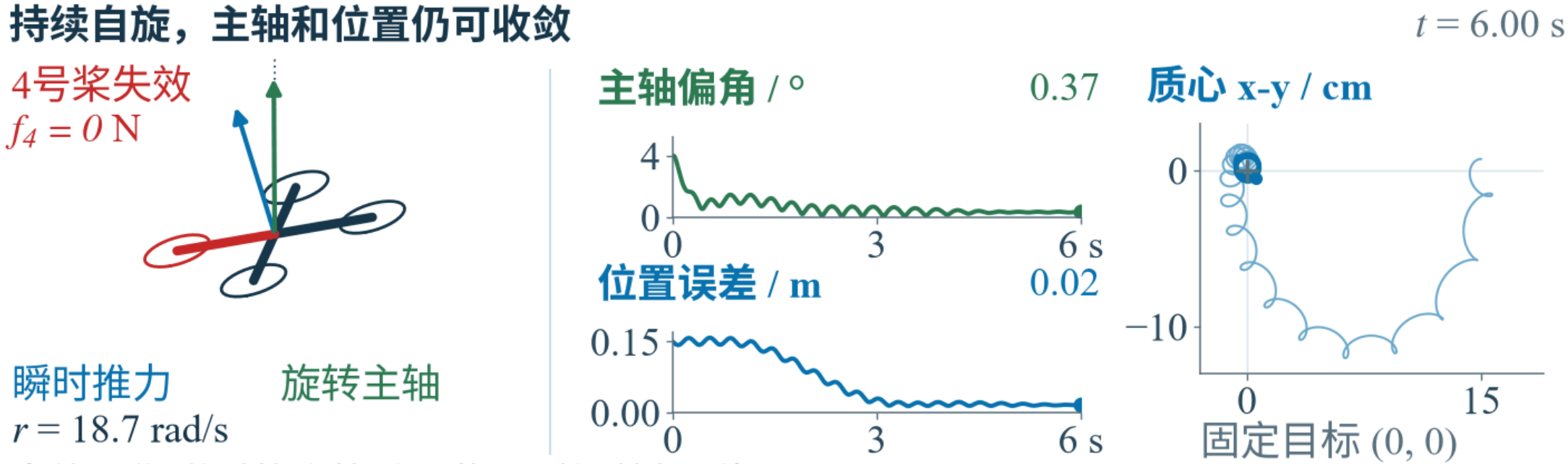
z轴角速度约20 rad/s；机体持续旋转。

图：同上论文Fig.8后两组曲线：同一失效试验的速度与旋转状态

7.6.1 处置与控制取舍：受控自旋

放宽悬停：持续自旋与主轴收敛

自旋平衡附近的小扰动数值算例；第4桨从初值起失效，不含切换过程。



图：Mueller与D’ Andrea动力学数值算例；非前页Nan等试验复现。

本章目录

安全决策与整机改进

第七章 健康管理与安全评估

7.1 健康管理与安全评估基础
7.2 统一故障建模框架
7.3 虚实安全强化与故障数据方法
7.4 故障危害与安全评估
7.5 故障诊断与健康预测
7.6 安全决策与整机改进
7.7 故障注入接口与算法开发框架
7.8 例程实验与虚实验证
7.9 习题与课堂讨论

本节 3 个教学单元

7.6.1 处置与控制取舍
教材对应：第7.1.5.5节

7.6.2 分层决策与训练
教材对应：第7.1.5.5节

7.6.3 整机回归与改进
教材对应：第7.1.5.6节

7.6.2 分层决策与训练：决策行为层与执行层

上层决定行为，底层落实动作并响应紧急故障

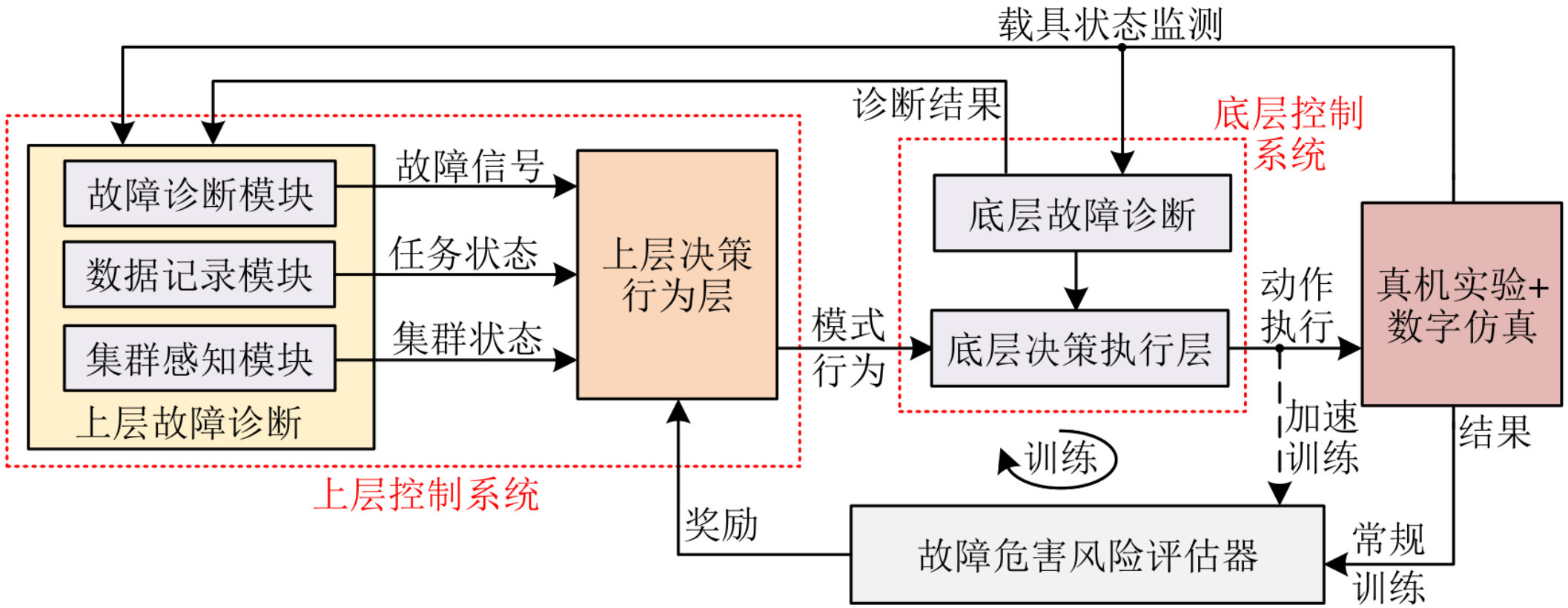


图7.11 安全决策执行层训练框架（随后分层拆读）

7.6.2 分层决策与训练：上层安全决策

上层融合故障、任务与集群状态，输出模式行为

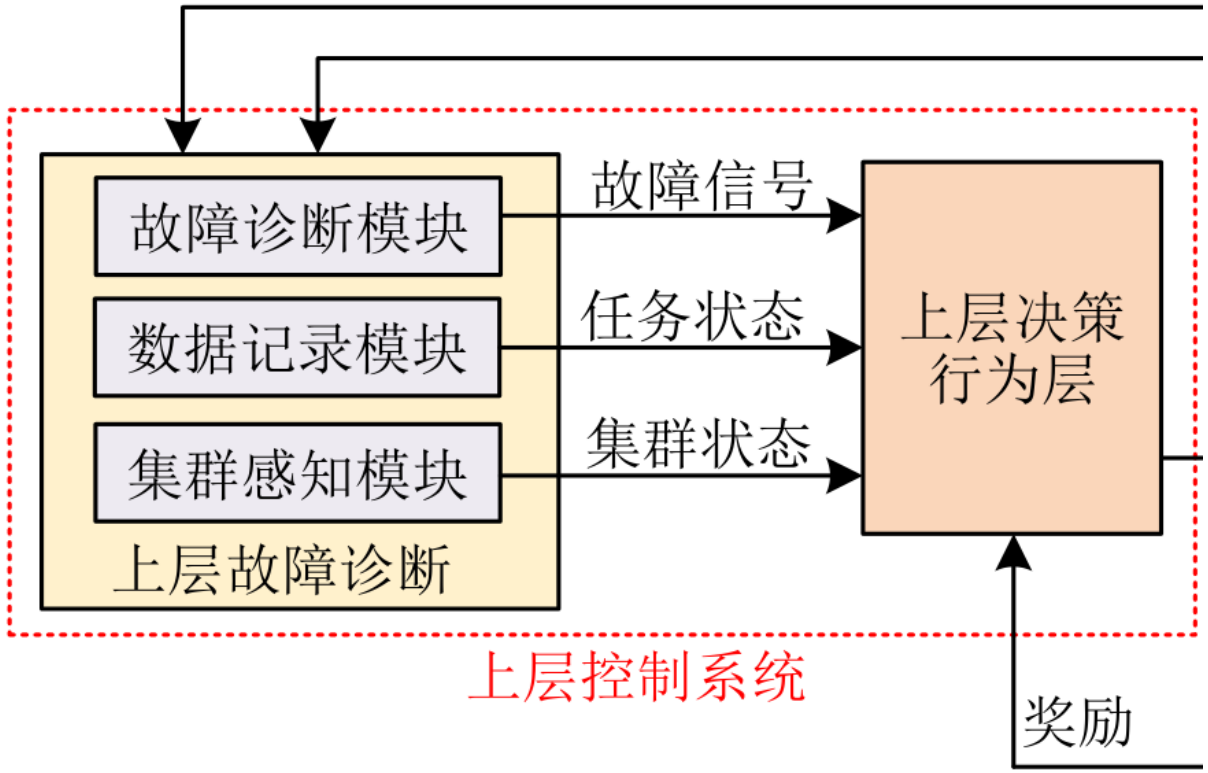


图7.11（上层）多类信息进入行为决策

输入融合：诊断说明异常，任务说明剩余目标，集群状态说明协同关系。

行为输出：返航、悬停、退出任务、等待接管或降落。

结果反馈：检查任务终态、危害变化与底层执行模式。

7.6.2 分层决策与训练：底层决策与风险反馈

底层保护直接改变执行，评估器再反馈结果

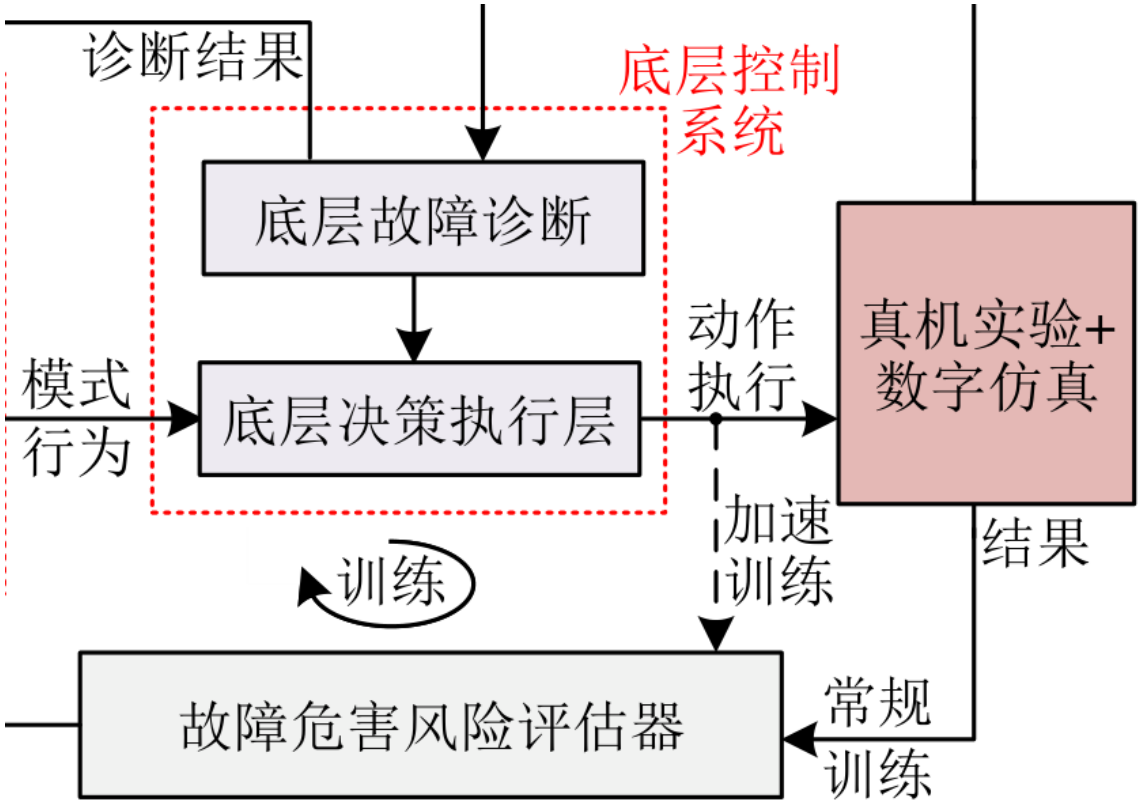


图7.11（右侧）执行、运行结果与离线预评估

底层规则：状态机快速切换保护；训练靠半实物或真机，难加速。

常规评价：运行后记录实际状态、模式和任务终态。

离线预评估：历史评估器先筛选动作；交互验证再检查模型误差。

7.6.2 分层决策与训练：分层响应时间算例

紧急故障可以绕开上层往返等待

教学设定：检测 0.15 s；上行 0.04 s；上层计算 0.05 s；下行 0.06 s；执行 0.01 s。

$$T_{\text{upper}} = 0.15 + 0.04 + 0.05 + 0.06 + 0.01 = 0.31\,\text{s}$$

$$T_{\text{local}} = 0.15 + 0.01 = 0.16\,\text{s}$$

两种路径都以故障可检测、保护可执行为前提；实际延迟由事件日志核对。

把紧急响应留在本地，可减少通信与上层计算带来的等待。

7.6.2 分层决策与训练：安全决策并行训练

并行节点扩展工况，反馈汇总推动统一更新

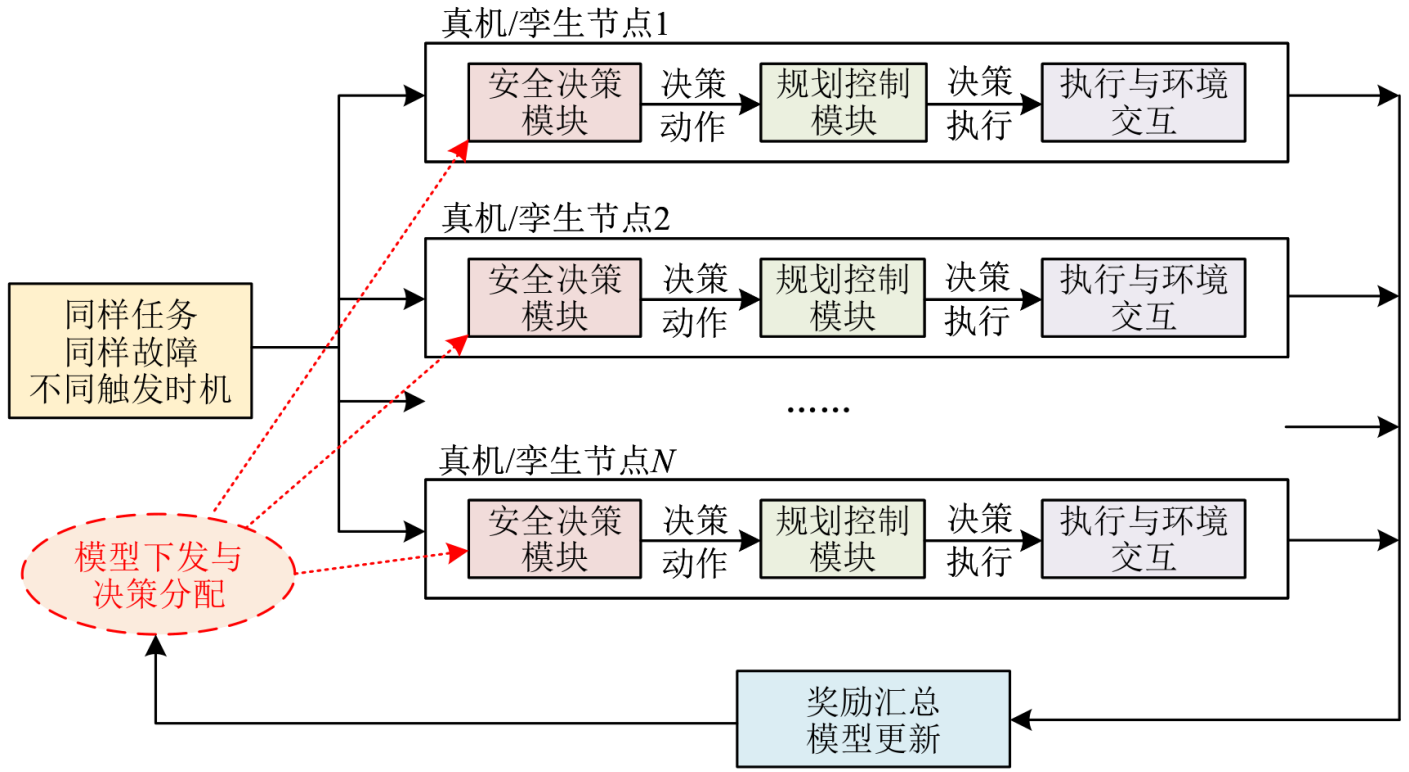


图7.12 安全决策算法并行训练框架

控制变量：
同任务、同故障。
改变触发时机或状态。

每个节点：选择行为，执行控制，观察环境反馈。

统一更新：汇总结果后更新模型，再下发同一版本。

7.6.2 分层决策与训练：并行节点内部机制

一个节点中的“决策、控制、交互”不可合并为一步

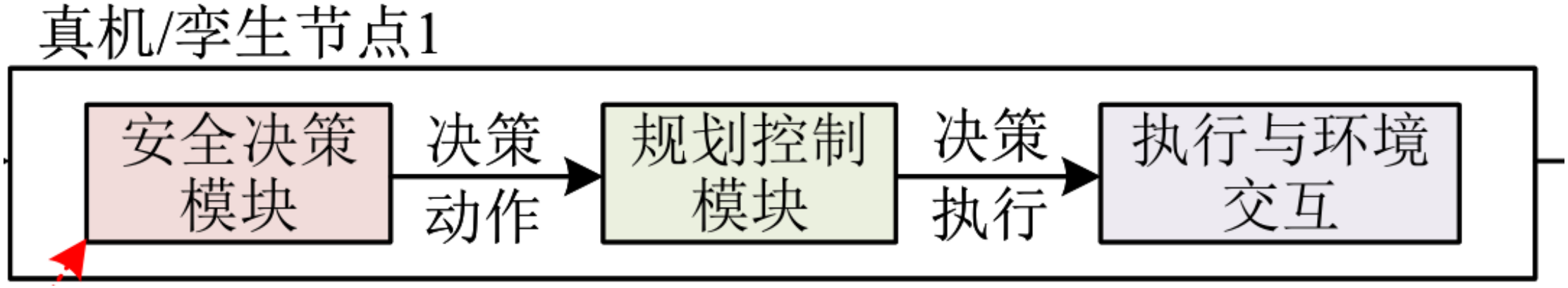


图7.12（节点1）行为选择经过控制与执行形成运动后果

行为：退出巡检，转向备用着陆区。

控制与结果：记录轨迹目标、推力/姿态输出、实际位置及是否完成着陆。

7.6.2 分层决策与训练：并行反馈与训练目标

奖励来自动作后果，模型更新与验收分开进行

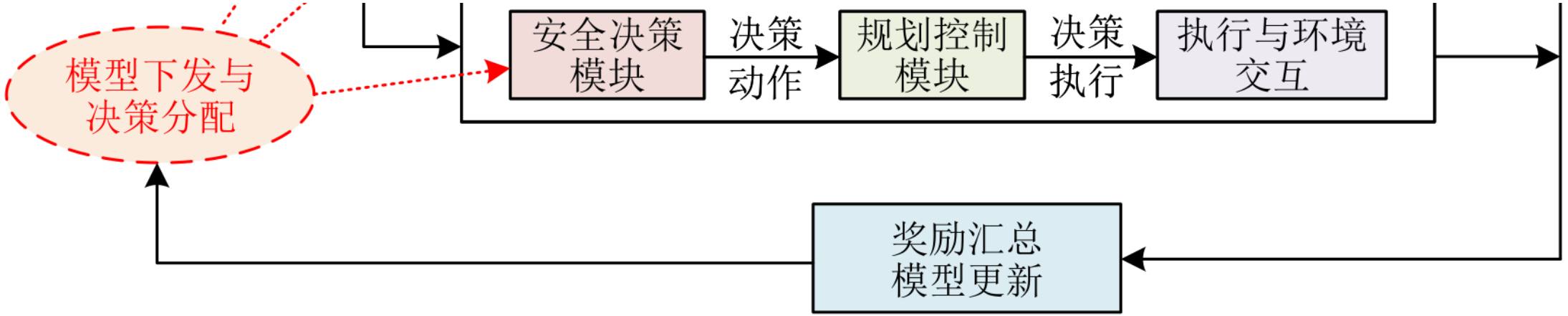


图7.12（下部）结果汇总、模型更新与下一轮模型下发

先检查可行性：导航、推力与可用着陆区决定动作集合。

再计算反馈：记录任务终态、危害变化与执行约束。

留出验收工况：保留未参与更新的任务、故障与触发状态。

7.6.2 分层决策与训练：大模型生成行为树

大模型离线生成行为树，测试反馈用于修订分支

输入：任务描述、故障案例、测试数据、预期目标。

大型语言模型按任务边界与允许动作生成行为树（BT），
以文本描述节点含义、触发条件与动作输出。

1 约束下离线生成

行为树结构与节点
候选条件与动作逻辑

→

2 逐节点审查与回放

检查条件、输出与分支
记录终态与保护响应

→

3 通过后验证与发布

独立工况验证并回归
符合发布判据后审定版本

未达标：将失败场景、触发节点与危害后果反馈给大模型，修订分支。
修订后的候选重新审查、回放、验证；符合判据后发布并保留版本。

测试结论限于已检验工况；生成或一次回放通过不代表保证安全。

7.6.2 分层决策与训练：混合方法与运行时监督

选学：监视器把高级策略限制在可回退范围内

运行时保障（RTA）：高级策略提出动作，监视器决定执行或交给备份控制。



备份控制的可用范围与切换条件先定义，再检验传感、延迟和执行余量。

NASA：运行时监督与备份控制的形式化框架

https://ntrs.nasa.gov/citations/20240006522

本章目录

安全决策与整机改进

第七章 健康管理与安全评估

7.1 健康管理与安全评估基础
7.2 统一故障建模框架
7.3 虚实安全强化与故障数据方法
7.4 故障危害与安全评估
7.5 故障诊断与健康预测
7.6 安全决策与整机改进
7.7 故障注入接口与算法开发框架
7.8 例程实验与虚实验证
7.9 习题与课堂讨论

本节 3 个教学单元

7.6.1 处置与控制取舍
教材对应：第7.1.5.5节

7.6.2 分层决策与训练
教材对应：第7.1.5.5节

7.6.3 整机回归与改进
教材对应：第7.1.5.6节

7.6.3 整机回归与改进：整机回归矩阵

策略更新后，要回到正常、故障与组合条件

回归工况	原策略与改进策略共同记录	主要判据
正常巡检	相同任务与初态；记录轨迹与控制量	任务完成； 正常性能不明显退化
电机效能下降	同注入参数、事件；诊断与模式时序	响应生效； 推力/姿态仍满足约束
故障＋定位受损	定位有效性、回退模式、水平漂移	使用具备前提的模式； 终态可接受
故障＋低电量	保护级别、目的地、能量与着陆状态	保护优先级正确； 避免继续耗尽

教学回归设计：相同初态与故障参数比较旧/新版本，门槛由任务自定。

7.6.3 整机回归与改进：完整链路的可观察证据

一次用例要追到动作生效和实际运动终态

故障实际生效 ➡ 诊断输出 ➡ 模式/目标改变 ➡ 运动与终态

日志证据	能定位的具体问题
故障与诊断事件	注入是否生效；误报、漏报与识别延迟
请求与实际模式	行为是否被接受；底层保护是否覆盖请求
控制输出与实际状态	饱和、推力不足、估计失真或模型误差
最终任务状态	是否完成退出/着陆；是否发生不可接受后果

7.6.3 整机回归与改进：安全强化的整机回环

未满足整机指标时，沿反馈分支改进设计

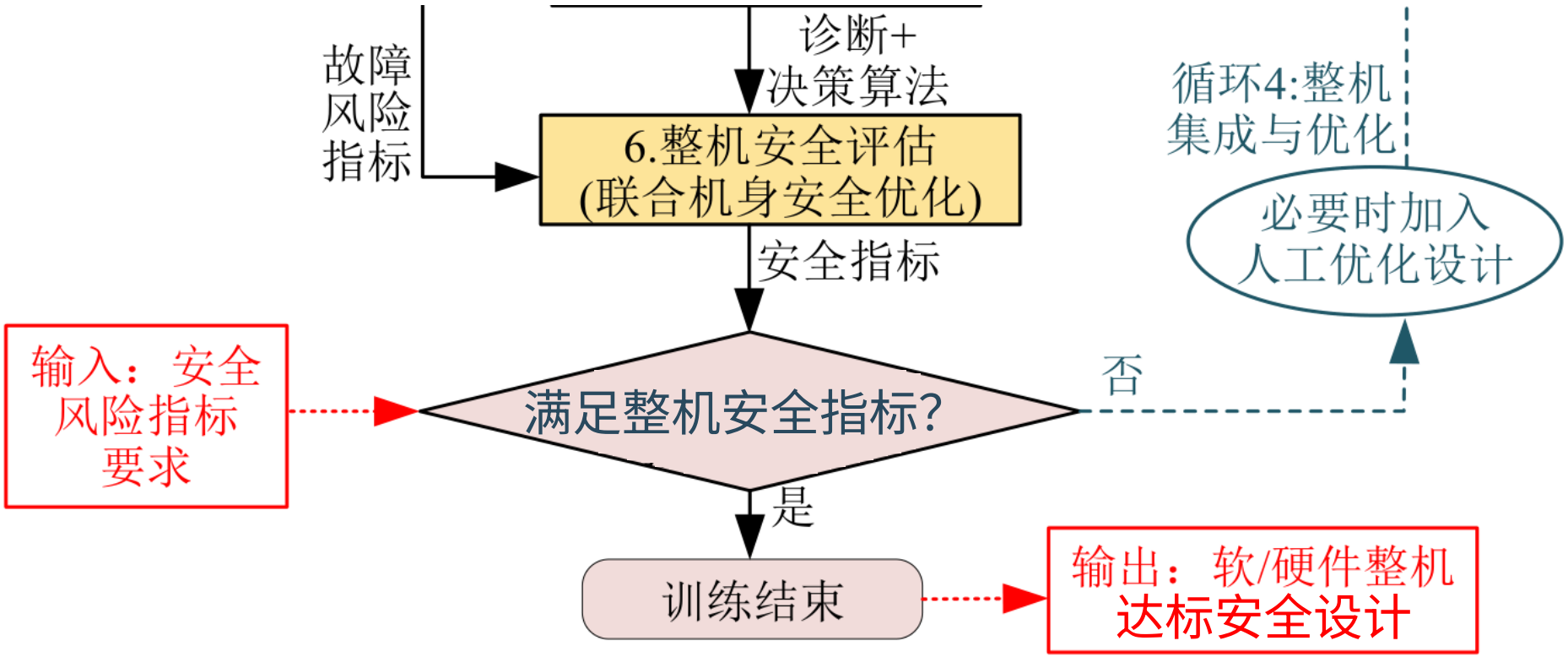


图7.6（下部）教材判断框写“达到最优或满足指标要求”，本课以整机达标为准

7.6.3 整机回归与改进：软硬件协同改进

降低故障发生、提高可检测性与剩余可控性

改进位置	巡检案例中的措施	验证内容与代价
故障源	部件选型、连接与环境防护	改变故障特征；检查维护与成本
检测链	专用检测、传感交叉校验	故障可见性、误报与检测延迟
剩余能力	执行/供电/感知冗余	可控性；重量与共因故障
软件策略	诊断阈值、保护优先级、控制重构	正常/故障回归；模式、版本一致

软件改进动作，硬件改变故障发生与可恢复的边界。

练习与讨论

练习：预测保护动作，并解释能力前提

v1.12.3：COM_RC_LOSS_T=1.0 s；COM_RCL_ACT_T=5 s；遥控丢失配置返航。
电量按20%/12%/6%依次警告、返航、降落；初始导航、返航与控制能力有效。
最后遥控输入在12.0 s；以下为纸面推演。

事件	问题
13.5 s：遥控仍未恢复	现在处于什么保护阶段？原定何时请求返航？
14.0 s：电量降至5%	继续保持等待，还是改变动作？为什么？
另设：定位失效	COM_POSCTL_NAVL=0，高度也失效：进入什么模式？

参考要点

参考：分开超时、优先级和导航回退

时序：13.0 s确认丢失；13.5 s原地保持；若无新条件，18.0 s请求返航。

电量竞争：14.0 s的5%低于6%，属紧急级，降落覆盖原先等待的返航。

定位回退：COM_POSCTL_NAVL=0、高度也失效，切换自稳模式，飞手操控。

结果判据：同时观察实际模式、下降响应与可接受着陆区；请求生效只是第一步。

练习与讨论

练习：估算分层响应时间，判断保护放在哪层

沿用分层响应时间算例：检测0.15 s；上层计算0.05 s；下行0.06 s；执行0.01 s。
上行通信由0.04 s增至0.2 s；某紧急故障要求保护在0.4 s内生效。
各项延迟均为教学设定；以下为纸面估算。

题目	问题
响应时间	上行增至0.2 s：两条路径各需多久？
时限判断	0.4 s时限：哪条路径满足？原算例呢？
分层归属	哪类故障留在底层先执行，哪类交给上层？

参考要点

参考：通信变慢后，紧急保护更要留在本地

上层路径： $T_{\text{upper}} = 0.15 + 0.2 + 0.05 + 0.06 + 0.01 = 0.47\,\text{s}$

本地路径： $T_{\text{local}} = 0.15 + 0.01 = 0.16\,\text{s}$

时限判断： 0.47 s超过0.4 s，只有本地路径满足；原算例上层0.31 s，尚在时限内。

留在底层： 特别紧急、实时性要求高的故障先执行后上报，如紧急电量降落。

交给上层： 需权衡任务进度、场地与集群状态的行为，如放弃任务或返航。

本节小结

用可执行动作和可追溯证据改进整机

诊断与条件 ➡ 剩余能力 ➡ 动作与执行 ➡ 结果与改进

处置：条件触发保护；导航、推力与场地决定动作是否可行。

分层：上层选行为，底层及时保护；反馈对应不同优化对象。

验证：同输入、同变量、可追溯版本；误差和容差限定结论范围。

回归：正常、故障、组合工况一起检验；软件与硬件协同改进。

易错点：模式请求≠运动受控；模型可信度≠安全概率；风扰表现≠故障容错。