

“远程协同打击攻防博弈挑战赛-暨弈起进化劳动竞赛” 预赛评分标准与细则

1、评分标准

预赛包含小组循环赛和大组晋级赛，其中，小组循环赛各队伍与同组其他队伍进行循环对抗，每两支队伍之间进行 1 场比赛，每场比赛分两局（分别执红方和蓝方），每个小组通过“机机对抗得分”选取第一名参与大组晋级赛。

大组内晋级赛评分标准包括“机机对抗晋级赛得分”和“智能体算法性能评分”两个部分。各参赛队伍的最终得分由两部分综合计算得出，其中“机机对抗晋级赛得分（记为 X）”占 70%，
“智能体算法性能评分（记为 Y）”占 30%。

2、机机对抗得分计算方式

根据平台培训后参赛队伍的资格审核结果，确定入围预赛队伍。队伍分为 A、B、C、D 四大组，每组含 2 个小组（A1、A2、B1、B2、C1、C2、D1、D2），每组队伍数量根据实际报名情况确定，分组通过抽签随机分配。

（1）对抗形式：每两支队伍进行 1 场比赛，每场比赛分两局，分别执红方和蓝方展开对抗。

（2）积分规则：采用“一场两局，大积分+小积分”规则。

“小积分”：单场两局对抗的得分总和（依据平台设定的任务完成度、目标毁伤率等量化指标）；

“大积分”：根据单场两局胜负确定的积分，具体为：两局全胜积 3 分；一胜一平积 2 分；一胜一负时，小积分较高积 2

分、小积分相同或较低积 1 分；两负积 0 分。

(3) 机机对抗得分的计算方式

$$Y = \frac{\sum_{i=1}^n A_i}{3*n} * 100 \quad (1)$$

其中，A 为单场比赛积分，n 为单支队伍参加的总场次，X 取值 0-100 分。

3、机机对抗晋级赛得分计算方式

大组内晋级赛由各小组循环赛排名第 1 的队伍参加，共 8 支队伍，分为 4 组对抗（A1 vs A2、B1 vs B2、C1 vs C2、D1 vs D2），分组为赛前抽签确定的小组对应关系。

(1) 对抗形式：每两支队伍进行 1 场比赛，每场比赛分两局，分别执红方和蓝方展开对抗。

(2) 积分规则：采用“一场两局，大积分+小积分”规则。

“小积分”：单场两局对抗的得分总和（依据平台设定的任务完成度、目标毁伤率等量化指标）；

“大积分”：根据单场两局胜负确定的积分，具体为：两局全胜积 3 分；一胜一平积 2 分；一胜一负时，小积分较高积 2 分、小积分相同或较低积 1 分；两负积 0 分。

(3) 机机对抗晋级赛得分 X 的计算方式

$$X = \frac{A}{3} * 100 * 70\% \quad (2)$$

其中，A 为单场比赛积分，X 取值 0-70 分。

4、智能体算法性能评分要点

由技术专家依据参赛队伍提交的智能体代码、设计说明文档进行评定，从 5 个维度打分，总分为 100 分，最终按 40%折算为 Y（Y 取值 0-40 分）。具体评分要点如下表所示。

表 1 辅助决策算法创新性评分要点

序号	评分要点	评判内容	评判依据	评分标准	分值
1	大模型战术决策与场景适配性	大模型对战场态势的语义理解精度、动态策略生成能力，及与非对称对抗场景（红方饱和打击 / 蓝方分层拦截）的适配性	1、智能体代码（大模型调用逻辑、态势解析模块） 2、设计说明文档（大模型决策流程部分）	能基于实时态势（如海面能见度变化、商船干扰）生成精准策略，红 / 蓝方角色切换时决策逻辑自洽，适配动态战场约束（如发射车间距、舰船拦截通道限制）	30
2	大模型与小模型协同机制	大模型（高层决策）与小模型（实时计算 / 执行）的分工合理性、协同效率，及策略从生成到落地的转化精度	1、智能体代码（大小模型接口、数据交互模块） 2、设计说明文档（协同架构部分）	大模型负责战略目标拆解与战术规划，小模型负责路径规划、火力分配等实时计算，两者数据交互延迟低，协同逻辑自洽	25
3	大模型策略生成的创新性	大模型基于博弈认知涌现的策略新颖性，及在	1、设计说明文档（大模型创新点部分）	策略突破传统规则局限（如红方非对称组合攻击、蓝方动态拦截资源调配），且能	20

		多轮对抗中的迭代优化能力		通过多轮对抗学习优化	
4	大模型提示词设计与指令执行效率	提示词对大模型决策的引导精度，及指令从生成到平台执行的准确性、实时性	1、智能体代码（大小模型接口、数据交互模块） 2、设计说明文档（协同架构部分）	提示词能精准约束大模型输出格式与决策逻辑，指令转换无歧义，执行延迟低	15
5	大模型相关代码与文档规范性	大模型调用代码的架构合理性、注释完整性，及文档对大模型工作流的阐述清晰度	1、智能体代码（大模型调用模块、注释） 2、设计说明文档（大模型实现细节部分）	代码模块化程度高，注释能清晰说明大模型输入输出逻辑；文档能完整阐述大模型从态势输入到决策输出的全流程	10

由相关技术专家根据智能体代码和设计说明文档进行打分，记为 B，取值 0 到 100 分，各支参赛队伍辅助决策算法创新性评分 Y 为：

$$Y = B * 30\% \quad (3)$$

其中，Y 取值 0-30 分。

5、总得分

各支参赛队伍的综合得分为： $Z = X + Y$ 。其中，Z 取值 0-100 分。

小组循环赛中，按 Z 值排名确定每组第 1 名；大组内晋级赛中，按 Z 值较高者晋级决赛。