

引用格式: 尹安. 基于墨子平台的水下多智能体协同占位决策算法设计[J]. 中国舰船研究, 2024, 19(增刊 1): 109–114.

YIN A. Design and research of underwater multi-agent cooperative getting-to-the-firing-position decision-making algorithm based on MoZi platform[J]. Chinese Journal of Ship Research, 2024, 19(Supp 1): 109–114 (in Chinese).

基于墨子平台的水下多智能体协同占位决策 算法设计

尹安*

中国舰船研究设计中心, 湖北 武汉 430064

摘要: [目的] 开展基于墨子平台的水下战场多智能体协同占位决策算法设计研究, 探索水下无人平台作战应用。[方法] 提出基于墨子平台的设计思路。针对算法处理确定性规则下的行动决策, 提出基于知识图谱的设计方法。采用 Neo4J 图数据库构建全局知识库, 并设计数据交换接口, 为智能算法模块提供确定性规则的辅助决策信息输入; 针对算法处理非确定性和不完全信息下的行动决策, 在设定场景下论述基于 MADDPG 的协同占位决策算法设计。提出状态空间、动作空间、奖励函数以及与知识图谱模块的交互设计思路。[结果] 基于上述思路, 完成了算法框架总体设计。在墨子平台中实现了知识图谱和智能算法的整合, 并通过墨子平台提供的可视化手段, 进一步验证了确定性规则下的行为约束与处理非确定性态势信息的智能算法进行整合的有效性。[结论] 仿真结果表明, 本文研究工作可为后续进一步开展多智能体协同占位算法设计提供参考。

关键词: 多智能体; 墨子平台; 知识图谱; MADDPG

中图分类号: U662.9

文献标志码: A

DOI: 10.19693/j.issn.1673-3185.03154

Design and research of underwater multi-agent cooperative getting-to-the-firing-position decision-making algorithm based on MoZi platform

YIN An*

China Ship Development and Design Center, Wuhan 430064, China

Abstract: [Objectives] This study focuses on the multi-agent cooperative getting-to-the-firing-position decision-making algorithm design of underwater battlefields based on the MoZi platform, and explores the combat application of underwater unmanned platforms. [Methods] A design idea based on the MoZi platform is proposed. For action decision-making under deterministic rules, the design of knowledge graph modules is discussed. The Neo4J database is used to build a global knowledge database, and a data exchange interface is designed to provide the auxiliary decision information input of deterministic rules for the intelligent algorithm module. For action decision-making under uncertain and incomplete information, the design of a cooperative space occupying a decision-making algorithm based on MADDPG is discussed under the set scenario, and the design concept of a state space, action space, reward function and interaction with the knowledge graph module is put forward. [Results] The overall design of the algorithm framework is completed on the basis of the abovementioned ideas. The integration of the knowledge graph and intelligent algorithm is realized on the MoZi platform, and the effectiveness of the integration of behavior constraints under deterministic rules and the intelligent algorithm for processing uncertain situation information is further verified through the visualization provided by the platform. [Conclusions] This paper can provide valuable references for the design of multi-agent underwater getting-to-the-firing-position cooperative algorithms.

Key words: multi-agent; MoZi platform; knowledge graph; MADDPG

收稿日期: 2022-11-01 修回日期: 2023-02-19

基金项目: 国家部委基金资助项目

作者简介: 尹安, 男, 1980年生, 博士, 高级工程师

*通信作者: 尹安

0 引言

近年来,世界主要军事强国都在大力加快无人化、智能化装备的研制和应用研究进度。无人装备具有自主化、智能化、多功能和低成本等优势,在未来战争中,可大幅提高作战能力。在新型装备和作战理念的双重驱动下,欧美等军事强国的水下作战能力可能会出现质的飞跃,甚至颠覆未来水下作战形态。

美国智库在2015年发布的《水下战新纪元》一文中即指出,随着无人装备的发展,潜艇作战样式需要从类似于飞机的前沿战术平台转变为类似于航母的协同平台,实现向潜艇-无人平台体系化转变。美国国防高级研究计划局(DARPA)在2016年开始以无人装备集群为切入点,布局立体化、无人化和分布式的作战模式创新项目,并持续投入经费研发水下无人装备^[1]。美国海军十分重视实装技术验证,2019年时曾经举行先进海军技术演习(ANTX),重点演示和验证了有人/无人平台协同作战的能力^[2]。我国也十分重视无人装备相关技术的发展,在无人装备智能算法研究方面,针对协同导航、协同探测和编队控制等领域也开展了大量研究工作。胡常青等^[3]提出了基于超短基线水声定位的无人水下航行器和无人水面艇的协同导航方法,可实现多源信息融合,提升水下航行器导航精度。蒲勇等^[4]开展了水下航行器协同探测问题的研究,基于Matlab平台完成了水下航行器集群对多目标协同探测算法的仿真。周宏坤等^[5]设计了基于主从网络结构的水下航行器协同探测方案,并对数据融合技术进行了探讨。王圣洁等^[6]开展了潜艇与水下航行器协同控制体系的研究,提出了动态性组织结构和系统决策分配模型。但国内外主要采用Matlab进行细分领域的分析研究,鲜有针对特定任务场景的水下协同决策类算法设计和测试分析平台应用的设计研究。但在算法、研究到装备应用方面,仍需进行深入研究。因此,建立算法测试平台和开展算法可视化分析工作是其中一个重要方面。

本文结合已开展的研究工作来论述水下占位智能算法设计研究。首先介绍基于墨子联合作战推演系统(后文简称“墨子平台”)的智能算法设计和测试总体思路;其次,针对算法处理确定性规则下的行动决策方面,论证知识图谱的设计工作,通过构建知识图谱为无人装备在规则性和任务边界明确下的感知和协同决策提供支撑;然后,针对算法处理非确定性和不完全信息下的行

动决策方面,以母艇与带缆型无人水下航行器(后文简称“带缆型UUV”)协同问题为切入点,以基于MADDPG的协同占位决策算法应用论述知识图谱和智能算法协同为设计思路;最后,介绍信息接口设计,实现知识图谱和智能算法在墨子平台环境中的整合,从而通过可视化手段进一步验证和分析算法合理性。

1 总体设计思路

作战行动和指挥决策需要考虑的因素大致可分为两类:一类是具有规则性和任务边界明确的决策要素,如我方平台的运动性能、能量储备等对战术机动能力的影响,以及武器射距与目标距本艇距离对可攻性判断的影响等;另一类是非确定性、不完全信息下的决策要素,如战场海洋环境信息的时变性、敌方平台对抗的不确定性等,这需要对战场态势变化信息进行收集、处理和整合,再综合应用多维信息要素和智能算法进行辅助分析,以提高决策可信度。

本研究采用知识图谱描述规则性和任务边界明确的知识,并与基于MADDPG的协同占位决策算法在墨子平台中进行整合。墨子平台具有模型丰富、功能完备和扩展性强等特点,结合当前深度强化学习等人工智能研究前沿技术,可支持基于TensorFlow, PyTorch等框架算法库的军事智能体开发,并提供任务想定管理、任务想定编辑、态势显示、仿真引擎、基础数据管理工具和人工智能开发接口等功能模块,能够支撑开展作战概念推演、新战法研究和装备体系论证。

基于墨子平台的算法设计总体设计思路框图如图1所示。

其中,对规则性和任务边界明确的决策要素采用知识图谱进行多维异构知识和作战规则的知识管理,包括知识存储子模块、知识抽取子模块、知识推理子模块和用于展示图数据库数据关系的可视化显示子模块,并根据输入的外部信息输出推理结果,为智能算法提供必要的信息支撑。对非确定性、不完全信息下的决策要素,以我方母艇携带两艘带缆型UUV协同场景为研究对象,设计开发协同占位决策智能算法模块,并接收知识图谱模块反馈结果,辅助开展智能分析。采用墨子平台作为算法测试验证和可视化展示的基础平台,通过自定义信息接口实现知识图谱模块和智能算法模块在墨子平台中的整合。在墨子平台中制定作战想定和红蓝双方初始态势,通过墨子平台提供的API接口获取目标跟踪、目标类型判

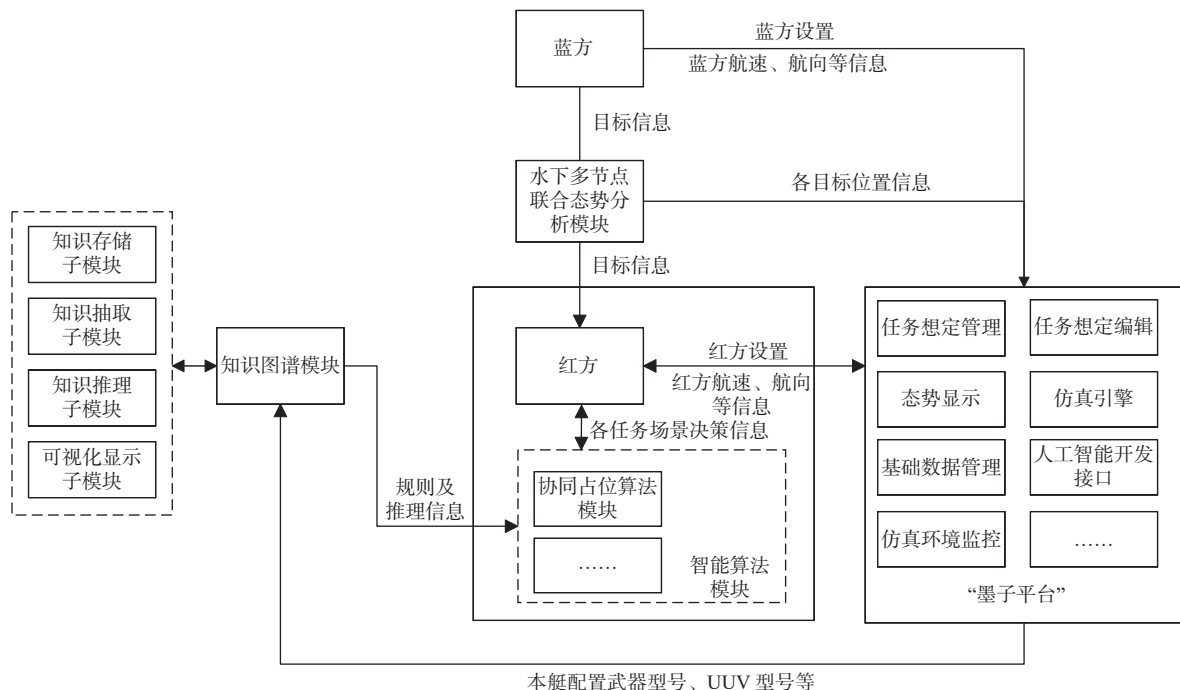


图 1 总体设计思路框图

Fig. 1 Overall architecture diagram

断等信息。蓝方也通过开发的水下多节点联合态势感知算法对目标进行协同估计,以增强红蓝双方的对抗效果。随着后续研究深入,在智能算法模块中可集成扩展和测试不同场景下的不同任务决策类智能算法。

2 主要模块设计

2.1 知识图谱模块设计

知识图谱是一种用图模型来描述知识和建模世界万物之间的关联关系的技术方法。知识图谱由节点和边组成。节点是实体或抽象的概念,边是实体的属性或是实体之间的关系。知识图谱旨在从数据中识别、发现和推断事物与概念之间的复杂关系,是事物关系的可计算模型^[7]。针对知识图谱在军事领域的应用,江志浩等^[8]提出了作战目标知识图谱构建的基本架构,探索在军事领域的应用;梅发国等^[9]提出了基于知识图谱的战场目标关系融合技术;邢萌等^[10]提出了军事领域知识图谱构建及应用技术架构,指出军事领域的知识图谱是链接各类作战要素的桥梁,是打通各军种不同业务间信息隔阂的重要手段。

在本研究中,基于知识图谱架构并采用 Neo4J 图数据库构建一个全局知识库,实现对已收集的我方平台信息、敌方平台信息、武器装备信息和作战规则信息等结构化、半结构化和非结构化知识的统筹管理,具备在多维度、大容量信息中,支

撑知识快速搜索和深度信息提取,并设计数据交换接口,为智能算法模块提供必要的辅助决策信息输入。Neo4J 图数据库采用节点和边来组织数据,以图数据库的形式管理知识,可快速查询长程关系,也可有利于进行知识推理,以挖掘知识中的隐藏关系,如基于规则的知识推理方法可从一个实体的关系图中,推理出与该实体存在多级关系的实体,从而分析出两个实体之间的组合关系,或采用最短路径推理和规则推理等技术来分析不同实体间的关系。

本研究中知识图谱应用框架如图 2 所示。

2.2 协同占位决策算法设计

海战场环境下实施武器打击一般包含目标跟踪、意图分析、态势分析、占位机动和武器参数设置等多个环节,单艇或多智能体协作下所需的算法和考虑因素各不相同。以我方母艇携带两艘带缆型 UUV 协同场景为研究对象,涉及智能体平台编队控制模型建立、编队队形生成和保持、目标跟踪模型建立、目标状态估计和协同决策算法设计等多个方面。本文将从基于 MADDPG 的协同占位决策算法应用为例论述设计思路。

2.2.1 MADDPG 算法

墨子平台中已提供了基于深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法的水面舰艇带直升机反潜和无人机突防应用范例,但仅是针对单个智能体的智能算法应用。在单个智能体强化学习任务当中,DDPG算法在单

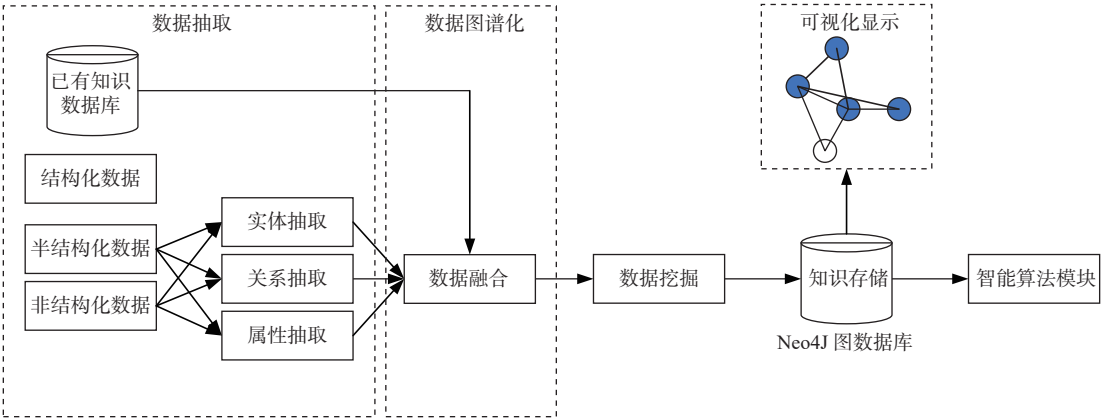


图 2 知识图谱应用框架
Fig. 2 Knowledge graph application framework

个智能体的环境下表现性能良好。本文研究对象为母艇与两艘带缆型 UUV 协同场景,属于多智能体范畴。在多智能体强化学习中,多智能体系统之间由于存在竞争与协作关系,相比较于单个智能体环境而言,算法设计难度更大。目前已提出的多智能体算法有 MADDPG(multi-agent deep deterministic policy gradient)^[11], COMA(counterfactual multi-agent)^[12] 和分布式 Q 学习(distributed Q-learning)^[13] 等。本文采用基于 DDPG 基本框架的 MADDPG 算法,将深度确定性策略算法应用到多智能体环境。MADDPG 算法为了获得多个智能体完成作业的最优策略,算法结构采用多个 Actor-Critic 网络结构,注重于智能体之间的交互以解决竞争或者合作环境中的协同任务。通过中心化的训练,去中心化的测试,将训练和测试过程分离,加快训练速度,也解决了环境的稳定性问题。

2.2.2 算法设计

以母艇与两艘带缆型 UUV 协同攻击中进行协同占位决策和保持队形控制为研究对象,分析基于 MADDPG 算法实际应用时的设计思路。其中,由墨子平台模拟提供智能体与战场环境交互。

1) 状态空间设计。

(1) 母艇状态空间设计。

母艇的目的是根据敌艇的运动状态占据攻击阵位,并指挥两艘带缆型 UUV 形成三角形的攻击阵型。将母艇运动状态在直角坐标系下进行分解得到母艇的位置 $[P_{x_red}, P_{y_red}]$, 根据母艇的速度与航向信息得到速度在横纵轴上的投影为 $[V_{x_red}, V_{y_red}]$ 。类似的,得到敌艇的位置坐标 $[P_{x_blue}, P_{y_blue}]$ 及敌艇的速度在横纵轴上的投影 $[V_{x_blue}, V_{y_blue}]$ 。以母艇与敌艇相对运动状态为分析基础,母艇的强化学习过程中状态空间为 $[P_{x_red} - P_{x_blue}, P_{y_red} - P_{y_blue}, V_{x_red} - V_{x_blue}, V_{y_red} - V_{y_blue}]$ 。

(2) UUV 状态空间设计。

由于 UUV 与母艇之间用缆进行连接,在保持队形中为了避免 UUV 与母艇的机动安全,需考虑 UUV 与母艇间的距离不能大于缆的最大长度,且 UUV 与母艇距离不能小于最小安全距离。根据第 i 个 UUV 位置与速度在横纵坐标上分解得到位置坐标 $[P_{x_uuv}, P_{y_uuv}]$, 速度投影为 $[V_{x_uuv}, V_{y_uuv}]$ 。因此,第 i 个 UUV 的在强化学习过程中的状态空间为 $[P_{x_red} - P_{x_uuv}, P_{y_red} - P_{y_uuv}, V_{x_red} - V_{x_uuv}, V_{y_red} - V_{y_uuv}]$ 。

其中,UUV 与母艇间的距离不能大于缆的最大长度,UUV 与母艇距离不能小于最小安全距离。上述约束条件属于确定性规则,由知识图谱提供输入。

2) 动作空间设计。

考虑各艇的动作空间是连续的,不考虑深度方向的机动变化,将各作战单元的动作空间分解为横轴上的速度 V_x 和纵轴上的速度 V_y 。

3) 奖励函数设计。

(1) 母艇奖励函数设计。

定义奖励规则如下:如果母艇距离占位点的距离小于最大误差,认为占位成功,奖励值为 5 分,完成占位;若当前母艇距占位点距离与上一步母艇距占位点距离之差小于设定的阈值,说明母艇正在接近占位点,奖励值为 1 分;若母艇距敌艇距离小于设定的最小安全距离,则认为母艇处于危险位置,奖励值为-5 分,并终止该轮次训练。

(2) UUV 奖励函数设计。

定义奖励规则如下:如果占位过程 UUV 距离母艇的距离小于缆长并大于最小安全距离,且 UUV 彼此间的距离大于最小安全距离并满足阵型要求的相对位置设定,奖励值为 1 分;若 UUV 距离母艇距离大于缆长,或小于最小安全距离,或 UUV 彼此间的距离小于最小安全距离,奖

励值为-5 分, 终止该轮次训练。

4) 占位实现。

利用深度强化学习进行占位攻击过程, 首先从知识图谱由武器性能等获得发射距离和发射舷角限制约束, 计算得到该占位点相对于敌艇的坐标位置, 并结合敌艇坐标以及敌艇航速航向信息计算得到占位点的理论坐标。在此基础上, 计算母艇距离该占位点的距离 d , 将该距离代入到母艇的奖励函数中, 给定母艇距离占位点可接受的最大偏移阈值 $self.dist_err_cm$, 若 $d < self.dist_err_cm$ 成立, 则认为占位成功。

基于 MADDPG 的协同占位决策算法流程图如图 3 所示。

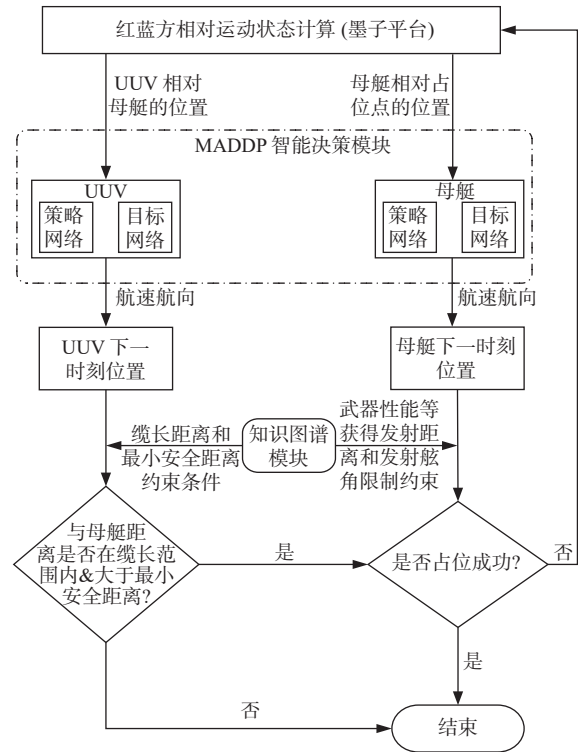


图 3 基于 MADDPG 的协同占位决策算法流程图
Fig. 3 Flow chart of collaborative occupancy decision algorithm based on MADDPG

3 算法验证

墨子平台和智能算法模块部署于同一台计算机中, 智能算法模块通过 `reset()`, `step()` 和 `get_agents_action()` 等函数实现与墨子平台环境的交互和训练控制; 知识图谱模块部署于另一台计算机中, 通过规则来查找实体间关系并按照接口协议与智能算法模块交互。

基于墨子平台验证本文提出的设计构想, 通过墨子平台本身(图 4)和在智能算法模块程序中记录运动轨迹(图 5 和图 6)两种手段可视化分析算法训练过程。

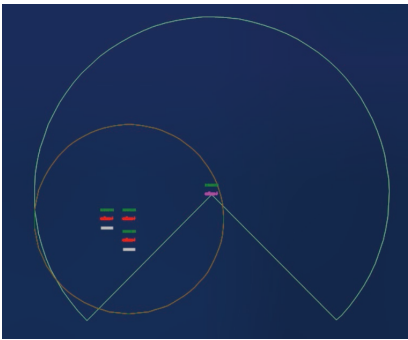


图 4 墨子平台可视化效果
Fig. 4 Visualization effect on MoZi platform

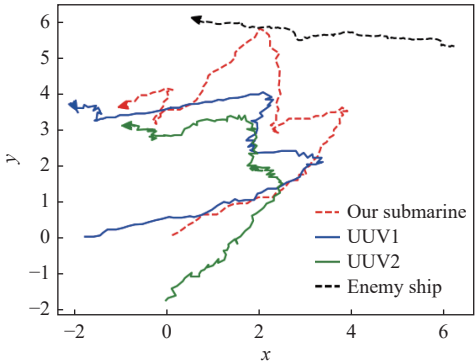


图 5 轨迹记录 1
Fig. 5 Moving track record 1

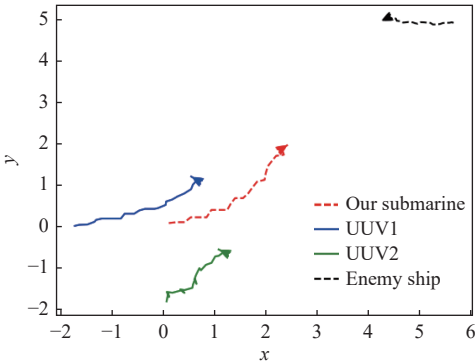


图 6 轨迹记录 2
Fig. 6 Moving track record 2

图 5 为训练在初始阶段中的母艇和携带 UUV 运动轨迹, 此时智能体仍在学习过程中, 在设计的状态空间和奖励函数的刺激下, 逐步具备了针对本文所述的特定任务的协同能力, 此时运动轨迹记录如图 6 所示。其中, 由知识图谱模块提供缆长距离和最小安全距离约束条件, 以及基于武器性能获得发射距离和发射舷角限制约束条件, 智能算法在奖励函数规则设计下对母艇和 UUV 行动进行激励。从母艇红色轨迹线和两个 UUV 分别对应的蓝色轨迹线和绿色轨迹线可以看出, 母艇按占位规则进行占位机动, 同时与两个带缆型 UUV 形成预期的三角形阵型; 两个带缆型 UUV 与母艇距离小于缆长并大于最小安全距离, 且两个 UUV 彼此间的距离大于最小安全距离并

满足阵型要求的相对位置设定。按照总体设计思路,在墨子平台中,实现了知识图谱和智能算法的整合,并可通过墨子平台可视化方式观察智能体行动,增强对算法执行过程的可解释性。

4 结 语

本文基于墨子平台开展了水下协同占位决策算法设计研究,构建知识图谱设计方案,为智能算法模块提供确定性规则的辅助决策信息输入,实现了知识图谱和智能算法的整合,并在墨子平台环境下完成设计和测试,可为后续进一步开展多智能体水下占位协同算法设计提供参考。

墨子平台提供了军事人工开发和训练的基础环境,但考虑水下战场弱联通、弱观测等特点,还需融入更贴近实战的水文信息模块、海底地形地貌模块和传感器探测模块等,并结合战术战法进行不断扩充和形成对多维跨域知识的推理。此外,还需考虑赋予蓝军更多对抗智能,以充分验证算法的适应性和鲁棒性。

进一步完善并形成一套水下战场智能体对抗通用模拟仿真和测试平台,加强对算法的可解释性分析和迭代优化设计,对探索水下占位智能算法应用意义重大。

参考文献:

- [1] 谢伟,杨萌,龚俊斌.水下攻防对抗体系及其未来发展[J].中国工程科学,2019,21(6):71-79.
XIE W, YANG M, GONG J B. Underwater attack-defense confrontation system and its future development [J]. Strategic Study of CAE, 2019, 21(6): 71-79 (in Chinese).
- [2] 王雅琳,刘都群,杨依然.2019年水下无人系统发展综述[J].无人系统技术,2020,3(1):55-59.
WANG Y L, LIU D Q, YANG Y R. Summary of the development of unmanned undersea systems in 2019[J]. Unmanned Systems Technology, 2020, 3(1): 55-59 (in Chinese).
- [3] 胡常青,文龙贻彬,张亚婷,等.基于超短基线水声定位的USV/UUV协同导航方法[J].中国惯性技术学报,2019,27(3):327-333.
HU C Q, WEN L Y B, ZHANG Y T, et al. Cooperative navigation for USV/UUV based on ultra-short baseline positioning system[J]. Journal of Chinese Inertial Technology, 2019, 27(3): 327-333 (in Chinese).
- [4] 蒲勇,袁富宇.UUV群探测系统仿真设计[J].计算机仿真,2012,29(1):143-147.
PU Y, YUAN F Y. Designed of cooperative detection simulation system of UUV group[J]. Computer Simulation, 2012, 29(1): 143-147 (in Chinese).
- [5] 周宏坤,葛锡云,邱中梁,等.UUV集群协同探测与数据融合技术研究[J].舰船科学技术,2017,39(12):70-75.
ZHOU H K, GE X Y, QIU Z L, et al. Research on UUVs cooperative detection and data fusion[J]. Ship Science And Technology, 2017, 39(12): 70-75 (in Chinese).
- [6] 王圣洁,康凤举,韩翊.潜艇与智能无人水下航行器协同系统控制体系及决策研究[J].兵工学报,2017,38(2):335-344.
WANG S J, KANG F J, HAN H. Research on control and decision-making of submarine and intelligent UUV cooperative system[J]. Acta Armamentarii, 2017, 38(2): 335-344 (in Chinese).
- [7] 王昊奋,漆桂林,陈华钧.知识图谱:方法、实践与应用[M].北京:电子工业出版社,2019:1.
WANG H F, QI G L, CHEN H J. Knowledge Graph: Method, Practice and Application[M]. Beijing: Publishing House of Electronics Industry, 2019: 1 (in Chinese).
- [8] 江志浩,周卿,石敏,等.作战目标知识图谱构建与应用[J].海军航空工程学院学报,2020,35(6):471-477.
JIANG Z H, ZHOU Q, SHI M, et al. Construction and application of knowledge graph for target in battlefield [J]. Journal of Naval Aeronautical and Astronautical University, 2020, 35(6): 471-477 (in Chinese).
- [9] 梅发国,戴大伟,张冀.基于知识图谱的战场目标关系融合技术[J].指挥信息系统与技术,2017,8(5):81-86.
MEI F G, DAI D W, ZHANG J. Battlefield target relation fusion technology based on knowledge mapping[J]. Command Information System and Technology, 2017, 8(5): 81-86 (in Chinese).
- [10] 邢萌,杨朝红,毕建权.军事领域知识图谱的构建及应用[J].指挥控制与仿真,2020,42(4):1-7.
XING M, YANG C H, BI J Q. Construction and application of domain-specific knowledge graph in military field[J]. Command Control & Simulation, 2020, 42(4): 1-7 (in Chinese).
- [11] LOWE R, WU Y, TAMAR A, et al. Multi-agent actor-critic for mixed cooperative-competitive environments [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: ACM, 2017: 6382-6393.
- [12] FOERSTER J N, FARQUHAR G, AFOURAS T, et al. Counterfactual multi-agent policy gradients[C]//Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence. New Orleans: AAAI, 2018: 363.
- [13] GALINDO-SERRANO A, GIUPPONI L. Distributed Q-learning for aggregated interference control in cognitive radio networks[J]. IEEE Transactions on Vehicular Technology, 2010, 59(4): 1823-1834.