

DOI: 10.3969/j.issn.1009-9492.2024.00089

王亦晨, 刘雪梅. 基于冲突搜索增强深度强化学习的多AGV路径规划方法 [J]. 机电工程技术, 2024, 53 (08): 23-27.

基于冲突搜索增强深度强化学习的多AGV路径规划方法*

王亦晨, 刘雪梅

(同济大学机械与能源工程学院, 上海 201804)

摘要: 针对多自动引导车 (Automated Guided Vehicle, AGV) 路径规划问题, 以最小化总行程时间为目标, 提出一种基于冲突搜索 (Conflict-Based Search, CBS) 算法增强的分布式独立Q学习 (Independent Q-Learning, IQL) 算法。首先, 采用栅格图法构建环境地图, 并对多AGV路径规划问题进行数学化描述, 涵盖了AGV间的碰撞类型和问题的基本假设。随后, 将问题转化为部分可观察马尔可夫决策过程 (Partially Observable Markov Decision Process, POMDP), 并详细定义了观测空间、动作空间及奖励函数。进一步地, 采用异步优先级经验回放架构, 将IQL方法扩展到分布式环境中, 并通过CBS算法引导Q网络, 优化AGV在拥挤环境中的决策过程。最后, 根据不同AGV数量设计了与其他深度强化学习算法的对比实验, 结果表明所提方法在成功率和平均步长关键性能指标上优于对照算法, 从而证实了所提方法的有效性和可行性。

关键词: 基于冲突搜索; 深度强化学习; 多AGV; 路径规划

中图分类号: TP249

文献标志码: A

文章编号: 1009-9492(2024)08-0023-05

Conflict-based Search Enhanced Deep Reinforcement Learning Approach for Multi-AGV Path Planning

Wang Yichen, Liu Xuemei

(School of Mechanical Engineering, Tongji University, Shanghai 201804, China)

Abstract: To address the path planning problem for automated guided vehicles (AGVs) with the objective of minimizing total travel time, an approach that enhances distributed independent Q-learning (IQL) with the conflict-based search (CBS) algorithm is proposed. Initially, a grid map method is employed to construct the environmental map, and the multi-AGV path planning problem is mathematically described, including assumptions about types of collisions between AGVs. Subsequently, the problem is transformed into a partially observable Markov decision process (POMDP), with detailed definitions of the observation space, action space, and reward function. Further, by utilizing an asynchronous priority experience replay architecture, the IQL method is extended to a distributed environment, and guided by the CBS algorithm to refine the Q-network, enhancing decision-making processes for AGVs in congested environments. Finally, comparative experiments with other deep reinforcement learning algorithms are designed based on varying numbers of AGVs. The results demonstrate that the proposed method outperforms the control algorithms in key performance indicators such as success rate and average step length, thereby validating the effectiveness and feasibility of the proposed approach.

Key words: conflict-based search; deep reinforcement learning; multi-AGV; path planning

0 引言

随着自动化技术的迅速发展, AGV已成为智能制造领域中一种关键的运输设备^[1]。多AGV系统的路径规划不仅要求为多辆AGV并行规划路径, 更需优化AGV间的协同行为, 以避免交叉冲突和环境障碍^[2], 确保路径规划的高效性和安全性。这一挑战对于提升生产效率和系统稳定性具有至关重要的意义。

目前, 学术界已将多AGV系统建模为多智能体路径搜索 (Multi-Agent Path Finding, MAPF) 问题^[3], 并采用将AGV视为智能体的方法进行求解。面对MAPF问题的高复杂性^[4], 研究者们已提出多种解决策略。例如, Sharon等^[5]提出的CBS算法通过顶层的约束树 (Con-

straint Tree, CT) 分裂策略和底层的无冲突路径搜索, 实现了多智能体系统中的高效路径规划。Li等^[6]进一步通过显示估计CBS算法 (Explicit Estimation CBS, EECBS) 使用在线学习技术优化顶层节点的选择, 显著加速了CBS算法的执行速度。此外, Tao等^[7]的交替迭代CBS算法 (Alternate Iterative CBS, AICBS) 通过将路径扩展和冲突解决交替进行的方式, 有效减少了因冲突引发的无效计划。刘恒麟^[8]则在智能仓储的多AGV路径规划中应用CBS算法, 并通过引入转弯因子和多线程并发处理冲突节点, 大幅度提高了路径求解的效率。Sartoret-ti等^[9]提出的PRIMAL框架通过结合强化学习和模仿学习, 实现了智能体间的有效协作。Ma等^[10]则结合多AGV

基金项目: 西门子 (中国) 有限公司合作项目 (kz0100020210591)

收稿日期: 2024-04-24

通信与深度Q学习,利用图卷积网络增强智能体的协作能力。Hu等^[11]在自动化集装箱码头的AGV防冲突路径规划问题中,采用了多智能体深度确定性策略梯度(Multi-Agent Deep Deterministic Policy Gradient, MADDPG)方法,并通过Gumbel-Softmax策略优化算法性能。

IQL^[12]是一种为多智能体系统设计的强化学习算法,适合处理多AGV路径规划问题^[13]。在此类问题中,各AGV之间不存在明显差异,每个AGV都独立地根据自身观察来做出行动决策。这种设置允许从单一AGV的视角出发,开发一个通用模型,而无需针对每个AGV定制不同的策略。在这种框架下,其他AGV被视为环境的一部分,从而使得从一个AGV学到的策略可以无缝地应用到其他AGV上。IQL通过使每个智能体独立地学习和更新其Q网络^[14],自然地适应于部分可观察的环境中分散的策略学习。此外,IQL的可扩展性允许它在不增加复杂性的情况下适应更多智能体,使其在大规模系统中特别有效^[15]。

鉴于CBS算法在解决冲突中显示出的强大能力,以及深度强化学习在处理复杂动态环境中的决策能力^[16],本文提出一种以最小化总行程时间为目标的CBS算法增强的分布式IQL算法(CBS-DIQL)进行多AGV路径规划。通过利用CBS算法指导拥挤状态下的AGV避免碰撞,并采用分布式IQL的方法使算法可扩展至大规模智能体系统。本文设计的实验将验证所提算法的有效性。

1 多AGV路径规划模型

1.1 环境地图建模

在多AGV路径规划问题中,精确的环境地图建模是设计有效解决方案的基础^[17]。本文采用图1所示栅格法对环境进行离散化建模,白色格子代表AGV可通行区域,灰色格子表示存在障碍物的不可通行区域,地图以外区域同样视为不可通行。每个AGV的位置由其所在栅格的坐标确定,不同颜色的格子区分不同的AGV。AGV的目标点分别由圆形标示,颜色与对应AGV的标识色一致。此外,每个AGV的移动仅限于相邻栅格之间。这种栅格化方法为复杂环境中的路径规划提供了清晰、易于执行的策略,能有效支持多AGV系统的高效运作。

1.2 问题描述

在本文研究的多AGV路径规划问题中,设有 n 个AGV,构成集合 $V = \{v_1, v_2, v_3, \dots, v_n\}$ 。每辆AGV需从其指定起点 b_i 移动至终点 e_i ,起点和终点分别组成集合 $B = \{b_1, b_2, b_3, \dots, b_n\}$ 和 $E = \{e_1, e_2, e_3, \dots, e_n\}$ 。路径规划在一个包含障碍物的二维栅格地图上进行,需要确保AGV在移动过程中避免与障碍物或其他AGV产生碰撞。

AGV之间的碰撞可分为两种主要类别,如图2所示。第一种是节点冲突,即两辆或更多的AGV在同一时间步尝试占据同一栅格;第二种是相向冲突,即两辆AGV试

图在同一时间步互换它们的位置。为了避免这些冲突,每辆AGV的路径 P_i 需规划得既独立又协调。

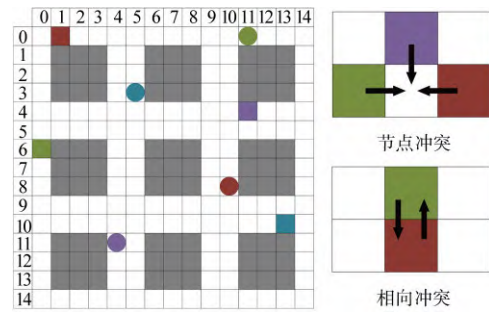


图1 环境地图模型

图2 AGV间的碰撞类别

本文路径规划的目标是最小化所有AGV路径的总长度,即

$$\min \sum_{i=1}^n \text{length}(P_i) \quad (1)$$

同时,也需要确保在任意时间步 t 内,对于任何两条路径 P_i 和 P_j (其中 $i \neq j$),不会发生节点冲突或相向冲突。通过图论方法,这个问题可以被视为在图的时间扩展版本中寻找一组不相交的路径。

不失一般性地,基于以下假设进行问题讨论:

- (1) AGV的每一步动作可以是上、下、左、右移动或静止不动,且不可沿栅格对角线移动;
- (2) 当AGV到达其终点后,其所在位置在下一步将被释放,即其他AGV可以占据此位置而不会导致碰撞;
- (3) AGV只能在地图内部移动,地图外的部分视为障碍物;
- (4) 所有AGV在整个路径中保持恒定速度,且转向时间忽略不计;
- (5) 所有AGV在每个时间步内同时进行决策和移动;
- (6) 环境中的障碍物及路径的可通行性在整个规划期间保持不变。

1.3 问题转化

在本文中,多AGV路径规划问题被形式化为一个部分可观察马尔可夫决策过程^[18],旨在有效应对路径规划中的不确定性以及环境状态的部分可观测性^[19]。此转化可以系统地运用决策理论来优化AGV行为,特别是在环境信息可能不完全可知或存在动态变化的情况下。

POMDP可以通过元组 $(S, A, T, R, \Omega, O, \gamma)$ 表示,其中 S 为状态空间, A 为动作空间, T 为转移模型, R 为奖励函数, Ω 为观测空间, O 为观测模型,而 γ (取值范围在0~1之间)为折扣因子。在每个时间步,环境处于某一状态 $s \in S$,AGV在当前状态 s 下采取动作 $a \in A$ 会过渡到状态 s' ,其概率由 $T(s, a, s')$ 确定。随后,AGV根据新状态 s' 接收到观测 $o \in \Omega$,此观测发生的概率由 $O(o, s', a)$ 给出,表明在执行动作 a 并转移到状态 s' 后观察到 o 的概

率。以下将对观测空间、动作空间及奖励函数进行更为详细的定义和描述。

1.3.1 观测表示

在部分可观察的设置中, 每个AGV的观测能力被限制在其视场(Field of View, FOV)内, 该视场通常设定为 $l \times l$ 的区域, 其中 l 一般为奇数, 以确保AGV位于FOV的中心位置。通过固定FOV的大小, 可以将策略推广至任意大小的地图, 并有效减少神经网络输入的维度, 从而简化模型的复杂性。

在有限的FOV内, 观测信息被组织到不同的数据通道中, 以便进行清晰的信息分析和处理, 如图3所示。具体来说, 每个观测由几个二进制矩阵组成, 这些矩阵分别标示出FOV内的障碍物、AGV的位置、当前AGV的目标位置, 以及其他可观测AGV的目标位置。若任一目标位置位于FOV之外, 则该位置会被映射到FOV的边界上。

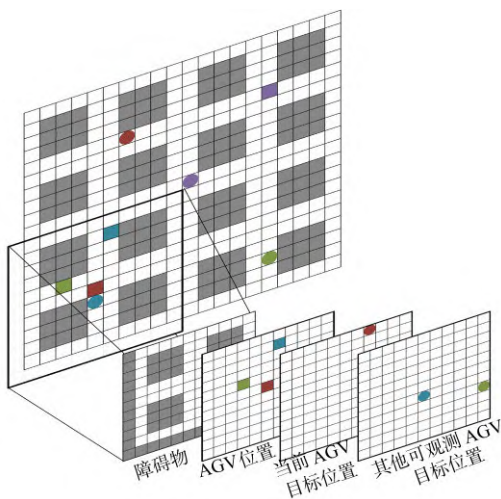


图3 FOV观测图

此外, 当AGV接近地图边缘时, 为了保持观测数据的一致性并防止信息丢失, 会在地图边界之外的区域添加虚拟障碍物。这种方法不仅提高了AGV在接近边界区域的导航能力, 还增强了模型在处理边界情况时的鲁棒性。这些设计考虑确保了AGV能够有效地接收和处理关键的环境信息, 即便在视觉受限的情况下也能保持高效的路径规划与决策执行。

1.3.2 动作表示

AGV在栅格地图上的移动行为被定义为一系列离散的动作, 包括向上、向下、向左、向右移动以及保持静止不动, 而不允许沿栅格对角线进行移动。在训练过程中, AGV可能执行一些无效动作, 导致它与障碍物或其他AGV发生碰撞。为了优化训练过程并增强模型的实用性, 当发生碰撞时, 会对AGV施加负奖励, 并将其状态恢复到碰撞前的位置, 以便继续后续的训练。此举是为了减少无效动作的发生频率, 提高训练效率。这种方法不仅有助于快速收敛到有效的策略, 而且还确保了AGV在实际操作中避免采取可能导致冲突的行动。

1.3.3 奖励函数

在本文中, 为AGV路径规划问题设计了一个综合的奖励函数体系, 如表1所示, 旨在通过奖励和惩罚机制指导AGV的行为。首先, 当AGV成功到达目标点时, 将授予一次性的到达奖励, 以表彰其完成任务。然而, 如果AGV在其行进过程中与障碍物或其他AGV发生碰撞, 将实施碰撞惩罚。进一步地, 为了鼓励AGV尽快到达目标点, 引入了步进惩罚, 即在AGV到达目标前, 每经过一个时间步将累积一定的惩罚。同时, 考虑到转向可能导致的额外时间消耗, 每次转向也将受到惩罚, 从而激励AGV采取尽可能直接的路径以减少不必要的转向行为。最后, 通过计算AGV执行动作前后与目标点的距离变化来调整奖励或惩罚: 如果AGV的动作使其与目标点的距离 d 减小, 将给予靠近奖励; 反之, 如果距离 d 增大, 则施加远离惩罚。这些措施综合考虑了效率、安全性和操作便捷性, 有效地引导AGV在复杂环境中做出最优的决策。

表1 奖励函数设计

行为	奖励值	行为	奖励值
AGV到达目标点	+50	AGV转向	-1
AGV发生碰撞	-50	AGV靠近目标点	$+10/\max(1, d)$
AGV移动或静止	-1	AGV远离目标点	$-10/\max(1, d)$

2 CBS算法增强的分布式IQL算法(CBS-DIQL)

2.1 分布式IQL架构的实现

Ape-X架构, 即异步优先级经验回放架构, 代表了一种先进的分布式深度强化学习框架^[20], 如图4所示, 其通过异步执行和优先级经验回放机制显著提高了并行训练的效率。在此架构中, 分散的执行者节点独立探索环境并生成数据, 这些数据随后被发送到集中的优先级经验回放缓冲区。这种方法加快了数据采集速度, 并根据时间差分误差(TD误差)调整经验的采样优先级, 有效地提升了训练的质量和效率。

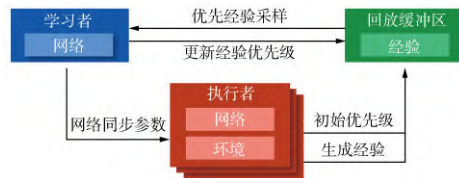


图4 Ape-X架构图

在多AGV路径规划任务中, Ape-X架构展现出显著的适应性和优势。借助其分布式和异步特性, 每个AGV作为独立执行者节点进行并行探索和学习, 优化路径选择并处理复杂交互。此外, 集中化的学习处理和优先级经验回放机制能够有效整合和利用来自各个AGV的经验数据, 生成更智能和效率更高的路径规划策略。在这一框架下, IQL方法可以被自然地扩展到分布式环境中。

分布式IQL架构在训练过程中将每个AGV视为一个独立的智能体, 其中每个智能体都维护着一个包含当前

Q网络的环境副本。这些智能体不断地生成新的状态转换,并为这些转换设定初始优先级。随后,来自所有智能体的转换数据被汇总到一个共享的优先级回放缓冲区中。智能体从此缓冲区中提取最具价值的经验,以便更新其网络和调整经验的优先级。通过这种共享的优先级系统,任何智能体探索到的有用经验都能提升整个Q网络的性能,从而提高了整体学习效率并加速了最优策略的形成。

2.2 CBS引导的Q网络优化

Q网络的随机探索对于丰富决策过程至关重要,然而在特定情境下,尤其是多AGV处于拥挤状态时,此类探索方法可能变得低效。在拥挤环境中,随机探索容易导致AGV间发生碰撞或死锁,降低了训练效率,也限制了神经网络模型的拟合能力。

为了优化拥挤状态下的搜索效率,本文采用CBS算法替代传统的随机搜索方法,实现更加智能化的路径规划。根据已建立的环境地图模型,定义拥挤状态为任意两个AGV之间的距离小于3个栅格,即一个AGV移动至另一AGV位置所需的最短时间步小于3。

在识别拥挤状态时,首先检测每个AGV,确定形成拥挤关系的AGV,并将它们组成初步拥挤集合。完成所有AGV的检测后,通过合并包含相同元素的集合并取并集,形成最终的最简拥挤集合,确保集合之间无重叠元素且集合内部无重复元素。例如,检测后可能形成的初步集合包括 $\{v_1, v_2, v_3\}$ 、 $\{v_3, v_4\}$ 、 $\{v_5, v_6\}$,合并后得到最简拥挤集合 $\{v_1, v_2, v_3, v_4\}$ 和 $\{v_5, v_6\}$ 。

在训练过程中进行动作选择时,首先为每个最简拥挤集合中的AGV指定动作。这包括确定集合中所有AGV的位置,并在栅格地图上标定一个能包含这些AGV的最小矩形框,随后扩大该框以形成足够操作空间的观察地图,其流程示意图如图5所示。在这一观察地图上,将集合内所有AGV的目标位置标注上去,如果目标位置超出了观察地图边界,则将其映射到边界上。观察地图最外圈栅格不设置障碍物,以确保目标位置可到达。然后,使用CBS算法在此地图上进行路径规划,以确定集合内每个AGV在当前时刻的最优动作。处理完最简拥挤集合中的AGV之后,对于其他AGV,继续采用Q网络进行动作选择。此方法不仅有助于避免AGV之间的潜在冲突,还可以推动处于拥挤状态的AGV向着解除拥挤的方向前进。

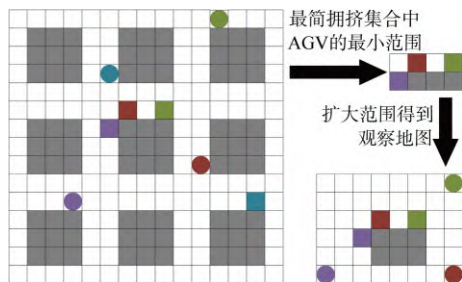


图5 获取观察地图的流程

2.3 训练过程及性能评估

本文采用分布式IQL框架来解决多AGV路径规划问题。训练环境构建于一个 30×30 的栅格地图上,其中包括障碍物以及随机指定的起点和终点。为了模拟不同程度的拥挤情况,每个训练周期内的AGV数量设置为5~10辆。

每个AGV作为一个独立的智能体进行操作,智能体的决策模型由一个三层全连接神经网络构成:输入层接收视野大小为 11×11 的局部栅格环境观察,隐藏层包含128个神经元并应用ReLU作为激活函数,输出层则预测每个可能动作(上、下、左、右、停止)的价值。模型的学习率设置为0.001,折扣因子为0.95,探索率初始为1并逐渐衰减至0.1。经验回放缓冲区容量为100 000,批次大小设为32。总训练时间步设置为3 000 000步,每个周期的最大时间步限制为200步。

为了精确评估模型的学习进度和效率,本文绘制了收敛情况、成功率和损失函数图。收敛情况如图6所示,展示了训练周期内智能体获得的平均奖励值变化趋势,奖励值的显著增加和稳定趋势表明模型成功收敛。

图7所示成功率直观地展示了模型在训练过程中性能的提升及稳定性,特别是模型确保所有AGV安全到达目标位置的能力。在每个训练周期结束时,记录所有AGV是否成功且无碰撞到达目的地,并在每100个周期结束时计算平均成功率。从图中可以观察得到,成功率最初较低,但随着训练的进行,逐渐增加至接近100%并稳定。这与收敛图的趋势一致,说明模型逐步提高了解决路径规划问题的能力,并确保了AGV的安全有效导航。

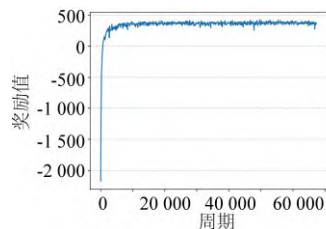


图6 收敛情况

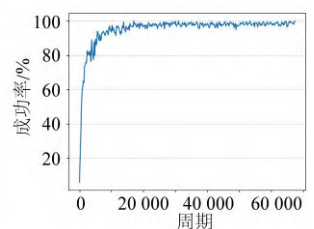


图7 成功率

损失函数如图8所示,展示了训练过程中损失值的变化趋势。损失函数采用均方误差(Mean-Square Error, MSE),衡量预测Q值与目标Q值之间的差异。损失值随训练逐渐下降并稳定,进一步验证了模型在逐步学习和性能优化方面的成效。

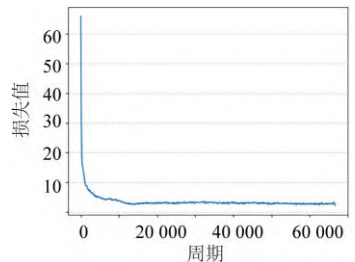


图8 损失函数

3 实验验证

3.1 实验设计

本文旨在验证CBS算法增强的分布式IQL算法(CBS-DIQL)在多AGV路径规划任务中的有效性,并将

其性能与先进的PRIMAL算法及未增强的分布式IQL(DIQL)算法进行比较。实验在一个包含30%随机分布障碍的40×40栅格地图上执行,以模拟复杂环境。为确保统计显著性,分别为4、8、12和16个AGV设置了200个独立测试用例,以评估不同规模AGV群体的表现。

性能基于两个指标评估:成功率和平均步长。成功率定义为所有测试用例中AGV全部成功到达目的地且无碰撞的比例,而平均步长记录了AGV从起点到终点的平均移动步数。为了与基准比较,采用集中式规划算法CBS计算理想情况下的步长。

此外,考虑到实际操作中的时间限制,每个测试用例的执行时间被限制在200步内,此措施旨在评估各算法在快速决策和高效路径规划方面的能力。

3.2 实验结果与分析

表2展示了不同AGV数量下CBS-DIQL算法与其他算法的成功率比较,从中可以观察到,CBS-DIQL算法在所有AGV规模中均表现出较高的成功率。具体来说,对于较小的AGV群体(4个AGV),CBS-DIQL的成功率接近100%,显著优于PRIMAL算法的84.5%和未增强DIQL的88%。随着AGV数量增加至8、12和16,尽管所有算法的成功率均有所下降,CBS-DIQL仍然保持领先。

表2 成功率对照表

AGV数量	PRIMAL /%	DIQL /%	CBS-DIQL /%
4	84.5	88	97.5
8	82	63.54	93.5
12	77.5	26.5	87.5
16	71.5	3.5	80.5

表3展示了不同AGV数量下各算法的平均步长对比。CBS-DIQL算法在所有配置中表现出更低的平均步长,显示其卓越的路径效率。在4个AGV的场景中,CBS-DIQL的平均步长较基准CBS算法仅增加了6.82%,相比之下,PRIMAL和未增强DIQL的步长分别增加了51.03%和37.28%。随着AGV数量增加,尽管所有算法步长均上升,CBS-DIQL的增幅最小,证明了其在管理大规模AGV群体时的优越路径优化能力。

表3 平均步长表

AGV数量	CBS	PRIMAL	IQL	CBS-DIQL
4	22.85	34.51	31.37	24.41
8	23.09	37.12	52.5	25.76
12	25.96	44.87	84.92	30.84
16	29.80	52.38	121.85	38.33

实验结果明确显示,CBS算法的整合显著增强了分布式IQL算法的性能,特别是在AGV数量增加时,CBS-DIQL算法展现了更稳定的成功率和卓越的路径效率。这对于实际部署AGV系统,特别是在复杂或动态环境中,

具有重要意义。这些结果不仅证明了CBS-DIQL在实际应用中的有效性,也为其在更具挑战性的环境中的应用提供了数据支撑。

4 结束语

本文提出了一种CBS算法增强的分布式独立Q学习(CBS-DIQL)算法,以最小化总行程时间为目标,旨在改进多AGV路径规划。该算法采用分布式IQL架构,提供了优异的可扩展性,并通过CBS算法引导Q网络,有效地指导AGV在拥挤环境中作出决策,从而促进其向流畅状态转移。通过与PRIMAL算法及分布式IQL算法的比较,针对不同AGV数量设计的实验结果显示,CBS-DIQL在成功率和平均步长这两个关键性能指标上表现优越,从而证实了其有效性。

本文不仅证明了CBS-DIQL算法在处理复杂环境下多AGV路径规划问题的潜力,还展示了其在未来路径规划技术发展中的应用前景,这些研究成果也为多AGV系统的路径优化提供了有力的理论和技术支持。

参考文献:

[1] REIS W P N,COUTO G E,JUNIOR O M. Automated guided vehicles position control: a systematic literature review[J]. Journal of Intelligent Manufacturing,2023,34(4): 1483-1545.

[2] ZHOU Z,XU L,QIN H,et al. A multi-AGV fast path planning method based on improved CBS algorithm in workshops[J]. Proceedings of the Institution of Mechanical Engineers,Part C: Journal of Mechanical Engineering Science,2024,238(5): 1507-1521.

[3] 刘志飞,曹雷,赖俊,等. 多智能体路径规划综述[J]. 计算机工程与应用,2022,58(20):43-62.

[4] MA H. Graph-based multi-robot path finding and planning[J]. Current Robotics Reports,2022,3(3): 77-84.

[5] SHARON G,STERN R,FELNER A,et al. Conflict-based search for optimal multi-agent pathfinding[J]. Artificial intelligence, 2015,219: 40-66.

[6] LI J,RUML W,KOENIG S. EECBS: A bounded-suboptimal search for multi-agent path finding[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2021,35(14): 12353-12362.

[7] TAO L,ZHANG S,CHEN S,et al. Multi-AGV pathfinding for automatic warehouse applications[C]//2021 China Automation Congress (CAC). IEEE,2021: 7194-7199.

[8] 刘恒麟. 多AGV路径规划算法的研究与实现[D]. 南京:东南大学,2020.

[9] SARTORETTI G,KERR J,SHI Y,et al. Primal: Pathfinding via reinforcement and imitation multi-agent learning[J]. IEEE Robotics and Automation Letters,2019,4(3): 2378-2385.

[10] MA Z,LUO Y,MA H. Distributed heuristic multi-agent path finding with communication[C]//2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE,2021: 8699-8705.

[11] HU H,YANG X,XIAO S,et al. Anti-conflict AGV path planning in automated container terminals based on multi-agent reinforcement learning[J]. International Journal of Production Research,2023,61(1): 65-80.

[12] ABED-ALGUNI B H,PAUL D J,CHALUP S K,et al. A comparison study of cooperative Q-learning algorithms for independent learners[J]. International Journal of Artificial Intelligence,2016, 14(1): 71-93.

[13] YE Y. A Review of Path Planning Based on IQL and DQN[C]// Proceedings of the 3rd International Symposium on Artificial Intelligence for Medicine Sciences. 2022: 190-195.

全生命周期智能制造试验线仿真系统对库位信息以及所生产印章所需的原料库存信息进行维护。通过建立仓库、WMS、method三个控件来定义仓储信息,仓储信息模型如图8所示。

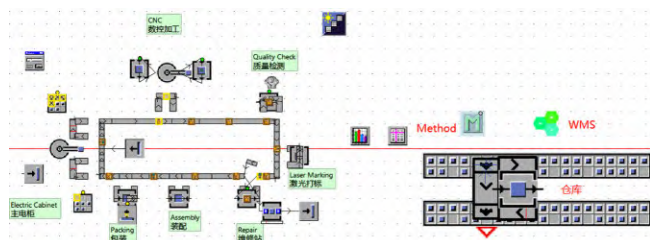


图8 定义仓储信息

在系统中定义章体、章饼和印章盖等内容的存储区域及存储策略,如先进先出,最短路径策略。仓储区域与存储策略如图9所示。

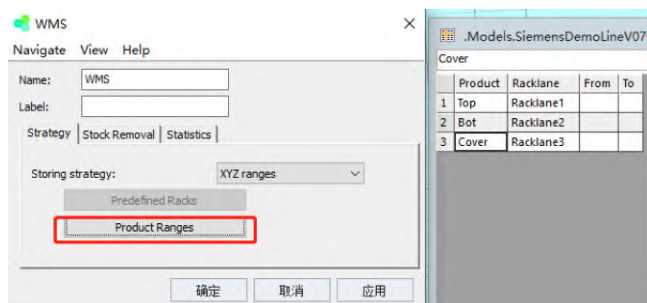


图9 定义存储区域和存储策略

6 结束语

本文提出了一种全生命周期智能制造试验线仿真平台,融合了物理资产、逻辑资产、生产过程的特点以及不同层次的需求,既考虑实体的建模方法又兼顾如何扩展,同时利用模型库的概念复用已有的信息达到快速建模的目的,并提出基于服务与应用软件进行交互的应用框架,实现上层信息系统对制造资源映射模型的有效访问。不仅实现资产等制造资源的管理,也可以通过不同虚拟模型组件的不同组合和编排实现柔性化生产仿真。基于本项目标准,不仅实现资产等制造资源的管理,也可以通过不同虚拟模型组件的不同组合和编排实现柔性化生产仿真。

参考文献:

[1] 周原令,胡晓兵,江代渝,等.基于改进NSGA-II的车间排产优

(上接第27页)

- [14] OROOJLOOY A, HAJINEZHAD D. A review of cooperative multi-agent deep reinforcement learning[J]. Applied Intelligence, 2023, 53(11): 13677-13722.
- [15] 熊丽琴,曹雷,赖俊,等.基于值分解的多智能体深度强化学习综述[J]. 计算机科学, 2022, 49(9): 172-182.
- [16] CHEN P, PEI J, LU W, et al. A deep reinforcement learning based method for real-time path planning and dynamic obstacle avoidance[J]. Neurocomputing, 2022, 497: 64-75.
- [17] 于琪,张静.基于融合任务规则优先级蚁群算法的多AGV路径规划实现研究[J]. 机电工程技术, 2024, 53(3): 190-194.
- [18] SKRYNNIK A, ANDREYCHUK A, YAKOVLEV K, et al. When to switch: planning and learning for partially observable multi-agent pathfinding[J]. IEEE Transactions on Neural Networks

and Learning Systems (Early Access), 2023: 1-14.

- [19] 黄武杰.基于深度强化学习的复杂场景下路径规划问题研究[D]. 北京: 北京交通大学, 2023.
- [20] DEGHANI M, YAZDANPARAST Z. From distributed machine to distributed deep learning: a comprehensive survey[J]. Journal of Big Data, 2023, 10(1): 158.
- 化算法研究[J]. 计算机工程与应用, 2021, 57(19): 274-281.
- [2] 王波,欧阳叶.浅谈发动机装配精益生产线的工艺设计规划[J]. 装备制造技术, 2019(9): 209-212.
- [3] 孙智超,高长水.基于Plant Simulation的汽车生产线仿真技术研究[J]. 机械制造与自动化, 2014, 43(6): 161-165.
- [4] 张浩,樊留群,马玉敏.数字化工程技术与应用[M]. 北京:机械工业出版社, 2006.
- [5] 张卫德,严洪森,徐成.基于Flexsim的生产线仿真和应用[J]. 工业控制计算机, 2005(9): 50-51.
- [6] 陈达.工业制造视域下的智能机器人生产线调度优化分析[J]. 装备制造技术, 2024(4): 144-146.
- [7] 石文玲,张喆,吴鹏,等.浅谈数字化仿真在吉利汽车焊装生产线规划上的应用[J]. 科学咨询(科技·管理), 2016(11): 37-38.
- [8] 孙智超,高长水.基于Plant Simulation的汽车生产线仿真技术研究[J]. 机械制造与自动化, 2014, 43(6): 161-165.
- [9] 彭飞,马志.基于SLP和Plant Simulation的车间工艺布局设计研究[J]. 现代制造工程, 2019(9): 28-36.
- [10] 刘煜辉,高艳强.多传感器融合的工业生产线故障预警技术[J]. 自动化与信息工程, 2024, 45(2): 41-44.
- [11] 童敏,徐锐,张剑.基于TIA博途平台的天然石膏粉生产线控制系统[J]. 自动化与信息工程, 2022, 43(1): 43-48.
- [12] 戴晓明,张洪渊,程海光,等.离散事件动态系统仿真工具DEDS Sim及其应用研究[J]. 计算机仿真, 2001(4): 38-40.
- [13] 戴晓明,冯瑞,樊向阳,等.基于模型组件技术的通用离散事件动态系统仿真模型[J]. 上海交通大学学报, 2002(4): 532-534.
- [14] 罗瑞平,盛步云,黄宇哲,等.基于数字孪生的生产系统仿真软件关键技术与发展趋势[J]. 计算机集成制造系统, 2023, 29(6): 1965-1982.
- [15] 马天牧,郑秉霖,柴天佑.面向服务的钢铁生产计划仿真平台[J]. 系统仿真学报, 2010, 22(4): 890-894.
- [16] 钟彩意,王欣,吴迪,等.四臂医疗转运机器人设计及仿真平台搭建[J]. 机电工程技术, 2021, 50(7): 20-24.
- [17] 黄凡林,万晨晖,胡友民,等.船管车间生产数据驱动仿真与调度优化[J]. 船舶工程, 2023, 45(6): 13-20.
- [18] 黄琦,刘丽兰,王森,等.基于动态数据驱动的钢铁生产仿真系统研究[J]. 制造业自动化, 2016, 38(7): 14-17.
- [19] 吴东隆,王向前.基于Flexsim模型的煤炭主生产物流系统仿真与优化[J]. 中国矿业, 2020, 29(9): 76-81.
- [20] 孙威,张浩,朱志浩,等.数字化工厂技术在生产线规划中的应用[J]. 计算机辅助工程, 2005(3): 43-46.

作者简介:

郭峻源 (1995—), 男, 硕士, 工程师, 研究领域为物流仿真与数字孪生技术, 已发表论文2篇。

牛鹏飞 (1984—), 男, 硕士, 高级工程师, 研究领域为智能制造生产管理系统开发与工厂, 已发表论文8篇。

郭志喜 (1995—), 男, 硕士, 工程师, 研究领域为离散物流仿真与规划, 已发表论文2篇。

作者简介:

王亦晨 (1997—), 男, 陕西西安人, 硕士研究生, 研究领域为生产线仿真、车间调度、AGV路径规划等。

刘雪梅 (1969—), 女, 湖南宁远人, 博士, 教授, 博士生导师, 研究领域为制造系统数字孪生、生产线规划与仿真优化、智能制造信息工程等。