

星际争霸视角的未来作战自主决策技术

黄彬城^{1,2}, 陈思^{1,2}, 高放^{1,2}, 葛建军^{1,2}, 吴雪玲³

1. 中国电子科技集团公司认知与智能技术重点实验室, 北京 100086

2. 中国电子科技集团公司信息科学研究院, 北京 100086

3. 北方自动控制技术研究所, 太原 030006

摘要 星际争霸游戏对于研究未来作战自主决策技术有重要参考价值。阐述了星际争霸游戏与自主决策过程的相似性, 分析了星际争霸战略、战术决策算法中面临的规划、学习以及不确定性等热点问题, 从决策技术复杂度入手, 讨论了未来作战自主决策技术面临的瓶颈问题, 并提出以打造大型战争游戏为手段, 重点从系统顶层架构、游戏AI建模技术、大型战争游戏引擎等关键技术出发, 试图指出未来作战自主决策技术发展着力点, 为自主决策系统的智能化技术开发和研究提供研究思路和理论基础。

关键词 智能化作战; 即时策略性游戏; 群体智能; 自主策略; 仿真博弈; 制胜机理; 人工智能

人工智能技术的发展给未来军事战争样式带来了革命性影响。无人化、智能化平台的大量使用, 给传统作战决策理念带来了方式创新。智能化战争的核心是夺取制智权, 未来作战自主决策技术是关键^[1]。未来作战自主决策技术应是在深刻理解未来战争集群化、体系化特点的基础上, 围绕以人为主导, 能够与真实战场态势实现双向映射, 辅助人、部分替代人的作战决策技术^[2]。

智能化作战残酷的战争本质虽然没有改变, 但作战域已从过去的物理空间、信息空间扩大到认知空间。“知识”的较量已成为战争主角, 其本质就是决策算法之间的较量。在当前阶段, 各类游戏依然是训练和验证智能决策技术的绝佳平台, 不仅因为

其对未来战争模拟程度非常之高, 更重要的是蕴含了当下人工智能在认知决策层面还没有很好解决的问题, 如在不完全信息下如何做推理与规划、多智能体如何协作完成复杂任务、短中长期收益平衡等, 这些基础性问题的突破对于未来群体智能研究将具有重要意义。

与未来战争相比, 目前在星际争霸中的游戏推演还是一个较为简单的过程, 但随着数据的积攒, 美国国防高级研究计划局(DARPA)预计在2040年将研制一个为军事演习各阶段提供帮助, 或在超人类层次上进行战争推演的系统。因此, 以大型即时策略游戏星际争霸为切入点, 通过研究未来自主决策技术面临的挑战, 提出打造大型战争游戏, 破解

收稿日期: 2020-04-27; 修回日期: 2020-11-03

作者简介: 黄彬城, 工程师, 研究方向为群体智能、多智能体决策, 电子信箱: huangbc1987@126.com

引用格式: 黄彬城, 陈思, 高放, 等. 星际争霸视角的未来作战自主决策技术[J]. 科技导报, 2021, 39(5): 117-125; doi: 10.3981/j.issn.1000-7857.2021.05.013

当前自主决策技术应用中缺乏数据和领域知识、科学验证评价手段等瓶颈问题,从系统架构、AI建模以及推演引擎等关键技术角度,阐述未来自主作战决策技术发展着力点,为智能化作战做好充分的认知储备。

1 星际争霸分析

1.1 概述

星际争霸是暴雪游戏公司在1998年开发的一款即时策略性大型战争游戏。主要通过在不同的地图条件下,由多名人类玩家控制3种角色,利用装备、资源以及士兵数量,以摧毁敌方大量战争资源为目的的高强度博弈对抗行为。

1.2 星际争霸决策算法

当前对星际争霸的研究主要集中在以多智能体深度强化学习方法作为分类手段,以解决特定战斗任务或者场景下的问题为牵引,旨在提升算法有效性。然而,基于算法分类方法,虽然能够提升优化算法的有效性,但为充分发挥星际争霸平台作为未来智能化战争指挥系统研制的先行者,难免会一叶障目,欠缺全局观。因此,本研究与传统按照单纯算法分类不同,拟从星际争霸任务层面,根据策略时效长短,将任务分解为2个层次,即战略类和战术类,与实际战场环境下的子任务群具有高度相似性。

战略和战术决策是实时战略游戏的两个关键方面。优秀的战略决策能帮助我方部队采集更多的资源、建立更多的战斗单位,而良好的战术决策能帮助我方部队在小规模战斗中赢得主动。

1.2.1 战略决策算法

战略决策包括3种方法,基于案例规划算法、分层规划算法以及目标驱动规划算法。

1) 基于案例的决策算法。

基于案例的决策算法是将当前状态与历史案例数据进行比对,找到适合于当前状态的最优规划结果。其中,如何降低计算空间的复杂性是主要的研究热点问题之一。2005年,有学者首次将基于案例的决策算法用于即时策略游戏,开发了一套可以基于智能体当前状态,自动求解战术和战略方案

的动态脚本。该方法面对动态对手,不需要额外的训练数据。特别是在建造顺序规划问题中,将规划空间抽象为一系列的动作,实现状态和动作空间复杂性的目的。另外一种方法是把模糊集合与基于案例的决策算法结合,与人的思维方式类似,将连续的特征集抽象为空间信息。其中空间信息可以由提取的作战单元与建筑物类型组成,然后生成5层动作规划。

2) 分层规划决策算法。

为有效降低状态空间的复杂性,分层规划决策方法可以在不同抽象程度下,单独处理子状态空间的规划问题。子目标的规划结果可以从宏观目标(如比赛胜负)到微观动作规划(如作战单元动作),并能很好地映射到全局规划结果中。分层任务网络结构构建是当前研究热点之一,是将复杂的任务分解为一系列顺序的简单子任务进行求解。因此,分层规划决策算法通常与其他方法结合使用。

3) 目标驱动决策算法。

目标驱动决策算法是根据智能体的任务目标及周围环境变化,构建定量变化关系模型。在星际争霸中,使用行为语言,搭建可以预测结果输出的目标驱动决策系统。当意外情况发生时,系统不但可以记录异常状态,还能针对异常状态,对系统做出反馈调整。显然,该方法在初始状态不能通过自主学习实现迭代更新,只能通过手动赋值完成系统初始化。应用于AIIDE星际比赛中,只击败了52%的人类选手。针对不能自主学习的问题,将基于案例规划法与强化学习结合,可以实现系统目标和领域知识的动态变化,迭代更新。这意味着人类可以无需特别指定目标、状态以及其他领域知识让系统自主学习。同样地,如果领域知识发生变化,系统也会自动适应。不足的是,如果领域知识过于简单,那么该方法的准确性不如目标驱动决策算法。

1.2.2 战术决策算法

战术决策是战略决策的重要保障,只有良好的战术决策,才能应对敌方骚扰和小规模进攻,为我方的宏观经济发展赢得时间。当前战术决策主要包括典型战术决策、建造顺序预测以及智能体反应控制3方面。

1) 典型战术决策算法。

典型的战术决策算法研究对象可以分为己方和敌方两方面。其中敌方研究较少,仅以经济状况建模方法以及对手策略预测为主,而已方研究成果较为丰富,涵盖了协同编队控制、协同攻击决策、整体对抗策略生成以及智能体快速运动控制等方面。

在协同编队控制方面,2011年,台湾中正大学

研究院提出一种基于遗传算法的编队策略。首先对战术编队参数化,然后使用遗传算法对参数进行优化,将多智能体编队分为固定式和参数式两类,其中固定式又分为直线分布、柱型分布以及V型分布;参数式通过定义相邻智能体的距离 α, β 实现,如图1所示。经实验验证,只有在小型平坦陆地环境下,控制效果良好^[3]。

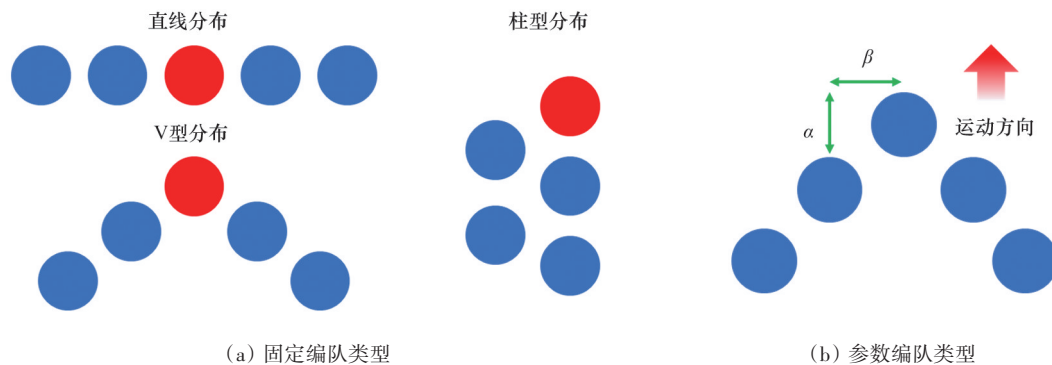


图1 编队控制结构示意图

在协同性方面,法国研究人员提出一种生成贝叶斯模型,可同时进行预测敌方攻击行为和对自己方的协同决策方法。该模型甚至还可以处理不确定性敌人信息。模型参数通过提取8000份1v1的比赛录像训练得到^[4]。显然,该方法如果在实战中应用,必须有交战数据库信息支持。随后,针对游戏中的巨大操作空间以及快速决策需求,韩国世宗大学提出基于人类选手比赛数据,使用模仿学习方法,实现AI的模仿行为。智能体根据录像,自动搜索与录像地图相比,得到优化移动路径^[5]。提高协同性的前提是对任务目标的有效分类,分布式多智能体系通常受到不确定性环境信息和指挥网络适用性等条件约束,如智能体的生命值、运行状态、目标状态以及周围环境信息等。2018年韩国西江大学又提出一种新的分布式任务分配策略,分配模型由所有智能体共享,通过网络连接相互传递任务信息,最终实现复杂任务的分配功能^[6]。

当前,计算复杂性制约着新方法的应用前景,特别是在星际争霸这类大型即时策略类游戏中。2017年,中国科学院自动化研究所针对巨大的状态和动作空间,构建了一种有效的状态表示方法以

降低计算复杂性。该方法考虑了独立单元状态和距离信息。为求解状态空间的稀疏性和延迟回报问题,引入每个时刻分步回报构成的新型回报函数。该回报函数能够平衡智能体攻击以及运动行为,并能鼓励智能体以团队形式完成任务。最后使用梯度下降法训练学习模型,让智能体通过战斗学习提高集群的协同性^[7]。

在整体对抗策略方面,巴西研究者探索了在非固定场景下,多智能体根据历史信息,选择对抗算法问题。在星际争霸实验中发现,最大最小 Q 值法特别适用于非固定场景以及固定场景下算法选择方法。搭载该算法的MegaBot在非固定场景中,表现出一定的鲁棒性。在2016年的AIIDE比赛中,排名达到前40%^[8]。2016年,美国内华达大学研究了利用启发式搜索算法生成高质量微观操作策略,以基本地图和隐藏地图实现对智能体的短期部署,智能体微观操作参数达到14个^[9]。

在智能体快速运动控制方面。2011年,荷兰学者率先提出一种使用神经网络方法在典型场景下的智能体控制策略。动作空间分为进攻和逃离两种情况,使用在线深度强化学习算法和短期记忆

回报函数构建控制模型,然后使用增量学习法进行大场景智能体训练,最后开发了一种使用单独视觉神经网络将环境数据提取输入至神经网络中的方法^[10];2013年,针对当前智能体在星际游戏中响应速度慢,仅限于局部应用等问题,提出一种每秒可控制百万个智能体算法^[11],随后,美国内华达大学又使用样本注入遗传算法实现高质量控制智能体行为^[12]。类似的,日本立命馆大学利用模糊积分和启发式搜索策略提高了对智能体运动控制的质量^[13]。

当前对敌方的研究较少,主要集中在对敌方经济状况预测以及敌方策略预测两方面。其中,纽约大学使用低维度经济学模型,定义资本/劳动比率 k ,科技/劳动比率 t ,使用基因算法,自动学习优化参数 w ^[14],实现对敌方经济状态的预测;也可通过建立非典型性对手策略数据库,使用机器学习在实战中预测对手策略^[15]。

通过以上分析发现,对于时间维度,研究重点逐步从己方的决策控制问题向敌方策略预测过渡;从研究对象上,小规模同构智能体向大规模异构智能体过渡。虽然取得了一定成果,但当前研究主要以理想固定场景为主,特定智能体为研究对象,对于混合异构智能体、大规模场景,特别是对不完全信息条件下的自主决策研究较少。

2) 建造顺序预测算法。

建造顺序是决定战争成败的核心因素。建造顺序包括3方面:建筑物种类、建造数量以及建造位置。建造顺序与己方经济发展速度、应对敌方战术袭扰密切相关。从研究方法上可以分为利用比赛录像和算法开发两类。

在研究方法方面,由于战争迷雾的存在,快攻是对手策略的典型特点。双方通常会派侦察兵探测对方的经济和军力发展情况。职业玩家会根据战况即时调整建造顺序,但对于AI机器人,其建造策略就比较固定。因此,可以使用比赛录像来预测对手,并决定实际比赛中的我方建造顺序。实验结果显示该方法能够准确预测对手策略,提高了比赛胜率^[16]。考虑使用神经网络Q学习(NNFQ)和卷积神经网络Q学习(CNNFQ)的方法,解决建造顺序以及建造类型手动编码缺陷问题^[17]。从提高快速

响应性出发,可将建造单元参数使用参数矩阵表示,对矩阵进行最优化求解就可得到最佳建造顺序^[18]。

3) 反应控制决策算法。

反应控制是指利用不同算法对作战单元实现移动、攻击、侦察、撤退功能。在算法层面,主要有基因算法、蒙特卡罗算法、势流法、贝叶斯法和强化学习。

在作战单元移动方面,根据敌方目标的数量和类型,可使用势流法实现对作战单元的编队控制。一旦作战单元发现目标,会与其他单元协同运动至目标距离一定范围处,对目标发起攻击^[19]。

在作战单元攻击方面,可简化为多目标连续优化问题,基于真实参数,构建仿真作战单元环境,使用基因算法控制作战单元实现最大化攻击和最小化伤害^[20]。而且能够实现作战单元依据环境条件,选择合适的动作行为,如打击、距离优先(打击距离近的单位)、打击优先(打击低血单元)等模式^[21]。强化学习技术可使作战单元能够根据敌方的行为做出相应决策,用6个参数表示状态空间:生命值、能够被攻击的敌方单元、可攻击己方的敌方单元、敌方攻击状态、相邻作战单元的距离、敌方生命值。作战单元动作分为攻击、支援和撤退,研究发现作战单元能够在不同的游戏阶段做出最佳决策。但该方法不具有通用性,仅适合特定场景下的决策^[22]。

在侦察方面,根据敌方目标的数量和类型,使用势流法对作战单元进行编队控制。一旦作战单元发现攻击目标,会与其他单元协同运动至距离目标一定范围处,对目标发起攻击^[23]。

未来战争制智权的夺取,决策算法是关键。星际争霸由于具有战争迷雾,以及作战空间和响应速度复杂等特点,与实际作战过程比较接近,是未来战争决策算法的孵化器。通过分析对星际争霸算法研究,对于搞清多智能体博弈对抗决策机理,以及未来作战自主决策具有重要意义。然而,除去算法本身之外,未来作战自主决策技术的落地应用,仍然面临一些挑战。

2 未来作战自主决策技术面临的挑战

2.1 未来作战自主决策技术AI复杂度分析

AI技术的发展给未来作战自主决策技术赋予新的发展前景^[24]。任何依靠AI手段解决实际问题,都可由算法选择、模型建立以及训练迭代3个步骤组成。以DeepMind的AlphaStar为例,研究者将游戏全程看作是一系列智能体进行长序列建模学习任务求解的过程,并将作战分解为子任务,各作战单位列表和属性作为子任务模型输入,输出作战单位动作指令。算法模型由一个神经网络模型、深度LSTM模型、具有指针网络的自动回归策略模型以及中心化评分基准模型混合而成。训练方法采用监督学习和自我博弈结合的方式,初期与AlphaGo类似,以历史比赛录像作为数据,通过监督学习方法让模型快速学习基础策略和微观操作技巧;中后期则以自我博弈对战为主,最后利用纳什分布采样得到最终智能体策略。

由以上分析可见,星际争霸决策问题通常是将复杂任务“分解”为单独任务,有针对性地进行求解。这种“分解”方法是否适合于实际战场的自主决策,则取决于算法对象的复杂性。AI算法模型对样本数量和复杂度具有强烈的依赖性,学习对象越复杂,对样本数量要求也越高。星际争霸地图尺寸通常为 256×256 ,每个玩家最多能控制30~35种建筑物种类以及200个作战单元。理论上,任意时刻就会有50~400个作战单元在地图上,如果仅考虑作战单元可能的移动位置,400个单元的复杂度就达到 $65536^{400} \approx 10^{1926}$,如果要考虑攻击类型、能源以及科技研发等因素,那么复杂度会更高,这也是星际争霸AI中采用任务分解求解方法的原因。

与星际争霸相比,真实战场的复杂度要高出多个数量级。一是作战单元种类和数量更多,与星际争霸200个作战单元相比,实际战场中的作战单元种类与数量都庞大得多;二是作战单元行为控制更加复杂,在游戏中,作战单元行为控制可通过键盘或者鼠标点击实现,而真实战场中作战单元的控制,往往需要设定大量的控制参数完成;三是作战域更加复杂,星际争霸作战域仅由陆、空组成,而随

着智能化战争的到来,真实作战域已扩展为海、陆、空、天、网、电磁以及意识域;四是评价标准更为复杂,与游戏相比,真实作战效果的评价结果不但有输赢,更包括一系列复杂的评价内容。

由此可见,未来作战自主决策技术具有很高的复杂性,并且真实战场数据的获取主要受限于成本,不能像星际争霸或者围棋那样相对容易地获取训练数据,所以自主决策AI所需的训练样本数量和质量都无法全面保证。因此,未来作战自主决策技术也可以尝试“分布式解决”模式,分解为具体的子问题,针对每个子问题运用对抗网络、模仿学习以及元学习等手段进行策略求解,最后将各个子策略进行集成。

2.2 未来作战自主决策技术瓶颈问题

通过研究分析,未来作战自主决策技术发展主要遇到3个瓶颈问题。

一是缺乏足够的领域知识和样本数据。战场样本数据通常是指在实际作战、演习训练以及仿真过程中,产生的情报、通信以及作业数据等。这些数据需要清洗加工和标注标签等工序才能够用在学习模型中。样本质量一方面是指数据本身的好坏,如分辨率、是否有噪声等,另一方面是指数据标签的准确性。当前自动化标注方法虽能够实现大量流水作业,但面对战场纷繁复杂的态势信息,与军事领域专家相比,灵活性不强。另外,当前数据的产生主要依靠演习训练和仿真试验,考虑到人员安全、成本等因素,战术对抗激烈程度和灵活程度以及装备使用程度,均不如实战,那么产生的数据自然会和实际作战场景具有相当差距。这也是当前用于模型训练的数据数量缺乏、形式单一、样本分布不均衡的原因。

二是缺乏科学的验证评价技术手段。不加验证的技术是把双刃剑,既会大力促进技术的发展,也会在实践中带来惨痛的教训。未来作战自主决策技术效果的好坏本质上是由算法模型质量决定的。模型通常需要使用大量的数据,以测试其泛化能力和可靠性。显然,如果将模型放在实际战场中去测试,代价太大,无法实现。因此,目前针对未来作战自主决策技术的效果验证,缺乏统一的量化指

标和标准,如态势感知准不准、战略决策对不对等问题。就算面对实际场景,这些问题都很难有统一的量化衡量标准。

三是缺少多学科融合人才。AI技术如同电流一样,用在不同的场景下,会产生不同的附加值。如电流经过点灯会带来光明,而流过计算机则会带来算力的提升以及生产力的解放。显然,“AI+X”才是未来作战自主决策技术的正确打开方式。在未来作战自主决策技术场景下,精通军事作战理论、熟悉装备性能、明确交战规则的领域专家,与掌握机器学习模型算法、强化学习训练技巧、深度学习网络构建经验的AI技术专家相结合,才能找到切入点和发力点,真正发挥未来作战自主决策技术效能。

3 基于大型战争游戏的未来作战自主决策技术发展着力点

本质上,未来作战自主决策技术的瓶颈问题是“人”和“物”的矛盾关系。缺乏军事决策技术领域与人工智能算法跨领域专家,这个“人”,以及可以有效训练学习模型数据,这个“物”。游戏不仅是游

戏,更是对未来战争的演练。通过对DARPA的“深绿”计划以及美国军方专用训练游戏“武装突袭”的研究发现,国外早已基于大型仿真游戏实践,提升士兵作战经验以及训练数据模型。大型战争游戏由于具有安全、经济和可定制等特点,适合作为自主作战决策技术发展的孵化器。因此,本研究从系统顶层架构、大型战争游戏AI建模技术、大型战争游戏引擎等关键技术出发,试图阐述未来作战自主决策技术发展着力点。

3.1 未来作战自主决策技术智能系统架构

未来作战自主决策技术智能系统架构如图2所示。主要由未来作战自主决策技术智能系统和作战单元两部分构成。未来作战自主决策技术与作战单元系统相互配合、相互协同,未来作战自主决策系统发出决策指令,作战单元执行,并将执行效果反馈给未来作战自主决策系统进行状态评估。所以,本文虽然重点讨论未来作战自主决策技术架构,但为更好描述作战回路,将作战单元也一并讨论。

未来作战自主决策系统主要由分布式虚拟资源层、中间件系统层、游戏推演引擎层、智能AI模型层、行动管理单元层、态势感知方案层、火力攻击方案层、行动协同方案层、用频策略方案层、综合效能评估单元和显示控制终端。

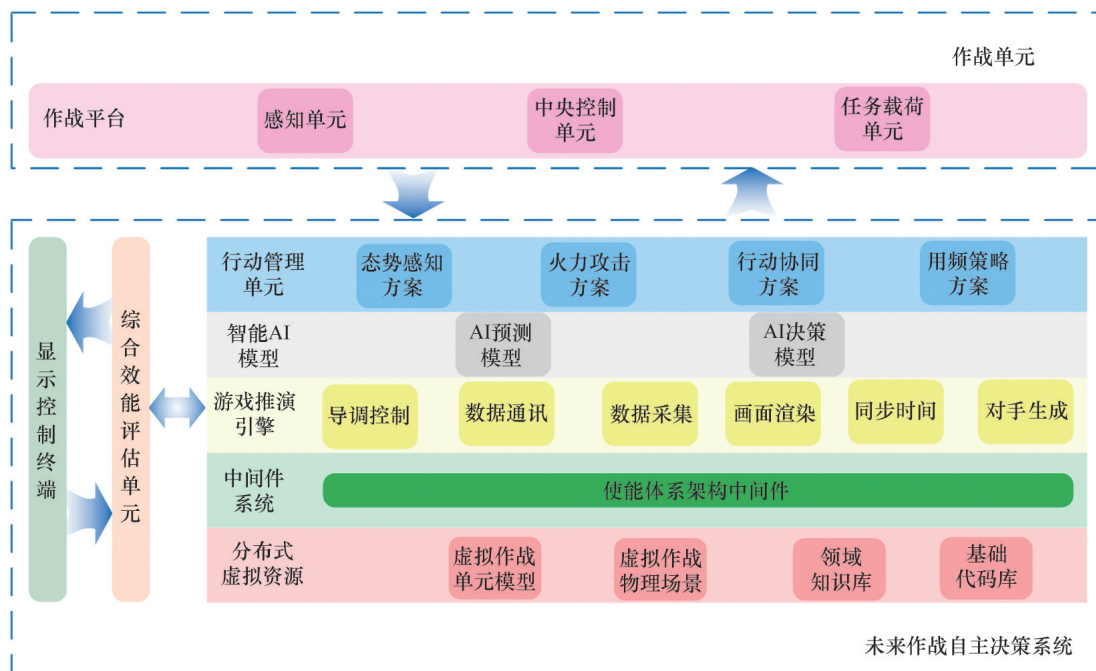


图2 未来作战自主决策技术架构

控制终端构成。其基本运行原理是通过对复杂战场环境进行建模,将战场资源虚拟化,构建虚拟-现实的实时映射关系,根据战场实际态势,并基于构建的虚拟化战场模型,利用AI预测、决策算法,得到适合于当前战场情况下有利于我方的作战决策行为。其中虚拟资源构建的真实程度,直接决定了智能AI模型的计算准确程度。

3.2 未来作战自主决策技术大型战争游戏AI虚拟建模技术

当前诸多大型战争游戏建模技术已较为成熟,具备精细化建模的能力,只是进一步提高逼真程度的问题。然而,该类游戏的目的在于提升玩家的体验感,除具备一定的物理模型外,更加侧重于对模型的渲染效果和外观逼真程度,与实战应用还有相当距离。一方面缺乏包含作战规则、裁决模型等决定实际战争效果的军事应用规则模型的嵌入,另一方面缺乏仿真模型与实际战场环境的交互能力,实际战场环境瞬息万变,只有通过平行仿真,才能量化地估计未来战场在某一时刻的战场态势。

从担负的任务和规模看,可将未来作战自主决策技术分为战术级、战区级、战略级3个层次。结合当前领域专家一致认为相当长的时间内,AI并不能完全替代人决策这个基本前提,考虑决策过程中的复杂度和算力作为约束条件,可先从战术级入手,在当前游戏建模技术的基础上,从作战概念模型和推演模型体系两方面,搭建适合于仿真推演的

虚拟模型^[29]。其中,作战概念模型包括认知过程模型、能力层级模型以及OODA环模型;推演模型体系包括作战决策模型、作战决策关系模型、侦察探测模型、通信传输模型、信息融合模型、兵力编成模型、交战模型以及战场环境模型等。

3.3 未来作战自主决策技术大型战争游戏推演引擎技术

由于战场态势的复杂度巨大,从而给战争游戏模拟带来巨大挑战,特别是多尺度推演(战略、战术和系统并行模拟)问题。对于战争游戏引擎而言,更重要的是应从设计理念、引擎模式、应用方式3方面进行技术研发。其中,设计理念是指一款战争游戏涵盖不了战争的方方面面,只能聚焦主要矛盾和矛盾的主要方面,这就要求以聚焦作战概念(比如敏捷精确作战、多域作战、蜂群作战等)为核心,针对当前引擎架构设计基本方法,突出显示需要修改开发的地方,支持作战概念,设计开发战争游戏引擎;引擎模式可以分为人在回路和人不在回路两种,人在回路是指作战人员可以指挥各个单元以及战场态势,在线实时修改游戏模型。人不在回路是指整个游戏推演过程完全由AI模型自主完成。应用方式是指战争游戏引擎可以支持调试运行模式,根据指战员训练、推演以及实战需要,根据战场全局的某个片段、某个子任务进行重复调用、重复训练指挥,推演引擎技术逻辑框架如图3所示。

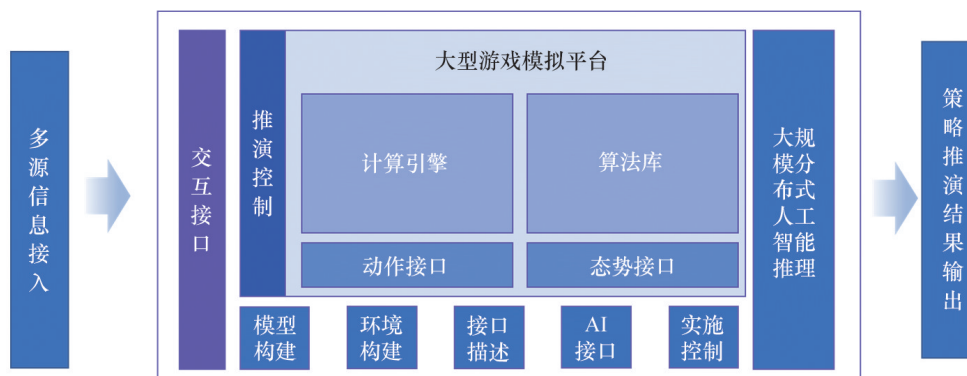


图3 未来作战自主决策技术大型战争游戏推演引擎技术逻辑框架

4 结论

星际争霸游戏是决策技术应用的一个重要领域,是算法开发、模型验证的重要平台,特别是近年来的 AAAI/AIHIDE、CIG 和 SSCAIT 等星际赛事吸引了学术界和工业界的持续研究与开发。与真实战场的复杂性和多变性相比,星际争霸直接应用于作战自主决策仍然有较大差距,但其具有高实时性、战争迷雾、状态和动作空间巨大等特点,在战场数据生成机理、战略战术算法开发和验证方面,能够为作战自主决策提供理论指导和技术支撑。并且博弈论将作为未来作战自主决策技术的理论牵引,模仿学习、迁移学习以及增量学习的探索,都会给该领域带来新的研究思路。

游戏不仅仅是游戏,更是对未来战争的演练。针对未来作战自主决策瓶颈问题,开展了星际争霸决策算法、未来作战自主决策技术复杂度分析以及瓶颈问题等研究,提出了基于大型战争游戏的未来作战自主决策技术的方法,突破智能系统架构、AI 虚拟建模方法以及游戏推演引擎构建等关键技术,形成未来自主作战算法生成的孵化器,为打通游戏环境与实际战场应用的藩篱提供理论基础。

参考文献(References)

- [1] 胡晓峰. 军事指挥信息系统中的机器智能: 现状与趋势[J]. 人民论坛学术前沿, 2016(15): 22-34.
- [2] 陶九阳. AlphaGo 技术原理分析及人工智能军事应用展望[J]. 指挥与控制学报, 2016, 2(2): 114-120.
- [3] Lin C S, Ting C K. Emergent tactical formation using genetic algorithm in real-time strategy games[C]//International Conference on Technologies & Applications of Artificial Intelligence. Taiwan: IEEE, 2012.
- [4] Synnaeve G, Pierre Bessière. Special Tactics: A bayesian approach to tactical decision-making[C]//Computational Intelligence and Games (CIG), 2012 IEEE Conference on. Granada: IEEE, 2012.
- [5] Oh I S, Cho H C, Kim K J. Imitation learning for combat system in RTS games with application to starcraft[C]//Computational Intelligence & Games. Dortmund: IEEE, 2014.
- [6] Park S, Park S, Lee H, et al. Collaborative goal distribution in distributed multiagent systems[C]//IEEE International Conference on Robotic Computing. Laguna Hills: IEEE Computer Society, 2018.
- [7] Shao K, Zhu Y, Zhao D. Cooperative reinforcement learning for multiple units combat in starCraft[C]//2017 IEEE Symposium Series on Computational Intelligence (SSCI). Honolulu: IEEE, 2017.
- [8] Rocha Tavares A, Vieira D K S, De Oliveira T N, et al. Algorithm selection in adversarial settings: From experiments to tournaments in StarCraft[J]. IEEE Transactions on Games, 2018, 11(3): 1-1.
- [9] Liu S, Louis S, Ballinger C. Evolving effective micro behaviors in real-time strategy games[J]. IEEE Transactions on Computational Intelligence and AI in Games, 2016, 12(3): 1-1.
- [10] Shantia A, Begue E, Wiering M. Connectionist reinforcement learning for intelligent unit micro management in StarCraft[C]//Neural Networks (IJCNN), The 2011 International Joint Conference on. San Jose: IEEE, 2011.
- [11] Churchill D, Buro M. Portfolio greedy search and simulation for large-scale combat in starcraft[C]//Computational Intelligence in Games (CIG), 2013 IEEE Conference on. Niagara Falls: IEEE, 2013.
- [12] Liu S, Louis S J, Nicolescu M. Using CIGAR for finding effective group behaviors in RTS game[J]. 2013, 15(2): 1-8.
- [13] Nguyen T, Nguyen K, Thawonmas R. Integrating fuzzy integral and heuristic search for unit micromanagement in RTS games[C]//2014 IEEE Congress on Evolutionary Computation (CEC). Beijing: IEEE, 2014.
- [14] Bryan S, Weber. Standard economic models in nonstandard settings-starCraft: Brood war[C]//Computational Intelligence in Games(CIG), 2018 IEEE Conference on. Maastricht: IEEE, 2018.
- [15] Certicky M, Sarnovsky M, Varga T. Use of machine learning techniques in real-time strategy games[C]//World Symposium on Digital Intelligence for Systems and Machines, Kosice: IEEE, 2018: 159-164.
- [16] Cho H C, Kim K J, Cho S B. Replay-based strategy prediction and build order adaptation for StarCraft AI bots [C]//Computational Intelligence in Games. Niagara Falls: IEEE, 2013.
- [17] Tang Z, Zhao D, Zhu Y, et al. Reinforcement learning for build-order production in StarCraft II[C]//2018

- Eighth International Conference on Information Science and Technology (ICIST), Cordoba: IEEE, 2018: 1–4.
- [18] Takino H, Hoki K. Human-like build-order management in StarCraft to win against specific opponent's strategies[C]//International Conference on Applied Computing & Information Technology/ International Conference on Computational Science & Intelligence. Okayama: IEEE Computer Society, 2015.
- [19] Chen C Y, Liao X L, Liao C C, et al. Pattern formation based on potential field in real-time strategy games[C]// Technologies and Applications of Artificial Intelligence (TAAI), 2012 Conference on. Tainan: IEEE Computer Society, 2012.
- [20] Horn J, Nafpliotis N, Goldberg D E. A niched pareto genetic algorithm for multi-objective optimization[C]//Evolutionary Computation, 1994. IEEE World Congress on Computational Intelligence. Proceedings of the First IEEE Conference on. Orlando: IEEE, 1994.
- [21] Hsu W L, Chen Y P. Learning to select actions in starcraft with genetic algorithms[C]//2016 Conference on Technologies and Applications of Artificial Intelligence (TAAI). Hsinchu: IEEE, 2016.
- [22] GPB Neto, CDA Siebra. Evolving the behavior of autonomous agents in strategic combat scenarios via SARSA reinforcement learning[C]//2014 Brazilian Symposium on Computer Games and Digital Entertainment, Porto Alegre: IEEE Computer Society, 2015, 12–19.
- [23] 陈承裕. 基于势的实时策略博弈模式生成[D]. 台湾: 中正大学, 2013.
- [24] 郭圣明, 贺筱媛, 胡晓峰, 等. 军用信息系统智能化的挑战与趋势[J]. 控制理论与应用, 2016, 3(12): 1562–1571.
- [25] 杨瑞平, 张兆峰. 指挥控制系统仿真[M]. 北京: 国防工业出版社, 2013.

On future combat autonomous decision technology for starcraft

HUANG Bincheng^{1,2}, CHEN Si^{1,2}, GAO Fang^{1,2}, GE Jianjun^{1,2}, WU Xueling³

1. Key Laboratory of Cognition and Intelligence Technology, China Electronics Technology Group Corporation, Beijing 100086, China
2. Information Science Academy, China Electronics Technology Group Corporation, Beijing 100086, China
3. CNGC North Automatic Control Technology Institute, Taiyuan 030006, China

Abstract StarCraft is an important game for studying the future combat autonomous decision technology. Similarities between StarCraft and the autonomous decision process are described. Planning, learning, and uncertainty in decision-making algorithms for StarCraft are also analyzed. Firstly, the key problem of future combat autonomous decision-making technology is discussed in terms of decision complexity. Then, the article proposes to create a large-scale war game to clarify the development of future battle autonomous decision-making technologies, such as system's top-level architecture, game AI modeling technology, large game engines, etc. in order to provide a useful reference for the development of autonomous decision system intelligent technology.

Keywords intelligent operation; real time strategic games; swarm intelligence; autonomous strategy; simulation game; winning mechanism; artificial intelligence ●



(责任编辑 傅雪)