

1-4-2019

Intelligent Technology of Command and Control System in the RTS Perspective

Wenfeng Wu

1. Joint Operations College Joint Operations Training Center, National Defense University, Beijing 100091, China; ;2. Military Operations Teaching and Research Office, Army Artillery and Air Defense Academy, Hefei 230031, China;

Zhang Yu

1. Joint Operations College Joint Operations Training Center, National Defense University, Beijing 100091, China; ;

Rong Ming

1. Joint Operations College Joint Operations Training Center, National Defense University, Beijing 100091, China; ;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Intelligent Technology of Command and Control System in the RTS Perspective

Abstract

Abstract: Real-Time Strategy (RTS) games have important reference value for studying the intelligent technology of command and control systems. The similarities between RTS games and the strategic battle level command and control systems are described according to the decision process. *The challenges brought by the problems of planning, learning, uncertainty and space-time reasoning in the intelligent technology of RTS games are analyzed. The key technologies and latest research progress of action sequence planning, plan recognition, state assessment, multi-agent collaboration and multi-scale AI are studied. The trend of intelligent technology development of strategic and operational level command and control systems is pointed out.* It provides a useful reference for the development and research of intelligent technology for command and control systems.

Keywords

Real Time Strategy, Command and Control System, Artificial Intelligence, Task Planning, Deep Learning

Recommended Citation

Wu Wenfeng, Zhang Yu, Rong Ming. Intelligent Technology of Command and Control System in the RTS Perspective[J]. Journal of System Simulation, 2018, 30(11): 4158-4171.

基于 RTS 视角的指挥控制系统智能化技术

伍文峰^{1,2}, 张昱¹, 荣明¹

(1. 国防大学联合作战学院联合作战演训中心, 北京 100091; 2. 陆军炮兵防空兵学院军政基础系军事运筹教研室, 安徽 合肥 230031)

摘要: 即时战略(Real-Time Strategy, RTS)游戏对于研究指挥控制系统智能化技术有重要参考价值。针对指挥决策过程阐述了 RTS 游戏与战略战役级指挥控制系统的相似性, 分析了 RTS 游戏中智能化技术面临的规划、学习、不确定和时空推理等问题带来的挑战, 研究了行动序列规划、敌方规划识别、状态评估、多智能体协作和多尺度 AI 等关键技术及最新研究进展, 指出了战略战役级指挥控制系统智能化技术发展的趋势, 为指挥控制系统的智能化技术开发和研究提供了有益借鉴。

关键词: 即时战略游戏; 指挥控制系统; 人工智能; 任务规划; 深度学习

中图分类号: TP391.9 文献标识码: A 文章编号: 1004-731X(2018)11-4158-14

DOI: 10.16182/j.issn1004731x.joss.201811014

Intelligent Technology of Command and Control System in the RTS Perspective

Wu Wenfeng^{1,2}, Zhang Yu¹, Rong Ming¹

(1. Joint Operations College Joint Operations Training Center, National Defense University, Beijing 100091, China;
2. Military Operations Teaching and Research Office, Army Artillery and Air Defense Academy, Hefei 230031, China)

Abstract: Real-Time Strategy (RTS) games have important reference value for studying the intelligent technology of command and control systems. The similarities between RTS games and the strategic battle level command and control systems are described according to the decision process. *The challenges brought by the problems of planning, learning, uncertainty and space-time reasoning in the intelligent technology of RTS games are analyzed. The key technologies and latest research progress of action sequence planning, plan recognition, state assessment, multi-agent collaboration and multi-scale AI are studied. The trend of intelligent technology development of strategic and operational level command and control systems is pointed out.* It provides a useful reference for the development and research of intelligent technology for command and control systems.

Keywords: Real Time Strategy; Command and Control System; Artificial Intelligence; Task Planning; Deep Learning

引言

指挥控制系统是信息化战争的必备手段, 虽然不直接完成火力打击、侦察和保障等任务, 但却是信息作战的倍增器, 地位作用日益凸显。美军对指

挥控制领域的智能化技术研究一贯非常重视, 从上世纪 70 年代的决策支持系统, 到 2007 年的深绿计划(Deep Green), 再到近期 DARPA 启动的 Insight、XDATA、Big Mechanism、Deep Learning、DEFT、PPAML 等大量基础技术研究项目和 Mind's Eye、TRACE、DBM、TEAM-US 等一系列面向实际作战任务背景的项目, 可以看出其已经从计算智能、感知智能发展到了认知智能^[1], 达到了一个前所未有的新高度。



收稿日期: 2018-05-31 修回日期: 2018-06-26;
基金项目: 中国博士后科学基金(2017M613318);
作者简介: 伍文峰(1982-), 男, 湖北钟祥, 博士, 讲师, 研究方向为体系评估、体系建模与仿真; 张昱(1979-), 男, 四川南充, 博士, 讲师, 研究方向为战争复杂系统建模与仿真、智能化指挥与控制。

<http://www.china-simulation.com>

• 4158 •

1 RTS 游戏和战略战役级指挥控制系统的关系

指挥控制系统有多种不同的分类方法,按照作战层次的不同,可以分为格斗级、战术级、战役级和战略级。不同层次关注的对象和问题不同。格斗级指挥控制系统主要针对单兵、单机或单舰等单个武器装备或简单武器装备系统之间的对抗,重点研究对格斗行为的观察、判断、决策和行动等的指挥控制;战术级重点关注战术作战行动中的行动路线选择、兵力分配、计划安排等问题的智能指挥与控制,比格斗级指挥控制的兵力数量和行动类型更多、态势也更加复杂;战略战役级则强调在更长的时间尺度和更大的空间尺度上综合考虑,决策过程定性成分较多,往往难以用精确地数值或定量模型进行描述。例如对于战略级指挥控制系统而言,需要考虑军事与经济、外交、科技、政治等的关系,是优先发展科技,还是发展军事,或是经济?对于战役级指挥控制系统而言,需要统筹考虑地域、时间、武器装备、作战力量等之间的协同。同时,因为决策是在对抗的环境中,敌对双方或多方之间会动态调整各自的策略,还可能会故意隐瞒自己的真实意图,所以决策是一个动态博弈过程,存在很大的不确定性,这些都对智能化战略战役级指挥控制系统的发展带来了巨大的挑战。本文主要研究战略战役级指挥控制系统。

游戏的特点决定了其对于人工智能技术研究有非常重要的参考价值:一是游戏一般采用固定规则集,且有环境约束,提供了接近真实世界的复杂测试环境,同时又没有真实世界中多种设施造成的混乱,可以专注于特定算法、特定环境的研究,研究出的结果可以应用于解决更复杂的真实世界中的问题;二是开放的游戏平台拥有无限的训练和比赛数据,记录了学习各种游戏玩家(包括不同等级的 AI 和人类玩家)的历史经验,便于研究人员进行并行的测试和研究;三是人工智能已经被广泛应用于棋牌类游戏(象棋、围棋、德州扑克、麻将)和对战电子游戏(街霸、Dota、Atari、CS 等),推动了

许多人工智能技术的开发。

即时战略游戏(Real-Time Strategy Game),简称 RTS 游戏,是策略游戏(Strategy Game)的一种。早在 2016 年 AlphaGo 取得成功时,DeepMind 公司就宣布将星际争霸游戏作为下一个挑战对象。之所以基于 RTS 游戏视角来研究指挥控制系统的智能化技术,是因为它与战略战役级的决策过程非常相似^[2-4]。

(1) 多方同时操作。通常在同一时间有多名玩家同时操作,这些操作都需要消耗时间,例如采矿需要派出采矿车,到达矿藏地点,返回矿石加工厂等,类似于战场上多方行动的独立性。

(2) 游戏实时进行。每个玩家决策的时间都是有限的,以确定下一部动作。与传统棋类游戏不同,在星际争霸中,游戏以每秒 24 帧的速度进行,即在游戏态势发生改变之前,只有 42 ms 的时间进行决策,游戏中的决策压力与实际作战过程非常类似,每一步的动作类似于指挥员制定下一步作战行动的决心。

(3) 不完全信息博弈。星际争霸具有“战争迷雾”,在双方没有直接交锋前,玩家只了解己方和友方力量范围内的环境,对敌方范围内或未侦察的区域中的经济、科技和军事发展的细节不清楚,需要派出专门的兵力侦察后才能掌握相应情况。Dota 在战争迷雾的基础上还有开雾、隐刀、真假眼等高级视野缠斗技巧。

(4) 强调长远战略与短期战术、格斗的协调。RTS 游戏中各类决策的影响长短不一,有的决策要过很长时间才能体现重要性,水平不高的甚至看不懂。例如,在星际争霸中,高水平的选手都会 hit-and-run (打了就跑)的骚扰战术,甚至上升到战略层面,如人族的雷车提速骚扰等。

(5) 状态空间和动作空间巨大。在每个决策周期中状态空间和可用的行动数量都很大,例如象棋的状态空间大约在 10^{50} 左右,德州扑克大约在 10^{80} 左右,Go 游戏大约在 10^{170} 左右,在典型环境下的星际争霸的状态空间达到了 10^{1685} ,比棋牌类游戏

高了很多个数量级。

当前军事指挥控制领域的人工智能技术发展很快,在格斗级和战术级等较低层次出现了基于遗传模糊树的 Alpha AI^[5]和基于增强学习的典型应用,但在战略战役级等较高层次仍然面临很多困难,公开的资料中还未见到能实用的系统^[6]。因此,本文期望通过对 RTS 游戏中人工智能技术的研究,重点围绕星际争霸游戏,探索战略战役级指挥控制系统智能化面临的挑战、关键技术和发展趋势。

2 战略战役级指挥控制系统智能化面临的挑战

智能化战略战役指挥控制系统的主要任务是根据任务目标,依据给定的资源和约束条件,运用智能技术,通过科学规划的方法,产生一系列的行动序列,以实现使命任务。RTS 游戏中的人工智能关注的研究问题与智能化战略战役决策有较大的相似性。早期主要集中于资源管理、不确定环境下的决策、空间和时间推理、多个 AI 之间的协同、对手的建模和学习、对抗式实时计划等问题^[2-3]。最近的研究进展又发现了一些新的挑战,例如,如何利用大量的领域知识等。

2.1 规划问题

RTS 游戏中状态空间的大小远远大于传统棋盘游戏,而且在给定的时间内可以执行的动作数量也大得多。因此,传统的规划方法,如游戏树搜索等,并不直接适用。但是可以将 RTS 游戏中的规划分为多个抽象层次:在较高层次,玩家需要进行长期规划,以便在游戏中发展强大的经济;在较低的层次,个体部队需要协调一致,在考虑到地形和对手的情况下进行战斗。通过采样或分层分解来解决这些大规模规划问题的技术尚不存在。

2.2 学习问题

鉴于通过直接使用规划技术玩 RTS 游戏比较困难,许多研究转向了学习问题的研究。RTS 游戏中的学习问题可分为以下 3 种:

(1) 游戏前学习:如何才能利用现有的数据,如现有的回放,或有关特定地图的信息,以便在开始之前学习适当的策略?目前已经有大量的工作已经朝着这个方向发展。

(2) 游戏内学习:机器人如何部署在线学习技术,让他们在玩游戏时提高游戏玩法?这些技术包括强化学习技术,以及对手建模技术。主要的问题是状态空间太大以及 RTS 游戏只是部分可观察的现实情况。

(3) 游戏间学习:从一场比赛中可以学到什么,可以用来增加下一场比赛胜利的机会?一些工作使用了简单的博弈论解决方案来选择一组预定义策略,但一般性问题尚未解决。

2.3 不确定性问题

在 RTS 游戏规模和不确定性条件下的对抗规划仍然是一个未解决的挑战。在 RTS 游戏中,存在两种主要的不确定性。首先,游戏是部分可观察的,玩家不能观察整个游戏地图,只有通过侦察才能看到对手在做什么。这种不确定性可以通过良好的侦察和知识表示来降低(以推断出什么是可能的)。其次,游戏具有对抗性,并且玩家无法预测对手将执行的动作,这也带来了不确定性。对于这种不确定性,作为人类参与者的 AI 只能建立一个对手可能做什么的合理模型。

2.4 时空推理问题

空间推理与地形开发的各个方面有关,涉及建筑物布局或基地扩建等任务。在建筑物布局方面,玩家需要认真考虑将建筑物定位到自己的基地,以便通过构建具有防御作用的墙来保护自己,并尽量避免可能卡住大型部队的错误配置。在基地扩张时,玩家必须选择好地点来建立一个新基地,涉及到自己的位置和对手的基地。最后,空间推理是战术推理的关键,玩家需要决定在哪里放置作战部队,例如当对手的单位被引入该地时便于己方攻击。空间推理的另一个例子是,当敌人处于低地时,在高地上总是有利的,因为低地的部队看不到高处

的部队。类似地, 时间推理是战役或战略推理的关键。例如, 定时攻击和撤退以获得优势。在更高的战略层面上, 参与者需要考虑何时执行长期影响的经济行为, 如升级、建筑施工、战略转换等, 这些行动的影响往往不是立竿见影, 而是长远的。

2.5 领域知识利用问题

在国际象棋等传统棋盘游戏中, 研究人员利用大量现有的领域知识创建了良好的评估函数。就 RTS 游戏而言, 目前还不清楚大量的领域知识(表格、战略指南或回放数据等)如何被机器人 AI 利用。这方面工作主要集中在两个方向: 一是寻找将现有策略硬编码为机器人的方法, 以便机器人只需决定部署哪种策略, 而不必解决完成每个时间步骤决定每个单元执行哪些动作的问题; 二是创建了大量的回放数据集, 这些数据集从策略、趋势或计划中尝试学习。然而, 星际争霸的游戏非常复杂, 如何自动从这些数据集中学习仍然是一个悬而未决的问题。

2.6 任务分解问题

出于上述多种原因, 现有的大多数 RTS 游戏玩法都是通过将大问题分解为一系列较小的问题来解决, 这些小问题都可以独立解决。常见的分解方法包括 5 个层次^[3]: (1) 战略类: 对应于高层决策过程, 这是游戏理解的最高抽象层次, 针对特定对手寻找有效策略或反策略是 RTS 游戏的关键, 涉及整盘游戏的输赢; (2) 战术类: 对应于战略战役的实施, 包括军队和建筑的定位、动作、时机等等, 涉及一组单位的输赢; (3) 反应控制类: 对应于战术的实施, 包括在战斗中移动、瞄准、射击、逃离、击打和跑动技术, 专注于控制特定的部队。(4) 地形分析类: 分析构成地图的区域, 例如阻塞点、矿物和天然气位置, 高和低的步行场地、岛屿等; (5) 情报收集类: 对应于收集对手的信息, 由于存在战争迷雾, 玩家必须定期派侦察兵监视和侦察敌方基地。

相比之下, 当人类玩星际争霸时, 他们通常以非常不同的方式划分他们的决策。星际争霸社区通常谈论两项任务: 一是微观任务, 即单独控制单位的能力(大致对应于上面的反应控制和部分战术), 一个好的微观操作玩家通常会让他们的队伍在更长的时间内保持活力。二是宏观任务, 即能够在适当的时间生产部队和进行扩张, 以保持部队的生产量(大致对应于上述除了反应式控制和部分战术之外的所有类型), 一个好的宏观操作玩家通常拥有较大的军队。任务分解带来的一个关键问题是体系结构的设计, 即如何解决处理上述各种子任务的 AI 之间的高效协同工作, 例如通信、冲突消解、资源使用排序等。

3 战略战役级指挥控制智能化关键技术

相对于 RTS 游戏中微观层面的行动控制, 宏观层面的智能化指挥控制与战略战役级指挥控制智能化具有更多的相似性, 因此本文重点关注宏观层面的活动控制。为了解决 RTS 游戏中的自动决策问题, 很多研究致力于创建可以在 RTS 游戏的战略层面进行智能行动的系统。典型的宏观任务包括战略决策(或者行动序列规划)、敌方规划识别、状态评估技术、多智能体协作和多尺度 AI 技术等^[2]。

3.1 行动序列规划技术

这是一个最具挑战性的问题, 因为不完整信息——“战争迷雾”, 掩盖了战场中远离友军单位的区域, 造成了巨大的状态和行动空间。行动序列规划技术期望 AI 能够达到类人的水平, 减少目前在游戏领域常用的硬编码开发工作量。

3.1.1 基于案例的规划

基于案例的规划(Case-Based Planning, CBP)的主要原理是通过寻找类似的过去情况, 为现状提供潜在的解决方案。但是 CBP 系统在战略层面上反应能力差, 在行动层面反应过度, 在低层次行动

失败之前不会对高层次的状态变化做出反应,或者由于单一行动失败而抛弃整个计划^[7]。早期的 CBP 在 RTS 游戏中应用一般通过创建“动态脚本”概念扩展的系统,根据当前情况选择战术和战略。通过将多个状态抽象为一个状态网格,和将行动抽象为一个战术集合来降低状态和动作空间的复杂性,其中状态网格中多个建设行为被组合成一个包含了一系列小的特征集合的游戏。这使得这个系统能够在多种游戏中,改善对每种情况下每种战术表现的评估,并最终学会始终如一地击败所有测试过的脚本对手。后续有将 ABL 中的行为、目标和生存条件的想法与 CBP 结合的研究。在执行计划时,可能会对其进行修改,以适应意外事件或补充未能实现目标的状态。还有研究分别将决策树、模糊集等与基于案例的推理(Case-Based Reasoning, CBR)结合,主要用于预测高级别的态势,确定用于案例选择的属性和权重,对连续特征值进行分组来提取状态信息,以简化状态空间。

3.1.2 分层规划

分层规划可以映射到 RTS 游戏中典型的目标和子目标的层次结构,从最高级别的目标(例如赢得游戏)直接映射到最低级别目标。通过分层次分解问题,规划系统能够在不同抽象层次分别处理局部情况,减少问题的复杂性,但在不同层次之间协调会产生一些潜在的新问题。分层规划经常被合并为其它方法的一部分。

一些研究人员将这个层次结构形式化为分层任务网络(Hierarchical Task Network, HTN)的明确结构,其中包含任务、任务排序以及实现它们的方法。HTN 中高级复杂任务可分解为一系列更简单的任务,这些任务本身可以分解,直到每个任务代表具体行动。HTN 允许 Agent 对问题进行动态反应,例如重建失去的建筑物或在需要时收集特定类型的额外资源,这与内置的脚本化 AI 不同。HTN 已被用于 RTS 游戏的战略决策,但未用于星际争霸。学习型 HTN 已经在相对简单的游戏中进行了

研究,但尚未直接用于 RTS 的 AI 领域,未来可能会有希望减少建造 HTN 所需的工作。

3.1.3 行为树

行为树是决策和行动节点的层次化结构,是游戏行业程序员和设计者用于定义 Agent 的行为的常用方法。因为它们可以使用可视化工具创建和编辑,使其对于非程序员更容易理解和接受,这与脚本不同,因此比较流行。而且,树的层次化结构鼓励重用,因为定义特定行为的树可以连接到另一个树上的多个位置,或者可以通过添加节点逐步定制。由于行为树是分层次的,因此它们可以涵盖从非常低级行为到战略级别决策等广泛的行为。Palma 等使用行为树实现了对基于案例的策划者行为的直接控制^[7]。借助该系统,机器学习可以用来通过规划人员创建复杂而强大的行为,同时允许游戏设计人员通过替换行为树来更改行为的特定部分,而不是行为或整个计划,即可以为特定想定义自定义行为,修复错误学习的行为,或根据需要调整学习行为。

3.1.4 目标驱动的自治

目标驱动的自治(Goal-Driven Autonomy, GDA)是一种模式,即一个 Agent 为其目标提出理由,确定何时需要更新,以及随后需要进行规划和执行时,对其进行更改或添加。GDA 解决了 CBP 遇到的高低层次反应性问题,并通过主动推理回溯目标成功或失败的原因。2010 年 Weber 等提出了一个使用 ABL 的星际争霸 GDA 系统,该系统能够形成预期结果的规划^[8]。如果发生意外情况或事件,系统可以将其标记为差异,为其发生原因作出解释,并形成修改规划的新目标,从而允许系统对意外事件做出适当反应,如图 1 所示。这个系统被用于 AIIDE 星际争霸 AI 比赛项目 EISBot 中,并且还通过在 International Cyber Cup 星际争霸天梯上与人类玩家对战来进行评估,结果表明它优于 48% 的竞争对手。

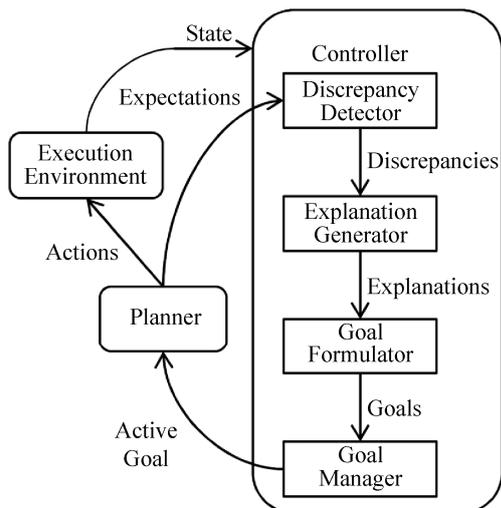


图 1 目标驱动的自治概念模型

Fig. 1 Goal-Driven Autonomous Concept Model

Jaidee 等将 CBR 和 RL 整合为一个 GDA 的学习版本, 允许该系统随着时间的推移改进其目标和领域知识^[9]。因为可以自动学习缺失的知识, 意味着人类专家需要更少的工作来指定可能的目标、状态和其它领域知识。同样, 如果底层领域发生变化, 学习系统也能够自动适应变化。

3.1.5 状态空间规划

状态空间的自动规划和调度是经典 AI 研究的一个分支, 启发式状态空间规划技术已应用于 RTS 游戏 AI。通常, Agent 被赋予一个开始和目标状态, 以及一组具有先决条件和效果的行动, 然后 Agent 找到一个行动序列来实现起始状态的目标。现有的 RTS 应用程序通过处理持续和并行动作、整数值状态变量和严格的时间限制等, 反而增加了基本问题的复杂性。

自动规划思想采用面向目标的行动规划 (GOAP) 体系结构, 已经成功应用于商业第一人称射击游戏。它允许 Agent 自动为它们当前的状态选择最合适的行为, 以满足一系列目标, 理想情况下导致更多样化、复杂和有趣的行为, 同时保持代码更便于可重用和维护。然而, GOAP 需要大量的领域工程来实施, 并且因为它将状态映射到目标而不是动作, 因此规划者无法判断实现目标是要即将开始规划, 或是即将失败, 还是已经失败。此外, 也

有人指出, GOAP 现在变成了一个死胡同, 因为学术界和产业界已经从 GOAP 转向支持分层规划, 以实现更好的性能和代码可维护性。但是, 仍然有大量研究使用类似于 GOAP 的自动规划方法来规划 RTS 游戏中的建设指令。

3.1.6 演化算法

演化算法通过评估不同的潜在解决方案, 并结合或随机化高适应度潜在解决方案的组分来寻找解决问题的有效方案, 以找到新的更好的解决方案。这种方法在 RTS 游戏 AI 领域很少使用, 但它已被有效地应用于星际争霸战术决策子问题, 以及在类似 RTS 游戏中学习战略知识。

Evolution Chamber 软件使用该技术来优化部分定义的构建指令, 如果在比赛的特定时间内产生一组目标单位、建筑物和升级, 则软件会搜索达到这些目标的最快速或最少资源消耗的方式。这个软件在星际争霸 2 早期就因为制定了一个不寻常的高效计划而备受关注。Ponsen 等使用演化算法在 Wargus 游戏中生成策略, 为了产生这些策略^[10]。演化算法通过对游戏中战术和战略级别动作序列进行交叉变异以形成脚本, 这些脚本打败了一组人造且非演化的脚本, 如图 2 所示。该方法的一个缺点是往往假设对手是静态的, 对动态变化的选手适应性不够。

3.1.7 认知架构

认知架构 (Cognitive Architectures) 通过对人类思考过程的推理机制进行建模, 实现战略级 RTS 游戏 AI, 这种方法可以使人们更好地理解人类的理性, 创造出更像人类的人工智能。2012 年该方法就已应用于星际争霸, 使用 Soar 认知架构, 采用 BWAPI 接口与 Soar 代理进行通信。它利用 Soar 的空间视觉系统来处理侦察活动和探索道路, 以及 Soar 的工作记忆来保存感知和推理的状态信息。然而, 目前仅限于玩星际争霸的部分游戏, 仅使用基本兵营和海军单位进行作战, 并使用硬编码地点进行建筑布置。

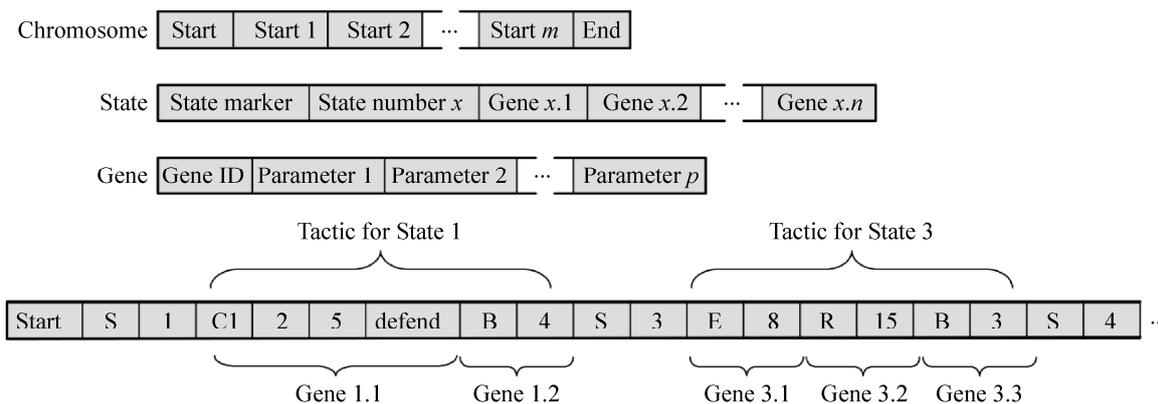


图 2 RTS 游戏 AI 演化算法的染色体设计
Fig. 2 Chromosome Design of RTS Game AI Evolutionary Algorithm

类似的方法是将 Soar 用于其它 ORTS, 而不是用于星际争霸。该方法将 Soar 认知架构与 ORTS 接口相连, 通过使用分组和抽象概念来降低问题的复杂性。Agent 可以将整个军队的所有单元视为一个实体, 但可以改变其注意力的重点, 使其能够一次感知一个地点的多个单元, 或在广阔区域内的多组单元^[11], 如图 3 所示。但是由于 Soar 的限制, Agent 一次只能追求一个目标, 这在星际争霸和大多数 RTS 游戏中都是受限的。

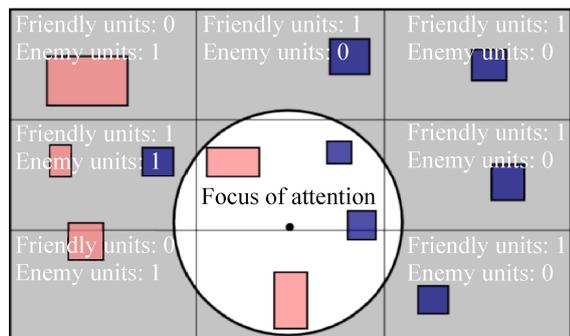


图 3 注意力限制了 Agent 收到的信息
Fig. 3 Attention Limits the Information Received by Agent

3.1.8 空间推理

RTS 游戏中 Agent 经常需要推断大量隐藏对象的位置和行为, 它们往往具有不同的属性、随时间变化、受动态环境中的对手控制。尽管这些问题非常复杂, 但人类可以快速而准确地推断出这些信息, 且只需很少的信息就可预测和拦截敌方攻击或逃跑的位置, 或利用地形以及己方作战单元和建筑

物的部署来获取优势。因此, RTS 是在受控环境下进行空间推理研究的极佳领域。

但是, 即使只是对 RTS 游戏中的地形进行分析, 想要忽略作战单元和建筑物也是一件不容易的事。Perkins 描述了 Brood War 地形分析器的实施和测试, 它已成为制作能够推理地形的星际争霸机器人的一个非常常见的库^[12]。该库使用地图的可行走瓦片信息, 创建并修剪了 Voronoi 图, 将节点识别为区域或阻塞点, 然后根据试错确定的阈值, 合并相邻区域以产生期望的结果。阻塞点节点被转换成分隔区域的线, 从而形成一组由区域点连接的区域多边形, 如图 4 所示。与人类识别的阻塞点相比, 其识别率较高, 但是分析所需时间仍需改进。一旦玩家能够简单推理地形, 就可以推理作战单元在该地形上的行动。此外, 势场(Potential Fields, PFs)和影响图也是两种主要的空间推理方法。

3.2 敌方规划识别技术

通过使用具有策略学习功能的 AI 系统, 可以让计算机对手更具挑战性、动态性和人性化, 同时也使它们更容易被创建。一些研究人员专注于识别对手战略的问题, 这在 RTS 游戏中特别困难, 因为“战争迷雾”隐藏了对手行为的不完整信息。大多数规划识别使用现有的规划库来匹配和尝试识别策略, 但也有些方法在没有任何预定义的情况下识别规划。

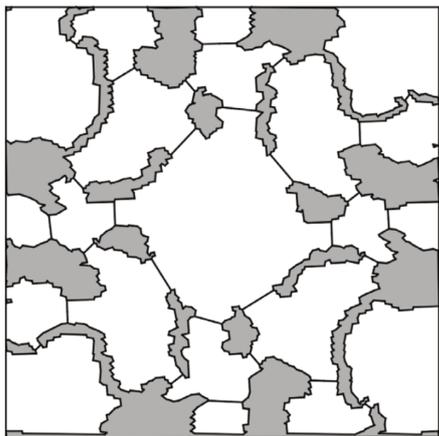


图 4 经 Brood War 地形分析器分析后的地形
Fig. 4 Terrain Analyzed by Brood War Terrain Analyzer

学习的数据是从专业人类玩家的广泛可用的回放文件中提取的。2012 年, Synnaeve 等创建了一个包含 7649 个回放文件的数据集; 2017 年, Facebook 开源了数据集 STARDATA, 包含大约 65 646 个回放文件^[13]; DeepMind 和暴雪公司合作, 开放了一个超过 50 万个回放文件的匿名的游戏回放数据集^[6]; 我国中科院自动化所近期在开放了训练 AI“大局观”的数据集 MSC(Macro-management in StarCraft)^[14]。规划识别与学习方法可以分为演绎式、回溯式、概率式和基于案例等 4 类。

3.2.1 演绎式

主要通过比较各种已知规划的预期行为假设情况来识别规划。通过观察特定的行为, 即使没有完整的知识, 也可以对正在进行的规划进行推断。一种是通过将对手的观察与所有可能产生这种情况的已知策略进行匹配, 然后模拟可能的规划, 以确定对手的未来预期行动, 根据新的观察结果判断规划的可能性, 并丢弃不匹配的规划。另一种方式是采用决策树机器学习方法。学习过程采用最近邻算法和非嵌套广义示例算法进行。由此产生的模型能够从不完整信息中预测构建顺序, 最近邻算法对不完整信息最有效。

3.2.2 回溯式

主要通过观察到的结果, 寻找关于状态最合理的假设来确定规划。2010 年, Weber 等描述的

GDA 系统是星际争霸规划中主动进行规划识别的一个例子^[15], 期望由行为的结果构成, 意外事件则被视为“差异”。通过从一组预定义的“解释”中选择来处理差异, 这些解释给出了差异的可能原因, 并创建新的目标来弥补假设状态的变化。这个系统需要实质性的领域工程, 来定义像星际争霸这样复杂的领域所需要的所有可能的目标、期望和需要的解释。2012 年, 增加了通过离线分析回放来学习星际争霸的领域知识的能力。

3.2.3 概率式

主要利用统计数据 and 预期概率来确定给定情况最可能的未来结果。Synnaeve 等在星际争霸中通过检查来自回放的修建指令进行 keyhole 概率规划识别, 而无需事先了解星际争霸构建命令^[16]。因为他们可以直接从回放中学习, 而无需任何人工输入, 所以只需要很少的工作就可以适应游戏中的变化或新情况。学习的模型可用于预测对手当前状态的未观察部分, 或者根据当前和过去的情况来预测玩家未来的战略方向。或者, 他们可用于识别游戏中使用的不寻常策略。这两种方法在使用的概率技术、应用范围以及系统的预测能力方面有所不同。Dereszynski 等使用隐马尔可夫模型来模拟玩家在一系列状态中的进展, 每个状态都有产生每个单位和建筑物类型的概率, 以及状态将转移到下一个状态的概率^[17]。在没有任何先验信息的情况下, 它能够学习一个与常用开放构建命令非常相似的状态转换图, 但没有提供对其预测能力的全面分析和评估。可以对上述方法进行拓展, 使用动态贝叶斯网络模型来识别星际争霸中的战略。

3.2.4 基于案例推理

基于案例推理(Case-Based Reasoning, CBR)通过存储代表问题和解决方案的具体知识案例, 并将新问题与过去案例进行比较以适应和重用过去的解决方案。通常用于学习 RTS 游戏中的战略游戏, 因为它可以捕获从具体经验中获得的复杂、不完整的情景知识, 试图推广一个非常大的问题空

间,而不需要转换数据。规划识别也可使用 CBR 作为基础。

虽然最近使用 CBR 进行 RTS 游戏的大部分工作都是从人类行为中学习,但也有人直接使用 CBR 而不观察人类游戏。该系统使用一组度量来衡量表现,以学习通过类似于 RL 的迭代过程来玩策略游戏 DEFCON。该系统使用过去玩过的游戏案例同时学习它应该做出的战略行动以及对手可能做出的动作。它通过使用影响图,并将单位划分为舰队和元舰队的方法获取某个地区的威胁和机会,从而抽象出有关单位和结构位置的较低级别信息。

3.2.5 通过观察学习

对于像 RTS 游戏这样复杂的领域,收集和维护专家知识或通过反复试验来学习它可能是一项非常困难的任务,但是游戏可以通过回放或跟踪提供对(某些)这些信息的简单访问。通过分析轨迹,可以从人类的正确行为演示中学习,而不要求程序员手动指定其行为。这种仅仅通过观察专家外部行为和环境的的学习通常被称为观察学习,类似于人们通过观察专家和模拟他们的行为来加速学习的方式。

通过观察学习特别适用于 CBR 系统,它可以在从专家提取知识或思考潜在案例并手动记录时,减少或消除对 CBR 系统设计人员的需要。通过检查玩家为响应情况和事件采取的行动,或者完成某些预定义任务,可以将回放转化为 CBR 系统的案例。通过观察学习也可以创建一个独立于领域的系统,在不知道任何关于其基本含义情况下,它可以很容易学会如何将一组知觉和行为联系起来。但是,如果没有领域知识来指导决策,要学习特定情况下的正确行动是非常困难的。为了弥补,系统必须使用特征自动加权和案例聚类等技术来处理和分析观察到的案例,以表达相关知识。

3.2.6 通过演示学习

从构建的案例中学习的方法被称为从演示学

习,不再从观察玩家与游戏互动的轨迹来学习,而是用额外的信息来注释轨迹(例如通常使用玩家的内部推理或意图),从而使演示更容易学习,并且提供关于所学的特定行为的更多控制。相应地,手动添加注释会使得演示更加耗时,已经有相关技术来自动执行此过程。

假设已知游戏中有关行动和任务的一些知识,可以使用多种方法从轨迹中提取案例,以用于通过观察或从示范系统中学习。案例获取技术包括基本的反应式和整体学习方法,更复杂的依赖关系图学习、时间跨度分析技术等。

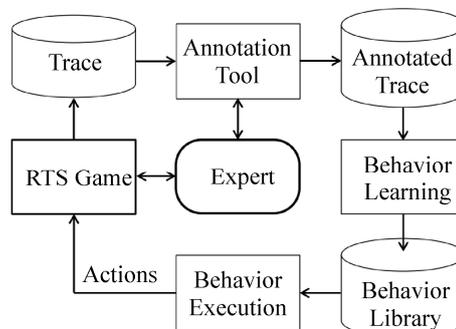


图5 学习型演示系统的一般结构
Fig. 5 General Structure of Learning Demonstration System

Mehta 等给出了一个能从有人注释的 Wargus 游戏回放中学习的 CBR 和规划系统,如图 5 所示^[18]。通过注释每个回放与玩家当时试图达到的目标,系统可以将动作序列分成实现特定目标的行为,并学习关于目标及其可能实现的命令的层次化结构。

3.3 状态评估技术

状态评估主要涉及标准化评估和全局状态评估两个方向。标准化评估是要解决统一的评估标准和评估指标的问题。尽管游戏是一个适合评估玩家效能和衡量表现的领域,但很难在大多数文献的结果之间做出公正的比较。首先,很多系统不适合玩星际争霸的整个游戏,而只是适合特定的子问题,因此不同的想定难以直接进行对比;其次,几乎每篇论文都有不同的方法、使用自定义指标进行评估,这使得相互之间几乎不可能进行比较。对此,一个可能的解决方案是选择一组通用的组件,可以

插入部分系统中, 并允许它们作为一个完整的系统进行测试。这可以通过将星际争霸 AI 竞赛中使用的开放源代码 AI 的部分区分开来, 或者可以针对部分系统运行一组通用测试。这些测试可以将 AI 系统的常见子问题(如战术决策、规划和规划识别)作为单独的测试套件进行检查。

在全局状态评估方面, 主要是面临评估方法的问题。很多研究致力于研究哪一种战略的胜率更高, 这种想法是简单系统的思维。战争系统是典型的复杂系统, 对抗过程存在很大的不确定性, 对手、战略、行动、地形等任何一个小的变化都可能会产生巨大的效果。目前可行的评估方法是在天梯比赛中自动测试机器人对抗其他机器人, 这也符合用“整体、动态、对抗”的思想研究体系能力的方法^[1]。例如在星际争霸比赛天梯中, 为了创建一个机器人能力基准, 可以从某顶级比赛的前 N 名机器人中选择一套比赛地图形成一套测试。这将提供足够的多样性来给出机器人能力的通用指标, 并且将允许在不同年份的论文之间比较结果。另一种方法是在与人类的比赛中进行测试。

3.4 多智能体协作技术

集体的努力往往可以解决个体无法解决的问题。即使像蚂蚁这样弱小的个体, 当其形成社会组织时, 也可以完成例如觅食、修建国家, 甚至发动战争这样有高度挑战性的任务。同样, 现实世界的 AI 应用程序也需要多个智能体协同工作。智能体对沟通和协调的有效学习是迈向通用人工智能(AGI)的过程中不可或缺的一步, 也是实现通用人工智能宏伟目标的基础。

RTS 游戏呈现出一个特定的复杂环境, 其规则和总体目标是固定的, 因此合作伙伴之间进行沟通的能力有限。因为在团队游戏中人们通常都会在整个游戏过程中通过协调行动相互帮助, 而大多数 RTS 游戏(包括星际争霸)中的 AI 玩家却是独立于队友的。智能体协作的一个可能方向就是研究为对

手建模的一些技术, 并将它们重用于建模盟友, 从而让玩家了解玩家应该如何与盟友进行协调。另一个方向是将其他领域的团队协作和协调方法适应或扩展用于 RTS 领域。尽管有研究强调合作是非常具有挑战性的人工智能研究问题, 但仍然有大量的探索研究。目前比较典型的研究是我国阿里巴巴认知计算实验室的研究和中科院自动化所的研究。

3.4.1 BiCNet 方法

阿里巴巴认知计算实验室与伦敦大学学院计算机系合作, 以“星际争霸 1”中的微观战斗场景为测试环境, 深入地研究了多个 AI 智能体之间的协作问题。以星际作为测试场景, 其任务是协调多个智能体作为一个团队来打败他们的敌人。为了保持可扩展但有效的通信协议, 引入了多智能体双向协调网络(BiCNet), 采用 actor-critic 表达的向量化扩展。实验表明, 在交战双方都有任意数量的 AI 智能体时, BiCNet 可以处理不同地形下的不同类型的战斗。其次, 如果没有任何诸如人类示范或标签数据的监管, BiCNet 可以学习与经验丰富的游戏玩家相似的各种类型的合作策略。此外, BiCNet 很容易适应异构智能体的任务^[4]。

双向协调网络(BiCNet)的结构如图 6 所示, 它由策略网络(actor)和 Q 值网络(critic)组成, 两者均基于双向 RNN, 策略网络用于独立智能体做出行动决策。因此, 独立智能体能够保持自己的内部状态, 同时能够与其他合作者共享信息。双向递归机制不仅是智能体之间的交流手段, 而且还作为本地记忆状态。该网络进一步加强了 actor 的分组概念, 这在社会行为中起着重要的作用。作为输入, 允许少量的智能体程序在进入递归层之前构建局部相关性。因此, 该模型的表现要好得多。实验发现, 基于分组的 actor 有助于学习社会活动, 如集中火力攻击。Q 值网络将策略网络的状态和行为作为输入, 向每个独立智能体返回估算的本地 Q 值。然后组合本地 Q 值, 提供全局回报的估计值。

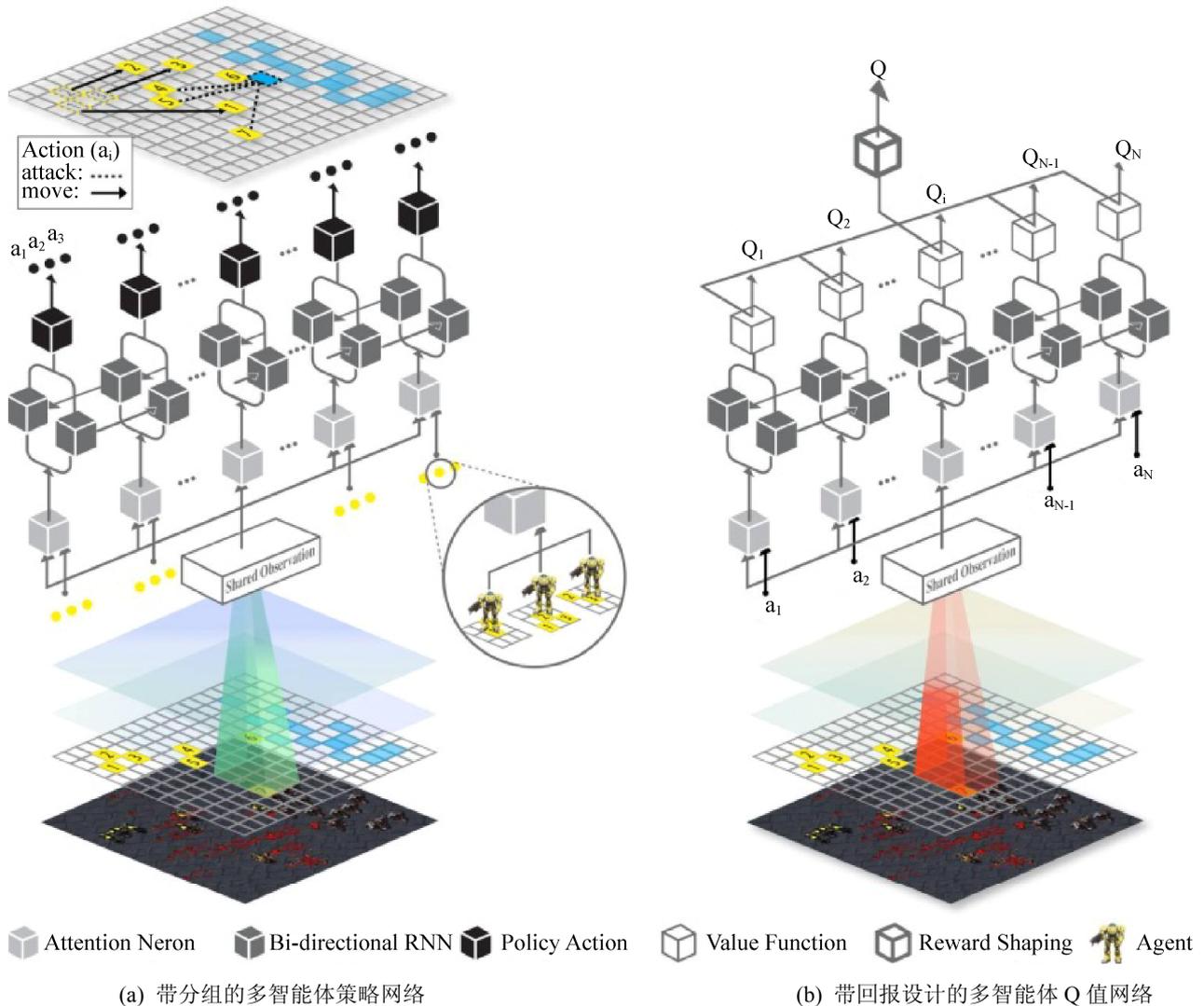


图6 双向协调网络(BiCNet)

Fig. 6 Bidirectionally-Coordinated Net (BiCNet)

3.4.2 PS-MAGDS 强化学习方法

中科院自动化所针对星际争霸中微观操作存在的状态、行动空间复杂和合作策略学习困难等问题,提出了一种强化学习算法用以解决星际争霸微操中的多智能体决策问题。研究提出用参数共享的多智能体梯度下降 Sarsa(λ)算法(PS-MAGDS)来训练作战单元。为了提高学习速度和跟踪奖励延迟的原因,在强化学习中引入了资格追踪(eligibility trace)的方法。作为强化学习的一个基本机制,资格追踪主要用于分配时序信用,它考虑了一组以前经历过的转换。即它不只是考虑最紧前时刻的

“状态-行动”对,还考虑了它所经历过的“状态-行动”对^[19]。

在实现 Sarsa(λ)的多个单元作战时,使用神经网络作为逼近器,并在所有单元之间共享网络参数。尽管只有一个网络来训练,但是不同单元之间可以有不同表现,因为每个单元接手不同的观察和行动作为其输入。这种方法能有效地训练同质 Agent,并且鼓励合作行为。同时,为了更有效地更新策略网络,采用梯度下降算法来训练 Sarsa(λ)强化学习模型,如图7所示。

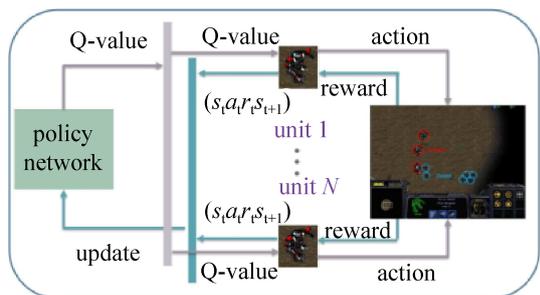


图 7 PS-MAGDS 强化学习过程

Fig. 7 PS-MAGDS Reinforce Learning Process

3.5 多尺度 AI 技术

由于 RTS 游戏的复杂性, 目前的机器人需要多种抽象和推理机制才能有效发挥作用。特别是大多数机器人在战术和战略不同层次对决策、资源管理、建设和侦察等采用不同的方式。这些问题中的每一个都涉及多个相关性问题, 例如在整个游戏过程中, 资源可用于改善资源生成、军队训练单元或构建新的基础设施, 每个选项都由不同模块控制, 自己不能评估其它模块。而人类似乎能够通过在线和离线、直接反应、协商和预测性推理等方式相结合, 很好地处理这些问题。

多尺度 AI 技术就是为了横跨多个抽象以实现并发和协调的目的, 涉及很多相关技术。一是工作记忆或“共享黑板”技术, 用于模块之间的间接通信, 其中每个模块公布其当前信念供其他人阅读; 二是允许将计划和推理模块生成的目标和计划插入到其核心的反应规划系统中, 并与当前的目标和规划并行执行; 三是改变激活行为的方法, 以便模块可以修改已定义行为的先决条件, 允许他们根据情况激活和停用行为。

一些简化的方法可能对 RTS 机器人中的部分有效。例如, 使用更高级别的战术命令^[16], 例如侦察员、持有位置、群体或战斗作为其微操控制器的输入之一; 或采用层次化结构进行单位控制^[20], 并配有总体游戏指挥官向战略战役指挥官和战斗指挥官发出命令, 他们再将命令传递给他们的下级指挥官。层次越低的指挥官越关注具体任务, 但是缺少对整体游戏状态的信息, 只能依赖于他们的上

级在较大的情况下采取行动。该系统允许底层指挥官将来自其上层的信息合并到层次结构中, 但无法直接与其他底层指挥官进行反应和协调以执行协作操作。通过分析星际争霸 AI 的大多数论文如何将决策分解为战术和战略决策就可以发现, 他们往往只关注人工智能的某一个方面, 从而避免了这个问题。

4 战略战役级指挥控制智能化技术发展趋势

RTS 游戏与作战决策有天然的相似性, 虽然战场环境、作战过程、作战效果等被大大简化, 但其对战略战役决策过程模拟得逼真度却相当高。通过分析当前人工智能技术在 RTS 领域的应用现状, 可以发现未来智能化战略战役辅助决策系统的发展将会有以下 3 种趋势:

4.1 宏观与微观管理的整合

当前在 RTS 游戏中, 微观任务管理研究的成果相对宏观层面更多, 客观原因是微观任务管理更加具体、更容易量化, 学习起来更直接, 因此效果相对较好。相反, 宏观层面的学习因为定性、模糊的成分和时延的因素, 不确定性很大, 因而学习效果相对较差。2017 年 DeepMind 公司训练了一个基于 A3C 算法的 DNN 网络, 不幸的是, 它在和星际争霸 2 中内置 AI 的对抗中完败。这不是偶然, 恰恰说明星际争霸游戏与战争具有天然的相似性, 都是典型的复杂系统。解决复杂系统必须要用其特有的自上而下和自下而上相结合的方法论, 即宏观与微观相结合的方法。因此, 未来一个可行的途径可能在宏观任务管理与微观相结合的方法上。2017 年哥本哈根大学 Niels 等进行的研究^[21], 从 2005 场顶级玩家 Replay 视频中学会了“大局观”, 新系统的表现显著超越了人族游戏内置 AI, 这是迄今为止直接从回放数据中学习宏观任务管理的第一个成功的案例, 必将为宏观与微观管理相结合的方法提供启发。

4.2 端到端的学习

自2016年AlphaGo在围棋游戏中大显身手以来,将端到端的深度学习应用到强化学习已经成为大家众相学习的典范。特别是在游戏领域,DeepMind、Facebook等大公司纷纷发布自己的平台^[6,13],以便开展进行端到端的学习。端到端的学习方式之所以流行,是因为强化学习通过时间连接模型来描述稀疏的数据和奖励之间的关系,与人解决推理问题的方式十分类似,而且互联网数据提供了学习的基础。如何将强化学习、迁移学习等多种学习方法进行整合,可能是端到端的学习方面一个重要的发展方向。但是我们也要警惕盲目的数据决定论,即只要有数据就能做出准确的决策^[1]。战争是典型复杂系统,战争迷雾将永远存在,不完全信息条件下的决策是作战指挥的本质特点。决策需要根据情况不断变化,数据要求也会不断发生改变。复杂性会导致决策本身不唯一,也就无法确定哪个正确。越低的层次,决策越接近简单系统,数据才越具有决定性。

4.3 探索新的学习范式

当前战略战役级指挥控制系统智能化还面临态势理解、体系能力评估、作战任务规划、样本数据稀缺等挑战,这些都能在RTS游戏的相关智能化技术中找到对应关系,例如态势理解与敌方规划识别对应、体系能力评估与全局状态评估对应、作战任务规划与行动序列规划对应,很多技术具有很强的通用性,例如关于样本数据稀缺的问题,目前生成对抗式网络技术已经日趋成熟,DARPA已经决定资助该技术应用于人工智能“造假比赛”,用于生成作战过程样本完全是有可能的^[22]。同时,越来越多的证据表明,当前主流深度学习方法也许并不是一项普适性技术,在更多类似星际争霸2等任务极其复杂、数据难以标注的领域,也许我们需要新的方法。Hinton最近公开号召摒弃现有深度学习(主要是反向传播、CNN)范式,重新寻找全新的道路。他认为,要想让神经网络能够自己变得智

能,即实现不依赖海量标注数据的“无监督学习”,意味着需要放弃反向传播等目前主流深度学习理念。还可以深入探索迁移学习、多任务学习、对抗学习、持续学习、小样本学习、元学习、神经推理等新的方法,解决从少量的数据中学习、从已有知识中学习、多种任务等带来的智能难题。

5 结论

RTS游戏是人工智能技术应用的一个重要领域,特别是近年来AIIDE、CIG和SSCAIT等星际AI赛事吸引了学术界和工业界的持续研究与开发。2018-06-26,OpenAI宣布其在Dota2游戏中5v5对抗中获胜,虽然离在军事指挥控制系统中使用仍有较大差距,但是人工智能在指挥控制决策智能化方面确实取得了新突破。本文围绕战略战役级指挥控制系统智能化面临的挑战,梳理了行动序列规划、敌方规划识别、多智能体协作和多尺度AI等关键技术,指出了宏观与微观管理的整合、端到端的学习、探索新的学习范式等发展方向。鉴于RTS游戏中与指挥控制系统的人工智能技术具有很好的通用性,以及真实的作战数据难以获取,RTS游戏必将成为军事指挥控制系统智能化技术研究的平台,其开源环境和海量数据将极大地促进人工智能技术的研究。

参考文献:

- [1] 胡晓峰,郭圣明,贺筱媛. 指挥信息系统的智能化挑战——“深绿”计划及AlphaGo带来的启示与思考[J]. 军用信息系统与技术, 2016, 7(3): 1-7.
Hu Xiaofeng, Guo Shengming, He Xiaoyuan. Challenge to Intelligent Command Information System: Reason and Revelation on Deep Green Plan and AlphaGo[J]. Command Information System and Technology, 2016, 7(3): 1-7.
- [2] Ontañón S, Synnaeve G, Uriarte A, et al. A Survey of Real-Time Strategy Game AI Research and Competition in Starcraft[J]. IEEE Transactions on Computational Intelligence and AI in Games (S1943-068X), 2013, 5(4): 293-311.
- [3] Glen Robertson, Ian Watson. A Review of Real-Time

- Strategy Game AI[J]. AI MAGAZINE(S0738-4602), 2014, 35(4): 75-104.
- [4] Peng Peng, Ying Wen, Yaodong Yang, et al. Multiagent Bidirectionally-Coordinated Nets: Emergence of Human-Level Coordination in Learning to Play Starcraft Combat Games[DB/OL]. (2017-03-29) [2018-05-31]. <http://cn.arxiv.org/pdf/1703.10069v4>.
- [5] Nicholas Ernest, David Carroll, Corey Schumacher, et al. Genetic Fuzzy Based Artificial Intelligence for Unmanned Combat Aerial Vehicle Control in Simulated Air Combat Missions [J]. Journal of Defense Management (S2167-0374), 2016, 6(1): 1-7.
- [6] Vinyals O, Ewalds T, Bartunov S, et al. Starcraft II: A New Challenge for Reinforcement Learning[DB/OL]. (2017-08-16) [2018-05-31]. <http://cn.arxiv.org/pdf/1708.04782v1>.
- [7] Palma R, Sánchezruiz A A, Gómezmartín M A, et al. Combining Expert Knowledge and Learning From Demonstration in Real-Time Strategy Games[C]// Case-Based Reasoning Research and Development, International Conference on Case-Based Reasoning, ICCBR 2011, London, UK, September 12-15, 2011. Proceedings. 2011: 181-195.
- [8] Weber B G, Ontañón S. Using Automated Replay Annotation for Case-Based Planning in Games[C]// ICCBR 2010 workshop on CBR for Computer Games (ICCBR 2010), 2010: 15-24.
- [9] Jaidee U, Aha D W. Integrated Learning for Goal-Driven Autonomy [C]// International Joint Conference on Artificial Intelligence. AAAI Press, 2011: 2450-2455.
- [10] Ponsen M J V, Spronck P, Aha D W. Automatically Acquiring Domain Knowledge for Adaptive Game AI Using Evolutionary Learning[C]// Conference on Innovative Applications of Artificial Intelligence. AAAI Press, 2005: 1535-1540.
- [11] Wintermute S, Xu J Z, Laird J E. SORTS: A Human-Level Approach to Real-Time Strategy AI [C]// Artificial Intelligence and Interactive Digital Entertainment Conference, June 6-8, 2007, Stanford, California, USA. DBLP, 2008: 55-60.
- [12] Perkins L. Terrain Analysis in Real-Time Strategy Games: An Integrated Approach to Choke Point Detection and Region Decomposition [C]// AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, Aiide 2010, October 11-13, 2010, Stanford, California, USA. DBLP, 2010: 223-231.
- [13] Yuandong Tian, Qucheng Gong, Wenling Shang, et al. ELF: An Extensive, Lightweight and Flexible Research Platform for Real-Time Strategy Games [C]// 31st Conference and Workshop on Neural Information Processing Systems (NIPS 2017), California, USA, December 4-9, 2017: 2656-2666.
- [14] Huikai Wu, Junge Zhang, Kaiqi Huang. MSC: A Dataset for Macro-Management in Starcraft II [DB/OL]. (2017-10-09)[2018-05-31]. <http://cn.arxiv.org/pdf/1710.03131v1>.
- [15] Weber B G, Mawhorter P, Mateas M, et al. Reactive Planning Idioms for Multi-Scale Game AI[C]// Computational Intelligence and Games. IEEE, 2010: 115-122.
- [16] Synnaeve G, Bessiere P. A Bayesian Model for RTS Units Control Applied to Starcraft[C]// Computational Intelligence and Games. IEEE, 2011: 190-196.
- [17] Dereszynski E, Hostetler J, Fern A, et al. Learning Probabilistic Behavior Models in Real-Time Strategy Games[C]// AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment. AAAI Press, 2011: 20-25.
- [18] Mehta M, Ontañón S, Amundsen T, et al. Authoring Behaviors for Games Using Learning From Demonstration[C]// Proc. of the 8th International Conf. on Case-Based Reasoning: Case-Based Reasoning Research and Development (ICCBR 2009). Berlin, Heidelberg: Springer-Verlag, 2009: 12-20.
- [19] Kun Shao, Yuanheng Zhu, Dongbin Zhao. Starcraft Micromanagement With Reinforcement Learning and Curriculum Transfer Learning[J]. IEEE Transactions on Emerging Topics in Computational Intelligence (S2471-285X), 2018 (99): 1-12.
- [20] Churchill D N, Saffidine A, Buro M, et al. Fast Heuristic Search for RTS Game Combat Scenarios[C]// National Conference on Artificial Intelligence, 2012: 112-117.
- [21] Niels Justesen, Sebastian Risi. Learning Macromanagement in Starcraft From Replays Using Deep Learning[C]// IEEE's 2017 Conference on Computational Intelligence in Games, August 22-25, New York, USA. 2017.
- [22] Will Knight. The US Military Is Funding An Effort to Catch Deep Fakes and Other AI Trickery[EB/OL] (2018-05-23) [2018-05-31]. <https://www.technologyreview.com/s/611146/the-us-military-is-funding-an-effort-to-catch-deepfakes-and-other-ai-trickery/>.