**人工智能在电子游戏上的应用综述**

## 摘要

电子游戏从出现开始就推动着计算机技术的发展，在过去的十年里，人工智能（AI）技术的发展也离不开游戏的研究。作为AI研究绝佳的测试平台，在人工智能方面从古老的棋盘游戏、之后又到经典的街机游戏、再到2016年AlphaGo战胜李世石，人工智能Agent都表现卓越。RPG游戏作为网络游戏的一支很重要的大军，势必要搭上人工智能的车重新占领游戏市场。本文分析了人工智能系统在电子游戏上的发展历史以及现有的人工智能游戏的研究方向，提出了人工智能在未来RPG游戏游戏上的发展方向，以期对今后人工智能在游戏上的应用的研发和创新方向给出一定的指导

## 关键词:游戏人工智能、电子游戏、RTS、MOBA、MMORPG

电子游戏一直是人工智能研究和测试的重要组成部分。在二十世纪九十年代中期，IBM 的超级计算机「深蓝」数次击败了国际象棋世界冠军卡斯帕罗夫。后来到了 2011 年，IBM 的沃森人工智能在游戏《危险边缘》就击败了最优秀的人类玩家，并展示了 IBM 在自然语言处理的进展。2015 年，DeepMind 就开始使用机器学习来训练人工智能玩一些复古的雅达利游戏（Atari games），并使其至少能做得和人类一样好。后来在 2016 年的三月份，DeepMind 通过另一种方法训练了 Alpha Go，并击败了围棋世界冠军李世乭。

在过去的十年里，人工智能以迅猛的发展之势横扫计算机界，而在有监督学习的情况下，机器在图像识别和语音识别能力上，已经有预期的超越了人类的认知水平。前不久，DeepMind在A3C上发表的论文中提出了一套UNREAL系统，在Atari游戏合集上的表现得分基本与人类玩家持平，而且有的还能超过人类。而阿里的新研究则通过BiCNet在《星际争霸》上实现了多兵种协同作战，该实验证明了这一系统在RTS（Real-Time Strategy Game）游戏上各兵种协调全作的有效性，以及指定奖励和学习策略存在相关性。谷歌DeepMind团队与暴雪开源了星际争霸2机器学习平台，提供了游戏输入输出API，DeepMind又做了基于Python的封装，持续发布基于replay的增强学习训练数据，此外两家公司还发布了上百万帧来自天梯专业选手的游戏记录，这将有助于基于平台算法应用到不同应用上。而在近日，研究者们又尝试使用在AlphaGo Zero中出现的蒙特卡洛树搜索等技术，在MOBA（Multiplay Online Battle Arena Game）游戏——《王者荣耀》上得到展现，证明了该项技术的深度神经网络实现可以创建一个适合《王者荣耀》游戏的有竞争力的AI智能体。

MMORPG（Massive Multiplayer Online Role-Playig Game）游戏在过去的十多年发展到现在，已经处在一个比较尴尬的状态，处于下滑态势。一方面由于游戏快餐化，娱乐模式呈现井喷，MMORPG又大体复制从出生就有的游戏模式，使得游戏在创新方面已无法满足玩家游戏乐趣。然而MMORPG停滞不前的十年正是人工智能发展的十年，随着人工智能不断的发展，游戏需要由低智能向高智能发展。以遗传算法和神经网络为代表的强人工智能将大大提高游戏的智能性。我们基于游戏人工智能在Atari游戏，RTS游戏，MOBA游戏等各类游戏上的研究成果，可以预测MMORPG游戏也将搭乘人工智能的势头再度引起市场的关注，而以上的研究成果也将为我们研究人工智能在MMORPG游戏的应用提供一定的参考。

#### 1.游戏人工智能的发展介绍与分析

1.1 Atrai游戏

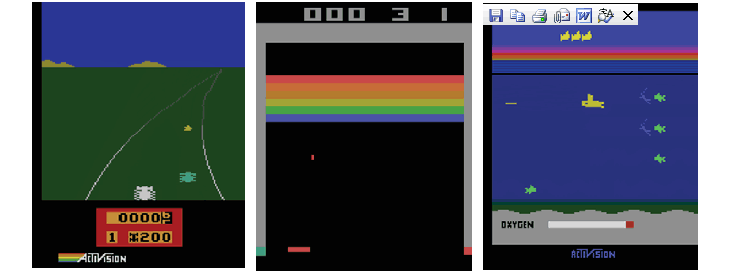
1977年，雅达利（Atari）发行了自家公司的新一代游戏机Atari2600，开创了现在游戏机的历史，同时也开创了强化学习领域在训练端到端的控制策略。《Playing Atari with Deep Reinforcement Learning》中说到，使用DQN网络来训练AI智能体，以达到游戏学习策略，这些策略的学习是通过计算机理解屏幕像素信息以及积分反馈来取得效果的，该网络最初只有两个卷积层，用来提取图像特征，然而却已经取得了很大的效果，AI智能体表现了不俗的能力。

图1.雅达利游戏深海游弋、摩托大战、打砖块

2010年，Lange提出Deep auto-encoder用于基于视觉的相关控制；在2011年，Cuccu（Jurgen Schmidhuber课题组）等人做出了与DRL相近的研究；同年，Abtahi等人使用DBN替代传统的强化学习中的逼近器；2012年，Lange提出Deep Fitted Q学习做车辆控制；基于以上发展，2013年，Deep Mind团队在nips上发表 《Human-level control through deep reinforcement learning》，工作人员在7个Atari游戏，分别是激光骑士（Beam Rider），打砖块（Breakout），摩托大战（Enduro），乓（Pong），波特Q精灵（Q\*bert），深海游弋（Seaquest），太空侵略者（Space Invaders），游戏过程中使用网络深层架构、学习算法，最后AI通过自学完成了49种Atari视频游戏的学习，并在23种游戏中击败了人类职业玩家。前不久DeepMind团队结合之前发表的论文中的成果，又提出了UNREAL（无监督强化和辅助学习/UNsupervised Reinforcement and Auxiliary Learning）代理，在一套57个Atari游戏全集和拥有13个级别的3D迷宫游戏Labyrinth中测试了一新系统，并在新系统中超过了人类玩家。

从以上Atari游戏中人工智能的发展看，我们开发的AI已经可以自主解决复杂问题，通过强化学习代理对大量数据进行长时间训练，提高我们的通用学习算法，使用无监督学习过程利用神经网络来教电脑玩视频游戏，可以使计算机在一系列复杂任务当中与人类表现相当，这是一个算法的突破点。

1.2 RTS游戏

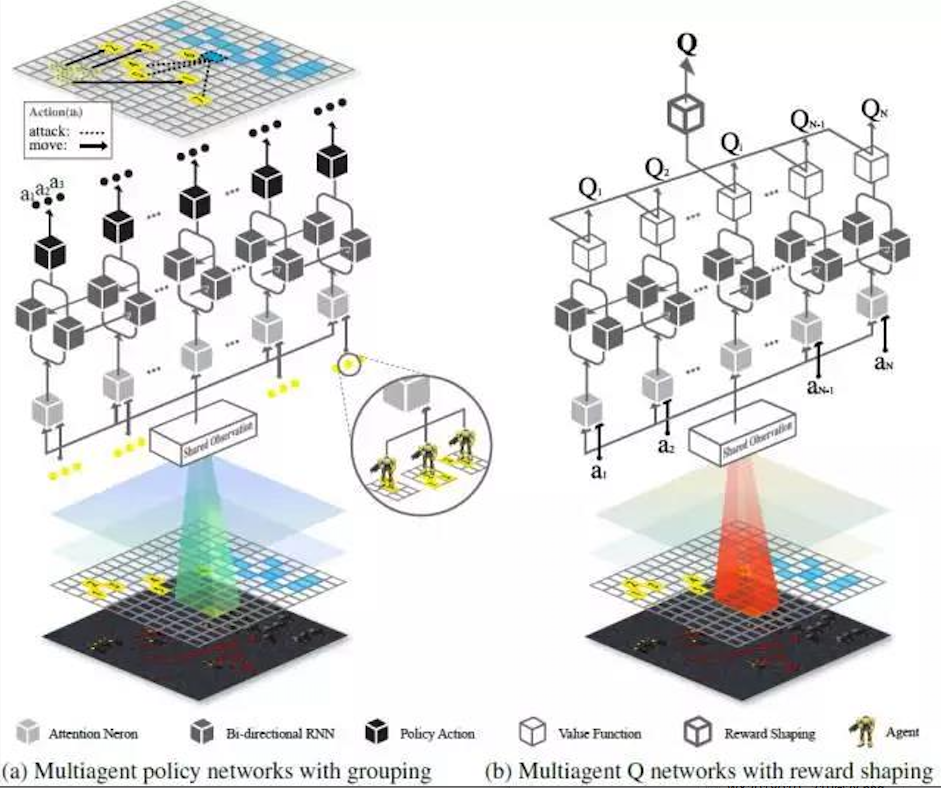
阿里巴巴和UCL的研究者们新提出了双向协调网络BiCNet来相互交流，通过评估-决策方式来完成星际争霸战斗任务。以 RTS 游戏《星际争霸》为测试场景，设定任务为多个智能体互相协作试图击败敌人。为了保证沟通方式有效且可扩展，他们引入了多智能体双向协调网络（BiCNet），它具有向量化扩展评价器（actor-critic）形式。验证了 BiCNet 可以协调不同兵种，在不同的场景和两方智能体数量任意的情况下正常工作。分析证明，在没有手动标记数据进行监督学习的情况下，BiCNet 可以学会多种有经验的人类玩家展示出的协调策略。而且，BiCNet 能够轻松适应异构智能体任务。在实验中，在不同的场景下用该新方法与不同的基准进行了对比；BiCNet 展现出了最先进的性能，它具有在现实世界大规模应用的潜在价值。

图2 双向协调网络示意图

BiCNet是一种利用双向神经网络的多智能体强化学习框架。通过构建矢量化的评估-决策方式，其中每个维度对应一个智能体Agent。Agent之前的协调通过内部双向通信。基于端到端学习，BiCNet可以学会多种有效的协同作战策略。BiCNet提出了一种强化学习与迁移学习的方法，从而可以解决星际争霸中多Agent控制问题。

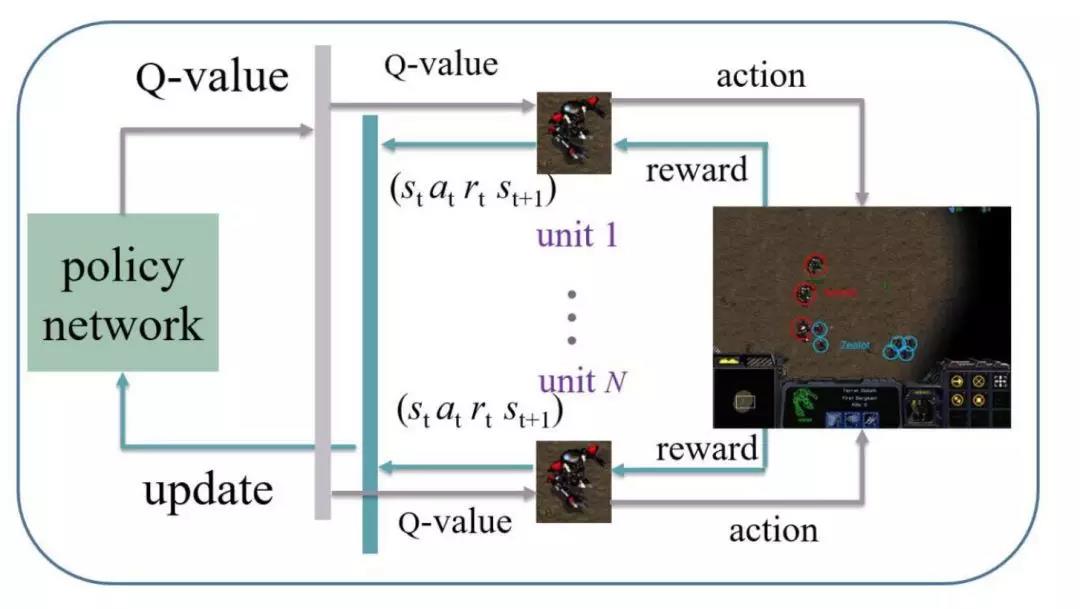
中国科学院自动化所也在近期提出的一种强化学习+课程迁移学习的方法，可以让AI Agent在组队作战过程中掌握微操作能力，控制多个单元。该研究定义了一种高效的状态表征，提出一个参数共享多智能体梯度下降Sarsa()(PS-MAGDS)算法训练单元，该方法使用一个神经网络作为函数近似器，以评估动作价值函数，还提出了一个帮助单元平衡智能体的移动和攻击。通过强化学习和课程迁移学习，智能体Agent可以在星际争霸的微操场景中学习合适的策略。

图3 StarCraft微操场景中的PS-MAGDS强化学习图示

在中科院的研究中，星际争霸微操被定义为多智能体强化学习模型，整个强化学习范式如图3。

去年第一季度,暴雪和 DeepMind开源了星际争霸2的机器学习平台，这个平台对于state-of-the-art的深度强化学习算法是极好的测试平台，此外，两家公司还发布了上百万帧的来自天梯上专业选手的游戏记录。在新界面里，《星际争霸 2》的图形被简化以便于机器学习系统进行识别，同时他们开放 API，允许系统读取游戏中的数据，实现原先电脑玩家的部分功能。暴雪还发布游戏 replay 文件数据集以供机器学习。DeepMind 的最终目标是让人工智能系统和人类玩家一样，通过处理视觉信息理解游戏。同时研究人员还开发出了一些更加有效的平台来推动RTS游戏探索多智能体的控制方向上的发展，其中包括TorchCraft、ELF和PySC2。

通过强化学习和课程迁移学习，我们的AI Agent能够在星际争霸微操场景中学习到合适的策略，并且在多单元协作行为上表现出众。DeepMind牵手暴雪，意在要让人工智能征服星际争霸，而这将为我们解决人类生活环境中的协同工作提供思路。当然我们有望可以先在游戏上应用上达到创新，以推动游戏的发展。

1.3 MOBA游戏

去年8月，由OpenAI设计的bot打败了Dota2世界顶级玩家，在人机匹配赛之前，OpenAI bot接受了一千次比赛的训练，而且bot通过自我对弈，可以预测其他玩家的下一步操作。从去年起，腾讯AI Lab一直致力于MOBA游戏上的AI研究。于是在今年前不久，匹茨堡大学、腾讯AI Lab等机构提交到ICML 2018大会一篇论文，展示了人工智能在王者荣耀中目前的能力。

2006年Remi Coulom首次介绍了蒙特卡洛树搜索（MCTS）并在2012年由Browne等人在论文中进行了详细的介绍。最近几年，MCTS在游戏AI的领域引起了广泛关注。而在近日的研究中，研究者们提出了一种新方法，在训练中充分发掘MCTS的局部特性，迭代更新所有状态的全局策略。其主要贡献为：

（1）提出了一个基于批量MCTS的强化学习方法，利用子节点鉴别器通过之前的树搜索结果进行更新生成更强大的树搜索；

（2）提供了方法的复杂度分析，表明足够大的样本规模和充分的树搜索可以使估计策略的性能接近最优；

（3）基于反馈的树搜索算法的深度神经网络实现在MOBA游戏《王者荣耀》上，且表现良好；

在《王者荣耀》的测试中，研究者添加了游戏内建的AI狄仁杰作为基准，选择六个内建AI狄仁杰能够打败的射手类英雄来对抗测试对手，在游戏中英雄对敌人造成伤害或者战胜敌人时，都会得到金币，经过对决，智能体在对决完成后金币比例总成达到一个高的值，表明其性能良好。



图4《王者荣耀》游戏模式截图

这项研究表明，MCTS已在人工智能领域取行成果，证明这项技术的深度神经网络实现可以创建一个适合MOBA游戏的有竞争力的AI智能体。这在某一种程度上提供了MOBA高阶AI智能体的可能性。

1.4 MMORPG

1.4.1 MMORPG发展历史

1980年第一款MMORPG（大型多人联机角色扮演）游戏诞生——《凯斯迈之岛》，后该类游戏逐渐进入玩家视野，从上世纪90年代末《网络创世纪》、《无尽的任务》到2004年在全球发售运营的《魔兽世界》。



图5 MMORPG发展历程

但是由于游戏市场的更迭以及各种类游戏对市场的逐渐分成，使得MMORPG在近年光环褪去，早在2010年，以《魔兽世界》为例，从曾一度拥有1200w续费玩家降到了780w，而且由于MMORPG游戏研发费用高昂，普通小型公司无法应付资金的巨大压力，而且大型公司则是由于多年市场经验不愿承担高额风险来换取该类型游戏的新鲜血液，导致MMORPG运转状况不佳，难以继续吸引大量新的玩家。



图6 MMORPG的黄金时代

Trion Worlds的《时空裂隙》和EA的《星球大战：旧共和国》都曾希望复制《魔兽世界》的成功模式，然而创新的玩法以及更高的智能体系统在没出现之前，旧模式始终无法玩家的游戏性。

《魔兽世界》截止目前仍然是最成功的MMORPG，2010年付费用户和收入到达峰值后，标志着欧美网游市场MMORPG开始走下坡路。

1.4.2 人工智能在MMORPG上的创新思路

基于Atria视频游戏，RTS战略游戏，以及MOBA对战游戏上算法的研究，将该算法思路运用到魔兽世界插件系统中，人工智能体在通过玩家操作进行训练，采用合适的以上算法训练智能体插件，再通过玩家对战进行实操，从而提供一套自学习的插件系统，智能插件系统将会辅助玩家进行预判，游戏决策以及团队决策等任务，这将成为游戏应用的一个创新设计。

## 2. 应用与创新实践

分析AI算法在MMORPG应用上的可行性:

首先World of Warcraft API是Blizzard WoW客户端提供的一套Lua函数与工具集，可以使用户与游戏客户端进行交互，同时用户操作的数据可以通过接口获取到；

其次，Lua作为一门轻量而快速的脚本语言，功能在高级动态语言中十分完备，对C API以及线程安全的VM支持，非常适合于处理业务逻辑；

本应用的几个难点在于：

（1）如何通过WoW API收集用户操作以产生可用的训练数据集；

（2）如何将可用的训练的数据运用到算法中，从而产生训练结果以及结果回传；

（3）如何进行训练算法的改进，提供完整的适用于玩家插件需求的结果数据；

最后，训练算法基于C++实现，Lua在C/C++上的数据传递以及封装上都比较方便，又魔兽世界又有开放的Lua脚本API，这将为该强化学习的系统在其上应用提供了保障。

## 3．总结与展望

基于星际争霸等游戏的机器学习和深度学习算法，使得游戏多智能体能够进行合作，以达到人类玩家微操的效果，这种强化学习以及迁移学习的方法，在Atari视频游戏，RTS战略游戏，MOBA对战游戏上的可行性得到了初步的证明。

而MMORPG游戏，特别是魔兽世界，在即时战略系统上也可以使用相同的原理，基于魔兽世界插件的开放接口，可以将该训练模型用于魔兽团队战略插件系统中。

这样的一次尝试将机器学习训练系统运用到MMORPG游戏中，这将进一步肯定AI算法在游戏中的应用，将有利于创造更完善的游戏AI系统，增加更智能的Boss体，以及更加智能的战斗游戏策略，以此来提高游戏体验。

## 参 考 文 献

[1] Mnih V, Kavukcuoglu K, Silver D, et al. Playing Atari with Deep Reinforcement Learning[J]. Computer Science, 2013.

[2] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, et al., "Human-level control through deep reinforcement learning," Nature, vol. 518(7540), pp. 529-533, 2015.

[3] Peng Peng, Quan Yuan, Ying Wen, Yaodong Yang, Zhenkun Tang, Haitao Long, Jun Wang, “Multiagent Bidirectionally-Coordinated Nets: Emergence of Human-level Coordination in Learning to Play StarCraft Combat Games,”, (1) Alibaba Group, (2) University College London, 2017.

[4] Kun Shao, Yuanheng Zhu, Dongbin Zhao, “StarCraft Micromanagement with Reinforcement Learning and Curriculum Transfer Learning,”,IEEE, 2018.

[5] Oriol Vinyals, Timo Ewalds, Sergey Bartunov, Petko Georgiev, Alexander Sasha Vezhnevets, Michelle Yeo, Alireza Makhzani, Heinrich Küttler, John Agapiou, Julian Schrittwieser, John Quan, Stephen Gaffney, Stig Petersen, Karen Simonyan, Tom Schaul, Hado van Hasselt, David Silver, Timothy Lillicrap, Kevin Calderone, Paul Keet, Anthony Brunasso, David Lawrence, Anders Ekermo, Jacob Repp, Rodney Tsing, “StarCraft II: A New Challenge for Reinforcement Learning,”, 2017.

[6] 陈钇似.基于机器学习的游戏智能系统研究与应用[D].电子科技大学.2017

[7] 郑其宝.人工智能影响下数字游戏智能化发展探究[D].南京艺术学院.2016

[8] 常迎春.基于分层时序记忆算法的游戏人工智能的异常检测[D].河北工业大学.2015

[9] 姜文军.网络游戏中人工智能的研究及应用[D].上海交通大学.2012

[10] <http://www.infoq.com/cn/articles/atari-reinforcement-learning>

[11] <https://blog.csdn.net/mydear_11000/article/details/51488118>

[12] <http://baijiahao.baidu.com/s?id=1566325283820033&wfr=spider&for=pc>

[13] <https://www.nextbigfuture.com/2014/12/deep-learning-and-deep-reinforcement.html>