

基于强化学习与深度强化学习的游戏AI训练

沈晨琦

深圳大学 广东深圳 518060

摘要: 游戏AI训练的提出主要是计算机与人工智能相结合的训练方式手段,是强化学习领域的一项主要载体环境。现阶段在游戏AI训练的环境下,面临道德困难问题和技术创新问题。主要集中在系数且延迟的反馈分析、高维状态动作的空间环境和不稳定的环境特点等方面。当前基于深度强化学习,我们就是需要通过对强化学习和深度强化学习的迈进,提出注意力机制为基础下的深度强化学习的基础框架结构,解决了复杂环境状态下的集群智能问题。
关键词: 游戏AI训练; 强化学习; 深度强化学习

Game AI training based on reinforcement learning and deep reinforcement learning

Chenyu Shen

Shenzhen University, Shenzhen, Guangdong 518060

Abstract: game AI training is a combination of computer and artificial intelligence. It is a main carrier environment in the field of reinforcement learning. At this stage, in the environment of game AI training, we are faced with moral difficulties and technical innovation. It mainly focuses on the feedback analysis of coefficient and delay, the space environment of high-dimensional state action and the characteristics of unstable environment. At present, based on deep reinforcement learning, we need to put forward the basic framework of deep reinforcement learning based on attention mechanism through the progress of reinforcement learning and deep reinforcement learning, so as to solve the problem of cluster intelligence in complex environment.

Keywords: game AI training; Intensive learning; Deep reinforcement learning

游戏AI训练的定义相对较广,方式可以在现有工作环境中产生出来适当水平的智能化,继而让游戏变得更加真实、有趣味性和挑战性就能被视为游戏AI,在整个游戏AI的构架中只有两种方式,一种是基于有限状态机或是行为树的特点的AI,其本身的形式表现可以被相关的技术所预测并进行分析,另一种是以神经网络为主,通过遗传学内的神经网络和数据计算分析得到非定性的AI,其本身的行为难以得到预测分析。

一、基于强化学习与深度学习的游戏AI训练的现状分析

在计算机学习技术的发展之中,我们可以利用强化学习理论设计游戏AI可以将游戏的智能角色在整个游戏

之中呈现出来的过程,转化为一个简单的马尔科夫的模型结构,智能角色利用他的智能感知系统或区域环境以及自身的工作形式,充分的结合自身的经验,从而有效的选择执行的操作行为分析,该行为的主要落实作用于游戏系统而实现了游戏进程的不断推进,同时环境形势下往往会对智能角色的执行行为进行反馈分析评价家判断。对于智能角色而言,如何获得最佳的行为措施,就可以获得环境最佳的反馈状态为主要的工作任务目标。智能化的角色在游戏环境之中执行这种操作形式,就可以构架成为一条基础的马尔科夫链条。循环持久之后所得到的最佳反馈效果就是当前工作的主要任务目标。传统的机器学习方式方法主要是以低维输入为基础,在这种环境下所收敛的实际效果都相对较为满意。但是随着不同硬件设施的实际性能逐渐的增强,同时随着神经网络本身的技术的创新,游戏智能所面临的问题也相对较

作者简介: 沈晨琦(2000-),男,本科,汉族,籍贯:上海,深圳大学在校学生,研究方向:计算机科学。

多,技术问题逐渐的凸显,现阶段主要集中在以下几个方面。

一是高维度模式下的空间和动作空间技术面临着一定的瓶颈问题,现阶段的行为智能技术智能完成较为简单地行走,不能实现有效的预测研判和决策等不同等复杂行为模式。二是游戏之中的反馈较为稀松,同时也有一定的延迟性效果,整个游戏的模式和流程在特殊的情况下难以对环境产生出较为直接的影响,导致智能对象的训练分析变得更加的困难。而针对此类问题的出现,神经网络模式优化成为了应对高维度动作空间管理技术调整的主要措施手段,强化学习为反馈系数以及延迟提供了潜在的应对路径。智能化的操作方式对于环境不稳定可以建立出一种通信高效的管理模式或是共享相关的参数,从而达到最佳的应对方式。

二、强化学习与深度强化学习的游戏AI训练实验基础分析

通过实验载体的筛选分析,初步拟定相关技术手段,在实验基础平台构建的同时形成实验的基础原型,从而实现了传统算法的预训练分析。通过调研多个与游戏智能相关强化学习训练平台的分析可知,我们用第一人进行设计类经典游戏为载体构建,从而加快载体结构的分析梳理,知名游戏引擎所开发的一种对战平台的构建,都是保证强化学习推进和游戏AI训练有序推进的基础。在当前工作的背景下,我们使用相关操作管理开发模式为基础,将其作为实验载体内容,以游戏引擎危机和赋予此平台稳定操作工作标准,同时可以高度的构建出实验环境以及版本迭代分析模式,底层采用合理化的控制优化措施进行分析,基于软件开发的实现方式,可以让相关的人员尽快地熟悉游戏AI训练的工作内容,保证最佳的游戏效果。

三、基于强化学习的游戏AI训练设计方案构建

强化学习主要是将整个智能体以及环境交互作为整体核心问题,通过选择性行为分析,实现最大化的累计奖励方式构建,从而不断地做好学习优化措施落实,继而从中形成最佳的工作序列。可以通过在游戏场景环境下进行强化学习,基于设定的奖励性训练得到整个复杂环境中的一种AI行为的策略手段。以此介绍初级中系统的强化学习模式。

深度强化学习的构建主要是结合了深度学习所形成的一种强化的学习能力基础,在很多的场合下都能彰显出最佳的工作效果,但是在面对长度决策的问题过程中,往往会呈现出不佳的表现形态。为此我们以深度强化学习理论、多智能体系统理论、强化分层思想理论

以及注意力机制为基础,继而实现工作模式的创新。强化学习也就是在全面观测系统的前提下,单个智能体的决策定义主要是构建一个科尔马夫链的过程。这个过程的实现主要是以一个元组结构构建形成的。在一个特殊的实践环境下,智能体始终处于一种稳定的状态环境下,通过策略执行动作的时限,继而获得环境反馈奖励收获,同时根据转移方程结构形态,进入到下一个状态模式。对于以智能个体构建为核心的马尔科夫链,定义环境下有损失系数的奖励反馈效果以及动作值定义。通过最大化数值的反应,我们可以获取一个最优的决策策略结果,因此采用的最优策略函数值是一定的。在强化学习的过程中,不会有一种固定的马尔科夫的过程,智能体需要通过和环境之间的交互学习达到最佳的策略。当前各种方式环境下我们广泛使用的措施手段就是结合深度学习而使用的,这种方式通过备份更迭的方式对核心数值函数进行估计。

在函数方面的研究发展较为迅速,这也得益于深度神经网络数据的研究分析和突破呈现,结合深度神经网络可以将一些高维的数据直接进行传输和方法。这些构建下的函数逼近方程的操作方式就在其中。现阶段我们所提及的Deep Q-Learning(DQN)就是一种被人们所广泛接受的一种处理方式和方法,当前已经集中使用到了不同的领域环境之中,包含了围棋、雅达利游戏等等。其中最为具体的呈现方式在于在第一次的迭代环境之中,从内存之中提取一定的经验元素,来实现其参数信息的更新传递,在更新的过程之中使用最小化的损失函数方程保证最佳的工作效果实现。经验内存采用的是最为先进的队列表现形式,其中存储的智能化的探索策略和经验元组内容数据,都能支撑后期运作的稳定推进。目标的相关网络参数在这样的形式下更新的速度和频率相对较低,这样所结合的经验回放学习机制,构建成为一种稳定的Deep Q-Learning关系。

为了更好地应对反馈奖励系数或是存在的延迟环境下并不能获得较佳的工作效果的问题,通过使用一种系统等专业化框架结构主体构建出一种强化学习模式,这项框架定义在每一个时间环境之中,智能体会选择出一种原始的动作或是涵盖多个步骤的措施手段进行分析。每一个策略的实施都涵盖了不同的原始动作或是其他的策略手段,同时根据随机的函数信息完成工作任务的推进。为此我们将传统的马尔科夫的计策过程将其拓展延伸成为半马尔科夫决策的过程手段,应对稀疏反馈以及延迟反馈到各种问题。

结合多智能体的训练分析,可以在现有的环境下构

建出一个独立的网络基线训练方式，单独为每一个智能对象构建出网络环境，将多个智能对象放置在一起完成训练操作。在此背景下通过智能对象的通信协议分析，可以让智能对象共享部分的观测信息化数据，达到精准决策的目的。在卷积神经网络之中进行特征的分析提取，之后对于智能对象进行研判，以沟通中心的分析判断实现情报数据的共享，再实现注意力机制的高效处理。现阶段为了分析深度强化学习方法下的动作类游戏的智能对象研究分析，我们可以综合上层结构输入小的工作特点，将其作为智能对象视野进行画面传输反馈，再进行卷积网络的特征提取分析，提取的特征数据信息通过沟通中心的构建向着智能对象完成共享分析，再将整个数据信息传输到主操作结构之中在进行加权以及追踪处理分析，再转移到网络训练模式下，就会出现多个实践规划目标结构。而下层结构输入多数都是智能对象为基础的视野和上层结构所形成的规划目标，主要是通过网络的训练分析之后输入相关的执行动作，再交给游戏的智能对象进行执行，实现与环境进行综合，继而达到任务目标的构建。这个过程不断地完成训练操作，最终期望可以通过分析达到最优协作处理以及对于策略参数下的最佳工作模式策略。

四、基于强化学习与深度强化学习的游戏AI训练实现

为了完成每一个子任务的工作目标，就需要加快不同子任务的分析处理，这样才能真正地完成任务内容。在基于随机神经网络下的强化学习任务目标所以出的用随机神经网络预先学习环境之中的通用技能手段，同时为每一个任务训练的落实推进以及单独的策略措施，调节不同的技能手段，这样所提出的强化学习模式中，设计了工作者、管理者的双层结构形式，工作人员按照每一个步骤的相关要求和迁移结果，将其作为管理者的迁移结果任务目标，由于管理者的实践性质较为突出，智能体的探索方向和能力也得到了全面地提升优化。

强化学习法在复杂的游戏设计环境下有着较为凸显的优势。诸如在不同的平台环境构建之中因为受到环境特定的长序贯、奖励系数等特征的影响，传统的强化学习方法本身并不能在游戏主体之中发挥出最佳的效果，另外一些诸如迷宫、蚂蚁寻物游戏也因为不同的场景环境的高复杂度状态，需要经过特殊的设计分析才能获得最佳的效果。但是通过层次实现强化学习的方法，就需

要在任务分解的背景下进行单独的子目标构建，并且基于一个层次的奖励分析，这样可以极大程度的优化提升学习效果，做好控制措施的落实推进，以达到最佳的任务水平为此强化学习方法就是通过学习的方式进行复杂性的游戏设计，从而形成潜力的优化应对方法。

虽然在当前的工作方式下强化学习可以有效地应对决策的核心问题，但这种结构形式多数都需要策略训练方法，样本利用效率相对较低，同时训练的实际难度较大。在此背景下将强化学习与异策略训练相结合，同时结合异策略修正方式进行工作落实，通过最大的估计分析，替换样本之中的子目标结构。试验的结果分析可知，基础算法的工作效果优于传统的算法结构。但是传统的算法使用随机的采样方法寻求一种近似的替换子目标计算方式，这种做法形态不能对子目标空间得到较好的限制和影响，子目标仍旧可以选择一种没有意义或是不能直接实现的方式方法，让下层学习存在的问题出现一种不稳定的工作效果。

随着产业结构的发展优化，游戏的工作机制和行为环境变得更加的真实和复杂多样，很多环境下都需要通过设计的复杂行为实现AI游戏之间的交互分析，这种技术水平较大程度的影响着游戏产品本身的质量和使用者们的游戏体验。这种AI游戏的设计往往费时费力，对于游戏的开发工作人员来说是一项巨大的挑战。这项发明以强化学习方法为基础，将原有的人工智能和环境交互的行为手段，分解成为了一系列的小型任务目标主题。在完成Q表的构建过程中，加快任务目标构建创新，保证各项任务的有序完成。

五、总结

为了达到复杂环境的分析，实现复杂行为决策问题的有效研判，就需要利用相关的设备引擎和插件信息，实现AI对抗智能系统的开发构建。在此背景下，本文所构建出的框架结构主体本身就有着较为鲜明的特点优势，只有做好了强化学习与深度强化学习的游戏AI训练，才能达到最佳的工作处理效果实现游戏的全面推进，降低不同隐患问题对于游戏实现所产生的直接影响，提升游戏体验。

参考文献：

- [1]周志华.机器学习[M].北京:清华大学出版社, 2016
- [2]周文吉, 俞扬.分层强化学习综述[J].智能系统学报, 2017 (05): 12-16.