随着机器人技术发展的愈发成熟，机器人在我们生活中扮演着越来越重要的角色。在机器人相关技术领域中，路径规划是非常重要的一项技术。在理论中，我们可以将地图数据传送给机器人进行分析，从而得到最佳路径，但是在实际应用中，机器人是不可能知道一个未知环境中的地理信息的，因此，对于在陌生环境中的，不具有先验知识的机器人，研究其探索和规划路径功能是具有非常重要的意义的。

本文以机器人在未知环境中的路径规划为背景，学习强化学习算法实现机器人的路径规划。在原有的强化学习算法Q-learning基础之上，提出改进，提高性能，完成规避障碍，求得最优路径功能。

With the development of increasingly robotics, robots play an increasingly important role in our lives. In the related field of robotics, the path planning is a very important technology. In theory, we can transport map data to the robot for analysis to obtain the best path. However, in practice, it is impossible to know the map data in an unknown environment, geographic information. Therefore, the study which a robotic without prior knowledge explore and find the best way in an unfamiliar environment is a very important significance.

In this paper, robot path planning in unknown environment as the background, we learn reinforcement learning algorithm to achieve robot path planning. Based on the original reinforcement learning Q-learning algorithm, we propose improvements to enhance performance, complete obstacle avoidance, get the optimal path function.

Q-learning，移动机器人，路径规划，避障，强化学习，最优路径

移动机器人是一种在复杂环境下工作，具有自行组织、自主运行、自主规划的智能机器人，它集中了传感器技术、信息处理、电子工程、计算机工程、自动化控制工程以及人工智能等多学科的研究成果，代表机电一体化的最高成就，是目前科学技术发展最活跃的领域之一。随着机器人性能不断地完善，移动机器人的应用范围大为扩展，不仅在工业、农业、医疗、服务等行业中得到广泛的应用，而且在城市安全、国防和空间探测领域等有害与危险场合得到很好的应用。因此，移动机器人技术已经得到世界各国的普遍关注。

移动机器人根据移动方式来分，可分为：轮式移动机器人、步行移动机器人单腿式、双腿式和多腿式、履带式移动机器人、爬行机器人、蠕动式机器人和游动式机器人等类型；按工作环境来分，可分为：室内移动机器人和室外移动机器人；按控制体系结构来分，可分为：功能式水平式结构机器人、行为式垂直式结构机器人和混合式机器人；按功能和用途来分，可分为：医疗机器人、军用机器人、助残机器人、清洁机器人等。

移动机器人的研究始于60年代末期。斯坦福研究院的和Charles Rosen等人，在1966年至1972年中研发出了取名为Shakey的自主移动机器人，目的是研究应用人工智能技术，在复杂环境下机器人系统的自主推理、规划和控制。我国机器人研究开始于世纪年代，至今已有多年。年，北京机械工业自动化研究所曹祥康在江苏嘉兴召开了全国性机械手技术交流大会，是我国第一个以机器人为主题的大型会议。比较有代表性的有：清华大学研制了智能移动THMR-III V型机器人；中科院自动化研究所研制出智能保安机器人，全方位移动式机器人视觉导航系统，智能移动机器人通用平台AIMR;香港城市大学研发了自动导航车及服务机器人；哈尔滨工业大学研制了导游机器人；国防科技大学研发成功了双足机器人等；“海龙2号”是我国自主研制的水下机器人，是我国目前仅有能在3500米水深、海底高温和复杂地形的特殊环境下幵展海洋调查和作业的最高精技术装备，它除了在潜水深度上的优势之外，还在国际上首次采用了一些自主研发的先进技术，包括虚拟控制系统和动力定位系统。

移动机器人导航移动机器人学的关键技术，是指按照预先给出的任务，根据已知的地图信息作出全局路径规划，并在行进过程中，不断感知周围的局部环境信息，自主地作出各种决策，随时调整位姿，引导自身安全行驶或跟踪已知路径，达到目标位姿，是移动机器人各项研究应用的基础和前提。自主导航是移动机器人导航的一项重要的基本能力，它主要解决的问题可以归纳为三个方面：“1.现在何处？2.要往哪去？3.如何去？”也就是移动机器人的地图构建、定位和路径规划题。地图构建是移动机器人在自主主导航的过程中，通过传感器感知环境并建立环境模型，这是述立环境模型最常用且有效的方法；定位移动机器人在工作环境中探索所处位置的过程，也就是机器人的全局坐标和姿态；路径规划技术就是移动机器人在具有障碍物的环境上按照一定的评价标准（如作代价最小、行走路径最短、行走时间最短等）、寻找一条从给起点到达目标终点的无碰路径。

但是不管采取何种导航方式，移动机器人上主要是完成路校规划、定位和避障等任务。路径规划问题是移动机器人自助式导航技术研究中的关键技术之，它是指移动机器人按照某一性能指标，搜索一条从起始位置到目标位置的最优或次优的无碰路径

随着计算机、传感器及控制技术的发展，特别是各种新算法的不断涌现，移动机器人路径规划技术已经取得了丰硕的研究成果，特别是在周围环境未知或部分未知的局部路径规划（人工势场法、模糊逻辑算法、神经网络法、进化算法、强化学习等），国内外学者已经作了大量的研究和改进，下面综述中将对与本文有关的一些方法进行讨论。

1986年，Khatib首次提出用人工势场法解决机器人避障问题，在此之后许多学

者不仅发展了这一方法并将其应用于机器人的实时路径规划和足球机器人的运动

控制。人工势场法可以实现快速的控制，所以人工势场法丨被广泛的应用于实

时的运动控制。

E.Shi等进一步对传统的人工势能场方法进行改进，他提出的方法虽然能够使移动机器人避幵障碍物，但其实施性并不强，虽然克服了零势能场（在这种情况下机器人无法到达目标位置）的情况，但又出现了很多问题，比如障碍物的影响范围、引力增益和斥力增益会影响到机器人运动过程中振动的程度、运动路径的平滑度、路径的长短及其运动过程中是否会碰撞到障碍物等，所以必须同时选择这三个系数的最优值，而在实验过程中是很难选取的，必须通过反复的实验才能选取较为合适的值；王奇志等提出了一种改进的人工势场法，通过排除一个距离机器人最远的障碍物，同时加一个同等大小反方向力来消除零势能点，从而实现消除零势能域，达到多障碍物情况下机器人运动规划的快速、实时、避障的效果，结果表明，该算法对多个障碍物和非静态的障碍物同样适用。陈华志等采用一种具有速度负反馈的基于模糊控制的移动机器人避障算法进行研究和仿真对移动机器人在实际环境中进行了实验，并证明了方法的可行性；Lee等使用模糊控制来决定移动机器人可能的十三个移动方向，使机器人在这种导航下避免机器人内部碰撞；Pradhan等在完全未知的环境中使用模糊逻辑对

多达1000个机器人进行了导航实验，在试验中，对四个输入变和两个输出变贤进行模糊化，最后证明了高斯隶属函数在移动机器人导航中最有效；Yang等设计了移动机器人的动态模型，根据多机器人的局部信息，建立了引力、斥力函数的模糊控制器，基于最优控制理论，来衡量系统的状态性能指标；刘利等针对海域不同的障碍环境，从运动学特性出发，提出了一种多主水航行器（AUV）编队避障策略。该策略是在领航跟随法的基础上，利用编队队形的几何关系，引入虚拟AUV，再采用模糊控制避障策略，使编队能顺利通过障碍物，到达目的地；Rigatos等研究了在包含固定障碍物和移动物体的部分未知的环境下，根据移动机器人的运动问题和控制规律，结合滑膜与模糊逻辑控制的基本原则，提出使机器人达到期望的运动效的控制器，即降低滑动模式模糊逻辑控制器的复杂度，具有稳定性和简易性；Montaner等为移动机器人导航设计了个模糊逻辑控制器，该控制器可以接收确定性和模糊的信息，并在一个由七个超声波传感器来获信息的环境中对移动机器人进行了文验研究了移动机器人在未知排集环

境屮反应式分航的题，文卜把反应式导航记义为感知数抓和命令的义系，

个反应式导航系统祁、于对机器人提供了个运动的地阁，在此味础卜出了

分种遍”的遍近方法——投糊逻樹，并解释了如何利门〗行为分解的；；法来

选抒校糊规则等提出了种“丨加速减速的模糊控制器校，使移动机器人

在动态环境屮〔主分航，括找丨：丨标、避幵障昭物和找优路找。

和提出了种离敗拓扑结构的祌经网络模项，移动机器人使这

种并联的祌经网络模在离散化的完全已知环境中进行路径规划，并假定机器

人的速度识定，这在真实的环境巾对移动机器人进行控制足没义的

等提出了种雄于模糊神经网络将传感器倍总；机器人的移动结合起来的导航力

法，这种力法可以使移动机器人充分感受周丨彳环境、丨彳〗避开静态和动态障物，

并在不同情况下避力：“死循环”产生到达的含观的轨迹，通过仿证明了该

方法的实性和彳性祌经元系统提出了种感知预测机器

人伙伴的控制方法，该预测方法对于减少计兑位、提収感知丨〖息足非常『取耍的

等在文中奶丨了遗传算法在移动机器人航巾的应川，使机器人“丨川业界

派调査，数掘采等，移动机器人可以在动态环境丨避免内部碰揽了这种兑