Introduction

**비디오 게임이 세상에 등장하고 그 형태는 현재와 비교했을 때 아주 볼품없어 보여도 누군가는 그 게임에 열광하기 시작했다. 아주 단순한 게임 속에서 사람들은 정적인 게임 환경에서 자신 또는 다른 사람의 점수로만 경쟁할 수 있었을 뿐, 다른 누군가와 동시에 게임을 한다는 느낌을 받을 수는 없었다. 더 시간이 흘러 컴퓨터 봇이 등장하게 되었고 비록 단순한 형태의 봇이지만 사람들은 자신이 아닌 누군가와 같이 게임한다는 느낌을 받으며 상대와 경쟁한다는 스릴감을 만들어냈다. 초기 게임 봇은 플레이어와 상관 없이 미리 정의된 패턴에 따라 일련의 동작만을 수행했기에, 그 스릴감은 오래 지속되지 못 하였다.**

기술이 발전함에 따라 게임 봇도 진화했습니다. 현대의 AI 봇은 이제 플레이어의 행동을 관찰하고, 전략에 적응하고, 심지어 고급 알고리즘을 사용하여 자신의 목표를 추구할 수도 있습니다. DeepMind가 개발한 AI 시스템인 AlphaGo가 세계 바둑 챔피언 이세돌을 물리친 것은 게임 AI 역사에서 중요한 순간이었습니다. 이 업적은 바둑 플레이어뿐만 아니라 전 세계 사람들을 놀라게 했으며, 복잡한 의사 결정 환경에서 AI 역량의 새로운 시대를 열었습니다.

**시간이 지나 기술이 발전함에 따라 게임 봇의 양상도 달라졌다. 이전에 플레이어의 입력과는 상관없이 움직이던 게임 봇에서 이젠 플레이어의 행동을 관찰하고, 전략을 세우고, 심지어 여러 알고리즘을 통해 원하는 목표를 아주 손쉽게 이루기도 합니다. DeepMind에서 개발한 AlphaGo가 세계 바둑 챔피언 이세돌을 이긴 순간은 바둑기사들뿐만이 아니라 바둑을 하지 않는 전세계인들이 주목하고 경악했던 이슈였다.**

**이렇듯 봇이 많이 발전해왔는데 이런 봇들에 대해 알아보고 직접 새로운 봇을 하나 만들어**

**그 performance를 측정해 볼 것임.**

**게임은 lunar lander로 할 것임. 이 게임은 달 표면에 우주선이 폭발하지 않고 잘 착륙하게 하는**

**단순한 게임. 그래서 딥러닝 기술을 이용해 ai agent를 만들어보고 그 성능을 점수와 시간으로**

**판단해볼것임.**

*이 프로젝트는 Lunar Lander* 게임을 위한 AI 게임 에이전트를 만들어 AI 게임 에이전트 개발을 탐구하는 것을 목표로 합니다 . 이 게임에서 플레이어의 목표는 충돌하지 않고 달 표면에 우주선을 안전하게 착륙시키는 것입니다. 게임의 목표는 간단해 보일 수 있지만, 우주선을 정밀하게 제어하고 연료 소비를 관리하는 것은 상당한 과제입니다. 이 프로젝트는 심층 강화 학습 기술을 사용하여 우주선을 자율적으로 탐색하고 착륙시키는 AI 에이전트를 개발합니다. 에이전트의 성과는 착륙 성공, 연료 효율성 및 완료 시간을 기준으로 평가됩니다. 이 프로젝트는 AI 에이전트를 만들어 게임 환경에 강화 학습을 적용하는 방법에 대한 이해를 심화하고 AI 기반 의사 결정의 발전에 기여하고자 합니다.

세상에 게임이 처음 등장하고, 사람들은 게임에 열광했지만 한 번 더 열광했음.

그 이유는 컴퓨터 봇의 등장. 이전까지 혼자 했던 게임들과 달리 봇의 등장으로 사람은

다른 무언가와 경쟁한다는 느낌을 받으며 그 매력에 빠짐.

분명 초기에는 user의 행동에 따라 달라지는 것이 아니라 그냥 자신의 정해진 일련의 행동을 하는

봇들이었음. 처음엔 이것조차도 사람들이 좋아했지만 사람같이 않은 봇에 그 열정은 식었을 것임.

그리고 이제 우리는 user의 행동을 보고 자신의 다음 행동을 하거나 아니면 그냥 정해진 목표를 이루기

위해 자신만의 방법을 쓰는 봇들이 사용되고 있음.

특히나 바둑에서의 이세돌과 싸웠던 알파고의 등장은 바둑기사들뿐만이 아니라 바둑을 하지 않는

전세계인들이 주목하고 경악했던 이슈였음.

최근 몇 년 동안, 게임에서 인공지능(AI)을 적용하는 것은 플레이어 경험을 향상시키고 복잡한 의사 결정을 내릴 수 있는 지능형 게임 에이전트를 만들 수 있는 잠재력으로 인해 상당한 주목을 받았습니다. 다양한 게임 환경 중에서 Lunar Lander는 AI 개발에 고유한 과제를 제공합니다. 이 게임은 연료 소비를 관리하고 동적 조건을 탐색하면서 달 표면에 우주선을 착륙시키는 작업을 시뮬레이션합니다. 목표는 제한된 리소스를 효과적으로 활용하여 지정된 착륙장에 달 모듈을 안전하게 착륙시키는 것입니다.

Lunar Lander 게임을 위한 AI 에이전트의 개발은 다양한 강화 학습 알고리즘을 탐색하고 구현할 수 있는 플랫폼을 제공합니다. 머신 러닝의 하위 집합인 강화 학습은 에이전트가 환경과 상호 작용하여 시행착오를 통해 최적의 정책을 학습할 수 있도록 합니다. Lunar Lander는 복잡성과 단순성의 균형으로 인해 이러한 알고리즘을 평가하기 위한 이상적인 테스트베드 역할을 하며, 다양한 학습 전략과 에이전트 성능에 미치는 영향을 분석할 수 있습니다.

이 논문은 Lunar Lander 게임을 위한 심층 강화 학습 기반 AI 에이전트를 설계하고 개발하는 것을 목표로 하며, 다양한 알고리즘과 하이퍼파라미터 튜닝을 통해 에이전트의 성능을 개선하는 데 중점을 둡니다. 심층 학습의 강점을 활용하여 에이전트의 의사 결정 능력을 향상시키고 궁극적으로 더 효율적이고 성공적인 착륙으로 이어지는 것이 목표입니다.

Problem definition & challenges

이 연구의 주요 목적은 Lunar Lander 게임에서 달 착륙선을 효과적으로 탐색하고 착륙시킬 수 있는 유능한 AI 에이전트를 개발하는 것입니다. 이 게임은 매력적인 도전을 제시하지만, 성공적인 에이전트 개발을 보장하기 위해 해결해야 할 몇 가지 문제와 과제도 제기합니다.

1. **상태 공간 복잡성** : Lunar Lander 게임은 모듈의 위치, 속도, 각도, 연료 수준을 포함한 여러 변수로 구성된 연속 상태 공간을 특징으로 합니다. 이 상태 공간의 복잡성으로 인해 에이전트는 다양한 시나리오에서 학습을 효과적으로 일반화해야 하며, 이는 기존 강화 학습 방법을 사용할 때 어려울 수 있습니다.
2. **액션 공간** : 에이전트는 제한된 이산 액션 세트에서 선택해야 합니다. 다른 방향으로 추력을 가하거나 추력을 전혀 가하지 ​​않는 것입니다. 이 제한된 액션 공간에서 에이전트는 달 착륙선의 하강을 제어하기 위해 정확하고 시기적절한 결정을 내려야 합니다. 에이전트가 다양한 착륙 조건에 대한 최적의 액션을 학습하도록 하는 것은 어려운 일입니다.
3. **보상 구조** : 적절한 보상 구조를 정의하는 것은 학습 과정을 안내하는 데 중요합니다. Lunar Lander 게임에서 에이전트는 안전한 착륙에 대해 긍정적인 보상을 받고 추락이나 과도한 연료 소비에 대해 부정적인 보상을 받아야 합니다. 이러한 보상을 균형 있게 조정하여 원하는 행동을 장려하는 동시에 최적이 아닌 전략을 방지하는 것은 상당한 과제입니다.
4. **안정성과 수렴** : 딥 강화 학습 알고리즘은 고차원 입력 공간과 학습 과정의 복잡성으로 인해 불안정성과 느린 수렴으로 어려움을 겪을 수 있습니다. 에이전트가 진동이나 발산 없이 시간이 지남에 따라 지속적으로 성능을 개선할 수 있도록 하는 것이 핵심 과제입니다.
5. **탐색 대 활용** : 에이전트는 탐색(새로운 행동을 시도하여 효과를 발견)과 활용(이전 학습에 따라 가장 잘 알려진 행동을 선택)의 균형을 맞춰야 합니다. 에이전트가 지역적 최소값을 피하면서 최적의 전략을 효과적으로 학습하려면 적절한 균형을 찾는 것이 필수적입니다.

본 연구는 이러한 과제를 해결함으로써 Lunar Lander 게임 환경에서 더욱 강력한 AI 에이전트의 개발에 기여하고, 궁극적으로 게임 환경에서 강화 학습 분야를 발전시키는 것을 목표로 합니다.

게임 규칙은 나중 논문에서 자세히 쓰는 듯.

"lunar lander"에서 game ai agent의 목표는 짧은 시간 안에 고득점을 내는 것.

점수 획득 조건 :

그러므로 착륙하는데 걸리는 시간이 짧을수록 점수를 많이 얻게 설정.

추가 점수를 얻게 하는데에는 착륙할 때의 각도가 수직에 가까울수록,

착륙 속도를 구간마다 나누어 점수 부과

착륙하는 플랫폼의 가운데일수록 점수 높게 부과.

남아있는 연료의 양으로 점수 부과.

게임 제약 :

연료 다 떨어지면 엔진 추진 불가능. 로켓 방향 조절만 가능.

게임 실패 조건 :

시간 초과

플랫폼 제외한 지면에 부딪힘

착륙속도가 얼마 이상일 때