



# 로봇팔을 활용한 체스게임

22 엄돈건(새내기과정학부) 22 최민준 (새내기과정학부)

## 초록



로봇팔(매니퓰레이터)을 활용하여 무언가를 집어서 옮길 수 있는 구조를 가진 로봇 시스템을 당초에 구현하고자 하였다. 특히 이전에 제작해둔 적이 있는 액추에이터를 활용하여 로봇팔을 구현하고자 하였으나, 여러가지 기술적 문제로 이미 있는 기성품 액추에이터를 활용하여 로봇팔의 각 관절을 구성하고자 하였다.

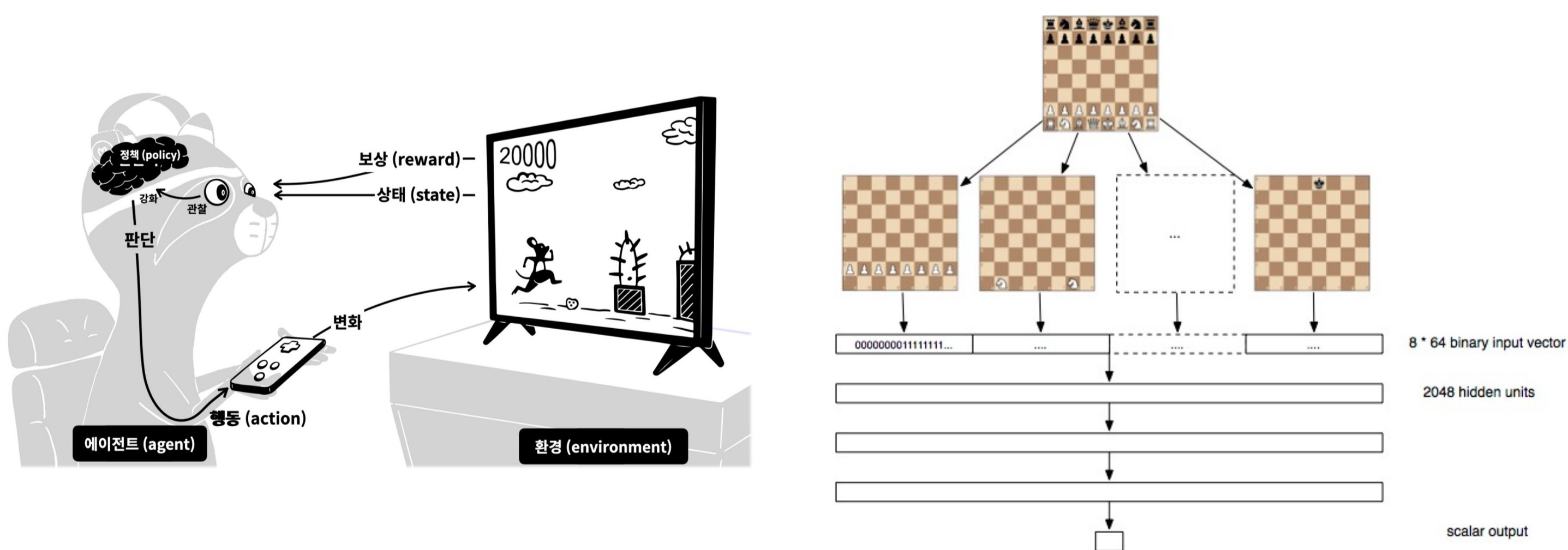
알파고의 원리는 딥러닝을 사용하여 실제 바둑 선수들의 기보를 따라하는 AI를 만들고 그 AI끼리 서로 바둑을 두며 강화학습을 시켰다. 이를 바둑과 비슷한 다른 게임에서도 사용할 수 있을 것이라 생각하여 체스를 두는 인공지능을 강화학습으로 학습시키고자 하였다.

이렇게 만든 로봇팔과 강화학습 인공지능을 활용하여 로봇팔을 활용한 체스 게임 시스템을 구현하고자 하였다.

## 이론적 배경



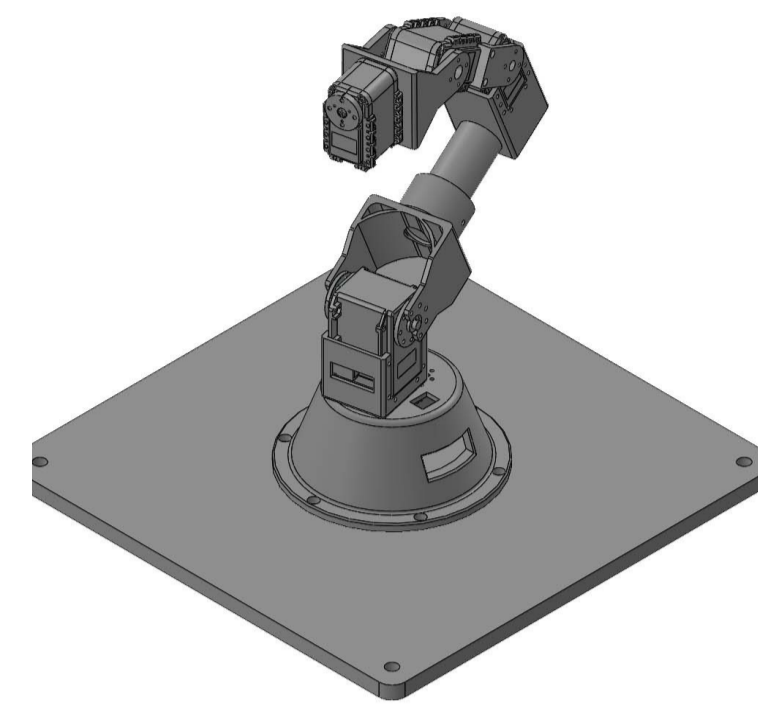
로봇 운영체제(ROS)는 Robot Operating System의 약자로 로봇 응용 프로그램을 개발할 때 필요한 하드웨어 추상화, 하위 디바이스 제어, 일반적으로 사용되는 기능의 구현, 프로세스간의 메시지 패싱, 패키지 관리, 개발환경에 필요한 라이브러리와 다양한 개발 및 디버깅 도구를 제공하는 프로그램을 의미합니다. 본 프로젝트에서는 강화학습을 통해 얻은 결과를 ROS를 사용하여 로봇팔의 모션을 구현하였다.



강화학습은 환경에 따라 주어지는 보상을 통해 정책을 수정하며 학습하는 방법이다. 심층신경망(neural network)로 이루어진 정책에 현재 환경이 입력되면, 해야 할 행동을 출력하게 된다. 이때 최종적으로 받는 보상이 좋을수록 현재 정책에 더 높은 가중치를 부여하여 그 정책을 자주 수행하게 하고, 보상이 나쁠수록 그 정책의 가중치를 낮추고 수정한다. 이를 반복할수록 최종적으로 좋은 보상을 받는 정책으로 변할 것이다.

체스의 경우에는 최종적인 결과가 보상으로 주어지게 된다. 게임에서 이기면 1점, 무승부(stalemate)는 0점, 지면 -1점을 부여한다. 게임에서 이긴다면 승부가 나기 직전에 한 행동은 매우 좋은 행동이므로, 그 상황에서 그 행동을 하는 정책의 가중치를 높인다. 그보다 더 과거에 한 행동도 좋은 행동이나, 그것이 최종적으로 결과에 미친 영향은 적기 때문에 가중치를 올린 것 하나 상대적으로 적은 양만큼 올리게 된다. 질 경우에도 같은 일을 수행하고, 이를 무수히 많이 반복하면 그 상황에서 이길 확률이 가장 높은 행동을 수행하게 될 것이다.

## 로봇 구조



```

===== Player A: f3g2
r . . q r . k .
p p p . . p p p
. . . . .
. . . p . . . .
. . . . .
. P . b b . P .
P . . . . . K P
n . . . . . N .

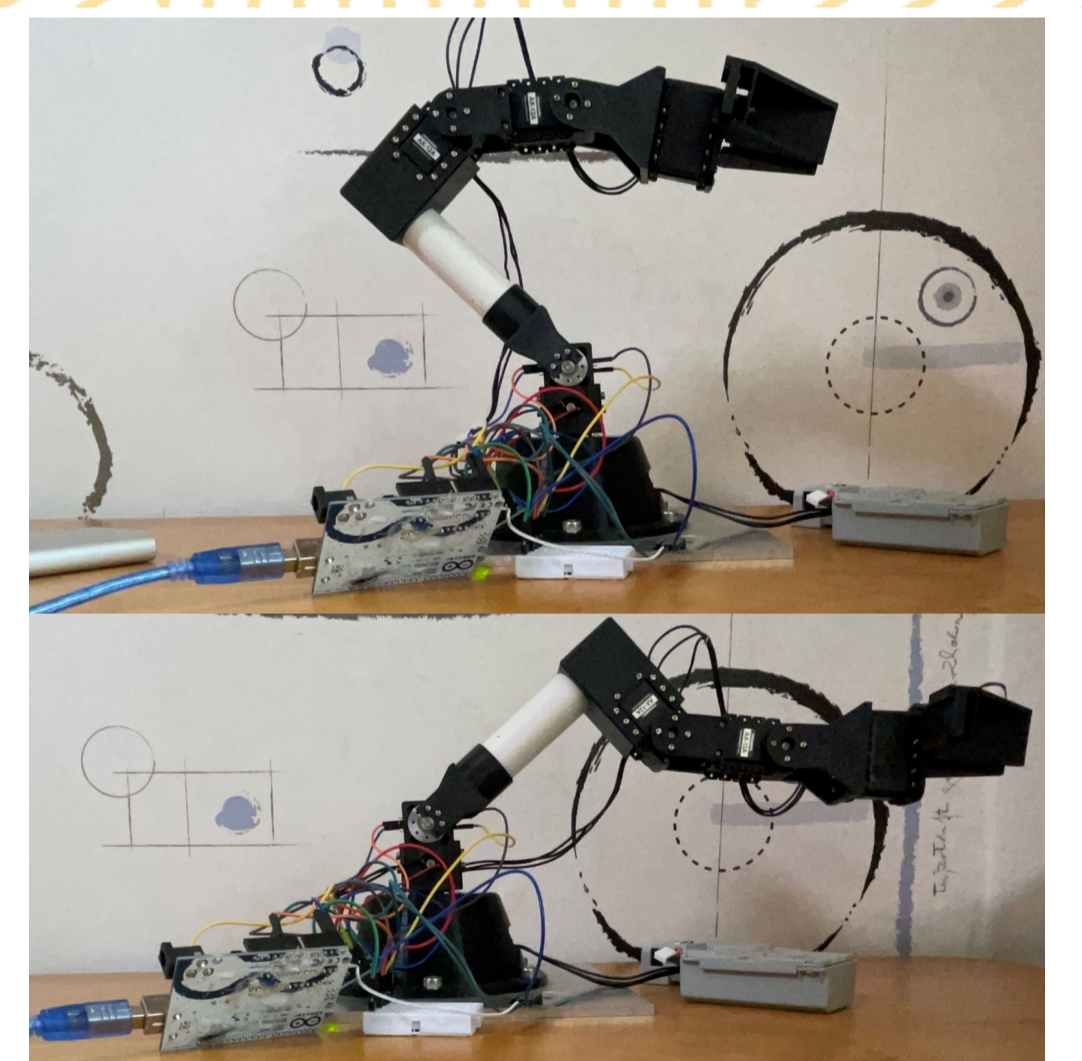
```

실제 로봇팔 제작에 앞서서 설계해둔 부품들을 Solidworks의 어셈블리 기능을 통해 가상으로 조립하여 각 관절의 가동 가능 범위에 문제가 없는지 조립 전 사전 확인하였고, 특이사항이 발견되지 않아 실제 조립을 진행하였다. 로봇팔은 총 5개의 액추에이터를 사용하여 제작하였고 가장 아랫 단 관절의 경우 가장 많은 토크가 걸릴 것으로 예상하여 다른 액추에이터들에 비해 상대적으로 토크 값이 높은 액추에이터를 배치하여 설계 및 제작을 진행하였다.

카메라로 체스판의 상황을 인식하면, 이를 python module인 chess 라이브러리 내의 Board 클래스로 변환한다. Board에는 현재 판의 상황이 기록되어 있고, 이전 상태의 Board와 tree와 같이 연결되어 있다. 이 Board를 심층신경망에 입력시키면, 출력으로 이동해야 할 기물의 위치와 이동시킬 위치를 a2a4의 형태로 출력하게 된다.

## 결과

로봇팔의 각 파트를 3D프린터로 출력하고 조립하였다. 사전에 Solidworks에서 확인한 바와 같이 각 관절 간의 간섭은 없었고, 운동범위도 예측했던 것과 같이 잘 구현되었다. 가장 아래에 위치한 관절에 적용한 액추에이터 이외에 나머지 액추에이터들은 토크가 낮아 상대적으로 로봇의 모션의 완벽히 자연스럽게는 않았지만 픽애플레이스를 진행하기에는 무리가 없는 수준이었다.



딥러닝을 사용하여 실제 진행된 경기를 심층신경망에 학습시켰고, 학습이 진행됨에 따라 loss가 줄어들어 확인되었다. 이를 바탕으로 강화학습을 수행하였고, 다른 체스 수행 인공지능인 sunfish와 비교하였을 때 승률이 올라감을 확인할 수 있었다.

Iteration	Learning rate	loss	
1	0.030000000	2.079441542	1 0.69698 0.50000
101	0.029931454	2.046203073	21 0.68430 0.55811
201	0.029865790	1.988807601	41 0.67678 0.60571
401	0.029736183	1.906336772	61 0.68258 0.57227
801	0.029479273	1.506403297	81 0.67094 0.58976
1601	0.028970367	1.448433834	101 0.68183 0.58102
2631	0.028333206	1.439250729	

## 최종 목표

앞서 완성한 체스 인공지능에서 a2a4의 형태로 출력한 값을 ROS를 사용하여 로봇팔에 체스 말의 위치값을 입력하여 체스 말을 이동시키는 시스템을 구현하고자 하였다. 이때 제작하면서 어셈블리를 한 파일을 사용하여 각 관절 간의 정확한 거리를 측정하고 이를 활용하여 보다 더 정확한 로봇팔 모션을 구현할 수 있도록 하였다.