

AI World Cup 환경을 이용한 멀티 에이전트 기반 지능형 가상 축구 에이전트 구현

이은후², 성현아¹, 정민지², 이혜인¹, 정진우^{1,2}, 이의철^{1,2}, 이지항^{1,2,+}

¹상명대학교 지능정보공학과

²상명대학교 휴먼지능정보공학전공

202132040@sangmyung.kr, leeunhoochu@gmail.com, 202132043@sangmyung.kr, jiminji98@gmail.com,

jioung@smu.ac.kr, eclee@smu.ac.kr, jeehang@smu.ac.kr

Developing artificial football agents based upon multi-agent techniques in the AI world cup

Eunhoo Lee², Hyeon-ah Seong¹, Minji Jung², Hye-in Lee¹, Jinoo Joung^{1,2}, Eui Chul Lee^{1,2}
and Jee Hang Lee^{1,2,+}

¹Department of AI and Informatics, Sangmyung University, Seoul, South Korea

²Department of Human-Centered AI, Sangmyung University, Seoul, South Korea

⁺Corresponding author: Jee Hang Lee

요 약

AI World Cup 환경은 다수 가상 에이전트들이 팀을 이뤄서 서로 상호작용하며 대전이 가능한 가상 축구 환경이다. 본 논문에서는 AI World Cup 환경에서 멀티 에이전트기반 학습/추론 기술을 사용하여 다양한 전략과 전술을 구사하는 가상 축구 에이전트 구현과 시뮬레이션 결과를 소개한다. 먼저, 역할을 바탕으로 협동하여 상대방과 대전할 수 있는 논리 기반 추론형 멀티 에이전트 기술이 적용된 Dynamic planning 축구 에이전트 9 세트를 구현하였다. 이후, 강화학습 에이전트 기반, 단일 에이전트를 조합한 Independent Q-Learning 방식의 학습형 축구 에이전트를 구현한 후, 이를 멀티 에이전트 강화학습으로 확장하여 역할 기반 전략 학습이 가능한 가상 축구 에이전트를 구현하고 시뮬레이션 하였다. 구현된 가상 축구 에이전트들 간 대전을 통해 승률을 확인하고, 전략의 우수성을 분석하였다. 시뮬레이션 예제는 다음에서 확인할 수 있다 (<https://github.com/I-hate-Soccer/Simulation>).

1. 서론

축구는 다수의 플레이어가 서로 협력/협동하고 상대팀과 경쟁하며 승리를 쟁취하는 운동이다. 하나의 환경 안에서 다수의 플레이어가 존재하고, 다수의 플레이어가 개개인의 역할에 따라 팀 내 협력과 팀 간 경쟁을 통해 승리에 이르는 목표를 수행해야 하므로 축구는 다른 게임 환경과 비교하여 고유의 복잡성과 불확실성이 존재한다. 이에 따라, 가상 축구 환경 및 축구 에이전트 구현은 딥마인드를 비롯한 인공지능 에이전트 연구자들의 주목을 받기 시작하였다 [1, 2].

AI World Cup 가상환경은 축구의 속성을 모두 포함하는, 축구 고유의 불확실성이 존재하는 환경이며 다수의 플레이어 에이전트를 활용하여 팀을 구성하고, 전략을 만들어 시뮬레이션이 가능한 환경이다 [3]. 본 연구에서는 이러한 AI World Cup 환경에서 멀티 에이전트

기반 학습/추론 기술을 사용하여, 복잡한 환경에서 다양한 상대에 적합한 전략과 전술 구사가 가능한 가상 축구 에이전트를 구현하고 시뮬레이션 하고 그 효과성을 확인하였다. 먼저, 역할을 바탕으로 팀간 협동을 통해 상대팀과 대전할 수 있는 논리 기반 추론형 가상 축구 에이전트를 구현하였다. 뿐만 아니라, 상대의 다양한 전략을 환경에 대한 불확실성으로 간주하고, 이에 따른 최적 전략을 학습하여 수행하는 강화학습 에이전트 기반 가상 축구 에이전트를 구현하였다. 처음에는 팀 간 협동을 고려하지 않는 단일 에이전트 조합형 강화학습 축구 에이전트를 구현하였고, 이를 멀티 에이전트 강화학습으로 확장하여 협동 및 전략 학습이 가능한 가상 축구 에이전트를 구현하고 시뮬레이션 하였다. 구현된 가상 축구 에이전트들 간 대전을 통해 승률을 확인하고, 전략의 우수성 및 효과성을 분석하였다.



(그림 1) AI World Cup 가상 축구 환경 [3]

2. AI World Cup 가상 축구 환경

그림 1은 AI World Cup 가상 축구 환경 화면이다. 각 팀당 가상 축구 에이전트 5명으로, 수비수 2명, 공격수 2명, 골키퍼 1명으로 구성된다. 상기 5개의 가상 축구 에이전트는 인공지능 알고리즘을 이용하여 제어되며, 상대팀과 대전하여 승리할 수 있는 전략을 개발하는 가상 축구 환경이다.

각 팀별 가상 축구 에이전트의 지능은 Python 프로그래밍 언어를 이용하여 구현한다. 축구 에이전트의 이동은 바퀴 속도 조절을 통해 이루어진다. 에이전트와 공의 좌표를 참고하여 각 에이전트의 바퀴 속도를 설정하고, 원하는 좌표로 이동하도록 한다. 바퀴 속도는 이동 속도 또한 결정한다. 에이전트의 공격과 수비 행위는 이동을 기본으로, 킥과 점프 동작을 조절해 설정할 수 있다. 축구 규칙은 현실 세계 축구 규칙 유사하나, 특이한 점은 여기서는 몸싸움에 대한 제재가 없다. 또한, 플레이어 에이전트의 에너지 소모는 없다.

3. 가상 축구 에이전트 구현

AI World Cup 환경에서 멀티 에이전트 기반 가상 축구 에이전트를 구현하기 위해 크게 Yi et al. [2]이 소개한 세 가지 접근 방법, Random walk, Dynamic planning 및 강화학습형 에이전트 방법으로 축구 에이전트를 구현하였다. 각 에이전트들의 구현 사례는 다음에서 확인이 가능하다 (<https://github.com/I-hate-Soccer/Simulation>).

Random walk 에이전트 아무런 제약 없이 순수 무작위로 행동하는 Random 에이전트 구현 방법이다. 이 접근은 개인 에이전트의 전략 자체가 없고, 주변 상황을 고려하지 않는다. 팀 멤버로서 서로 간의 협력도 없다. 오직 무작위로 선정한 행동을 통해 경기에 참여한다. 세 가지 접근 중 가장 지능적이지 않은 에이전트이지만, 다른 접근 방법의 에이전트들을 훈련하고 평가하는 데에 유용하다. 완전한 무작위 행동을 수행하므로, 상대방과 대전 시 의외성을 줄 수 있으며, 이에 따라 행동에 대한 불확실성을 제공할 수 있다. 또한, 무작위적으로 생성하는 개인 전략, 팀 전략은 예측이 매우 어려운 복잡성을 내포할 수 있으므로, 대전 시 상대방 팀들이 이러한

불확실성과 복잡성을 대비하는 전략 학습에 도움이 될 수 있다. 이에 따라, 복잡성/불확실성에 대응하는 능력을 측정하는 실험군으로 사용하였다. Random walk 축구 에이전트의 행동은 에이전트의 양측 바퀴 속도를 random으로 결정하여 무작위로 움직이도록 하였다.

Dynamic planning 에이전트[4] 논리 기반 추론형 가상 축구 에이전트를 구현 방법이다. 인공지능 디자이너가 보유한 통찰과 전문 지식을 바탕으로 전략을 설계하는 방식으로, 개별 에이전트들에 역할을 부여하고, 상대팀과 공의 위치에 따라 전략과 행동을 직접 디자인하는 방식을 말한다. 이 접근은 빠르고 효율적으로 전략을 구현할 수 있는 장점이 있다. 복잡한 상황과 불확실성을 미리 짐작하여 빠르게 프로토타입을 생성할 수 있고, 대전 상대에 따라 즉각적으로 전략을 수정하고 미세 조정할 수 있으며, 다수의 팀을 빠른 시간 안에 구현할 수 있는 장점이 존재한다. 다만, 빠르게 변화하는 환경에서 복잡성과 불확실성에 대응하기 어렵고, 학습하여 최적 전략으로 대처하는 것이 어렵다. 본 논문에서는 실세계 축구에서 사용된 전술을 모델링하기 위해 본 접근을 사용하였다.

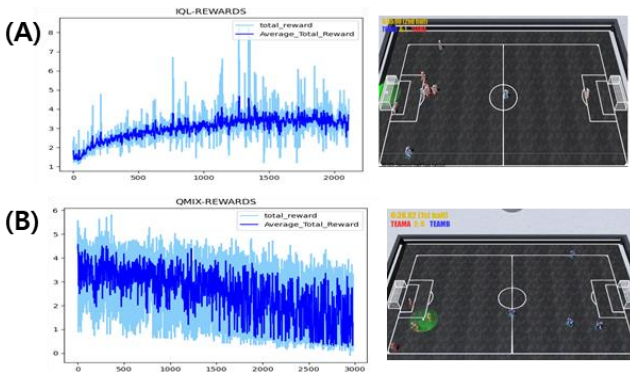
여기서는 총 9 가지 전략을 구현하였다 (표 1). 먼저 ‘Standard role-based players’는 공격수, 수비수, 골키퍼가 미리 정의된 영역에서 공수 역할을 수행하는 표준 전략이다. ‘Gegenangriff’는 골키퍼를 제외한 수비수 2명, 공격수 2명이 오직 공격만 시도하는 전략이다. ‘Gegenpressing’은 팀 내 수비수 2명이 상대팀 공격수를 방어하고, 팀 내 공격수 1명은 상대팀 수비수를 방어하며, 팀 내 나머지 공격수는 상대팀 골대에 돌진하는 전략이다. ‘All Becoming goalies #1’는 팀 내 5명이 상대팀 골대에 일렬로 서서 상대팀 골대를 원천적으로 차단하는 동시에 팀 골대는 완전히 비워 두는 전략이다. ‘All Becoming goalies #2’는 이전과는 반대 전략으로 팀 골대에 일렬로 서서 팀 골대를 원천적으로 차단하고, 다른 공격/수비 전술은 수행하지 않는 전략이다. 다음은 더 빠른 속력을 갖는 ‘Becoming sprinters’, 우리 팀 골대를 Penalty area 바로 앞 바깥 영역에서 방어하는 ‘Becoming defenders in the penalty area’, 상대팀을 따라다니는 ‘Becoming man-to-man defenders’, 무조건 공만 따라다니는 ‘Becoming ball-hunters’를 전략으로 구현하였다.



(그림 2) Dynamic planning 에이전트 구현 사례 및 스크린 샷. (A) Gegenangriff: 골키퍼를 제외한 모든 에이전트가 공격수인 전략. (B) Gegenpressing: 강력한 몸싸움기반 수비형 전략.

강화학습 에이전트 강화학습은 환경에 대한 상태 정보가 주어지고, 에이전트가 활용 가능한 행동이 주어졌을 때, 목표 완수에 따른 보상을 최대화하는 방향으로 행동을 조합하여 최적 전략을 도출하는 학습 방법을 말한다 [5]. 이는 불확실한 상태 변화에 따라 상대방 전략에 대응하는 법을 학습하고 일반적인 상황에 적용이 가능한 전략을 도출할 수 있는 에이전트 구현 접근 방법으로 볼 수 있으며, 딥러닝과 결합하여 인간의 능력을 뛰어넘는 성능을 보여주고 있다 [6, 7]. 최근에는 단일 에이전트 뿐 아니라, 멀티 에이전트 환경에서 다수의 팀 구성원들이 서로 협동하고 협업하여 목표를 성취하는 다양한 멀티 에이전트 강화학습 알고리즘이 제안되고 있다 [8, 9, 10]. 본 연구에서는 멀티 에이전트 강화학습 에이전트를 활용하여 다양한 상대 (예: Random walk, Dynamic planning 에이전트) 들과 대전하여 최적 대응 전략을 학습하고, 일반적이며 표준적인 축구 전략을 배워서 다양한 게임에서 활용 가능한 사례를 중점적으로 구현하고 시뮬레이션 하고자 하였다.

여기서는 Independent Q-Learning(이하 IQL)[11]과 QMIX [10], 두 가지 알고리즘을 사용하였다. 학습은 Random walk 기반 가상 축구 에이전트를 사용하였으며, 보상함수는 ‘Standard role-based players’ 를 기준으로 구현하였다. 각 에이전트가 미리 정의된 영역에서 움직이면서 역할에 맞게 공/수를 잘 수행했을 경우 높은 보상을 얻게 된다. 강화학습의 학습 결과는 그림 2 와 같다.



(그림 3) 강화학습 에이전트 학습 결과 및 전략 예시. (A) IQL 강화학습 에이전트 학습 결과(좌) 및 전략 수행 중 스크린 샷(우). (B) QMIX 멀티에이전트 강화학습 에이전트 학습 결과(좌) 및 전략 수행 중 스크린 샷(우). 학습은 Random walk 에이전트를 상대로 수행되었으며, 3000K 에피소드동안 진행함.

4. 실험 및 결과

3 장에서 제시한 구현 방법을 바탕으로 총 12 개 전략을 수행하는 가상 축구 에이전트 준비하고 대전 시뮬레이션을 수행하였다 (표 1). 실험은 리그 방식으로 진행, 모든 가상 축구 에이전트가 자신을 제외한 다른 에이전트들과 대전하도록 설정하였다. 1 차 시뮬레이션을

이용하여 승률이 높은 5 개 에이전트 (Random walk, Gegenangriff, Gegenpressing, IQL, QMIX)를 선정하고 최종 실험을 진행하였다. 총 200 번 대전 (= 5 팀 × 4 대전상대, 각 대전별 10 게임 수행) 을 통해, 각 에이전트들의 득점과 승률을 확인하고, 분석하였다.

<표 1> 가상 축구 에이전트 구현 리스트 (굵은 글씨는 대전 시뮬레이션 대상 전략을 의미함)

접근	전략 이름	특징
Random walk	Random walk (RW)	무작위 움직임
	Standard role players	미리 설정된 규칙대로 움직임
Dynamic planning	Gegenangriff (GGA)	골키퍼 제외, 모두 공격
	Gegenpressing (GGP)	수비수와 공격수 1 이 수비, 공격수 2 가 공격
	Becoming sprinters	더 빠른 속력을 가짐
	Becoming defenders in the penalty area	우리 팀 페널티 영역에서 수비
	Becoming man-to-man defenders	상대 에이전트를 따라다님
	Becoming ball-hunters	무조건 공만 쫓음
	All Becoming goalies #1	상대 팀 골대 앞에 서서 수비
	All Becoming goalies #2	우리 팀 골대 앞에 서서 수비
RL	IQL	개별 에이전트들의 행동을 Q-Network 로 결정. 협업 없음
	QMIX	공동의 행동을 목표로 다른 에이전트의 행동을 고려하여 학습

<표 2>는 시뮬레이션을 통해 얻은 Score 결과이다. 득점은 두 에이전트 간 대전에서 기록한 골 숫자를 말한다. A 는 성능 측정을 위한 에이전트, B 를 대전 상대 에이전트, N 을 총 대전 수 (N=10)이라고 했을 때, A 와 B 의 평균 득점 차이를 Score 로 정의하고 $\frac{\text{득점}_A - \text{득점}_B}{N}$ 으로 계산하였다.

<표 2> 가상 축구 에이전트의 평균 Score 결과

A\B	RW	GGA	GGP	IQL	QMIX	평균
RW		-4.9	-3.8	-6.1	-2.1	-4.225
GGA	+4.9		-3.2	+4.8	+6.6	3.275
GGP	+3.8	+3.2		+2.6	+1.4	2.75
IQL	+ 6.1	-4.8	-2.6		-0.8	-0.525
QMIX	+2.1	-6.6	-1.4	+0.8		-1.275

평균 Score 를 놓고 볼 때, GGA 가 보편적으로 가장 우수한 성능을 보였다. 극단적 몸싸움 전략 GGP 에만 열세를 보였을 뿐, 대다수의 전략을 압도하는 결과를 보였다. 수비수 2 명과 공격수 2 명인 다른 전략들과 다르게 공격수만 4 명이라는 특성을 갖기 때문에 score 를 많이 낼 수 있었을 것으로 예상된다. 다만, GGA 는 score 는 높을 지언정 승률면에서는 GGP 에 열세로 나타났다.

비록 평균 score 측면에서 2 위를 차지했으나, GGP 는 가장 높은 승률을 기록하였다. 가장 경쟁력 있는 승리

전략으로 볼 수 있다. 극단적 몸싸움을 바탕으로 상대 에이전트들을 몰아 세우고, 공격적 에이전트가 재빨리 골을 넣는 전략이 가장 효과적으로 승리하는 전략임을 내포하고 있다. 분산 측면에서도 GGP는 가장 작은 분산 값을 보이며, 모든 상대에 대해 균일한 경쟁력을 보였다. 축구가 score 보다 승률이 더 중요한 게임인 바, GGP가 대전한 전략 중 가장 우수한 전략으로 판단된다.

예상외로, 강화학습형 에이전트들의 Score는 하위권에 포진해 있다. 학습형, 적응형의 강점이 있을 것으로 예상하였으나, 보상함수 자체가 역할별 충실도를 바탕으로 구현되어 있고, 특별한 전략을 shaping 하는 것이 아니므로 경쟁력이 떨어지는 것으로 사료된다. 뿐만 아니라, 강화학습 에이전트들이 random walk와 학습한 상태인 바, 불확실성/의외성에 대해 대비는 되었을 지라도, 경쟁력있는 공격/수비 전술을 경험하지 못하여 GGA, GGP 및 상대 강화학습 에이전트에 대해 열세를 보인 것으로 예상된다.

흥미로운점은, 강화학습형 에이전트의 경쟁력이 그리 높지 않음에도 불구하고, 가장 강력한 GGP에 대해서는 가장 우수한 성능을 보였다. 특히 멀티에이전트 강화학습 기반 QMIX의 경우 모든 전략 중 가장 높은 score를 보였는데, 에이전트 간 협력 행동이 고려된 바, GGP대전 상대 중 GGP의 전략을 가장 효과적으로 방어하고 대처한 것으로 예상된다. QMIX의 전략 차별점은 추후 승률 관점에서 추가 분석하고, 대전 내 trajectory들을 분석하여 효과를 확인하고자 한다.

5. 결론 및 향후 연구

본 연구에서는 AI World Cup 환경에서 멀티 에이전트기반 학습/추론 기술을 사용하여 다양한 전략과 전술을 구사하는 가상 축구 에이전트 구현과 시뮬레이션 결과를 소개하였다. 해당 연구를 통해 Dynamic planning 에이전트들을 설계하고, 멀티 에이전트 강화학습 기술을 적용한 가상 축구 에이전트를 구현하였다. 구현된 에이전트들 간 대전을 통해 학습된 강화학습 기반 에이전트들과 Dynamic planning 에이전트들 간의 대결 결과를 분석함으로써 더 좋은 성능을 낼 수 있는 에이전트를 설계하는 알고리즘을 파악할 수 있었다.

학습을 통해 상황에 적합한 최적 전략을 수행하는 멀티 에이전트 강화학습 에이전트들이 기술적으로 가장 진보된 형태임에도 불구하고, 개별 에이전트들에 역할을 부여하고, 상대 팀과 공의 위치에 따라 전략과 행동을 직접 디자인하는 방식을 취한 Dynamic planning 에이전트들이 더 높고 균일한 성능을 발휘하는 것을 확인할 수 있었다.

향후에는 축구 에이전트를 더욱 다양하고 복잡한 전략의 대전 상대 에이전트들과 학습시켜 실제 축구

경기에 적용하는 연구를 해 볼 예정이다. 이를 통해 실제 경기의 결과 뿐 아니라, 상대 팀의 경기 전술을 예측하는 것 또한 가능할 것으로 예상된다. 멀티 에이전트 강화학습 측면에서는 보다 다양한 상대와 전략을 인식하고, 맞춤형 최적 전략을 구사할 수 있는 목적 기반 메타 강화학습을 적용하여 보다 높고 균일한 성능을 보일 수 있도록 연구를 진행할 예정이다.

Acknowledgement

이 성과는 정부(과학기술정보통신부)의 재원으로 한국 연구재단의 지원을 받아 수행된 연구임 (No. 2020R1G1A1102683). 본 연구는 삼성미래기술육성센터의 지원을 받아 수행하였음 (No. SRFC-TC1603-52). 본 결과물은 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 사회맞춤형 산학협력 선도대학(LINC+) 육성사업의 연구결과임.

참고문헌

- [1] Tuyls, Karl, et al. "Game Plan: What AI can do for Football, and What Football can do for AI." *Journal of Artificial Intelligence Research* 71 (2021): 41-88.
- [2] Yi, Sanghyung, et al. "A Competitive Path to Build Artificial Football Agents for AI Worldcup." *IEEE/IEIE International Conference on Consumer Electronics (ICCE) Asia*. IEEE, IEIE (2018).
- [3] http://aiworldcup.org/ai_soccer (retrieved 20210302)
- [4] Brooks, Rodney A. "Intelligence without representation." *Artificial intelligence* 47.1-3 (1991): 139-159.
- [5] Sutton, Richard S., and Andrew G. Barto. *Introduction to reinforcement learning*. Vol. 135. Cambridge: MIT press, (1998).
- [6] Mnih, Volodymyr, et al. "Playing atari with deep reinforcement learning." *arXiv preprint arXiv:1312.5602* (2013).
- [7] Silver, David, et al. "A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play." *Science* 362.6419 (2018): 1140-1144.
- [8] Wang, Tonghan, et al. "Rode: Learning roles to decompose multi-agent tasks." *arXiv preprint arXiv:2010.01523* (2020).
- [9] Wang, Jianhao, et al. "Qplex: Duplex dueling multi-agent q-learning." *arXiv preprint arXiv:2008.01062* (2020).
- [10] Rashid, Tabish, et al. "Qmix: Monotonic value function factorisation for deep multi-agent reinforcement learning." *International Conference on Machine Learning*. PMLR, (2018).
- [11] Tan, Ming. "Multi-agent reinforcement learning : Independent vs. cooperative agents." *International Conference on Machine Learning* (1993).