

인공지능 멀티플레이어 게임을 위한 성능평가 기술 동향

• 김승현 (한국전자통신연구원)

I. Introduction

인공지능 기술이 활발히 개발되고 있다. 실생활에서의 인공지능 도입을 위해서는 게임과 같이 제한된 환경에서 기술을 개발하고 성능을 검증하는 작업이 필요하다. 이 때문에 체스, 바둑, 포커 등 전통적인 게임에서 많은 연구가 이루어졌다. 온라인 게임에서도 OpenAI의 Atari[1]같은 단순한 게임뿐만 아니라, 스타크래프트나 DoTa 등 복잡한 실시간 전략 게임에서도 인공지능 에이전트를 활발히 연구 중이다.

인공지능 에이전트를 연구하기 위해서는 정확한 성능 평가가 필수적이다. 성능평가 시스템의 결과가 정확할수록 플레이어의 평가 결과와 실제 실력이 동기화되고, 성능 평가에 대한 신뢰도가 증가한다. 개발한 인공지능 에이전트의 성능 수준을 정확히 파악함으로써 올바른 성능 개선 및 객관적인 성능 비교가 가능해진다. 또한 플레이어의 수준과 성향에 맞는 다른 플레이어 또는 인공지능 에이전트와 함께 게임을 진행할 수 있기 때문에, 플레이어에게 흥미로운 경기 경험을 제공할 수 있다.

대부분의 게임은 플레이어의 성능을 판단하기 위해 게임의 목표가 충족되었는지 여부를 고려한다. 1인 플레이어의 경기는 플레이어가 획득한 점수로 쉽게 성능을 평가할 수 있지만, 두 명 이상의 플레이어(멀티플레이어)가 포함된 경기의 경우는 복잡하다. 주어진 환경, 우리편 플레이어의 수준, 상대방 플레이어의 수준 등 다양한 변수를 고려해야 한다. 멀티플레이어 게임에는 종종 경기 참가자의 실력을 순위로 매기는 방법이 사용된다. 순위가 높은 플레이어는 순위가 낮은 플레이어 이길 가능성이 더 높고, 순위차이가 커질수록 그 가능성이 더욱 커진다.

본 논문에서는 멀티플레이어를 위한 대표적인 성능평가 기술인 ELO, Glicko, TrueSkill을 살펴본다. 그리고 성능평가 기술에 대한 장단점을 논의한 후 결론을 제시한다.

II. 대표적인 성능평가 기술

1. ELO

ELO[2]는 미국의 물리학 교수이자 체스 플레이어인 Arpad Elo 박사가 만든 시스템으로, 경쟁 게임에서 가장 많이 사용된다. ELO 시스템은 각 경기에서 플레이어가 승리할지 여부를 0~1 사이의 기댓값으로 표현하고, 실제 경기 결과에 따라 플레이어의 레이팅을 보정한다.

플레이어의 승리 기댓값은 수식 (1)로 계산한다. R은 플레이어의 레이팅으로 처음에는 미리 지정된 값이 할당되며, 특정 범주를 가지는 정규 분포 또는 로지스틱 분포를 따른다. N은 한 플레이어가 다른 플레이어보다 10배 잘하는 경우의 기댓값 차이를 의미하며, 경기 종목에 따라 다르게 설정된다.

$$E_A = \frac{1}{1 + 10^{\frac{R_B - R_A}{n}}} \quad (1)$$

경기 결과에 따라 업데이트 되는 플레이어들의 레이팅은 (2)로 계산한다. S는 실제 경기 결과로, 1은 승리, 0.5는 무승부, 0은 패배이다. K는 경기 결과가 레이팅에 미치는 영향을 나타내는 지표로, 경기 특성에 따라 다르다. 가령 야구나 체스같이 많은 게임이 진행되는 종목은, 매 경기의 중요도가 떨어지기 때문에 K가 낮다. 하지만 미식축구같이 게임 횟수가 적은 종목은, 매 경기가 중요하기 때문에 K가 높다.

$$R'_A = R_A + K(S_A - E_A) \quad (2)$$

ELO 랭킹 시스템은 세계 축구 리그, 미국 체스 연맹 또는

- 1) 레이팅(rating): 성능을 표현하기 위한 수치로, 레이팅이 높을수록 성능이 뛰어나다

세계 체스 연맹 등 2인전 경기를 중심으로 다양한 리그에서 성공적으로 사용되고 있다. 온라인 게임에서도 워크래프트3, 스타크래프트 2, 리그오브레전드 등 실시간 멀티플레이어 전략게임에서 매치 밸런스에 활용된다. 인공지능 에이전트 개발의 경우, Google 답마인드가 알파고 개발의 성능지표로 ELO를 활용한 것이 대표적이다[3].

2. Glicko

Glicko[4] 시스템은 ELO 시스템을 확장한 것으로 Glickman가 개발했다. Glicko 시스템은 플레이어의 실력 분포가 가우시안 분포를 따르며, 플레이어의 실력 또한 일관적이지 않다고 가정한다. 단순히 플레이어의 레이팅만 계산하는 것이 아니라, 레이팅에 대한 신뢰도 또한 계산한다. 이 신뢰도를 레이팅 분산(RD, Rating Deviation)이라고 부른다.

Glicko의 레이팅 알고리즘은 RD 결정, 새 레이팅 결정, 새 RD 결정의 3단계로 구분된다. 1단계는 RD 결정이다. 새 플레이어는 350을 할당하고, 기존 플레이어는 경기를 마친 뒤, 경과한 t 시간과 상수 c를 고려하여 새로운 레이팅 편차를 수식 (3)으로 계산한다. T는 얼마나 자주 경기가 시작되는지에 따라 수 분 또는 몇 개월이 될 수도 있다. 등급 편차는 불확실성을 나타내기 때문에, 게임 결과가 누적될수록 줄어들고 비활동 시기에는 늘어난다.

$$RD = \min(\sqrt{RD_0^2 + c^2 t}, 350) \quad (3)$$

2단계는 새 레이팅 결정이다. 레이팅 기간에 발생한 m건의 게임 결과를 토대로, 수식 (4)에 의해 새 레이팅이 결정된다. r은 플레이어의 레이팅을 나타낸다. s는 각 게임의 결과를 나타내며, 승리는 1, 무승부는 0.5, 패배는 0이다.

$$r = r_0 + \frac{q}{RD^2 + \frac{1}{d^2}} \sum_{i=1}^m g(RD_i)(s_i - E(s|r_0, r_i, RD_i)) \quad (4)$$

여기서,

$$g(RD_i) = \frac{1}{\sqrt{1 + \frac{3q^2(RD_i^2)}{\pi^2}}}$$

$$E(s|r_0, r_i, RD_i) = \frac{1}{1 + 10^{\frac{(g(RD_i)(r_0 - r_i))}{400}}}$$

$$q = \frac{\ln(10)}{400} = 0.00575646275$$

$$d^2 = \frac{1}{q^2 \sum_{i=1}^m (g(RD_i))^2 E(s|r_0, r_i, RD_i)(1 - E(s|r_0, r_i, RD_i))}$$

3단계는 새 RD 결정이다. 1단계의 RD 계산은 경기가 없는 기간이 길어질수록 플레이어의 레이팅 수준이 더욱 불확실해지는 특성을 고려한 것이다. 레이팅 기간 이후에 레이팅 편차를 수식 (5)로 업데이트한다.

$$RD' = \sqrt{\left(\frac{1}{RD^2} + \frac{1}{d^2}\right)} \quad (5)$$

Glicko 랭킹 시스템은 포켓몬, Chess.com, online-go.com, 카운터 스트라이크, Dota 등의 온라인 경쟁 게임에 활용되었다.

3. TrueSkill

TrueSkill[5]은 Microsoft에서 개발한 레이팅 시스템으로 베이저안 추론을 사용한다. TrueSkill의 경우, 각 플레이어의 스킬과 실력은 가우시안 분포를 따르며, 각 팀의 실력은 팀 플레이어들의 실력의 합이라고 가정한다. 이 시스템은 거의 모든 경기 구성(예: 개인, 팀 대 팀 또는 free-for-all)에 사용할 수 있다. TrueSkill 시스템은 플레이어의 평균 실력, 플레이어의 기술에 대한 불확실성이라는 두 수치를 특징으로 한다. 불확실성이 작아질수록 플레이어의 기술이 평균 실력에 가깝다는 확신을 가진다.

TrueSkill 모델에서 경기 데이터는 일련의 경기 결과로 구성되며 각 경기 결과에는 관련된 플레이어, 팀 할당, 경기 시작 시간 및 길이, 각 플레이어가 플레이 한 시간 및 최종 점수가 나열된다. 이 데이터를 해석하기 위해 TrueSkill은 일치 결과의 확률적 생성 모델을 가정하고, 스킬과 경기 결과를 생성하는 직관적인 무작위 프로세스를 구성한다. 이 모델은 경기 데이터와 비교 확인한 뒤, 발견된 불일치를 기반으로 구체화된다. 데이터에 대한 적합성이 만족스러우면 수식 (6)의 베이저안 추론을 적용하여 스킬 레이팅 업데이트를 위한 최적의 알고리즘을 얻는다.

$$p(s|r, A) = \frac{p(r|s, A)p(s)}{p(r|A)} \quad (6)$$

TrueSkill은 헤일로 3부터 포르자 모터스포츠 7까지 다양한 Xbox 라이브 게임에서 플레이어를 순위와 매치하는데 사용되었다.

4. 각 방식의 장단점

ELO는 다음과 같은 특징을 가진다.

- 수식이 간단하고 직관적인 이해가능
- 매 경기마다 레이팅 갱신 가능
- 1:1 경기에 최적화
- 비활동으로 인한 실력저하 고려 못함

Glicko는 다음과 같은 특징을 가진다.

- RD를 적용하여 레이팅 정확도 개선
- 1:1 경기에 최적화
- 일정 경기마다 레이팅 갱신 가능
- 각 경기의 레이팅 영향 파악 어려움
- 경기 상대에 따른 레이팅 변경 문제

TrueSkill은 다음과 같은 특징을 가진다.

- 여러 타입의 경기에 적용 가능
- 매 경기마다 레이팅 갱신 가능
- 새로운 플레이어의 레이팅 파악 쉬움
- Microsoft의 독점기술
- 환경 설정 및 파라미터 튜닝 어려움

M., “Playing atari with deep reinforcement learning,” arXiv preprint arXiv:1312.5602., 2013.

- [2] Elo, Arpad E., “The rating of chessplayers, past and present,” Arco Pub., 1978.
- [3] Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... & Dieleman, S., “Mastering the game of Go with deep neural networks and tree search,” nature, 529(7587), pp. 484–489, 2016.
- [4] Glickman, M. E., & Jones, A. C., “Rating the chess rating system,” CHANCE-BERLIN THEN NEW YORK-, 12, pp. 21–28, 1999.
- [5] Minka, T., Graepel, T., & Herbrich, R., “TrueSkill(TM): a Bayesian skill rating system,” Adv. Neural Inf. Process. Syst. 20, pp. 569 - 576, 2007.

III. Conclusions

본 논문에서는 인공지능 에이전트 연구에 필수적인 성능 평가 기술과 관련하여, 대표적인 성능평가 기술인 Elo, Glicko, TrueSkill을 살펴보았다. 각 기술은 저마다의 장단점을 가지고 있다. 최고의 기술은 없으며, 각 기술의 특성을 고려해서 사용 목적에 따라 선택해야 한다. 또한 레이팅이 안정적으로 수렴하기 위한 시간이 필요하며, 도메인 지식에 기반한 튜닝이 요구된다. 향후에는 다양한 타입의 경기를 지원하며, 빠르고 정확하게 플레이어의 성능을 파악할 수 있는 성능 평가 기술의 개발이 필요하다.

REFERENCES

- [1] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller,

저 자 소개



Seung-Hyun Kim received the B.S. in computer engineering in 2002 from Kumoh National University of Technology, Rep. of Korea, and the M.S. in computer engineering in 2004 from POSTECH, Rep.

of Korea. He received his Ph.D. in computer science in 2017 from KAIST, Rep. of Korea. Since 2004, he has been a senior member of research staff with ETRI, Daejeon, Rep. of Korea. His main research areas are ID management, mobile security, privacy, and machine learning.