

심층 강화학습 기반 조세 정책 최적화 시뮬레이션 환경 분석 및 실험

허주성¹, 최요한¹, 석영준¹, 윤주상², 한연희^{1*}

한국기술교육대학교 컴퓨터공학과 미래융합공학전공

{chil1207, yowief, dsb04163, yhhan}@koreatech.ac.kr¹, jsyoun@deu.ac.kr²

Deep Reinforcement Learning-based Tax Policy Optimization Simulation Environment Analysis and Experiment

Joo-Seong Heo¹, Yo-Han Choi¹, Yeong-Jun Seok¹, JooSang Youn², Youn-Hee Han^{1*}

Future Convergence Engineering, Dept. of Computer Science and Engineering, Korea University of Technology and Education¹, Dong-eui University²

요약

AI는 실제로 많은 영역에서 꾸준히 발전되고 있지만, 특히, 경제 분야는 다양한 환경, 변수로 인해 AI를 통해 정책을 실험하고 평가할 수 있는 방법이 제한적이다. 본 논문에서는 Salesforce 팀의 AI 기반 경제 시뮬레이션 환경인 AI Economist를 활용하여 경제 활동 에이전트와 정부 에이전트 간의 2단계 학습을 통하여 조세 정책 최적화를 위한 심층 강화학습 기반 조세 정책 최적화 시뮬레이션 환경 분석 및 실험을 진행하였다.

I. 서론

전통적인 경제학에서 경제 정책은 정부와 다양한 경제 주체들 간의 상호 작용을 고려하여 많은 실험과 공식, 계산 등의 복잡한 절차를 통해 연구, 수립된다. 수많은 경제 활동, 경제 활동 주체들 간의 상호 작용 등을 고려해야 하기 때문에 강화학습을 통해 경제 정책을 수립한다는 것은 현실적으로 많은 제한사항이 있다. 선행 연구에서는 Salesforce 팀에서 개발한 AI Economist 환경을 사용하여 경제 활동 에이전트들을 학습시키고 에이전트의 정책을 최적화하기 위한 실험을 진행하였다 [1][2]. 또한, 실험을 통해 세금 정책이 없는 자유시장 환경에서 경제 활동 에이전트인 Resident들이 각각 다른 스킬 레벨에 따라 자원 수집, 거래, 건축 등의 서로 다른 행동을 통해 각자의 보상을 최대화하기 위한 최적 정책을 학습하는 것을 확인하였다 [3]. 본 논문에서는 선행 연구에 이어 2단계인 Planner 에이전트의 조세 정책 학습을 통해 생산성과 불평등의 균형 즉, 사회복지를 최대화하기 위한 조세 정책 최적화 실험을 진행하였다.

II. 본론

1. 선행 연구

앞서 진행했던 선행 연구에서는 AI Economist 시뮬레이션 환경에서 경제 활동 에이전트 정책 최적화를 위한 1단계 학습을 진행하였다. 시뮬레이션 환경은 25X25 크기의 2차원 그리드 공간의 Gather and Build 환경으로 Resident들은 이동하며 자원을 수집하고 거래 혹은 건축을 통해 보상을 얻게 된다[4]. 이때, Resident들은 스킬 레벨에 따라 차등된 보상을 받게 되며, 세금 정책이 없는 자유시장 환경에서 Resident들은 각자의 보상인 유틸리티를 최대화하기 위하여 다양한 행동 양식을 보였다. [그림 2]는 step별 학습 결과를 나타낸 것으로 스킬 레벨이 높은 Resident가 더 많

은 집을 짓고 상대적으로 스킬 레벨이 낮은 Resident는 건축이 아닌 거래를 통해 유틸리티를 높이는 쪽으로 행동하는 것을 알 수 있었다.



[그림 1] step별 학습 결과(왼쪽부터 step-0, step-300, step-999)

2. 조세 정책

2단계 학습은 정부의 조세 정책을 의미한다. Planner 에이전트는 각 Resident들에게 100 time step마다 소득 구간별 세율을 통해 세금을 걷고, 이를 균등하게 다시 재분배 할 수 있다. 이때, 세율이 높아지면 평등해지지만 생산성이 감소하고, 반대로 세율이 낮아지면 생산성은 올라가지만 불평등 격차가 심해지는 문제가 발생하므로 Planner는 수식 1과 같이 보상으로 사회복지 함수를 활용하여 서로 트레이드오프 관계인 생산성과 불평등의 균형을 맞추는 것을 목표로 조세 정책을 학습한다. 본 논문에서는 AI Economist, Saez, US Federal, Free Market 총 4가지 조세 정책에 대해 2단계 학습을 진행하고 비교 분석하였다. 첫 번째는 자유시장으로 세금 정책이 없어 1단계 학습과 차이가 없다. 두 번째는 US Federal 정책으로 2020년 미국의 소득 구간별 세율을 그대로 적용하여 실험하였다[5]. 세 번째는 Saez 세금 정책으로 경제 개념 중 Saez 공식을 통해 계산된 소득 구간별 세율을 적용하였다. 마지막 네 번째는 AI Economist로 세율 학습을 통해 소득 구간별 사회복지를 최대화하기 위해 학습된 모델로 실험을

* 교신저자: 한연희

진행하였다.

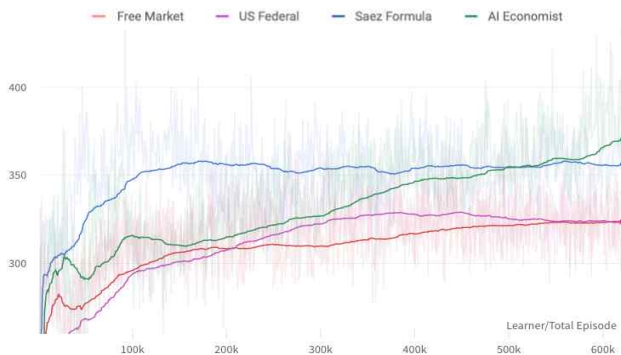
$$swf_t = Eq(coin_t) \cdot Prod(coin_t)$$

$$Eq(coin_t) = 1 - gini(x^c), \quad Prod(coin_{i,t}) = \sum_{i=1}^N coin_{i,t}$$

이때, 소득 구간은 [0.1, 0.12, 0.22, 0.24, 0.32, 0.35, 0.37]로 US Federal의 소득 구간을 Coin에 맞도록 1,000으로 스케일링한 것을 활용하였다.

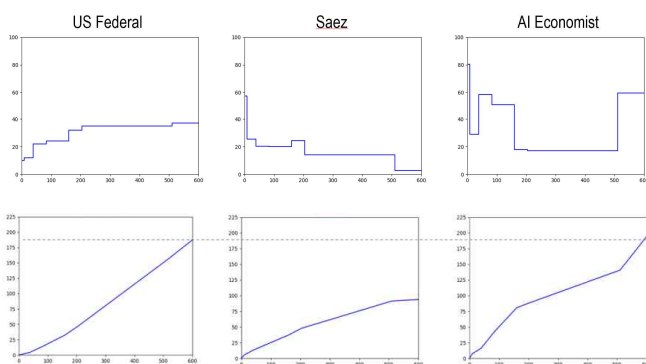
3. 실험 및 평가

[그림 2]는 정책별 2단계 학습 결과를 나타낸 그래프이다. 본 논문에서는 총 60만 번 에피소드를 학습했으며, 기존 다른 정책 대비 학습을 통한 AI Economist 조세 정책이 사회복지 지수가 가장 높은 결과를 보였다. 자유 시장의 경우 세금을 걷지 않기 때문에 생산성이 가장 높지만, 불평등이 가장 심해 사회복지 지수가 낮았고, US Federal은 소득이 높을수록 세금을 많이 걷어 생산성이 가장 낮았다. Saez는 AI Economist보다 종합적인 사회복지 지수는 낮았지만, 생산성이 상대적으로 높았다.



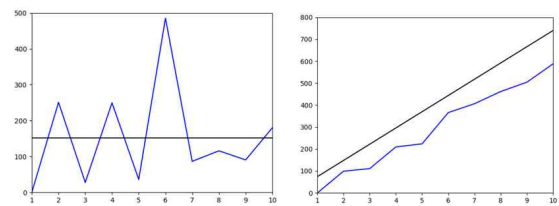
[그림 2] 정책별 2단계 학습 결과 (사회복지 지수)

[그림 3]은 정책별 소득 구간별 세율과 누적 세금을 비교한 그래프이다. 자유시장은 세금이 없어 제외하였으며, US Federal과 Saez는 각 정책에 대한 세율을 그대로 적용하였다. 소득이 올라갈수록 세율이 올라가는 US Federal이나 반대로 세율이 낮아지는 Saez와는 달리 학습된 AI Economist는 소득 구간별로 지그재그 형태로 세율이 높아졌다 낮아지는 경향을 확인하였다. 특히, 가장 소득이 낮은 구간과 가장 소득이 높은 구간에서 세율이 높고 다시 세율이 낮아지는 형태로 학습된 것을 알 수 있었다. 반면, 정책별 누적 세금 비교 결과를 통해 지그재그 형태의 AI Economist 정책이 다른 정책과 비교하여 누적 세금이 적지 않다는 것을 알 수 있었다.



[그림 3] 정책별 세율, 누적 세금 비교 결과

[그림 4]는 AI Economist 정책에서 Resident의 소득과 납부해야 할 세금을 비교한 그래프이다. 파란색은 실제 Resident가 행동한 결과이며, 검정(1)색은 평균 값을 나타낸다. 학습된 정책은 [그림 3]과 같이 소득 구간에 따라 세율이 높아졌다 낮아지기를 반복하는 것을 알 수 있는데, 이에 Resident는 세금이 걷히는 100 time step을 주기로 소득을 열심히 얻다가 얻지 않는 형태의 행동 양식을 보였다. 이러한 행동은 Resident의 전체 수입을 기준으로 평균 값과 납부할 세금을 계산해보면 확실한 차이가 나는 것을 알 수 있다. Resident가 세금을 최대한 적게 납부하기 위해 AI Economist 조세 정책에서의 세율이 낮은 소득 구간에 해당하기 위해 또한, 본인의 유틸리티를 최대화하기 위하여 소득을 조절한다는 것을 유추할 수 있다.



[그림 4] Resident의 소득과 세금 비교 그래프

III. 결론

본 논문에서는 전통적인 경제 개념과 공식을 적용하여 다양한 경제 주체들 간의 상호작용 속에서 조세 정책을 실험 및 평가하기 위해 AI Economist를 활용하여 심층 강화학습 기반 조세 정책 최적화 시뮬레이션 환경을 분석 및 실험하였다. 경제 활동 에이전트인 Resident와 조세 정책을 통해 사회복지지를 실현하기 위한 Planner 에이전트를 2단계로 나누어 학습하였으며, 실험 결과 각 Resident의 스킬 레벨과 정부의 세금 정책에 따라 각 에이전트들이 서로 다른 행동 양식을 보이고, 실제 현실의 경제 환경을 일부 반영할 수 있다는 것을 확인하였다.

또한, 향후 소득 구간, 보상, 경제 활동 등 각종 변수들과 강화학습 알고리즘을 개선하여 다양한 경제 정책을 적용할 수 있도록 개선할 계획이다.

참 고 문 헌

- [1] <https://www.salesforceairesearch.com/projects/the-ai-economist>
- [2] <https://github.com/salesforce/ai-economist>
- [3] Joo-Seong Heo, Yo-Han Choi, Yeong-Jun Seok and Youn-Hee Han, "Experiment and Analysis of Policy Optimization for AI-based Economic Agents" KICS Fall Conference 2022
- [4] S. Zheng et al., "The AI Economist: Improving Equality and Productivity with AI-Driven Tax Policies." arXiv, 2020.
- [5] S. Zheng, A. Trott, S. Srinivasa, D. C. Parkes, and R. Socher, "The AI Economist: Optimal Economic Policy Design via Two-level Deep Reinforcement Learning." arXiv, 2021.