

AI기반 경제 활동 에이전트 정책 최적화 실험 및 분석

허주성, 최요한, 석영준, 한연희*

한국기술교육대학교

{chil1207, yowief, dsb04163, yhhan}@koreatech.ac.kr

Experiment and Analysis of Policy Optimization for AI-based Economic Agents

Joo-Seong Heo, Yo-Han Choi, Yeong-Jun Seok, Youn-Hee Han*

KOREATECH

요약

4차 산업 혁명으로 AI가 사회 전반에 걸쳐 상용화되고 꾸준히 발전하고 있지만, 경제 분야는 여전히 데이터 부족, 다양한 환경, 변수 등으로 AI 적용이 어렵다. 실제 사회 경제적 문제를 해결하기 위해선 경제 주체 간 다양한 환경과 상호작용 요인들을 확인하며 경제 활동 및 정책 수립 과정을 설계하고 테스트해야 하지만, 경제 관련 데이터가 부족하고, 실제 정책을 실험하기 위한 환경 구성이 어렵다. 본 논문에서는 Salesforce 팀의 AI 기반 경제 시뮬레이션 환경인 AI Economist를 활용하여 AI 기반 경제 활동 에이전트 정책 최적화 실험 및 분석을 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. 예를 들어, 경제 성장을 위한 조세 정책 수립을 한다면, 기본적으로 국가 재정 운용 계획에 따른 재정 운용 전략뿐만 아니라 세목별 조세 정책 방향, 소득 구간별 조세 부담 형평성 등 심지어 경제, 복지, 인구 구조, 통일, 환경까지 고려해야 할 사항이 많다. 실제로도 전례 없는 코로나 19 상황으로 인해 우리나라뿐만 아니라 세계 경제가 급변하고 있는 만큼 효율적인 경제 정책을 수립하는 것은 어려운 일이다.

본 논문에서는 이러한 다양한 경제 활동, 정책 수립에 따른 상호 요인 분석 등을 실험하기 위해 Salesforce 팀에서 개발한 AI Economist 환경을 사용하여 경제 활동 에이전트들을 학습하고 최적화하기 위한 실험을 진행하였다.

II. 본론

1. AI Economist

Salesforce는 고객 관계 관리를 위한 플랫폼 서비스를 제공하는 IT 기업으로, 최근 Salesforce Research 팀은 주요 글로벌 이슈 중 하나인 경제적 불평등 개선을 목적으로 AI 기반 경제 정책 설계에 도움을 줄 수 있는 프레임워크를 개발하였다[1][2].

AI Economist는 경제 활동 에이전트와 정부 에이전트를 통해 강화 학습 알고리즘을 적용하여 학습한 조세 정책이 실제 경제적 불평등을 개선할 수 있는지 평가하기 위한 시뮬레이션 환경을 제공하며, 실제 AI Economist가 학습한 조세 정책은 기존의 잘 알려진 세금 정책과 비교했을 때 개선된 성능을 보여준다[3].

일반적으로 정부의 세금 정책은 부의 재분배를 통해 불평등 해소에 중요한 역할을 한다. 하지만, 불평등 해소를 위해 세율을 올리게 되면

생산성이 감소하고 반대로 세율이 낮아지면 생산성은 올라가지만 불평등 격차는 심해지게 된다. AI Economist는 경제 모델의 가정이 아닌 관찰된 데이터를 활용하여 주어진 환경에서 불평등과 생산성의 균형을 위한 최적의 세금 정책을 시뮬레이션 할 수 있다[4].

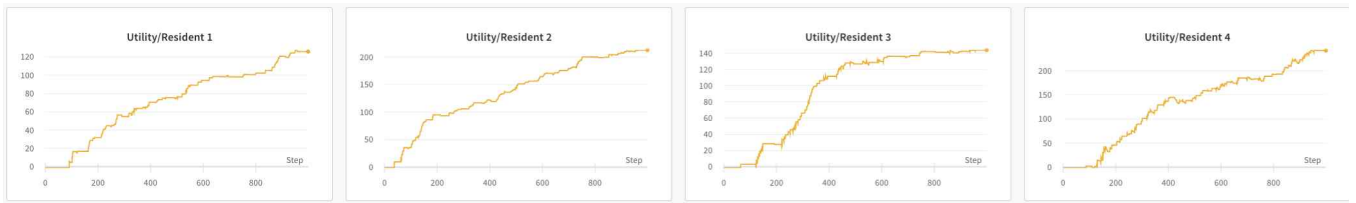
2. 시뮬레이션 환경

AI Economist는 경제 시뮬레이션을 위한 25X25 크기의 2차원의 그리드 공간의 Gather and Build 환경을 제공한다. 환경에서 Resident(경제 활동 에이전트)들은 이동하며 자원인 돌과 나무를 수집하고, 자원을 활용하여 집을 지어 코인을 얻거나, 다른 에이전트 간의 거래를 통해 코인으로 교환을 하는 다양한 경제 활동을 할 수 있다. 또한, Planner(정부 에이전트)는 생산성 향상과 불평등 해소의 균형을 맞추기 위해 소득 구간별 세율을 조정할 수 있다. 초기 Resident들은 서로 다른 스킬 레벨을 가진 채 4구역으로 분리된 공간에 위치한다. 맵 임의의 공간에 생성된 돌과 나무를 수집 혹은 거래를 통해 모아 집을 지을 수 있으며, 돌과 나무 하나씩 소모하여 집 한 채를 짓고 보상으로 코인을 받는다. 이 때, 스킬 레벨에 따라 자원의 수집량과 집을 짓고 받는 코인은 다르게 지급된다.

3. 행동 및 관찰 정보

Resident의 행동은 상하좌우 이동을 위한 행동 4개와 집 짓기 1개, 거래를 위한 행동 4개 마지막으로 아무 행동도 하지 않는 것까지 총 50개로 스택별 하나의 행동을 선택할 수 있다. 자원의 수집은 돌과 나무 위로 Resident가 위치하면 자동으로 수집되며, [표 1]은 Resident의 행동 분류를 나타낸다. Resident의 관찰 정보는 총 4가지를 포함한다. 첫 번째, 맵은 전체 맵에서 각 Resident를 중심으로 11X11 크기만큼의 주변 정보를 관리하며, 수집 가능한 돌과 나무, 집의 개수, 맵 정보 등이 포함된다.

* 교신저자: 한연희



[그림 1] Resident별 유틸리티 비교 그래프

두 번째, 거래소는 자원별 거래 시세와 내역, 거래된 매수, 매도 호가를 관리한다. 세 번째, Resident가 보유하고 있는 코인과 자원의 양을 관리하는 자산정보와 네 번째, 수집, 집 짓기에 사용될 스킬 레벨 정보를 관리하고 있다.

행동		개수
이동		4
집 짓기		1
거래	Buy / Sell	2
	Wood / Stone	2
	Coin(0 ~ 10)	11
	거래 행동합계	44(=2*2*11)
건너뛰기		1
Resident 행동합계		50

[표 1] Resident의 행동 분류

반면, Planner는 7X22개의 행동이 있다. 사회복지자를 실현하기 위한 조세 정책을 위한 행동으로 총 7개의 소득 구간과 각 구간별 세율을 위한 22개의 행동으로 나뉜다. (0.05 단위로 0부터 100을 나누기 위한 21개와 건너뛰기 1개) Planner의 관찰 정보는 Resident와 달리 부분이 아닌 전체 맵에 대한 정보를 관리한다. 전체 맵에서 Resident별 스킬 레벨 정보를 제외한 Resident의 위치, 집의 개수, 거래 정보 등을 포함한다.

4. 보상

Resident의 보상은 한계효용을 적용한 유틸리티 함수를 사용한다. Resident의 1차 목표는 자원 수집, 거래, 집 짓기 등을 통한 자산의 최대화지만, 자산이 늘어갈수록 한계 효용 체감 법칙에 따라 만족감은 감소한다는 경제학 개념을 적용하였다. 따라서, Resident마다 각각 스킬 레벨이 다르고, 레벨에 따라 얻는 자원, 코인의 양 또한 다르기 때문에 노동 대비 얻는 자산의 양을 아래 수식 1과 같이 정의하여 단순히 자산을 최대화 하는 것이 아닌 한계 효용이 적용된 노동대비 자산의 최대화를 목표로 한다.

$$u_i(x_{i,t}, l_{i,t}) = \text{crra}(x_{i,t}^c) - l_{i,t}, \quad \text{crra}(z) = \frac{z^{1-\eta} - 1}{1-\eta}, \quad \eta > 0 \quad (1)$$

위 식에서 i 는 Resident, t 는 타임스텝, x^c 는 자산, l 은 노동을 의미하며, crra는 Constant Relative Risk Aversion로 경제학에서의 효용함수를 나타낸다.

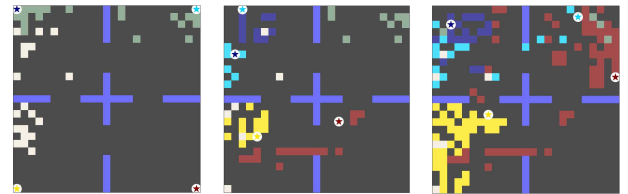
Planner의 보상은 사회복지 함수를 사용한다. 사회복지 지수는 다양한 것이 있지만 본 논문에서는 서로 상관관계에 있는 생산성과 불평등 지수를 활용하며, 사회 복지 함수를 아래 수식 2와 같이 적용하였다.

$$\text{swf}_t = \text{Eq}(\text{coin}_t) \cdot \text{Prod}(\text{coin}_t) \quad (2)$$

$$\text{Eq}(\text{coin}_t) = 1 - \text{gini}(x^c), \quad \text{Prod}(\text{coin}_{i,t}) = \sum_{i=1}^N \text{coin}_{i,t}$$

위 식에서 Eq는 평등지수로 자산에 대한 지니 계수를 사용하며, Prod는 생산성으로 모든 Resident의 코인의 양을 사용한다.

5. 실험 및 평가



[그림 2] step별 학습 결과(왼쪽부터 step=0, step=300, step=999)

본 논문에서는 세금이 없는 자유시장 환경에서 Resident에 대한 정책 최적화 실험을 진행하였다. 세금이 없기 때문에 Planner는 학습 하지 않았다. [그림 1]은 Resident별 유틸리티 비교 그래프로 스킬 레벨에 따라 유틸리티도 차이가 나는 것을 확인 할 수 있다. [그림 2]는 step별 학습 결과를 시각화한 것으로 초록색은 나무, 하얀색은 돌, 파랑, 하늘, 노랑, 빨강 별은 Resident와 각 Resident 색의 사각형은 집을 의미한다. 스킬 레벨이 높을수록 더 많은 집을 짓고 상대적으로 낮은 Resident는 집을 짓는 것보다 자원 수집과 거래를 통해 유틸리티를 높이는 쪽으로 학습되는 것을 확인하였다.

III. 결론

본 논문에서는 Salesforce에서 제공하는 AI Economist 환경에서 AI기반 경제 활동 에이전트 정책 최적화 실험을 진행하였다. 세금 정책이 없는 자유시장 환경에서 Resident들은 각각 다른 스킬 레벨에 따라 자원 수집, 거래, 건축 등 행동을 통해 유틸리티를 최대화하기 위한 최적 정책을 학습하는 것을 확인 할 수 있었다.

또한, 향후 2단계인 Planner 에이전트의 세금 정책 학습을 통해 생산성과 불평등의 균형 즉, 사회 복지를 최대화하기 위한 실험을 진행할 계획이다.

참 고 문 헌

- [1] <https://www.salesforceairesearch.com/projects/the-ai-economist>
- [2] <https://github.com/salesforce/ai-economist>
- [3] S. Zheng et al., "The AI Economist: Improving Equality and Productivity with AI-Driven Tax Policies." arXiv, 2020.
- [4] 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.