

월드모델을 통한 ARC의 핵심 지식 추출

이지환¹ · 이승필² · 김세진¹ · 김선동^{1*}

광주과학기술원 AI대학원¹, 광주과학기술원 전기전자컴퓨터공학부²

{jihwan.lee, seungpil}@gm.gist.ac.kr, {sejinkim, sundong}@gist.ac.kr

Extracting the core knowledge of ARC with the World Model

Jihwan Lee¹ · Seungpil Lee² · Sejin Kim¹ · Sundong Kim^{1*}

GIST AI¹, GIST EECS²

요 약

현재 인공지능 분야는 대량의 데이터를 학습하는 방식을 통해 다양한 상황에 대한 학습을 수행하지만, 이는 학습하지 않은 상황에서의 대처가 어렵다는 근본적인 한계가 존재한다. 이러한 한계를 극복하기 위해 제안된 인공 일반 지능(AGI)은 인간과 같이 다양한 업무를 수행할 수 있는 인공지능을 의미한다. 인공 일반 지능을 달성하기 위해서는 경험을 직접적으로 학습하는 방식보다는, 마치 인간과 같이 경험을 통해 얻을 수 있는 핵심 지식(core knowledge)을 간접적으로 학습하는 방식이 필요하다. 본 연구에서는 월드 모델이 ARC 데이터셋으로부터 핵심 지식을 추출해낸 것을 확인하는 실험을 설계하고, 이렇게 추출한 핵심 지식을 활용하기 위한 향후 연구 방향을 제안한다.

1. 서 론

최근 인공지능의 큰 발전은 대량의 데이터로부터 학습을 하는 데이터 기반 방식과 함께 이루어졌는데, 트랜스포머 아키텍처를 기반으로 하는 대규모 언어 모델이 대표적이다. 그러나 이러한 학습 방식은 사람이 환경과의 상호 작용 혹은 관찰을 통해 간접 학습하는 방식과는 상당한 거리가 있다. 사람은 경험으로부터 세계의 물리적 혹은 사회적 법칙 등을 학습할 수 있고[1], 이러한 법칙들의 집합을 핵심 지식 이라고 정의할 수 있다. 사람은 습득한 핵심 지식을 기반으로 미래를 예측하며, 이를 통해 새로운 상황에서도 적은 수의 경험만으로도 효과적으로 문제를 해결할 수 있다.

핵심 지식을 어떻게 구축할 것인가에 대한 연구[2, 3, 4, 5, 6, 7, 8]들은 이미 활발하게 진행되어 왔다. 하지만, 기존의 연구들은 대량의 데이터를 저장해두는 방식이라는 점에서 사람이 핵심 지식을 습득하는 방식과는 거리가 있다. 이러한 기존 연구들의 한계를 극복하기 위해서, 인간이 핵심 지식을 학습하는 방식을 모방한 월드 모델(world model)[9]이 제안됐다. 월드 모델은 연속으로 주어지는 입력으로부터 환경(environment)의 규칙을 위한 핵심 지식을 학습하고, 이를 순환 신경망의 은닉 상태(hidden state)에 저장한다. 이러한 방식으로 학습된 월드 모델은 행동(action)이 주어질 때 환경과 상호 작용하지 않아도 다음 상태(next state), 다시 말해서 미래를 예측할 수 있다. 그리고 이러한 학습 방식은 월드 모델이 해당 환경의 핵심 지식을 학습했다고 말할 수 있다.

그러나 월드 모델이 학습하는 핵심 지식은 특정 도메인(domain)에 한정된 규칙이기 때문에, 다른 도메인에 대해서 추가적인 학습을 하는 과정에서 이전 도메인에 학습한 정보를 파괴적 망각(catastrophic forgetting)을 하게 될 가능성이 존재한다[9]. 파괴적 망각은 인공지능 모델이 새로운

작업을 학습하는 과정에서 이전에 학습된 내용을 잊게 되는 현상으로, 다양한 도메인에 대해서 학습할 수 있는 인공지능을 위해서는 필수적으로 해결해야 하는 문제이다. 이를 극복하기 위해 연속 학습(continual learning)[10]과 같은 방식이 제안되었으며, 월드 모델에도 향후에 이러한 방식 등을 통해서 파괴적 망각에 대한 문제를 해결해야 한다.

본 연구는 월드 모델을 이용한 핵심 지식 구축 및 평가를 위해 ARC 데이터셋[11]을 사용할 것을 제안한다. ARC 데이터셋은 Francois Chollet가 만든 데이터셋으로, 서로 다른 규칙을 가진 400개의 훈련(training) 문제와 400개의 평가(evaluation) 문제로 이루어져 있다. 모든 문제는 규칙을 추론하기 위해 2~5개의 예제 입력-출력 격자쌍이 주어지며, 이를 통해서 학습한 규칙을 적용해서 1~2개의 입력 격자의 알맞은 출력 격자를 예측해야 한다. 모든 문제의 규칙들은 국가, 인종, 성별 등에 관계 없이 모든 인류에게 공통적으로 존재하는 핵심 지식을 기반으로 만들어졌다. 따라서 ARC 데이터셋이 내포하고 있는 핵심 지식을 월드 모델을 이용해 학습한 후, 이를 지식 데이터베이스로 구축함으로써 인공 일반 지능 발전에 기여하고자 한다.

2. 관련 연구

2.1 핵심 지식

핵심 지식은 사람이 살아가는데 있어서 상황마다 필요한 지식을 의미하며, 이미지, 언어, 수학, 물리, 사회와 같은 다양한 분야[1]에 대해 존재할 수 있다. 사람은 모든 기억을 저장할 수 없기 때문에 경험들로부터 이러한 핵심 지식을 추출해서 기억의 형태[9]로 가지고 있다. 주어진 대량의 데이터로부터 지식 데이터베이스를 구축해서 이를 활용하는 기존 연구로는 Cyc[2], ATOMIC[3], WebChild[4], ConceptNet[5], WordNet[6], FrameNet[7], Visual Genome[8] 등 다양하게 존재한다. 하지만 기존 연구들은 지식 데이터베이스를 구축하는 과정에서 핵심 지식을 추출하는 것이 아니라, 이를 활용하는 과정에서 축적된 대용량의 지식 데이터베이스로부터 핵심 지식을

¹ 이 논문은 과학기술정보통신부의 재원으로 한국연구재단과 정보통신기획평가원의 지원을 받아 수행된 연구임 (RS-2023-00240062, RS-2023-00216011, 2019-0-01842)

추출하는 방식을 취하고 있다. 이러한 방식은 인간의 지식 습득 및 활용 방식과는 거리가 멀다는 한계가 존재한다.

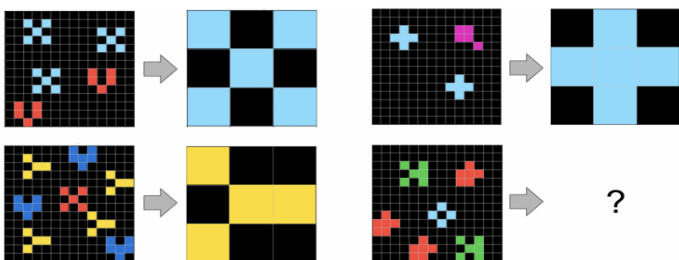
2.2 월드 모델

월드 모델[9]은 인지심리학을 기반으로 한 모델 기반 강화학습(model-based RL) 방법론으로, 이미지를 포함하는 상태, 행동, 보상 등으로 구성되는 오프라인 데이터로부터 환경에 대한 정보를 학습한다. 월드 모델은 비전 모델, 메모리 모델, 컨트롤러 모델로 이루어져 있다. 우선, 비전 모델에서 입력으로 주어지는 이미지를 입력받아 이를 압축한 잠재 벡터(latent vector)를 생성한다. 다음으로, 메모리 모델에서 비전 모델이 생성한 잠재 벡터와 현재 행동, 은닉 상태를 통해 다음 잠재 벡터를 예측하는 확률 모델을 학습해 다음 은닉 상태를 생성한다. 마지막으로, 컨트롤러 모델에서 비전 모델이 생성한 잠재 벡터와 메모리 모델이 생성한 은닉 상태를 통해 다음 행동을 예측한다.

월드 모델의 후속 연구로는 PlaNet[12], Dreamer[13], Dreamer-v3[14] 등이 있다. 특히 Dreamer-v3는 월드 모델을 학습한 후 액터-크리틱 방식을[15]을 통해 행동 정책(policy)를 학습해, 고정된 하이퍼파라미터(hyperparameter)로도 Atari, DMLab, Mine-RL 등 다양한 환경에서 좋은 성능을 보였다. 하지만 Dreamer-v3 연구는 각 작업 별로 별도의 모델을 생성했다는 한계가 존재한다. 따라서 저자는 월드 모델이 핵심 지식을 이전할 수 있는 모델이기 때문에, 충분히 큰 모델을 생성하면 다양한 도메인에서 다양한 작업을 하는 모델을 생성할 수 있다고 언급했다.

2.3 ARC

ARC 데이터셋은 인공지능 모델의 능력 습득 효율을 측정하기 위해 제안된 벤치마크 데이터셋[11]이다. 모델이 효율적으로 능력을 습득하는 것을 측정하기 위해, ARC의 문제들은 아래 [그림 1]과 같이 2~5개의 예제 입력-출력 격자쌍과 1~2개의 입력 격자라는 한정된 양의 데이터만 주어진다. 모델은 이렇게 매우 적은 양의 데이터만으로 규칙을 추론해야 하며, 추론한 규칙을 적용해서 주어진 입력 격자들에 대한 출력 격자를 예측해야 한다.



[그림 1] ARC 데이터셋의 문제의 예시. 이 문제는 3개의 예시 입력-출력 격자쌍과 1개의 입력 격자로 이루어져 있으며, 3개의 예시 격자쌍들로부터 학습한 규칙을 1개의 입력 격자에 적용해 적절한 출력 격자를 예상해야 한다.

ARC 데이터셋은 모든 문제가 다른 규칙으로 구성되어 있으며, 이 규칙들은 객체 개념, 목적 의식, 기하학, 위상수학과 같은 개념들을 기반으로 생성되어 있다. 인간은 비교적 적은 수의 예제로부터 규칙을 찾아내 문제를 푸는 과정을 수월하게 처리 가능하여 평균적으로 약 80%의 정확도를 보이는 것에 비해[11],

인공지능 모델들은 현재까지 최대 30%의 정확도를 달성하는데 불과했다. 이마저도, 이동, 회전, 대칭과 같은 기본적인 규칙들을 도메인 특화 언어(DSL)의 함수 형태로 직접 정의한 하드코딩 접근법이 현재까지 가장 높은 성능을 보여주고 있다[16]. 최근 인공지능 분야에서 각광받고 있는 ChatGPT를 활용한 모델들 역시 이 성능을 넘지 못하는 한계를 보였다. 이는 트랜스포머 기반 모델의 추론 능력이 부족하다는 최근 연구 결과[17]에 따른 것으로 해석된다.

3. 월드 모델을 이용한 ARC 데이터셋의 핵심 지식 구축

3.1 ARC 데이터셋의 적합성

앞에서 언급한 내용들과 같이, 현재의 인공지능 모델이 갖고 있는 대표적인 한계는 학습되지 않은 상황에서 모델이 원래의 성능을 보여주지 못하는 것이다. 이러한 문제를 해결하기 위해서 메타 강화 학습(meta-RL)[18, 19]과 같은 연구가 이루어지고 있으나, 메타 강화 학습 또한 하나의 도메인에서 다양한 작업을 수행하는 멀티 태스크((multi task) 상황에서는 큰 진전을 보이는 반면, 도메인 자체가 다양한 멀티 도메인(multi domain) 상황에서는 여전히 그 한계를 보이고 있다. 이러한 문제를 해결하기 위해서는 다수의 도메인에서 공통적으로 적용될 수 있는 핵심 지식을 구축할 필요가 있다.

ARC 데이터셋은 이러한 멀티 도메인 상황을 훈련하고 테스트하기 적절한 데이터셋이다. ARC 데이터셋의 문제들을 적용되는 규칙들은 사람이 기본적으로 가지고 있는 핵심 지식으로 구성되어 있다. 그리고 이러한 핵심 지식들은 다수의 ARC 문제에서 공통적으로 적용될 수 있으며, 더 넓게 보아서 다른 분야에서도 충분히 활용할 수 있는 개념들이다. 따라서 월드 모델을 통해 이러한 핵심 지식을 추출하는 것이 중요하다.

3.2 월드 모델을 통한 핵심 지식 구축 확인

ARC 데이터셋에 존재하는 400개의 훈련 문제를 학습한 하나의 거대 월드 모델이 있다면, 이 모델의 은닉 상태에는 그동안 풀었던 문제들에 대한 개념들이 학습되어 있을 것이다. 이러한 상황에서 새로운 문제가 주어졌을 때 이 문제는 기존 문제들 중 일부와 공통된 핵심 지식이 필요할 가능성이 높다. 따라서 400개의 훈련 문제를 학습한 거대 월드 모델이 아무 문제도 학습하지 않은 새로운 월드 모델 보다 새로운 문제들을 더 빠르게 해결하는 것을 실험을 통해 보일 수 있다면, 거대 월드 모델이 400개의 훈련 문제를 학습하는 과정을 통해 핵심 지식을 구축하는데 성공했다고 말할 수 있다. 또한, 이렇게 구축한 핵심 지식이 새로운 문제를 해결하는데 도움이 되었다고 추론하는 것 역시 타당하다.

하지만 이와 같이 학습된 거대 월드 모델이 ARC 데이터셋 이외에 다른 문제를 해결할 수 있는지에 대해서는 아직 기대하기는 어렵다. 이는 ARC 데이터셋의 문제 해결이라는 하나의 공통적인 환경에서 수행된 과정이기 때문이다. 만약 위에서 설명한 거대 월드 모델에서 ARC 데이터셋 자체에 대한 핵심 지식을 제거해낼 수 있다면, 다른 환경에서도 모델 훈련에 도움을 줄 수 있는 핵심 지식을 추출해낼 수 있을 것이다. 또는 학습 과정에서 ARC 데이터셋 이외에 다양한 데이터셋에서 훈련이 가능한 광범위한 환경을 구축한다면, 보다 광범위한 핵심 지식을 추출할 수 있을 것으로 예상된다.

4. 결론

본 논문에서는 인공지능 기술이 인공 일반 지능수준에 도달하기 위해서 핵심 지식 추출은 반드시 해결해야 할 문제임을 밝혔다. 이를 위해 인간의 학습 하는 방식을 모방한 월드모델에 대한 소개와, 월드 모델이 가지는 핵심 지식 추출 능력에 대해 알아보았다. 하지만 월드 모델을 통해 다양한 환경에 적용할 수 있는 핵심 지식을 얻기 위해서는 추가적인 연구가 필요할 것으로 예상된다.

참고문헌

- [1] Anna Dawid, Yann LeCun. "Introduction to Latent Variable Energy-Based Models: A Path Towards Autonomous Machine Intelligence." *arXiv preprint arXiv:2306.02572v1*(2023)
- [2] Lenat, Douglas B. "CYC: A large-scale investment in knowledge infrastructure." *Communications of the ACM* 38.11 (1995): 33-38.
- [3] Sap, Maarten, et al. "Atomic: An atlas of machine commonsense for if-then reasoning." *Proceedings of the AAAI conference on artificial intelligence*. Vol. 33. No. 01. 2019.
- [4] Tandon, Niket, Gerard De Melo, and Gerhard Weikum. "Webchild 2.0: Fine-grained commonsense knowledge distillation." *Proceedings of ACL 2017, System Demonstrations*. 2017.
- [5] Speer, Robyn, Joshua Chin, and Catherine Havasi. "Conceptnet 5.5: An open multilingual graph of general knowledge." *Proceedings of the AAAI conference on artificial intelligence*. Vol. 31. No. 1. 2017.
- [6] Fellbaum, Christiane, ed. *WordNet: An electronic lexical database*. MIT press, 1998.
- [7] Baker, Collin F., Charles J. Fillmore, and John B. Lowe. "The berkeley framenet project." *COLING 1998 Volume 1: The 17th International Conference on Computational Linguistics*. 1998.
- [8] Krishna, Ranjay, et al. "Visual genome: Connecting language and vision using crowdsourced dense image annotations." *International journal of computer vision* 123 (2017): 32-73.
- [9] Ha, David, and Jürgen Schmidhuber. "World models." *arXiv preprint arXiv:1803.10122* (2018).
- [10] Chen, Zhiyuan, and Bing Liu. *Lifelong machine learning*. Vol. 1. San Rafael: Morgan & Claypool Publishers, 2018.
- [11] Chollet, François. "On the measure of intelligence." *arXiv preprint arXiv:1911.01547* (2019).
- [12] Hafner, Danijar, et al. "Learning latent dynamics for planning from pixels." *International conference on machine learning*. PMLR, 2019.
- [13] Hafner, Danijar, et al. "Dream to control: Learning behaviors by latent imagination." *arXiv preprint arXiv:1912.01603* (2019).
- [14] Hafner, Danijar, et al. "Mastering diverse domains through world models." *arXiv preprint arXiv:2301.04104* (2023).
- [15] Konda, Vijay, and John Tsitsiklis. "Actor-critic algorithms." *Advances in neural information processing systems* 12 (1999).
- [16] Hodel, Michael (2023, July 7). *ARC: Where do we stand today?*. Lab42. <https://lab42.global/arc-article/>
- [17] Dziri, Nouha, et al. "Faith and Fate: Limits of Transformers on Compositionality." *arXiv preprint arXiv:2305.18654* (2023).
- [18] Wang, Jane X., et al. "Learning to reinforcement learn." *arXiv preprint arXiv:1611.05763* (2016).
- [19] Duan, Yan, et al. "RL²: Fast reinforcement learning via slow reinforcement learning." *arXiv preprint arXiv:1611.02779* (2016).