

월드 모델을 이용한 ARC 문제의 사전 지식 평가*

이승필⁰¹ 이지환² 김선동²

¹광주과학기술원 전자전기컴퓨터공학부 ²광주과학기술원 AI 대학원

iamseungpil@gm.gist.ac.kr, jihwan.lee@gm.gist.ac.kr sundong@gist.ac.kr

Evaluating Prior Knowledge of ARC Using World Models

Seungpil Lee⁰¹ Jihwan Lee² Sundong Kim²

¹GIST EECS ²GIST AI

요 약

오늘날 AI 연구들은 특정 과제를 해결하는 데엔 성과를 거둔 반면, 일반화된 문제를 해결하는 방향으로 큰 진척을 보이지 못하고 있다. 보편적인 인지 능력을 시험할 방법이 부족한 것이 그 이유 중 하나이다. ARC 벤치마크는 추론 능력을 정확히 측정하기 위해 새로이 제시된 문제 집합이다. 하지만, 아직까지 ARC 벤치마크를 해결할 때 사전 지식이 주는 영향과, 필요한 사전 지식의 종류가 분석되지 않아 연구가 어려운 상황이다. 본 연구에서는 월드 모델(World Model) 알고리즘을 사용해 ARC 벤치마크를 풀 때 사전 지식이 미치는 영향과 ARC에 포함된 사전 지식의 종류를 분석하는 새로운 방법을 제안한다.

1. 서 론

보편적인 문제를 해결할 수 있는 모델을 개발하는 것은 오늘날 인공지능 연구의 당면 과제 중 하나이다. 자연어 처리와 이미지 분석과 같이 하나의 특정 과제를 해결하기 위한 모델들은 깊이 연구됐지만, 유연한 추론 능력을 갖춘 모델에 관한 연구는 큰 진척이 없는 실정이다. 이처럼 보편적인 문제를 해결할 수 있는 일반 인공지능(AGI) 연구가 어려운 이유는, 추론 능력이 어떤 요소들로 이루어지는지, 이를 어떻게 평가할지 예상하기 어렵다는 데 있다.

Abstraction and Reasoning Corpus(이하 ARC) 벤치마크[1]는 추론 능력을 정확히 측정하기 위해 개발됐다. 추론 능력만을 평가하기 위해 문제 해결에 필요한 사전 지식의 양과 데이터 개수를 최대한 줄인 점이 ARC의 특징이다. 그러나, ARC 벤치마크를 해결할 때 1) 사전 지식이 미치는 영향이 어느 정도인지 2) 어떤 종류의 사전 지식이 필요한지 구체적으로 연구된 바 없다. 그 탓에 ARC가 실제로 추론 능력만을 평가할 적절한 데이터셋으로 기능하는지, 또 각 ARC 문제 별 난이도를 구분하기 어렵다는 문제가 존재한다.

본 연구에서는 ARC 벤치마크에 내재하는 사전 지식을 분석하기 위해 월드 모델(World Model)[2] 알고리즘을 사용하는 새로운 접근 방법을 제안한다. 월드 모델은 환경에 내재하는 사전 지식을 추출하여 행동 결정에 활용하는 강화학습 알고리즘이다. 이 특성을 활용하여 ARC 벤치마크의 사전지식을 분석하는 다음 두 가지 방법을 제안한다.

1. 월드 모델을 사용했을 때와 사용하지 않았을 때의 문제 해결 성능을 비교한다.
2. 월드 모델이 각 ARC 문제에서 추출한 특징들을 클러스터링한 후 유사도를 비교한다.

일련의 과정을 통해 ARC 벤치마크를 푸는 데 사전 지식이 어느 정도 영향을 미치는지, ARC 벤치마크에 어떤 종류의 사전 지식이 포함 되는지 분석할 수 있길 기대한다.

2. 연구 배경

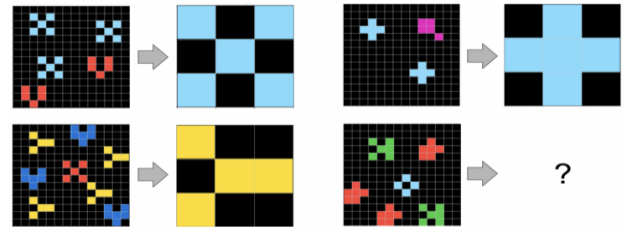
2.1 ARC

ARC 벤치마크는 인공지능의 지능을 측정하기 위해 만들어진 벤치마크이다. Francois Chollet 는 지능 I 를 다음과 같이 정의한다[1].

$$I = \text{Avg}[\alpha \sum_{c \in C_{ur}} [Pr(C) \cdot \frac{GD_C}{P+E_C}]] \cdots (1)$$

식(1)에서 α 는 가중치를, C 는 하나의 커리큘럼(훈련 데이터)를, $Pr(C)$ 는 커리큘럼이 나타날 확률을, GD_C 는 풀어낸 문제의 난이도(Generalization Difficulty)를, P 가 사전 지식을, E 가 모델이 한 경험을 각각 의미한다. 즉, 위 수식은 지적 능력이 풀어낸 문제의 난이도에 비례하고, 사전 지식과 경험의 양에 반비례함을 뜻한다. 이러한 정의에 따라서 Francois Chollet 는 모델의 지능을 정확히 측정하기 위해 사전 지식과 경험을 줄인 ARC 벤치마크를 제안한다.

ARC 벤치마크의 한 문제는 2~5 개 내외의 입력-출력 쌍과 출력이 주어지지 않는 하나의 입력으로 이루어진다. 문제를 해결하기 위해선 주어진 입력-출력 쌍으로부터 규칙을 도출해, 입력에 대해 예상되는 출력을 맞춰야 한다.



[그림 1] ARC 문제의 한 예

ARC 벤치마크는 인간과 인공 지능의 차이를 포착해내는 데 일정 수준 성공을 거두었다. 사람이 평균 80%의 정확도를 보이는 반면, 현재 가장 좋은 결과를 보인 모델은 30%, Transformer 기반 모델은 불과 10%의 정확도에 그쳤다[3].

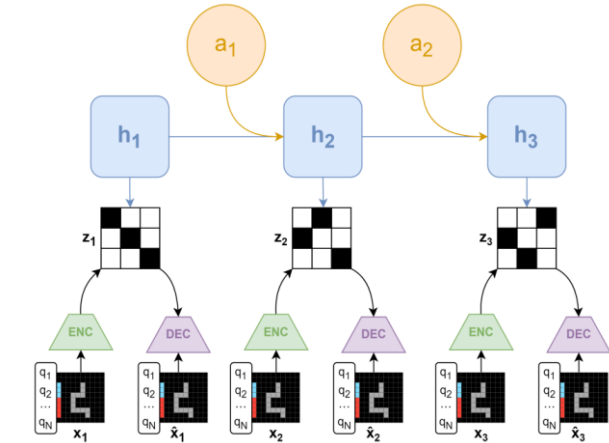
* 이 논문은 과학기술정보통신부의 재원으로 한국연구재단과 정보통신기획평가원의 지원을 받아 수행된 연구임 (RS-2023-00240062, RS-2023-00216011, 2019-0-01842)

그럼에도 불구하고 현재 ARC 벤치마크는 각 문제를 푸는데 필요한 최소의 사전 지식이 무엇인지 제시하지 못한다는 단점이 존재한다[1]. 그 탓에 모델이 문제를 풀기 전에 주는 사전 지식을 최소화해야 한다는 Chollet의 주장을 엄격하게 만족시키진 못하고 있다. 실제로 ARC 벤치마크를 해결하는데 GPT-4로 대표되는 대형 언어 모델을 사용하는 방법들이 제안됐으나, 이러한 접근법은 모델에 필요 이상의 사전 지식을 제공할 수도 있다는 점에서 Chollet가 정의한 지능의 접근 방법과는 상이하다. 따라서 ARC 벤치마크가 보편적인 지능의 평가 척도로 사용되기 위해선 각 문제에 포함된 사전 지식을 분석할 새로운 방법이 필요하다.

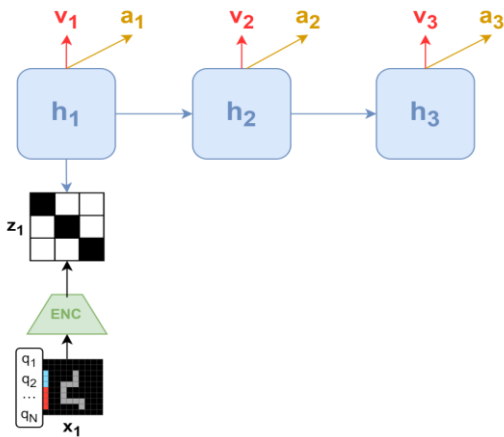
2.2 월드 모델(World Model)

월드 모델[2]은 복잡한 환경이 주어졌을 때 미래 상황을 빠르게 예측하기 위해 제안된 모델이다. 기존 강화학습 모델들은 다음 상태를 예측할 때마다 입력 값으로부터 계산해야 했다. 그 탓에 입력이 복잡해질수록 행동을 결정하는 데 오랜 시간이 걸렸다. 이를 해결하기 위해 월드 모델은 환경으로부터 단순화된 특징 벡터를 학습하고, 이 벡터로부터 다음 상태를 빠르게 예측해낸다.

DreamerV3[4]는 월드 모델에 기반한 강화학습 모델이다. DreamerV3는 월드 모델과 액터-크리틱(Actor-Critic) 두 부분으로 이루어졌다. 월드 모델은 이미지에서 해당 도메인에 대한 사전 지식을 추출하는 한편, 액터-크리틱은 추출한 사전 지식에 바탕해 실제 행동을 결정하는 역할을 나누어 맡는다.



[그림 2] DreamerV3의 월드 모델이 학습하는 방법을 나타내는 그림



[그림 3] DreamerV3의 액터-크리틱이 학습하는 방법을 나타내는 그림

DreamerV3의 핵심은, 학습 과정에서 [그림 3]과 같이 행동 정책을 이미지 입력 대신 h_t 벡터로 표현되는 사전 지식으로부터 학습한다는 데 있다. 이후 일어날 상황을 단순화된 벡터에서 바로 추측하는 방식 덕에 계산량이 줄어들어, 자연히 복잡한 문제를 빠르고 정확하게 예측할 수 있게 된다[5].

DreamerV3의 월드 모델은 Recurrent State-Space Model[5]을 이용해 환경에 대한 사전 지식(Latent Dynamics) h_t 를 학습한다.

$$\text{RSSM} \begin{cases} \text{Sequence model:} & h_t = f_\phi(h_{t-1}, z_{t-1}, a_{t-1}) \cdots (2) \\ \text{Encoder:} & z_t \sim q_\phi(z_t | h_t, x_t) \cdots (3) \\ \text{Dynamics predictor:} & \hat{z}_t \sim p_\phi(z_t | h_t) \cdots (4) \end{cases}$$

h_t 는 시간 t 에서의 사전 지식을, z_t 는 이미지로부터 추출한 임베딩을, \hat{z}_t 는 h_t 에서 예측한 이미지 임베딩을 각각 의미한다. 이 때, 월드 모델은 실제 값인 z_t 와 예측 값인 \hat{z}_t 사이 차이를 최소화하는 것을 목표로한다.

$$L_{pred}(\phi) = -\ln p_\phi(x_t | z_t, h_t) \cdots (5)$$

식(5)는 DreamerV3 월드 모델의 손실함수이다. ϕ 는 파라미터를 의미하며, $p_\phi(x_t | z_t, h_t)$ 는 실제 이미지 임베딩 z_t 와 사전 지식 h_t 가 주어졌을 때, 이미지 x_t 를 바르게 예측할 확률로, 이 값이 1에 가까워 질수록 사전 지식 h_t 로부터 예측한 값이 정확하다는 사실을 뜻한다.

DreamerV3는 고려해야할 정보(중력, 다양한 사물, 상호작용)가 많고, 빠르게 변화하는 탓에 기존 인공지능 방법론으로는 풀지 못했던 마인크래프트 과제를 해결했다. 이처럼 DreamerV3가 환경에 내재하는 사전 지식을 추출하는데 강력한 성능을 보여준 바, 여러 월드 모델 구현 중 DreamerV3를 분석에 채택하고자 한다.

2.3 시간-비의존적(Time-Invariant) DreamerV3

일반적으로 모델이 추출한 특징 벡터는 공분산(covariance)이 커 분석하기 쉽지 않다고 알려져 있다. 특징 벡터에서 사람이 이해할 수 있는 의미를 찾아내기 위해선 선형적으로(linearly separable) 특징을 추출하는 별도의 학습 방법이 필요하다[6]. 이와 같이 특징 벡터를 선형적으로 추출하는 학습 방식을 Disentanglement Representation Learning(이하 DRL)이라 부른다. 기존의 VAE나 GAN 모델에 DRL을 적용한 β -VAE[7], InfoGAN[8] 등이 대표적인 예이다.

Christopher Reale 등은 기존 월드 모델의 손실 함수에 시간 비의존 성질을 추가할 경우 특징 벡터를 선형적으로 추출할 수 있음을 보였다[9]. 시간을 뒤섞은 손실함수를 추가함으로써, 시간이 지나도 변하지 않는 환경만의 고유한 특징을 추출할 수 있기 때문이다.

본 연구에서도 추출한 특징 간의 원활한 비교를 위해 DreamerV3에 시간 비의존성을 더한 손실함수(식(7))를 제안한다.

$$L_{invariant}(\phi) = -\ln p_\phi(x_t | z_t, h_{m(t)}) \cdots (6)$$

$$L_{mod}(\phi) = \frac{L_{pred} + L_{invariant}}{2} \cdots (7)$$

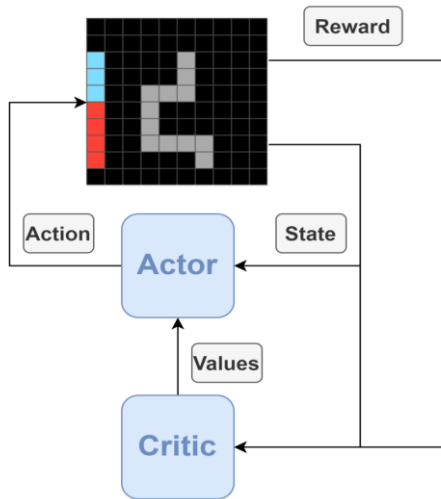
$m(t)$ 는 시간 집합 $\{0, 1, \dots, T_d\} \rightarrow \{0, 1, \dots, T_d\}$ 로 일대일 대응되는 함수로, 스텝(time-step)을 임의로 섞는 역할을 한다. 즉, 변경된 손실함수(식(7))는 시간 순차적인 기존 손실함수(식(5))에 시간을 임의로 섞은 손실함수(식(6))를 추가했음을 의미한다.

3. 실험 제안

3.1. 실험 방법

본 연구는 1) ARC 문제를 풀 때 사전 지식이 주는 영향을 분석하고, 2) ARC 문제에 포함되는 사전 지식의 종류를 확인하는 데 목표를 둔다. 이를 위해 두 가지 실험을 제안한다.

먼저, 1) DreamerV3 모델(월드 모델 + 액터-크리틱)과 순수 액터-크리틱 모델의 ARC 문제를 푸는 성능을 비교하여 월드 모델이 추출한 사전 지식이 문제 해결에 어느 정도 도움을 주는지 확인하고자 한다. 비교에 사용하는 액터-크리틱은 사전 지식인 h_t 벡터를 사용하지 않고 학습한다는 점에서 DreamerV3의 액터-크리틱과 다르다. 사전 지식에 큰 영향을 받는 문제일수록 두 모델의 정확도 및 풀이 시간에 차이가 커질 것으로 예상된다.



[그림 4] 순수 액터-크리틱이 학습하는 방법을 나타내는 그림

다음으로, 2) DreamerV3가 추출한 특징을 클러스터링한 후 그 유사도를 비교하여, 어떤 문제들이 유사한 풀이법을 공유하고, 전체 문제를 몇 가지 유형으로 나눌 수 있을지 확인하고자 한다. DreamerV3가 추출한 특징 벡터가 과제의 사전 지식을 의미하기 때문에, 이 과정을 통해 ARC 벤치마크를 풀 때 필요한 사전 지식의 종류를 확인할 수 있을 것으로 기대한다.

실험 세부 사항은 아래와 같다.

1. 성능 비교 실험 세부
 - A. 순수 액터-크리틱과 DreamerV3, 시간 비의존적 DreamerV3 모델을 ARC 전체 400 문제에 학습시킨다.
 - B. 정확도 차이와 해결한 문제에 대해 걸린 시간을 비교한다.
2. 특징 분석 실험 세부
 - A. 순수 액터-크리틱과 DreamerV3, 시간 비의존적 DreamerV3 모델을 각 ARC 문제 별로 학습시킨다.
 - B. 각 문제 별로 추출한 특징 벡터를 클러스터링한 후, 실루엣 계수가 가장 높을 때의 군집에 대해 특징을 분석한다.

3.2. 실험 환경

ARC 벤치마크를 강화학습을 통해 학습하기 위해서는 모델이 학습할 수 있는 환경을 구성하는 것이 중요하다. Arcle[10]은 ARC 벤치마크를 강화학습에 사용할 수 있는 형태로 제공하는 라이브러리로, ARC와 Mini-ARC 문제를 모두 제공한다. 모델은 주어진 환경에서 ARC 문제의 격자 중 일부를 고르고, 고른 격자에 색칠, 회전, 복사 등 34개의 상호작용 중 하나를 적용하는 행동을 취할 수 있다.

4. 결론

본 논문에서는 ARC 벤치마크에 사전 지식이 미치는 영향과 포함된 사전 지식의 종류를 분석하기 위해 1) 월드 모델이 ARC 문제를 푸는 성능을 확인하고, 2) 월드 모델이 ARC 문제로부터 추출한 특징 벡터를 클러스터링하는 두 가지 방법을 제안했다. 이 연구는 다음 두 가지 시사점을 가진다.

1. ARC 벤치마크를 푸는데 사전 지식이 얼마나 필요할지 분석함으로써 사전 지식만을 이용하겠다는 목적에 얼마나 부합하는지 확인한다.
2. 각 ARC 문제 별로 필요한 사전 지식을 분석하여 ARC 문제를 소분류하는 새로운 기준을 제공한다.

참고문헌

- [1] François Chollet, On the Measure of Intelligence. arXiv, 2019.
- [2] David Ha, World Models, arXiv, 2018
- [3] Aysja Johnson, Fast and Flexible: Human Program Induction in Abstract Reasoning Tasks, arXiv, 2021
- [4] Danijar Hafner, Mastering Diverse Domains through World Models, arXiv, 2023
- [5] Danijar Hafner, Learning Latent Dynamics for Planning from Pixels, arXiv, 2018
- [6] Xin Wang, Disentangled Representation Learning, arXiv, 2023
- [7] Irina Higgins, beta-VAE: Learning Basic Visual Concepts with a Constrained Variational Framework, International Conference on Learning Representations, 2017.
- [8] Xi Chen, InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets, arXiv, 2016
- [9] Christopher Reale, Learning and Understanding a Disentangled Feature Representation for Hidden Parameters in Reinforcement Learning, arXiv, 2022
- [10] "Arcle", <https://github.com/ConfeitoHS/ar>