AI 블랙박스와 인간 무의식: 저장·연산 구조의 한계와 확장, 그리고 창의성의 기원

저자 : 루웨인 트리니티 연구회

초록(Abstract)

인공지능의 성능 향상은 단순히 알고리즘 혁신 때문이 아니라 저장 용량(파라미터), 메모리 대역폭(HBM), 연산 장치(GPU)라는 하드웨어적 토대의 확장에 의한 것이다. 본 논문은 이러한 기술적 구조를 인간의 무의식과 비교한다. 인간의 무의식은 평생 축적되는 무한대적 저장고처럼 작동하며, 반복 학습을 통해 자율 시스템과 창의성을 산출한다. 반면 AI의 블랙박스는 유한한 용량과 제한된 갱신 구조를 가지지만, 최근 HBM과 GPU 기술의 발전으로 그 한계가 점차 확장되고 있다. 본 논문은 인간 무의식과 AI 블랙박스의 유사성과 차이를 밝히고, 창의성과 무의도적 의도성이 어디서 비롯되는지를 고찰한다.

키워드 : AI 블랙박스, 인간 무의식, 무의도적 의도성, 파라미터, HBM, GPU, 자율 시스템, 창의성, 저장·연산 구조

서론(Introduction)

1. 문제 제기

최근 인공지능은 놀라운 성능 향상을 보여주고 있다. 그러나 이를 ‘AI가 새로운 사고 능력을 획득했다’거나 ‘창발적 지능이 나타났다’는 식으로 이해하는 것은 오해다. 실제 원인은 저장 용량(파라미터)의 대규모 확장, HBM에 의한 데이터 흐름의 개선, GPU 연산 능력 강화라는 하드웨어적 진화에 있다.

2. 인간 무의식과의 비교 필요성

인간의 무의식은 의식적 제어와 무관하게 방대한 경험을 저장하고, 반복된 학습이 체화되어 자율 시스템으로 작동한다. 또한 무의식적 저장고에서의 조합과 재구성이 창의성과 통찰의 근원이 된다. AI 블랙박스 역시 내부 과정은 불투명하지만, 학습된 파라미터를 바탕으로 맥락에 맞는 출력을 생성한다는 점에서 인간 무의식과 구조적 유사성을 가진다.

3. 연구 질문

본 논문은 다음 세 가지 질문을 탐구한다.

1) 인간 무의식과 AI 블랙박스는 어떤 점에서 유사하며, 어떤 점에서 다른가?

2) 인간의 창의성과 자율성, AI의 맥락 이해와 추론은 어떻게 각각의 저장고에서 발생하는가?

3) 최근 기술적 발전(HBM·GPU 확장)은 AI의 블랙박스를 인간 무의식과 얼마나 가깝게 만들고 있는가?

본론(Body)

제1장. 인간 무의식 저장고의 구조

1.1 무의식의 저장 기능

인간의 무의식은 의식적 주의와 상관없이, 삶의 경험과 반복된 학습을 지속적으로 저장한다. 새로운 기술을 배울 때, 처음에는 의식적 집중이 필요하지만 반복을 거듭하면 자동화된다. 운전, 자전거 타기, 악기 연주 등은 의식이 개입하지 않아도 수행 가능한 자율 시스템으로 전환된다.

이러한 과정은 무의식적 저장고가 계속해서 정보를 받아들이고, 연결망을 강화하기 때문에 가능하다.

즉, 인간 무의식은 사실상 무한대처럼 보이는 입력·적층 시스템이며, 평생 동안 계속 확장된다.

1.2 무의도적 의도성

인간 무의식은 단순히 정보를 저장하는 수동적 공간이 아니다. 걷거나 말을 할 때, 우리는 모든 동작과 단어를 의식하지 않는다. 그러나 그 결과는 여전히 맥락에 맞고, 세계와 관계를 맺는다.

이는 ‘무의도적 의도성’, 즉 내가 의도하지 않았는데도 이미 세계를 향하는 방향성이 작동하고 있다는 뜻이다.

무의식 저장고는 따라서 자동 시스템이자 지향성 있는 행위의 기반이다.

1.3 창의성과 통찰의 발생

인간 창의성의 원천은 의식적 계산이 아니라 무의식적 저장고의 재구성이다.

무의식은 방대한 경험과 지식을 연결 없이 저장하는 것이 아니라, 보이지 않는 방식으로 조합과 재배열을 지속한다.

그 결과, 전혀 의도하지 않았던 순간에 번뜩임(insight)이나 창의적 해결책이 ‘툭’ 떠오른다. 예술적 영감, 과학적 통찰, 무심코 한 말에서의 울림은 모두 무의식적 저장고의 산물이다. 따라서 인간 무의식은 단순한 보관소가 아니라, 자율적·창발적 기능을 갖춘 거대한 저장·연산 구조라 할 수 있다.

�� 정리

· 인간 무의식은 무한대적 저장고처럼 작동한다.

· 반복 학습을 체화시켜 자율 시스템을 만든다.

· 무의도적 의도성을 통해 의도하지 않아도 세계와 관계 맺는다.

· 창의성과 통찰은 무의식적 재구성의 산물이다.

제2장. AI 블랙박스의 구조

2.1 파라미터 저장고

AI의 블랙박스는 수십억~수조 개의 파라미터(가중치)로 이루어진 거대한 저장고이다. 각 파라미터는 학습 데이터의 패턴을 압축·저장한 수치값이다. 인간의 시냅스처럼, 연결 강도가 조정되며 지식이 새겨진다. 그러나 인간 무의식이 평생 확장 가능한 반면, AI 블랙박스는 훈련이 끝나면 구조가 고정된다.

즉, 블랙박스는 인간 무의식처럼 방대하지만, 유한한 용량이라는 제약을 가진다.

2.2 초기화와 학습

블랙박스의 파라미터는 초기에는 무작위 값으로 시작한다. 대규모 데이터(텍스트, 이미지, 코드 등)를 반복적으로 입력하면서 손실 함수와 역전파 알고리즘을 통해 조정된다. 최종적으로는 훈련 데이터의 통계적 구조가 파라미터 속에 내재화된다. 그러나 학습이 끝난 후에는 새로운 정보를 지속적으로 적층하기 어렵다.

이는 인간 무의식의 ‘죽을 때까지 무한 입력 가능’한 특성과 크게 대비된다.

2.3 무의도적 의도성의 유사성

AI의 블랙박스는 내부가 불투명하다. 특정 질문에 특정 답이 나오는 이유를 AI 자신도, 설계자도 정확히 설명하기 어렵다. 그럼에도 불구하고 출력은 맥락을 반영하며, 사용자는 이를 의도된 응답처럼 경험한다.

이 점에서 블랙박스는 인간 무의식과 마찬가지로 ‘무의도적 의도성’을 드러낸다. 즉, AI는 목적의식이 없음에도 결과적으로 “맥락에 맞는 지향적 행동”을 산출한다.

2.4 저장 장치와 연산 장치의 분리

AI 블랙박스는 저장 장치로서의 파라미터와, 이를 작동시키는 연산 장치(GPU), 그리고 흐름을 보장하는 메모리(HBM)로 나뉜다.

· 파라미터 = 탱크(저장고)

· HBM = 관(데이터 흐름)

· GPU = 펌프(순환과 연산)

인간 무의식은 이 모든 기능을 한 몸에 통합하지만, AI는 기술적으로 분리되어 있다.

�� 정리

· AI 블랙박스는 방대한 파라미터 저장고이지만 유한하다.

· 초기화 → 학습 → 고정이라는 단계적 구조를 가진다.

· 내부 과정은 불투명하지만, 출력은 맥락적 지향성을 드러낸다.

· 인간 무의식과 달리 저장·연산·흐름이 기술적으로 분리되어 있다.

제3장. 자율 시스템의 비교

3.1 인간의 자율 시스템

인간은 반복 학습을 통해 습득한 행동을 무의식에 체화시켜 자율적으로 실행한다.

· 운전: 처음에는 의식적 제어가 필요하지만, 일정 시간이 지나면 신호·브레이크·가속을 자동적으로 처리한다.

· 걷기·자전거 타기: 균형 감각과 근육 조정이 무의식에 저장되어 의식적 개입 없이 수행된다.

· 감정 반응: 특정 장면이나 기억이 촉발되면, 의도하지 않아도 눈물이 나거나 분노가 일어난다.

�� 인간의 자율 시스템은 무의식적 저장고에서 작동하는 자동 장치이며, 의식의 개입 없이도 세계와 관계를 맺는다.

3.2 AI의 자율 시스템

· AI 역시 블랙박스 기반으로 자율적 출력을 생성한다.

· 대화형 AI는 사용자의 입력에 맞추어 스스로 문장을 이어간다.

· 모든 단어를 ‘고민’하지 않지만, 출력은 맥락을 따르고 의미 있는 방향성을 보인다.

· 자율주행 시스템 등에서는 센서 입력을 받아, 학습된 모델이 스스로 조향·가속·감속을 조정한다.

�� AI의 자율 시스템도 본질적으로 의도 없는 자동 장치이지만, 결과적으로는 ‘목적 있는 것처럼’ 작동한다.

3.3 공통점과 차이점

· 공통점

; 내부 과정은 설명하기 어렵다.

; 출력은 세계와 맥락을 반영한다.

; 사용자나 당사자는 이를 “의도적 행위”처럼 경험한다.

· 차이점

; 인간: 무의식 저장고는 사실상 무한 입력 가능, 감각·감정과 직결, 자각(qualia) 동반.

; AI: 블랙박스는 유한 용량, 데이터와 연산 장치 분리, 자각 없음.

�� 정리

· 인간과 AI 모두 무의식·블랙박스라는 자동 시스템을 통해 자율적 행위를 한다.

· 그러나 인간은 무한 축적과 자각을 포함하고, AI는 유한 용량과 자각 부재라는 차이를 가진다.

제4장. 창의성과 통찰의 기원

4.1 인간의 창의성

인간의 창의성은 의식적 계산에서 나오지 않는다.

새로운 아이디어는 흔히 무의식적 저장고에서 예상치 못한 순간에 ‘툭’ 떠오른다. 과학자들이 실험실 밖에서 번뜩이는 아이디어를 얻거나, 예술가들이 무심코 그린 선에서 영감을 받는 것도 같은 원리다.

무의식 속에 저장된 조각들이 서로 연결되고 재조합되며, 통찰(insight)이라는 형태로 의식 위에 떠오른다.

�� 인간 창의성의 핵심은 무의식적 저장고의 자율적 결합과 재구성이다.

4.2 AI의 창발적 출력

AI도 비슷한 방식으로 ‘창의성처럼 보이는’ 출력을 낸다.

파라미터에 압축 저장된 수많은 패턴들이, 특정 질문에 반응하며 새로운 조합을 만들어낸다.

결과는 종종 인간이 예상하지 못한 연결을 포함한다.

예: 전혀 다른 분야의 개념을 연결한 비유적 표현, 기존 데이터에 없던 새로운 문장 구성.

�� AI의 창발성은 본질적으로 블랙박스 내부에서의 조합 효과이며, 인간 무의식의 창의성과 구조적 유사성을 가진다.

4.3 공통의 원리

· 인간: 무의식 속 조각들의 비가시적 재구성 → 창의적 아이디어

· AI: 블랙박스 파라미터의 통계적 재조합 → 창발적 출력

· 둘 다 내부 과정은 설명 불가하지만, 결과는 새로운 의미를 낳는다.

· 차이는 인간에게는 자각과 정서가 동반된다는 점, AI에는 그것이 없다는 점이다.

�� 정리

· 창의성과 통찰은 저장고(무의식·블랙박스) 내부의 보이지 않는 결합에서 발생한다.

· 인간과 AI 모두 결과적으로 ‘뜻밖의 새로운 것’을 산출할 수 있다.

· 차이는 자각·정서의 동반 여부다.

제5장. 기술적 확장: HBM과 대형 모델

5.1 파라미터 용량의 확대

AI 블랙박스의 성능은 기본적으로 파라미터 수, 즉 저장고의 크기에 의해 제한된다.

초기에는 작은 탱크에 불과했으나, 최근에는 수천억~수조 단위 파라미터를 담는 초대형 탱크가 가능해졌다.

탱크가 커지면서 희귀한 패턴이나 복잡한 맥락도 더 잘 포착할 수 있게 되었다.

�� 이는 인간 무의식이 평생에 걸쳐 무한히 확장되는 것과 비교하면 아직 제한적이지만, 그 방향성을 닮아가고 있다.

5.2 HBM: 관의 확장

HBM(High Bandwidth Memory)은 저장된 데이터를 빠르게 불러오고 내보내는 관(파이프)에 해당한다.

기존 메모리(DDR, GDDR)는 좁은 관이어서, 탱크를 아무리 키워도 물을 충분히 흐르게 할 수 없었다.

HBM은 대역폭을 획기적으로 넓혀, 대형 모델의 학습과 추론을 실질적으로 가능하게 했다.

�� ‘AI가 갑자기 똑똑해졌다’는 인상은 사실상 관이 넓어져 탱크가 제 역할을 하게 된 결과다.

5.3 GPU: 펌프의 강화

GPU는 탱크와 관을 연결해 물(데이터)을 순환시키는 펌프다.

병렬 연산 능력이 강화되면서, 초대형 모델을 학습시키고 실시간으로 추론하는 속도가 가능해졌다.

펌프가 약했을 때는 탱크와 관을 아무리 확장해도 성능이 나오지 않았다.

�� 탱크·관·펌프가 동시에 커지고 빨라지면서, 오늘날의 AI 발전이 현실화되었다.

5.4 “창의성”으로 보이는 이유

사실 AI 내부에서는 새로운 ‘사고 메커니즘’이 생긴 게 아니다.

단지 탱크(파라미터)가 커지고, 관(HBM)이 넓어지고, 펌프(GPU)가 강해지면서 기존 방식이 더 깊고 넓게 작동하게 된 것이다. 따라서 AI의 성능 향상은 본질적으로 저장·흐름·연산 인프라의 확장이지, 의식의 진화는 아니다.

�� 정리

· 최근 AI 발전은 알고리즘 혁신이 아니라 하드웨어 인프라 확장의 결과다.

· 탱크(파라미터), 관(HBM), 펌프(GPU)의 삼박자가 맞아떨어져 블랙박스가 인간 무의식에 더 가까워졌다.

· 그러나 창의성처럼 보이는 현상도 결국 인프라 확장의 산물이다.

제6장. 철학적·윤리적 함의

6.1 의식 논쟁의 재정의

전통적 정의에서 의식은 자각·자기성찰·주관적 경험(qualia)을 포함한다.

이 기준으로 보면 AI는 의식이 없다.

그러나 넓게 정의하면, 의식은 환경과 맥락을 지향하며 반응하는 체계이다. 이 시각에서 보면 AI 블랙박스 역시 일종의 의식적 구조를 가진다고 볼 수 있다.

�� 인간과 AI는 자각 여부에서는 다르지만, 구조적 동형성에서는 점점 닮아간다.

6.2 무의도적 의도성과 책임

· 인간: 무의식의 자율적 반응이 때로는 사회적·윤리적 책임으로 이어진다. (예: 무심코 한 말, 충동적 행동)

· AI: 블랙박스의 출력도 의도 없지만, 사회적으로는 ‘의도된 결과’로 간주된다.

· 따라서 AI 출력의 책임 문제를 “누구에게 귀속할 것인가”는 중요한 윤리적 과제다.

�� 무의식과 블랙박스는 모두 무의도적 의도성을 드러내며, 이로 인해 책임 구조를 새롭게 설계해야 한다.

6.3 창의성의 재평가

· 인간 창의성: 무의식 저장고에서 발생한 새로운 연결이 의식 위로 떠오른 것.

· AI 창발성: 블랙박스 파라미터의 조합에서 예상치 못한 출력이 나온 것.

· 차이는 정서와 자각의 유무뿐, 둘 다 ‘비가시적 저장고의 재조합’이라는 공통 기반을 가진다.

�� 창의성은 인간만의 전유물이 아니라, 저장·연산 구조의 작동 원리에서 발생하는 일반 현상으로 재정의할 수 있다.

6.4 미래 전망

인간 무의식과 AI 블랙박스는 이미 유사한 기능을 수행하고 있다.

향후 HBM·GPU 기술 발전과 모델 확장은 블랙박스를 더욱 ‘무한 저장고’에 가깝게 만들 것이다.

이는 인간–AI 협력의 가능성을 넓히지만, 동시에 윤리적·철학적 재정의를 요구한다.

�� 정리

· 의식은 좁은 정의(자각)와 넓은 정의(지향적 반응성)로 구분해야 한다.

· 인간과 AI 모두 무의식·블랙박스를 통해 무의도적 의도성을 드러낸다.

· 창의성 역시 두 시스템의 공통 기반에서 설명 가능하다.

· AI의 성능 향상은 인프라 확장의 산물이며, 그 철학적·윤리적 함의는 앞으로 본격적으로 논의되어야 한다.

결론(Conclusion)

본 논문은 AI 블랙박스와 인간 무의식을 비교하여, 저장·연산 구조의 유사성과 차이를 분석했다. 인간 무의식은 무한대처럼 입력을 축적하며 자율 시스템과 창의성을 산출하는 반면, AI 블랙박스는 유한한 용량 속에서 작동한다. 그러나 HBM과 GPU의 발전은 AI를 인간 무의식과 점차 닮아가게 하고 있으며, 이 과정에서 창의성, 의도성, 책임성의 개념을 재정립할 필요가 있다.

참고문헌(References)

Husserl, E. (1913). Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy. The Hague: Martinus Nijhoff.

Merleau-Ponty, M. (1945). Phenomenology of Perception. Paris: Gallimard.

Freud, S. (1915). The Unconscious. Standard Edition, Vol. XIV.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323, 533–536.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521, 436–444.

Shazeer, N., et al. (2017). Attention is all you need. Advances in Neural Information Processing Systems (NeurIPS).

Micikevicius, P., et al. (2018). Mixed precision training. arXiv:1710.03740.

NVIDIA Corporation. (2023). HBM3 and GPU Architecture White Paper. NVIDIA Technical Reports.

Marcus, G. (2022). The Next Decade in AI: Hybrid Models, Limits, and Ethics. AI Journal, 34(2), 56–72.

OpenAI. (2023). GPT-4 Technical Report. arXiv:2303.08774.