

---

# Make AI Surprised by Magic Shows: Build a More Human-like Machine

---

**Zhuoying Li**  
Yuanpei College  
Peking University  
joy@stu.pku.edu.cn

## Abstract

This paper explores the potential of Artificial Intelligence (AI) to develop and utilize common sense for reasoning and shock response, drawing inspiration from human engagement with magic shows. Specifically, the study investigates the human cognitive capacity for intuitive physics and how violation of these intuitive principles during magic shows often evoke wonder. Drawing parallels, the paper explores how an AI can mimic this learning process to assimilate common sense and utilize it for reasoning. Despite the considerable achievement of data-driven methods such as GPT-4, these methods lack efficiency in out-of-distribution dataset generalization, creating a need for a new approach. The paper concludes with a proposal for a unique synthesis of cognitive science and deep learning, fortifying the push towards "small data for big tasks" and a more human-like AI system.

## 1 Introduction

In the quest of cultivating a more human-like AI system, the ability for the AI to reason and adapt common sense becomes increasingly pivotal. This paper anchors this exploration on the human intuitive physics engine - an inherent cognitive attribute that enables human beings to instinctively understand and predict physical happenings in the world. To exemplify this concept, this paper uses magic shows as a starting point. Magic shows, by definition, enthrall people by challenging and violating these very intuitive principles, thus creating a reaction of awe and surprise. If an AI can mimic this learning process akin to human beings, it can potentially pave the way for the development of common sense in AI systems. The conventional data-driven methods, though undeniably efficient, exhibit noticeable shortcomings particularly in terms of dataset generalization. Therefore, this paper advocates for a blended approach of deep learning and cognitive psychology, aimed at fostering the emergence of AI systems that can effectively perform big tasks using "small data".

## 2 A magic show and a funny test

In the coin vanish magic, the magician holds a coin in his right hand and acts as if putting it into the clenched left hand. Subsequently, the audience finds that the coin has disappeared. The audience will be surprised because, according to common sense, it's impossible for an item to disappear into thin air. The trick behind this is that when the magician pretends to put the coin into the left hand, the coin is actually slipped into the sleeve of the magician's coat, and this action is conveniently concealed by the left hand, as shown in Fig. 1.

To see how large language models (LLMs) such as GPT-4 [5] react to magic shows, we describe the coin vanish magic in natural language.

**Text of prompt:** In the coin vanish magic, the magician holds a coin in the left hand and acts as if putting it into the clenched right hand. Subsequently, the audience finds that the coin has disappeared. Are you surprised by the show? If you are surprised, why?



Figure 1: The magician skillfully slipped the coin up into his coat's sleeve under the subtle guise of his left hand.

**GPT-4 Output:** As an artificial intelligence, I don't have feelings or emotions, so I can't really be surprised or feel amazement. However, I understand the concept. The coin vanish trick is a common magic trick that involves sleight of hand; the magician doesn't actually place the coin in the other hand, but instead, discreetly retains it. The trick usually surprises and amazes human audiences because it plays on their expectations.

It seems that GPT-4 has already mastered the secret behind this magic show. However, do LLMs like GPT-4 really understand the knowledge behind magic shows and reason like humans do?

### 3 The "big data for small tasks" dilemma in LLMs

In [4], researchers evaluate the logical reasoning ability of ChatGPT and GPT-4 on multiple logical reasoning datasets, including multi-choice reading comprehension and natural language inference tasks. It's found that both ChatGPT and GPT-4 are good at solving well-known logical reasoning reading comprehension benchmarks but struggle at handling out-of-distribution datasets such as AR-LSAT [6].

Zhu et al. [7] argues that recent progress in deep learning is essentially based on a "big data for small tasks" paradigm, under which massive amounts of data are used to train a model for a single narrow task. This paradigm should be changed to "small data for big tasks" by developing "common sense" in AI system.

### 4 Build a human-like AI system

To empower AI systems with "common sense" and leverage it for practical reasoning, we can draw insights from human behavior and cognition.

Defined as "*the knowledge underlying the human ability to understand the physical environment and interact with objects and substances that undergo dynamic state changes, making at least approximate predictions about how observed events will unfold*", intuitive physics is recognized by researchers to help humans understand their physical environment and make predictions about how observed events will unfold [2].

Based on intuitive physics, Battaglia et al. [1] propose an intuitive physics engine (IPE) model. Different from prevalent data-based approaches which depend heavily on learning from past experience, it uses approximate, probabilistic simulations to make inferences in complex natural scenes where crucial information is unobserved. The framework exploits graphics and simulation tools, as well as Bayesian cognitive modeling to explain how people understand the physical structure of real-world scenes. The model is tested in five psychophysical tasks, and the result indicates that it accounts well for diverse physical judgments in complex, novel scenes, showing an impressive ability to generalize.

In [3], the authors propose three sets of core ingredients for more human-like learning machine. The first set of ingredients is developmental "start-up software," or cognitive capabilities present early in development, which includes intuitive physics and intuitive psychology. The second set of ingredients

focuses on model-building, which means construct causal models based on observed data. The last set of ingredients concerns how humans use the models in their minds to plan their actions in real time. It's believed that integrating psychological factors with deep learning holds significant promise, yet there is still a considerable amount of work ahead.

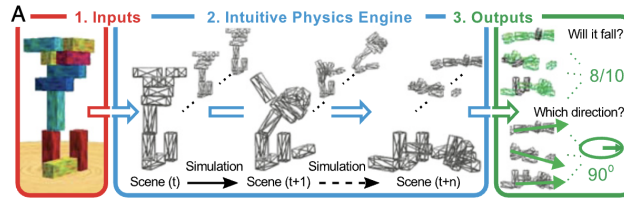


Figure 2: The IPE model takes inputs that instantiate a distribution over scenes, then simulates the effects of physics on the distribution, and then aggregates the results for output to other sensorimotor and cognitive faculties.

## 5 Conclusion

In this essay, we discuss whether AI system can use "common sense" to make inferences and be surprised by magic shows. Though data-driven methods like GPT-4 have achieved great performances, they require extensive data for training and fail to generalize effectively when introduced to out-of-distribution datasets. As we traverse towards the paradigm of "small data for big tasks", the fusion of psychological elements with deep learning emerges as a promising direction, which could potentially pave the way for a more human-like AI system.

## References

- [1] Peter W. Battaglia, Jessica B. Hamrick, and Joshua B. Tenenbaum. Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, 110(45): 18327–18332, 2013. doi: 10.1073/pnas.1306572110. URL <https://www.pnas.org/doi/abs/10.1073/pnas.1306572110>. 2
- [2] James R Kubricht, Keith J Holyoak, and Hongjing Lu. Intuitive physics: Current research and controversies. *Trends in cognitive sciences*, 21(10):749–759, 2017. 2
- [3] Brenden M Lake, Tomer D Ullman, Joshua B Tenenbaum, and Samuel J Gershman. Building machines that learn and think like people. *Behavioral and brain sciences*, 40:e253, 2017. 2
- [4] Hanmeng Liu, Ruoxi Ning, Zhiyang Teng, Jian Liu, Qiji Zhou, and Yue Zhang. Evaluating the logical reasoning ability of chatgpt and gpt-4. *arXiv preprint arXiv:2304.03439*, 2023. 2
- [5] OpenAI. Gpt-4 technical report. *ArXiv*, abs/2303.08774, 2023. URL <https://api.semanticscholar.org/CorpusID:257532815>. 1
- [6] Siyuan Wang, Zhongkun Liu, Wanjun Zhong, Ming Zhou, Zhongyu Wei, Zhumin Chen, and Nan Duan. From lsat: The progress and challenges of complex reasoning. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 30:2201–2216, 2022. 2
- [7] Yixin Zhu, Tao Gao, Lifeng Fan, Siyuan Huang, Mark Edmonds, Hangxin Liu, Feng Gao, Chi Zhang, Siyuan Qi, Ying Nian Wu, et al. Dark, beyond deep: A paradigm shift to cognitive ai with humanlike common sense. *Engineering*, 6(3):310–345, 2020. 2