Intuitive Physics Engine Is Not So Intuitive

Zhancun Mu
YuanPei College
Peking University
yhbylch@stu.pku.edu.cn

Abstract

It is beyond doubt that intuitive physics exists in human cognition. However, it remains unclear how physical knowledge is learned, used and generalized to novel situations. Among all the existing approaches to this problem, intuitive physics engine (IPE) is an appealing hypothesis. In this essay, we will discuss main existing approaches and their feasibility. Then we will take a deeper look at IPE, trying to show IPE is not a satisfying method.

1 Introduction

Humans excel at understanding environmental dynamics and making predictions about the future. These skills are essential for cognition, reasoning, task planning, and so on. We do not need to be physicists to have these abilities. Instead, we rely on our intuition, make approximate predictions, and still make some mistakes. The knowledge underlying this process is defined as *intuitive physics* [6]. It is definitely important to equip AI with such skills.

However, it remains unclear how physical knowledge is learned, used and generalized to novel situations. Research on intuitive physics has shown surprising results. For example, infants can show awareness of unexpected events [2]. While again highlighting the importance of intuitive physics, these results raise questions about the learning process of physical knowledge. Another phenomenon is that misconceptions widely exist in explicit reasoning [9]. For the question in Fig. 1, even many adults believe the ball will move in a curved path. However, if researchers provide animated displays, or even when the situation is replaced by water exiting a curved hose, human errors are greatly reduced [5].

The general picture that emerges from previous research is that humans understand and learn about the world *rapidly, automatically, with misconceptions, biases, and generalizing ability*. To reach human-level performance in physical reasoning, a method needs to take all these features into account. Inspired by these phenomena, a variety of frameworks and hypotheses have been proposed. Next, we will discuss some of them and further analyze the reasonableness of an appealing hypothesis, IPE.

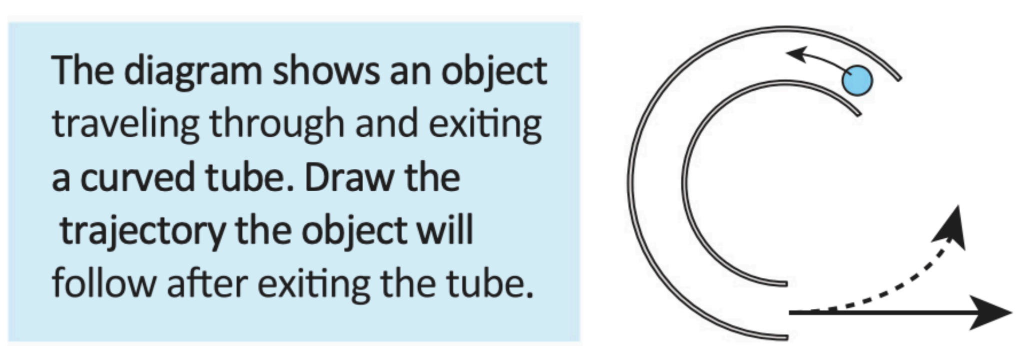

Figure 1: Predict trajectory of the ball. Image credit: Kubricht et al. [6]

2 Approaches to learn intuitive physics

The complexity of humans' physical reasoning process has led to two divergent approaches, which can be broadly categorized into two paths: model-free methods and model-based methods.

2.1 Model-free methods

Heuristic methods: Inspired by our prior knowledge of physics, rule-based models have been proposed. We can design heuristic rules for a certain problem. For example, to solve the problem in Fig. 2, we could set up a series of rules, *e.g.*, if ball 1 bounces back, then $m_2 > m_1$. Formally, experts can design a symbol system according to Newtonian mechanics, and each problem can be calculated via symbolic logic [10]. This method can be extremely fast, which corresponds with the rapidness of human reasoning. However, this is inconsistent with the fact that humans are born with core knowledge of physics, as discussed above. Besides, this method relies heavily on hand-crafted rules, which can hardly be generalized to other tasks. It is also inefficient due to the complexity of rules needed for a simple problem. Besides, it is impossible to code all the rules like Maxwell's demon. Hence, the granularity of rules must be carefully considered.

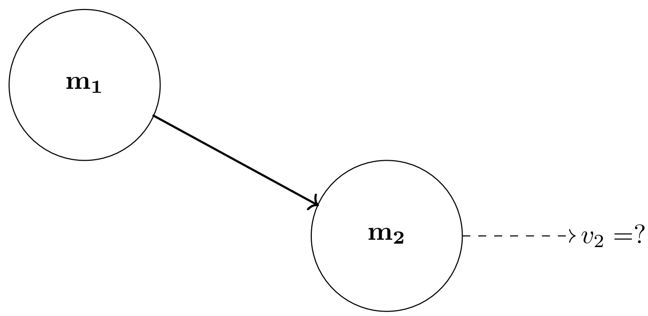

Figure 2: Which object is heavier?

Learning-based methods: Powered by deep learning and new model architectures like transformers [12], learning-based methods have achieved great success in physics reasoning. A major problem for both humans and machines is that the physical attributes of objects are mostly unobservable, like mass, friction, *etc*. Through learning from data, deep neural networks such as CNNs [7] can take an image as input and estimate the physical attributes, which are important for further reasoning.

Neural networks can also be directly applied to predict future dynamics. For example, PhysNet by Lerer et al. [8] has achieved success in predicting trajectories for simplistic block tower scenarios. It is capable of predicting future trajectories of both artificial and real block towers, yet still struggles to generalize to unseen situations. It also requires thousands of examples to learn. These disadvantages contradict the fact that even infants can learn from just a few examples. Other limitations include the diversity of situations and shortage of high-quality data, which are also faced by current deep learning methods.

Another direction is to emulate physical principles through learning. The NeuroAnimator model by Grzeszczuk et al. [4] can learn state transition patterns by viewing transformation examples. This is a promising step towards a learnable physical engine.

2.2 Model-based methods

Someone proposes that an intuitive physics engine exists in our brain. It builds representations of the world and its dynamics, predicting the future in a simulating way. This is supported by brain imaging studies [3] which show some brain regions are more active when making physical inferences than nonphysical inferences. This engine is supposed to perform a task explicitly, which is different from deep learning methods. Also, considering the misconceptions and biases in human reasoning, this physical engine is supposed to be probabilistic and uncertain. In general, this engine has the following properties [14]: 1) It performs physical judgments by running simulation directly; 2) the simulation is stochastic and uncertain. Inspired by this hypothesis, probabilistic simulation methods have been proposed and exceed the performance of model-free methods.

One successful method is the Noisy Newton model proposed by Sanborn et al. [11] and extended by many researchers, e.g. Battaglia et al. [1]. It assumes that people integrate noisy inputs with prior knowledge, and models the simulation with Newtonian mechanics. Predictions are made by thousands of simulations according to these constraints. The uncertainty from noise as well as unobservable variables helps align the model with human behaviors. Besides, uncertainty in the inference stage can also provide a better fit to human performance. Fig. 3 illustrates the pipeline of IPE. When predicting the future state of the block tower, the model performs epochs of simulation. In each simulation, information of perceptual variables sampled according to the data distribution is passed to the model which integrates physical knowledge and utilizes the Monte Carlo method for uncertainty. Then results from each simulation are aggregated to produce a prediction distribution, which is consistent with the distribution of results across humans.

It is also worth noticing that this method can be combined with morden deep learning methods. As we talked above, deep learning methods can be used to infer physical attributes of visual inputs. Hybrid approaches like [13] have had some success in physical prediction like problems in Fig. 2.

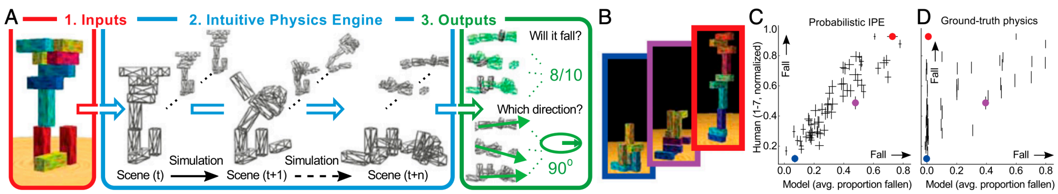

Figure 3: The intuitive physics engine simulation process. Image credit: Battaglia et al. [1]

3 Is IPE satisfying enough?

Though the idea is appealing and has achieved great performance, the disadvantages and limitations are also obvious. Major limitations include but are not limited to:

- Like heuristic methods, IPE also needs to manually feed physical knowledge into the model, limiting its generalization ability.
- It is computationally inefficient. The repeated simulation process is inelegant and inconsistent with the rapidness of human reasoning.
- Learning ability still needs to be integrated into the system.

Besides, this pipeline has some critical flaws. The introduction of noise is indeed useful to model human behaviors, but this only models uncertainty for humans as a whole, not for individuals. The uncertainty for individuals should be represented in prior knowledge, simulation processes guided by cognition, and so on.

The hypothesis itself is also questionable. Despite being able to model future trajectories, there must exist other mechanisms underlying intuitive physics. A good example is a player shooting a basketball. It is not true that we simulate every situation in which the basketball's initial angle and velocity vary. Instead, we do that in an automatic and heuristic way, and use the future state to choose the current state. And IPE's prediction ability is not enough for this situation. We need a kind of "feeling" to hit the basket, which is accumulated through gradual practice. Formally, we may model a latent variable z from past experience and context, thus directly predicting the future state using $p(s|z)$. This variable is a high-level representation of the future and dynamics and greatly reduces computation cost. And we can sample our movement through the Bayesian rule, where s_f represents the desired goal and s is the action we take.

$$z \sim p(s_f|z)p(z), z \sim p(z|s)p(s)$$

We believe a similar procedure exists in our brain, despite the possible "intuitive physical engine." However, our explanation is too idealistic and needs further investigation.

4 Conclusion

In conclusion, a variety of methods have been proposed to simulate human performance in intuitive physics. However, the true mechanisms and intuitions underlying our reasoning, predicting, and inference abilities still need to be further investigated. More inspirations should be drawn from human experiments and neuroscience to shed light on how people make physical judgments and predictions. We believe that a thorough research effort into modeling intuitive physics will greatly benefit the development of more human-like artificial intelligence.

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. 3

[2] Susan Carey. The origin of concepts. *Journal of Cognition and Development*, 1(1):37–41, 2000. 1

[3] Jason Fischer, John G Mikhael, Joshua B Tenenbaum, and Nancy Kanwisher. Functional neuroanatomy of intuitive physical inference. *Proceedings of the national academy of sciences*, 113(34):E5072–E5081, 2016. 2

[4] Radek Grzeszczuk, Demetri Terzopoulos, and Geoffrey Hinton. Neuroanimator: Fast neural network emulation and control of physics-based models. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 9–20, 1998. 2

[5] Mary K Kaiser, Dennis R Proffitt, Susan M Whelan, and Heiko Hecht. Influence of animation on dynamical judgments. *Journal of experimental Psychology: Human Perception and performance*, 18(3):669, 1992. 1

[6] James R Kubricht, Keith J Holyoak, and Hongjing Lu. Intuitive physics: Current research and controversies. *Trends in cognitive sciences*, 21(10):749–759, 2017. 1

[7] Yann LeCun, Yoshua Bengio, et al. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10):1995, 1995. 2

[8] Adam Lerer, Sam Gross, and Rob Fergus. Learning physical intuition of block towers by example. In *International conference on machine learning*, pages 430–438. PMLR, 2016. 2

[9] Michael McCloskey, Alfonso Caramazza, and Bert Green. Curvilinear motion in the absence of external forces: Naive beliefs about the motion of objects. *Science*, 210(4474):1139–1141, 1980. 1

[10] Allen Newell. Physical symbol systems. *Cognitive science*, 4(2):135–183, 1980. 2

[11] Adam N Sanborn, Vikash K Mansinghka, and Thomas L Griffiths. Reconciling intuitive physics and newtonian mechanics for colliding objects. *Psychological review*, 120(2):411, 2013. 3

[12] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017. 2

[13] Jiajun Wu, Ilker Yildirim, Joseph J Lim, Bill Freeman, and Josh Tenenbaum. Galileo: Perceiving physical object properties by integrating a physics engine with deep learning. *Advances in neural information processing systems*, 28, 2015. 3

[14] 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