Lili FAN, Chao GUO, Yonglin TIAN, Hui ZHANG, Jun ZHANG, Fei-Yue WANG, 2024. Sora for foundation robots with parallel intelligence: three world models, three robotic systems. *Frontiers of Information Technology & Electronic Engineering*, 25(7):917-923. https://doi.org/10.1631/FITEE.2400144

Sora for foundation robots with parallel intelligence: three world models, three robotic systems

Key words: Sora; Foundation models; Infrastructure robots; Parallel intelligence; ChatGPT

Corresponding author: Fei-Yue WANG
E-mail: feiyue.wang@ia.ac.cn
ORCID: https://orcid.org/0000-0001-9185-3989

Motivation

- Following the emergence of foundation models such as ChatGPT and Sora, there is a focus on developing technology that leverages these advancements. Sora not only understands the elements of prompts given by users, but also comprehends how these elements interact in the physical world. Several initial reports have been written on Sora; however, the potential impact of foundation models on robotics has not been explored in depth.
- Recent technologies such as foundation models are being used to transform scientific and industrial research into more virtual or data-driven forms. This can be thought of as augmenting traditional methods with artificial counterparts, or introducing digital beings and robots to parallel biological beings. With this in mind, we discuss the potential impact of newly developed foundation models on robotics and automation, and propose a framework by which robotic systems could be automatically trained by such models.

1) Big robotic world models

- We define three types of robots, physical, human, and digital, as a means to describe the integration of AI technologies. Accordingly, we develop the concept of “foundation robot” based on foundation models and parallel intelligence, which could potentially enhance automation of tasks and development of machine intelligence.

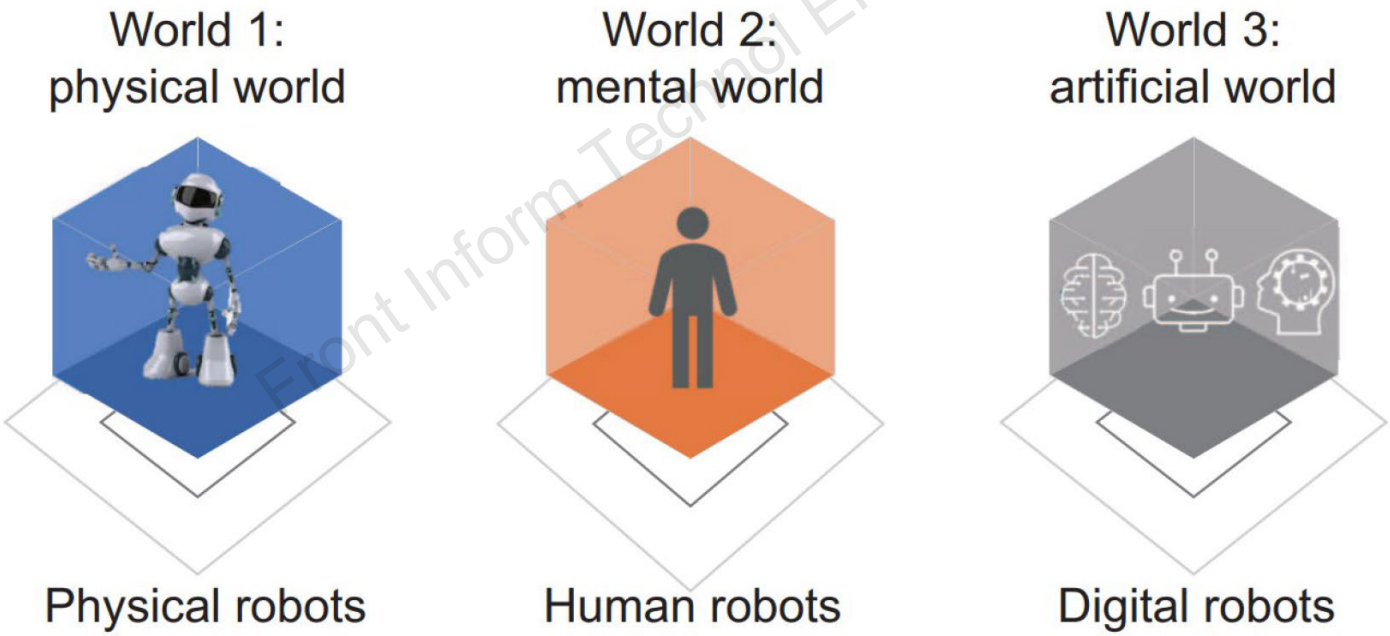

Fig. 1 Three worlds, three types of robots

2) Foundation models for foundation

- Based on a framework of parallel robotics and foundation models, we define this cooperative system as a “foundation robot.”
- Digital and physical robots could operate autonomously for most (say 80%) of the time while providing necessary monitoring information to human robots. In a minority (say 15%) of the time, the system could operate in a parallel mode, with human robots intervening to guide the other robots to resolve exceptions. In exceptional cases (say 5% of the time), an expert/emergency mode could be triggered, with humans directly handling issues until the system returns to a normal autonomous operation.

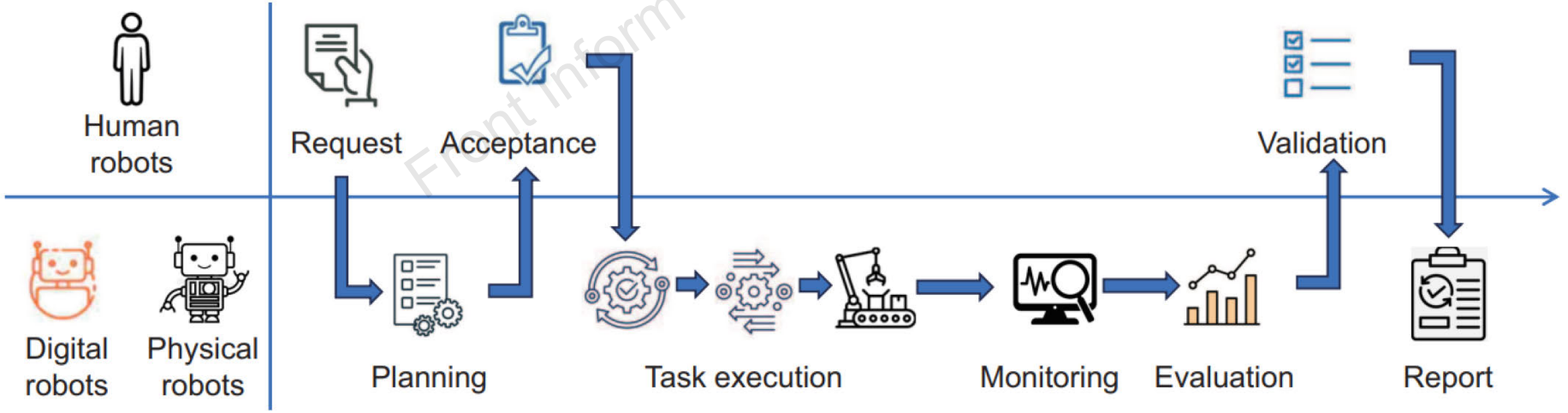

Fig. 2 Framework for foundation robots based on parallel robotics and foundation models

3) Digital robotic schools with curriculum learning

- To improve the learning and generalization abilities of robotic systems, we propose the concept of “digital robotic school.” Fig. 3 shows a proposed framework for a digital robotic school. This school can be thought of as three parts: digital teaching, curriculum learning, and parallel learning.

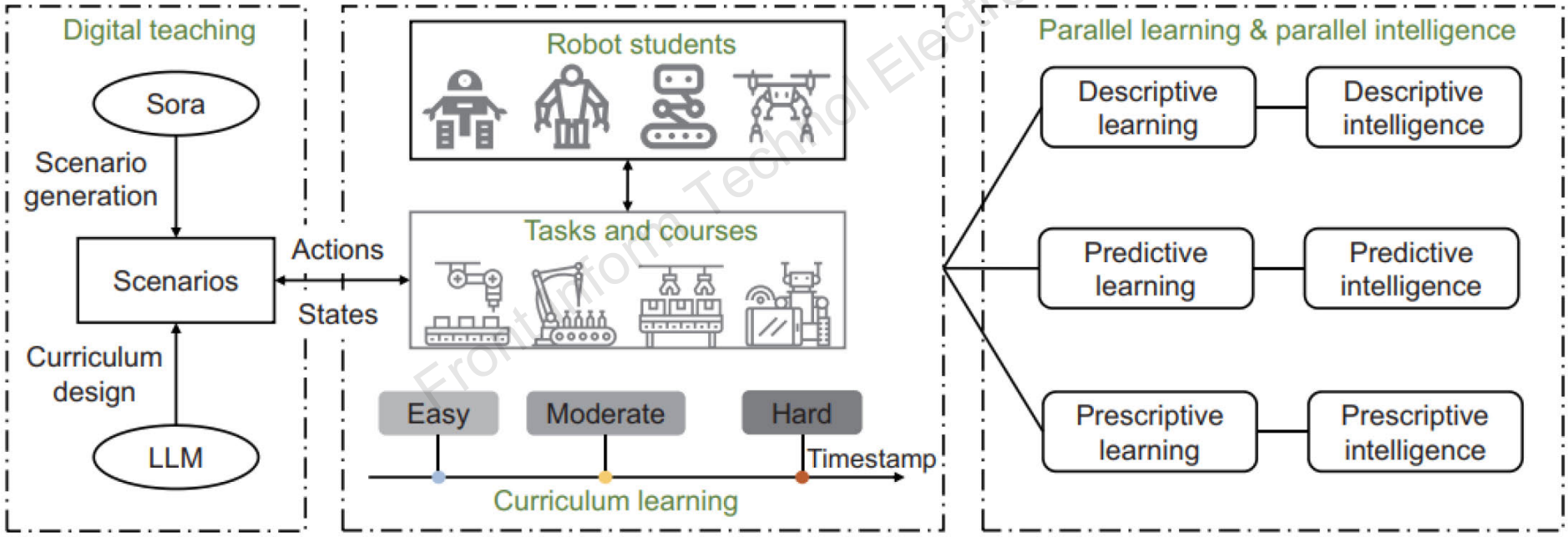

Fig. 3 Proposed framework for a digital robotic school

Conclusions

- The development of foundation robots for specific functions across various fields could lead to more generalized robotic systems. Considering future improvement of computing power and further optimization of algorithmic models, foundation robots will likely possess enhanced learning and reasoning abilities, enabling them to extract more accurate insights from larger-scale data.
- This infrastructure robot could then be used as a basis for further transportation applications. Accordingly, we would likely see infrastructure robots for educational, medical, logistical, and a multitude of service and industrial applications. With increasing advancements in AI technologies, foundation and infrastructure robots could become the new basic machines upon which smart cities and societies are developed.