

# 问答 ChatGPT 之后: 超大预训练模型的机遇和挑战

卢经纬<sup>1,2</sup> 郭超<sup>1</sup> 戴星原<sup>1</sup> 缪青海<sup>3</sup> 王兴霞<sup>1,3</sup> 杨静<sup>1,3</sup> 王飞跃<sup>1,2</sup>

**摘要** 超大预训练模型 (Pre-trained model, PTM) 是人工智能领域近年来迅速崛起的研究方向, 在自然语言处理 (Natural language processing, NLP) 和计算机视觉等多种任务中达到了有史以来的最佳性能, 促进了人工智能生成内容 (Artificial intelligence-generated content, AIGC) 的发展和落地. ChatGPT 作为当下最火热的 PTM, 更是以优异的表现获得各界的广泛关注. 本文围绕 ChatGPT 展开. 首先概括 PTM 的基本思想并对其发展历程进行梳理; 接着, 详细探讨 ChatGPT 的技术细节, 并以平行智能的视角阐述 ChatGPT; 最后, 从技术、范式以及应用等多个方面对 PTM 的发展趋势进行展望.

**关键词** 预训练模型, ChatGPT, Transformer, 人工智能生成内容, 平行智能, 社会化大闭环

**引用格式** 卢经纬, 郭超, 戴星原, 缪青海, 王兴霞, 杨静, 王飞跃. 问答 ChatGPT 之后: 超大预训练模型的机遇和挑战. 自动化学报, 2023, 49(4): 705–717

**DOI** 10.16383/j.aas.c230107

## The ChatGPT After: Opportunities and Challenges of Very Large Scale Pre-trained Models

LU Jing-Wei<sup>1,2</sup> GUO Chao<sup>1</sup> DAI Xing-Yuan<sup>1</sup> MIAO Qing-Hai<sup>3</sup>  
WANG Xing-Xia<sup>1,3</sup> YANG Jing<sup>1,3</sup> WANG Fei-Yue<sup>1,2</sup>

**Abstract** In recent years, very large scale pre-trained models (PTMs) have become a rapidly rising research direction in artificial intelligence, achieving state of the art in most tasks, especially natural language processing (NLP) and computer vision, and speeding up the development and implementation of artificial intelligence-generated content (AIGC). ChatGPT, as the hottest PTM, has been brought to the fore on account of its excellent performance. This paper is organized around ChatGPT. First, we outline the basic idea of PTM and review its development history. Then, the technical details of ChatGPT are explored, and ChatGPT is revisited from the perspective of parallel intelligence. Finally, the development trends of PTMs are presented in terms of technologies, paradigms, and applications.

**Key words** Pre-trained models (PTMs), ChatGPT, Transformer, artificial intelligence-generated content (AIGC), parallel intelligence, grand socialization closed-loop

**Citation** Lu Jing-Wei, Guo Chao, Dai Xing-Yuan, Miao Qing-Hai, Wang Xing-Xia, Yang Jing, Wang Fei-Yue. The ChatGPT after: Opportunities and challenges of very large scale pre-trained models. *Acta Automatica Sinica*, 2023, 49(4): 705–717

2022 年底, 一款名为 ChatGPT (Chat generative pre-trained Transformer) 的人工智能聊天程

序引爆全球, 成为了各行业关注的焦点. ChatGPT 火热之余, 不禁让众人思考: ChatGPT 有何奥秘? “ChatGPT 是一种基于人工智能的自然语言处理技术, 能够根据给定的问题和语境生成符合语言习惯和逻辑的回答. 这种技术可以应用于各种领域, 如客服、智能助手、教育、医疗等, 能够帮助人们更方便、高效地获取所需信息. 当前, ChatGPT 在各个领域的应用越来越广泛, 因为它可以实现智能化、个性化、高效化的服务, 得到了广大用户的认可和喜爱. 随着技术的不断进步和完善, ChatGPT 在处理语言难度、语境理解、知识获取和表达等方面都得到了不断的提升, 使得它在应用领域的效果和准确度都越来越好. 因此, 当下 ChatGPT 的火热现象主要是因为它能够提供高效便捷的智能服务, 并

收稿日期 2023-03-05 录用日期 2023-03-27

Manuscript received March 5, 2023; accepted March 27, 2023

国家自然科学基金 (U1811463), 行动元联合研究项目: 伺服驱动系统的基础建模和平行驱控研究资助

Supported by National Natural Science Foundation of China (U1811463) and Motion G, Inc. Collaborative Research Project for Foundation Modeling and Parallel Driven/Control for Servo-Drive Systems

本文责任编辑 袁勇

Recommended by Associate Editor YUAN Yong

1. 中国科学院自动化研究所复杂系统管理与控制国家重点实验室 北京 100190 2. 青岛智能产业技术研究院 青岛 266114 3. 中国科学院大学人工智能学院 北京 100049

1. The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190 2. Qingdao Academy of Intelligent Industries, Qingdao 266114 3. School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049

且在技术上不断突破和进步,能够越来越较好地满足人们的需求.同时,随着 ChatGPT 技术的发展,它还将继续在各个领域发挥重要作用,成为人们生活中不可或缺的一部分.”<sup>1</sup>就这一问题,ChatGPT 自己如是说道.由 ChatGPT 的回答,我们不难发现 ChatGPT 的论述条理清晰、井然有序,更难得的是符合大众日常的语言习惯,这让该问题的答案不言而喻.ChatGPT 的横空出世势必会给各行业带来巨大的冲击和挑战,同时也意味着新机遇的诞生.因此,人们需对 ChatGPT 有全面且深刻的认识.

ChatGPT 属于自然语言处理 (Natural language processing, NLP) 领域的超大预训练模型 (Pre-trained model, PTM)<sup>[1-2]</sup>,这类超大 PTM 也称基石模型<sup>[3-4]</sup>、大模型<sup>[5-7]</sup>.简单来说,超大 PTM 旨在设计具有超大参数量的深度神经网络 (Deep neural network, DNN),在海量未标记的数据上对其进行训练.利用超大参数量 DNN 强大的函数近似能力,预训练可使超大 PTM 在这些海量数据上提取共性特征.再根据下游任务,对超大 PTM 进行微调或上下文学习,使最终的模型可在具有一定相关度但不同的任务中获得优异的表现.目前,国内外众多科研机构、公司研发的超大 PTM 已在各领域取得了巨大的突破,引领了新一轮的人工智能科技竞赛.

为进一步推进以 ChatGPT 为代表的超大 PTM 技术的发展和落地,加速人工智能生成内容 (Artificial intelligence-generated content, AIGC) 落地,本文首先梳理了超大 PTM 的经典模型,并进行简

要介绍.其次,详细地介绍 ChatGPT 中的关键技术——Transformer,探讨 ChatGPT 的设计与实现,同时以平行智能的视角解读 ChatGPT.在综合分析 ChatGPT 和其他 PTM 的基础上,我们进一步从技术、生态、范式以及应用等多个方面探讨超大 PTM 的发展趋势.

## 1 超大预训练模型概述

本节简要梳理超大 PTM 的经典模型.为方便起见,下文 PTM 均指超大 PTM.

### 1.1 超大预训练模型

目前,PTM 的发展呈现出大数据驱动、小样本或零样本领域适应、跨模态关联的趋势,近年来的典型 PTM 的发展历程如图 1 所示.首先,与特定任务的模型相比,训练 PTM 所需的数据量要大得多,对并行计算的软硬件性能提出了极高的要求.其次,由于 PTM 通常基于大规模数据进行训练,能够掌握相关领域内的通用知识,因此表现出容易适应多种下游任务的能力,往往只需要基于少量样本进行微调或者通过提示的方式实现零样本迁移.最后,PTM 逐渐呈现出跨模态关联的趋势.随着语言 PTM 的成功应用,PTM 的方案逐步迁移应用于其他模态,从而建立包括文本、图像、语音、序列数据在内的跨模态通用特征表示,获取跨模态知识关联,提升模型的泛化能力和对任务的通用性.尤其是在视觉-语言模型方面,完成了由文本描述生成高质量图像并实现了大规模商用.此外,由于 PTM 训练所需的数据量远超出传统监督学习所需标注数据量

<sup>1</sup> 引号内文字由 ChatGPT 生成 (<https://chat.openai.com/chat/>)

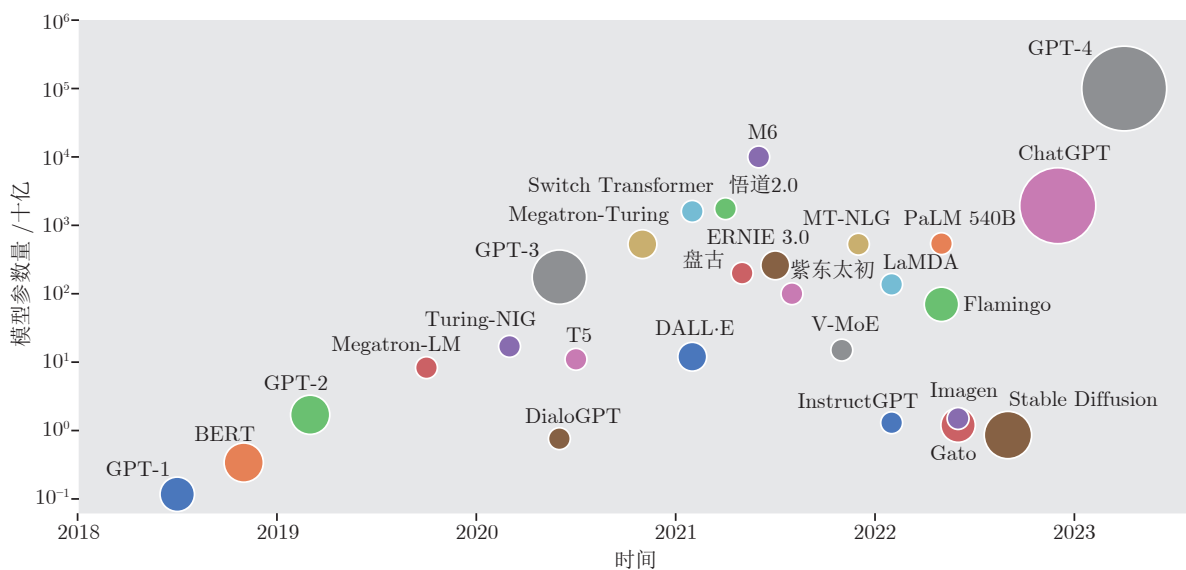


图 1 典型超大预训练模型的发展历程

Fig.1 Typical very large scale PTMs

的范围, 因此 PTM 往往通过自监督学习的方式获得强大的表征能力. 目前 PTM 已经在 AIGC 得到了大规模应用, 一般根据生成内容的数据形态包括文本、图像、音频、视频、3D 模型等. 典型的应用案例包括 ChatGPT 用于文本生成, DALL·E 2<sup>[8]</sup> 和 Stable Diffusion<sup>[9]</sup> 用于绘画和设计图像生成等. 这些生成模型使得普通大众能够借助人工智能的力量产生符合自身需求的内容, 大幅提升了工作效率.

将 PTM 按照处理的数据模态可以分为语言模型、视觉模型、语音模型、多模态模型等, 其中多模态模型一般指建立跨两种或以上数据模态特征表示的模型, 例如视觉-语言模型等. 在多种模态任务中, 目前最高水平 (State of the art, SOTA) 的 PTM 几乎都采用 Transformer 的模型架构实现<sup>[2-3, 10]</sup>.

## 1.2 语言超大预训练模型

PTM 首先在 NLP 领域取得突破性进展, 一定程度上得益于大规模文本数据的易得性. 目前, 自然语言 PTM 在诸如文本生成、问答、阅读理解、逻辑推理等领域取得了满意的效果. 其中代表性的模型有 ELMo (Embeddings from language models)<sup>[11]</sup>、BERT (Bidirectional encoder representations from Transformers)<sup>[12]</sup>、GPT (Generative pre-trained Transformer)<sup>[13-15]</sup> 等. 根据预训练方式, 主流模型主要有自编码、自回归和编码器-解码器模型等类型. 自编码模型通过上下文预测遮蔽数据, 多用于文本理解和分类. 而自回归模型主要通过上文预测当前词汇进行学习, 与编码器-解码器模型类似具有序列生成能力, 多用于文本生成. 但自回归的方式更难训练, 通常需要更大规模的训练数据.

BERT 是由 Google 提出的自编码训练方式中遮蔽语言模型的典型代表, 通过构建对遮蔽数据的预测进行模型参数优化. BERT 为基于 Transformer 的双向语言 PTM, 能够利用当前词的上下文信息进行特征提取, 并通过不断地进行遮蔽词预测进行训练, 学习上下文、语法结构等较为全面的文本特征. 该模型可在微调后显著提升文本分类、抽取式问答、文字标注等下游语言任务的性能, 推动了语言 PTM 的广泛应用. 此后, 大量研究从预训练任务、模型结构、遮蔽方式、模型蒸馏等方向对其进行探索和改进, 不断提升在各类语言任务中的性能.

GPT 是 OpenAI 提出的自回归训练方式中因果语言模型的代表, 通过预测下一位置数据进行模型参数优化. 在 GPT-1 中, 首先在无标签数据上进行无监督学习, 之后在文本分类、问答等有监督任务中进行模型微调. GPT-2 在 GPT-1 的基础上, 采

用了更大的模型参数规模和训练数据, 提升了该模型的表征泛化能力, 并表明了基于超大模型参数和海量数据所训练的词向量模型能有效迁移到相关任务中. GPT-3 更进一步大幅提升了模型容量, 并得益于上下文学习, 在常规绝大多数零样本或小样本学习任务中实现了 SOTA 性能, 并在部分任务超越了微调后 PTM 的性能. 本文主要关注的 ChatGPT, 则以 GPT-3.5 为基础, 并通过人类反馈的强化学习 (Reinforcement learning from human feedback, RLHF) 实现高性能的对话问答<sup>2</sup>.

基于 BERT 和 GPT 有很多改进的衍生模型, 这里不再赘述. 此外, T5<sup>[16]</sup>、Switch Transformer<sup>[17]</sup> 等模型也分别在 NLP 相关任务上取得了不错的效果, 极大地推动了语言 PTM 的发展.

## 1.3 多模态超大预训练模型

相比语言 PTM, 多模态 PTM 输入数据由单一的文本延伸到了包括文本、图像、音频、视频、数值序列等多模态数据形式. 典型的多模态 PTM 有 ViLBERT<sup>[18]</sup> (Vision-and-language BERT)、CLIP (Contrastive language-image pre-training)<sup>[19]</sup>、Flamingo<sup>[20]</sup> 等视觉-语言模型和 Gato<sup>[21]</sup> 决策大模型. 同时, 诸多国内团队和公司在研发核心多模态大模型的基础上, 构建了对应的人工智能服务平台和产品生态, 逐步实现多模态 PTM 的大规模应用. 典型代表有华为研发的盘古<sup>[22]</sup>、智源人工智能研究院研发的悟道<sup>[23]</sup>、阿里巴巴研发的 M6<sup>[24]</sup>、中科院自动化所研发的紫东太初<sup>[25]</sup>、基于 ERNIE 的百度文心<sup>[26-27]</sup> 等.

CLIP 是由 OpenAI 提出的基于对比学习的图像-文本 PTM. 该模型由文本编码器和图像编码器构成, 使用互联网上大量存在的图像文本数据对进行训练. CLIP 能够学习到针对图文数据统一的隐空间表示, 具备了优秀的图像、文本多模态表征和关联能力, 因此被广泛用于构建文本和图像之间的关联约束, 或者通过计算文本与图像的相似度对下游任务进行零样本推理. 例如, 将 CLIP 中的编码器用于下游任务的图像和文字编码, DALL·E 2 实现了惊艳的文生图效果.

Flamingo 由 DeepMind 提出, 在文本补全、视觉问答等任务上呈现出了强大的上下文学习能力. 该模型通过视觉编码器提取视觉特征, 并经由重采样嵌入到 ChinChilla 自回归语言模型<sup>[28]</sup>, 实现图像、视频等视觉信息与文本等语言信息的交叉融合, 并充分利用语言模型的推理能力. 这样, 不仅可以通

<sup>2</sup> <https://openai.com/blog/chatgpt/>

过输入的语言提示来区分不同任务,而且可以同时交叉处理图文信息.此外,通过上下文学习的方式,Flamingo 根据给定的少量任务实例理解任务,并给出问题的答案,表现出良好的模型迁移能力.

Gato 由 DeepMind 提出,实现了包含连续控制(机器人等)、离散控制(Atari 游戏等)、自然语言对话、图像描述生成等在内的多种任务.该模型将文本、图像、离散与连续数据分别编码到统一的特征空间,采用监督学习的方式训练. Gato 同样通过语言提示进行任务区分,并通过预留的输入位置融合文本外其他各模态信息进行编码,不再需要模型微调即可适应下游任务.该模型主要针对强化学习(Reinforcement learning, RL) 决策问题,将现有 RL 中 SOTA 算法的结果和图文数据作为训练数据进行学习,并没有在单一任务上显著提高性能,但大幅提升了模型的通用性,是对通用人工智能的尝试.

## 2 ChatGPT

本节将详细分析 ChatGPT. 首先概述了 ChatGPT, 其次介绍了 ChatGPT 核心技术 Transformer 的原理, 然后根据相关资料和技术原理探讨了 ChatGPT 可能的实现途径, 并以平行智能 Hanoi (Integrated human, artificial, natural, and organizational intelligence) 的视角解读了 ChatGPT<sup>[29-35]</sup>.

### 2.1 ChatGPT 概述

ChatGPT 主要以文字方式交互, 可以和人类进行自然对话, 还可以从事复杂的语言工作, 包括自动文本生成、自动问答等多种任务. 在自动文本生成方面, 可根据用户要求生成论文、代码、剧本、企划等; 在自动问答方面, ChatGPT 可自动理解用户问题, 并给出相应答案. ChatGPT 的知识范围覆盖了自然科学、人文社科、社会科学、文体娱乐以及奇闻趣事等等, 同时能掌握用户的语气, 正确分析文本情感, 并联系语境进行交流. 图 2 简要展示了 ChatGPT 的若干功能. 尽管在一些专业知识方面, ChatGPT 仍会一本正经地说些不着边际的话, 但却提供了堪比真人的交流对话体验, 是之前自然语言 PTM 所无法企及的.

PTM 在预训练阶段会遇见各类数据, 可能是人类交流的数据, 也可能是机器所产生的数据, 当然也无法避免一些错误的表述. 因此就用户具体的提问, PTM 的输出是难以确定的. 预训练阶段中的数据模式并非都符合人类交流的语言习惯, 因此需规范 PTM 的“言行举止”. 规范 PTM“言行举止”的

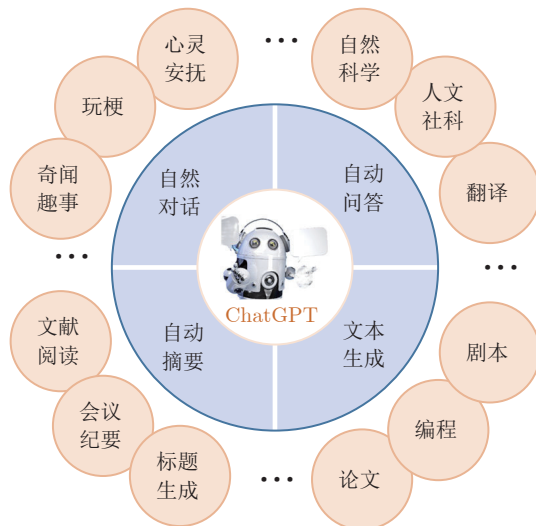


图 2 ChatGPT 的功能

Fig.2 The functions of ChatGPT

过程即为人工智能对齐, 也可称为社会化, 是 AI 技术社会化的积极探索. 需特别说明的是, 目前 OpenAI 尚未发表关于 ChatGPT 的论文或开源 ChatGPT 的代码, 本文对 ChatGPT 技术细节的探讨主要参考 ChatGPT 的官方博客<sup>3</sup>、InstructGPT<sup>[36]</sup> 和相关技术理论.

### 2.2 ChatGPT 关键技术: Transformer

基本的 Transformer 包含编码器与解码器两部分<sup>[10, 37]</sup>, GPT 采用了解码器, 其可视为一个自回归模型, 基于输入序列逐步生成输出序列, 即在每一步预测输入序列的下一个元素, 并将该元素作为附加输入以得到下一步预测. 在具体实现上, GPT 采用的 Transformer 解码器中输入与输出序列等长, 其结构如图 3 所示, 这种结构使得模型能够实现大规模预训练并在测试阶段生成变长的文本序列.

在预训练阶段, 对于一段长度为  $(L + 1)$  的文本序列  $[x_1, \dots, x_{(L+1)}]$ , GPT 输入前  $L$  个元素构成的序列  $\mathbf{x} = [x_1, \dots, x_L]$ , 预测  $\hat{\mathbf{x}} = [x_2, \dots, x_{(L+1)}]$  的概率. 需要说明的是, 对于序列  $\hat{\mathbf{x}}$  的预测为一次性输出, GPT 使用掩码注意力层使得其在预测第  $l$  个元素  $x_l$  的过程中不会使用  $x_l$  之后的元素信息. 本节以词为序列元素, 介绍 GPT 的运算过程.

首先, 输入序列  $\mathbf{x}$  中的每个词符  $x_l$  经过学习的文本嵌入层被编码为词向量, 每个词向量与词的位置编码向量相加, 进而得到词的最终编码  $\mathbf{y}_l \in \mathbf{R}^d$ . 位置编码描述了词在序列中的位置信息, 其引入动机是为了弥补 Transformer 的注意力机制没有显式

<sup>3</sup> <https://openai.com/blog/chatgpt/>



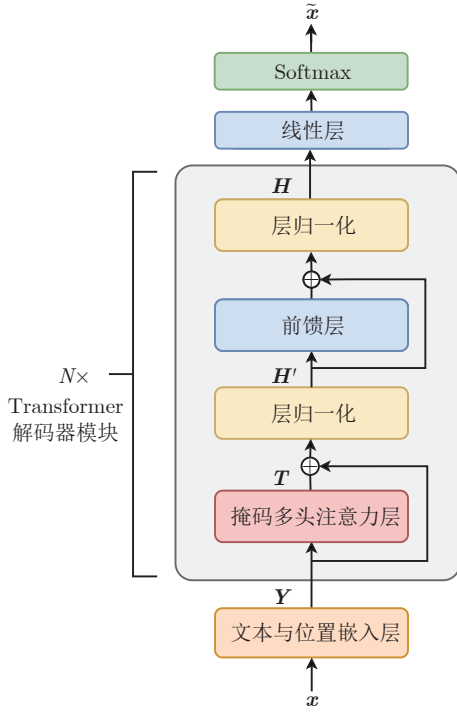


图 3 ChatGPT 采用的 Transformer 解码器结构

Fig.3 The Transformer decoder for ChatGPT

地考虑序列中词之间顺序关系的问题, 进而使得 Transformer 能够更好地捕捉输入序列中的局部与全局信息. 经过文本与位置编码, 输入序列  $x$  被转化为  $L$  个包含位置信息的词嵌入构成的矩阵  $Y = [y_1, \dots, y_L] \in \mathbf{R}^{d \times L}$ . 然后,  $L$  个词嵌入经过多个 Transformer 解码器模块得到文本的高层表征. 每个解码器模块包含掩码多头注意力层、前馈层与层归一化操作, 以下以文本与位置嵌入层相连的解码器模块为例, 介绍 Transformer 解码器的运算过程.

文本嵌入  $Y$  被输入到掩码多头注意力层, 这也是 Transformer 的核心结构. 掩码多头注意力层包含  $H$  个可并行执行的掩码注意力函数, 每个函数关注不同子空间的注意力权重, 进而提升了模型的表征能力与泛化能力. 对于第  $h$  个掩码注意力函数, 其首先将  $Y$  通过线性变换转化为三个维度均为  $(d/H) \times L$  的矩阵: 查询矩阵  $Q^{(h)}$ , 键矩阵  $K^{(h)}$  与值矩阵  $V^{(h)}$ . 然后, 基于掩码点积注意力函数得到考虑词之间关联信息的文本表征

$$Z^{(h)} = \left[ z_1^{(h)}, \dots, z_L^{(h)} \right] = \text{MaskedAttentionHead}(Q^{(h)}, K^{(h)}, V^{(h)}) = V^{(h)} \cdot \text{softmax} \left( \frac{1}{\sqrt{p}} \left( Q^{(h)} \right)^T K^{(h)} + M \right) \quad (1)$$

其中, 掩码矩阵  $M \in \mathbf{R}^{L \times L}$  通过将当前位置之后的信息掩盖, 使得模型只能关注到之前生成的部分序列, 其表示为

$$M(l, m) := \begin{cases} 0, & m \leq l \\ -\infty, & m > l \end{cases} \quad (2)$$

设当前词在序列中的序号为  $l$ , 则其之前的词  $x_{m_1}$  ( $m_1 \leq l$ ) 对当前词  $x_l$  的掩码值为 0, 这不会影响基于查询与键矩阵计算的词  $x_l$  对词  $x_m$  注意力权重. 与之相反, 当前词之后的词  $x_{m_2}$  ( $m_2 > l$ ) 对  $x_l$  的掩码值为  $-\infty$ , 这会使得  $x_l$  对  $x_{m_2}$  的注意力权重为 0, 因此在预测下一个词时不会考虑  $x_{m_2}$  信息.

经过掩码自注意力层输出的  $H$  个注意力值矩阵  $\{Z^{(h)} \in \mathbf{R}^{(d/H) \times L}\}_{h=1}^H$ , 经过合并与线性变换得到矩阵  $T \in \mathbf{R}^{d \times L}$ , 然后采用残差连接将  $T$  与  $Y$  相加并进行层归一化操作. 之后, 再经过一次基于前馈层的残差操作得到一个 Transformer 解码器模块的输出  $H$ . 综上, 一次 Transformer 解码器模块的运算可描述为

$$\begin{cases} H' = \text{LayerNorm}(\text{MaskedAttention}(Y) + Y) \\ H = \text{LayerNorm}(\text{FFN}(H') + H') \end{cases} \quad (3)$$

GPT 使用多个层叠的 Transformer 解码器模块提取文本的高层表征, 最后一个解码器模块的输出经过线性层与 Softmax 函数将得到对  $\hat{x}$  的预测.

### 2.3 ChatGPT 的设计与实现

如前所述, ChatGPT 的出色表现得益于其成功地引入了人类的价值偏好. 不同于其他 PTM, ChatGPT 采用 RLHF 的方式将人类的语言习惯引入模型中, ChatGPT 实现的基本流程如图 4 所示, 可大致分为如下 4 步:

**步骤 0: 预训练 GPT.** 基于大规模语料库, 以自监督学习的方式预训练 GPT 模型. 使 GPT 在大规模语料库上提取自然语言共性特征.

**步骤 1: 监督微调 GPT.** 基于真人标注偏好的答案初步引入真人价值偏好, 根据人工示例监督微调 GPT.

**步骤 2: 奖励模型设计.** 基于真人对模型输出排序的数据, 监督训练获得奖励模型, 使奖励模型学习到真人的价值偏好.

**步骤 3: RL 反馈优化 GPT.** 基于奖励模型并采用近端策略优化 (Proximal policy optimization, PPO) 算法<sup>[38]</sup>, 闭环反馈优化监督微调后的 GPT, 获得 ChatGPT.

经步骤 0 后的 GPT, 在具体任务上表现并不一定出色, 但已具备相当潜力, 通过微调或者上下

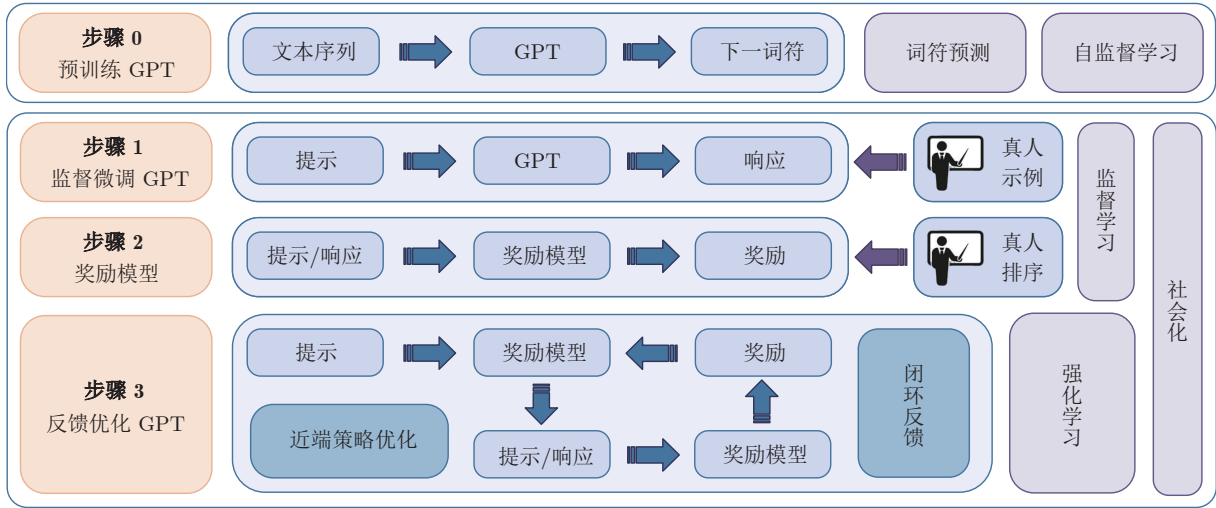


图 4 ChatGPT 的实现流程

Fig.4 The implementation process of ChatGPT

文学习的模式即可在多种任务中获得优异表现. 步骤 0 的介绍可参考第 2.2 节. 而步骤 1 至步骤 3 是 ChatGPT 的关键步骤, 这些步骤成功地将人类因素引入到了 GPT 中. 换一个角度来看, 步骤 1 至步骤 3 也是 RL 的标准流程.

RL 常用于处理多级决策问题, 其研究一般基于马尔科夫决策过程 (Markov decision process, MDP)<sup>[39]</sup>. 典型的 MDP 可由如下五元组构成:  $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma \rangle$ . 其中,  $\mathcal{S}$  为状态  $s$  的集合;  $\mathcal{A}$  为动作  $a$  的集合;  $\mathcal{P}: \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{S}$  为状态转移概率, 即状态  $s \in \mathcal{S}$  在给定动作  $a \in \mathcal{A}$  下, 转移到状态  $s' \in \mathcal{S}$  的概率;  $\mathcal{R}: \mathcal{S} \times \mathcal{A} \rightarrow \mathbf{R}$  是由状态和动作确定的奖励  $r$  的集合;  $\gamma \in [0, 1]$  是为长期回报的折扣因子. RL 的目标是使智能体学习到合适的策略  $a = \pi(s)$ , 极大化如下累积奖励:  $G = \sum_{k=0}^N \gamma^k r_k$ , 其中  $r_k$  为决策过程中第  $k$  步获得的奖励,  $N$  为最大步长, 若无限时域问题, 则  $N$  应为  $\infty$ . 那么在 ChatGPT 的实现过程中, ChatGPT 可视为智能体, 提示可视为状态  $s$ , ChatGPT 的响应可视为动作  $a$ , 奖励  $r$  由奖励模型给出, 对话环境决定了状态转移概率  $\mathcal{P}$ . 图 5 给出了 RL 视角下实现和训练阶段的 ChatGPT. 之所以称为实现和训练阶段的 ChatGPT, 是因为 ChatGPT 采用了迭代部署, 包括初步部署、对齐、评估和迭代部署、部署和持续评估以及下游评价. 上述各步环环相扣, 实际部署后也会根据反馈不断调整以应对各方面的问题<sup>4</sup>. 因此, ChatGPT 本质上构建了社会化大闭环, 这和我们之前论述的 PTM 研发思路一致<sup>[40]</sup>, 是 Hanoi 框架下 PTM 设计与实现的标准流程<sup>[31]</sup>. 我们首先介绍步骤 1 至步骤 3, 并在

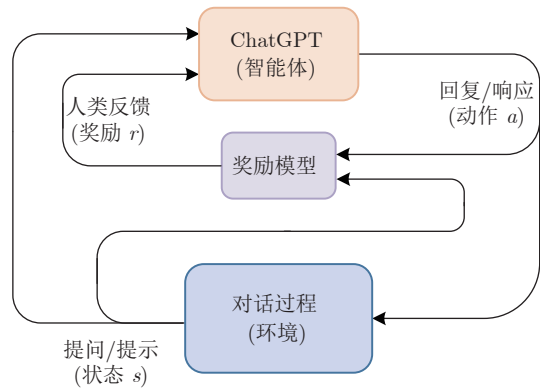


图 5 强化学习视角下的 ChatGPT

Fig.5 ChatGPT from the perspective of RL

第 2.4 节讨论 Hanoi 视角下的 ChatGPT.

首先是步骤 1 (监督微调 GPT). 该步应视为模仿学习, 根据人类给出的示例, 监督微调 GPT. GPT 的响应 (动作) 由策略函数产生, 该策略函数可称为对话策略. GPT 对话策略由 DNN 近似, 即  $a = \pi(s; \theta_a)$ , 其中  $\theta_a$  为 DNN 参数. 该步的目的是, 根据真人示例数据集  $\mathcal{D}_h$ , 监督微调  $\theta_a$  获得一个初步具有真人对话体验感的对话策略  $\pi^{\text{SFT}}(\cdot)$ .

一般在复杂 RL 问题中, 通常首先会根据专业的示教数据对初始策略进行监督训练, 获得一个具有一定能力但并不一定表现出色的初始策略. 例如在 DeepMind 研发的 AlphaGo, AlphaGo 首先基于 KGS 围棋的专业棋谱训练其初始策略<sup>[41]</sup>, 再采用自我博弈的方式进行反馈优化. 这样做的原因可归纳为如下两点: 1) 忽略这些已经高度提炼的专业知识十分浪费; 2) 复杂 RL 任务中策略提升效率低,

<sup>4</sup> <https://openai.com/research/language-model-safety-and-misuse>

基于一个完全随机的初始策略开始学习, 很难快速且有效地习得一个合适的策略。

其次是步骤 2 (奖励模型设计). 在正式进入反馈优化前还需考虑的一个重要的问题——奖励  $r$  的设计. 在 RL 任务中, 奖励设计直接影响学习结果<sup>[39]</sup>. ChatGPT 也不例外, 只有获得一个足够符合人类语言习惯的评价标准 (奖励模型), 才能确保 ChatGPT 的表现. 该步的实现方式为, 首先用 DNN 构建奖励模型  $r = R(s, a; \theta_r) \in \mathbf{R}$ , 其中  $\theta_r$  为 DNN 参数; 再基于真人偏好数据训练参数  $\theta_r$ , 使奖励模型习得真人价值偏好. 根据 InstructGPT, 奖励模型的损失函数为

$$\mathcal{L}(\theta_r) = -\frac{1}{\binom{K}{2}} \mathbb{E}_{(s, a_w, a_l) \sim \mathcal{D}_r} \left[ \sigma(R(s, a_w; \theta_r) - R(s, a_l; \theta_r)) \right] \quad (4)$$

其中,  $a_w$  是同一个提示的两个响应  $a_w$  和  $a_l$  中人类更偏好的一个,  $\mathcal{D}_r$  是人类偏好的数据集.

最后是步骤 3 (RL 反馈优化 GPT). 该步采用 PPO 算法优化监督微调后的 GPT. PPO 是由 OpenAI 提出的一种基于策略的 RL 算法. 相比于以深度 Q 网络为代表的基于值的 RL 方法<sup>[41]</sup>, 基于策略的 RL 具有良好的收敛性和训练稳定性, 因此在优化大规模网络参数时具有一定优势. 该步通过极大化如下联合目标函数调整对话响应策略的 DNN 参数

$$\phi(\theta_a) = \mathbb{E}_{(s, a) \sim \mathcal{D}_\pi} \left[ R(s, a; \theta_r) - \beta \log \left( \frac{\pi(a|s; \theta_a)}{\pi_{\text{SFT}}(a|s)} \right) \right] + \gamma \mathbb{E}_{s \sim \mathcal{D}_{\text{pretrain}}} [\log(\pi(s; \theta_a))] \quad (5)$$

其中,  $\beta$  是  $K$ - $L$  奖励系数,  $\gamma$  是预训练损失系数,  $\mathcal{D}_{\text{pretrain}}$  是预训练分布.

综上, RLHF 的目标是, 在任意提示下, 使 ChatGPT 的响应尽可能和真人的响应一致, 即 ChatGPT 的对话策略尽可能地逼近真人的对话策略. 由于通过奖励模型引入了人类反馈, 该技术因而得名 RLHF. 另外, ChatGPT 中 MDP 关于动作的设置, 我们认为有两种可能: 1) 按 InstructGPT 的说法, 步骤 3 的 RL 处理的是赌博机问题, 那么 GPT 的整个回复应视为一个动作, 优化一次问答可看作一个单步优化问题. 2) 如想实现更精细化的模型优化, 可将 GPT 输出的单个词符作为一个动作, 那么该优化问题就是一个多步优化问题, 可结合 RL 中的稀疏奖励进行优化. 类似地, 还可以按句子、段落等定义动作. 需说明的是, 若按后者方式进行学习, 势必会提升学习难度. 若优化技术和计算资源无法支持, 并不能取得比前者更好的效果.

## 2.4 Hanoi: ChatGPT 的成功之道

正如前文所述, ChatGPT 采用了迭代部署的方式, 本质上构建了社会化大闭环, 是 Hanoi 在 PTM 设计与实现中的一次成功示范<sup>[31, 40, 42-43]</sup>. 图 6 给出了 Hanoi 框架下的 ChatGPT 全流程示意图.

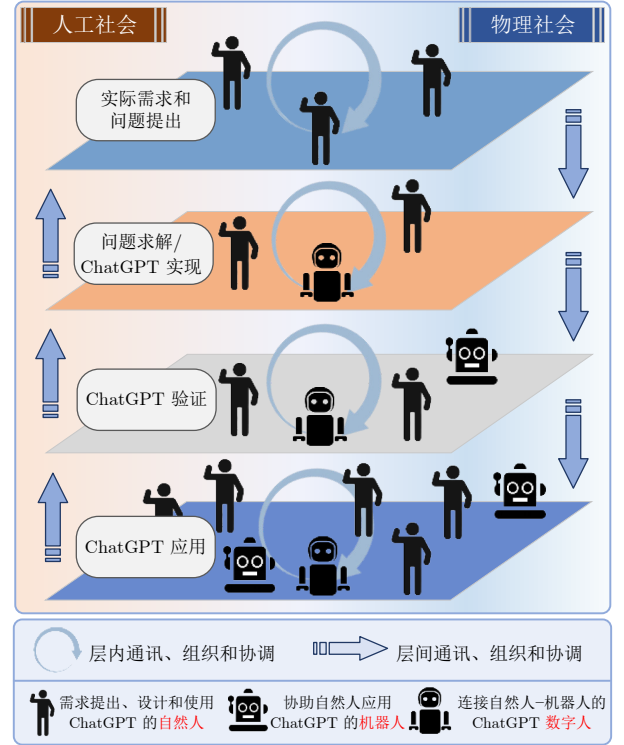


图 6 社会化大闭环下的 ChatGPT

Fig.6 ChatGPT in the grand socialization closed-loop

ChatGPT 的设计、实现与持续优化是一个典型人机混合的过程, 可以大致分为: 实际需求和问题提出、问题求解/ChatGPT 实现、ChatGPT 验证、ChatGPT 应用. 参与其中的人包括: 需求提出、设计和使用 ChatGPT 的自然人, 协助自然人应用 ChatGPT 的机器人以及连接自然人-机器人的 ChatGPT 数字人. 自然人按实际需求提出问题并寻求答案; 机器人协助自然人在物理社会完成繁杂的任务; 数字人则是自然人和机器人在虚拟空间中的映射, 完成人机交互、任务协调、计算实验等功能, 拓展对解决方案的探索空间. 三类人基于 DAO (全中心化自主组织及全中心化自主运行) 的框架进行通信、组织和协调<sup>[42]</sup>, “层”与“层”之间分工明确而又环环相扣, “人”与“人”之间工作独立而又紧密相连, 人工社会和物理社会的各要素通过数字人实现虚实交互, 协同推进完成任务, 形成了分布式、去中心化、自主性、自动化、组织化与有序化的工作框架, 构成了人在回路的社会化大闭环.



Hanoi 框架能够为 ChatGPT 的改进提供有力的支撑. 首先, 针对 ChatGPT 自身存在编造的问题, Hanoi 可从三个要素入手, 一是“a”, 也就是基础大模型, 在语言智能范畴内, 基础模型是对现实世界“n”的逼近; 二是通过“a”引入人的知识, 对回答进行约束; 三是通过“o”将人有效引入学习闭环, 建立更有效的指示和引导, 避免违背法律和伦理的有毒输出. 其次, Hanoi 基于 DAO 的组织架构, 能够将分散的人类组织起来, 更好地量化人的因素在指示学习和提示学习中的贡献, 同时可通过区块链技术的追溯性保证指示和提示内容的规范性.

### 3 超大预训练模型的挑战和发展

本节将从研究范式、平行智能化以及决策智能等方面探讨 PTM 的发展趋势.

#### 3.1 超大预训练模型面临的挑战

以 ChatGPT 为代表的 PTM 为诸多领域提供新发展机遇的同时也带来了许多新的挑战. 首先, PTM 的高性能是以高算力为代价. OpenAI 在 2018 年发布的报告中指出, 自 2012 年以来, AI 训练的算力呈指数级增长<sup>5</sup>, 这意味着 PTM 在提升性能的同时也消耗了更多算力. 其次, PTM 的置信度有待提升. PTM 的准确度由训练样本的数量和质量共同决定, 因此在处理一些复杂问题时准确度会降低, 甚至出现一些完全错误的答案, 不恰当的使用会导致严重的损失. 也因此, PTM 难以在工业领域应用, 控制和决策类的 PTM 也很少见. 其次, PTM 在创新能力方面还存在很大的上升空间. 观察由 ChatGPT 生成的相应文案可以发现, 其生成的文本在格式方面都大同小异, 缺乏多样性和创新性. 最后, 由于人类在此类模型中扮演了开发者和使用者的角色, PTM 在给人类带来便利的同时也带来了额外的法律和道德问题, 因此如何正确使用科技带来的便利也成为亟待解决的问题. 为迎接这些挑战, PTM 需进一步发展, 本节剩余部分将讨论 PTM 的发展方向.

#### 3.2 研究范式升级

近年来, 从以 AlphaGo 为代表的算法智能 (Algorithmic intelligence, AI) 到以 ChatGPT 为代表的语言智能 (Linguistic intelligence, LI), 人工智能研究的范式快速转换. AI 的代表 AlphaGo 将监督学习、RL、蒙特卡洛树搜索结合, 通过自我博弈实现策略提升. LI 的代表 ChatGPT 重新树立人在智

能系统中的作用, 以指令学习和 RLHF 的方式引导预训练 PTM 释放生成能力, 在交互问答方面实际已通过图灵测试.

人工智能范式升级的下一步是什么? 想象智能 (Imaginative intelligence, II) 已经初现端倪<sup>[1, 42]</sup>. 图 7 展示了所设想 PTM 研究范式的发展. II 强调在模拟、模仿或者参考现实场景的基础上, 使用想象力和创造性的方法来产生全新的信息和概念. II 的核心在于多模态大模型, 其目标不仅是学习现有数据中的规律, 而是进一步生成新的数据, 创建新的概念, 提出新的方案. 具体来说, II 可以根据不同场景和需求创造性地生成全新的文字、图像、视频、3D 动态场景等内容, 在模拟人类的想象力和创造性思维方面具有广阔的应用前景和潜力.

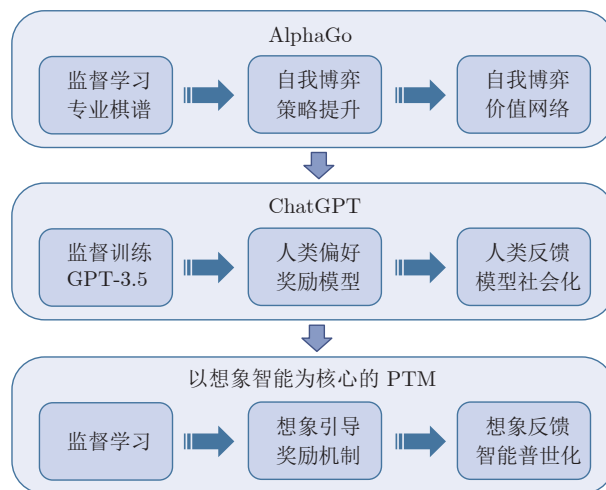


图 7 PTM 研究范式

Fig. 7 Research paradigms of PTMs

#### 3.3 深度平行化的超大预训练模型

平行智能的特点是真实世界与虚拟世界的交互, 包含描述、预测、引导三大模块. 即使再大规模的数据量, 也不能完全描述一个开放的真实世界. 而且事实和知识具有动态时间属性, 不是一成不变的. 为保证回答以及决策的与时俱进, PTM 必须与外部环境交互, 源源不断地接受外部知识的更新. 另外, 数据表示的不完备性, 也需要 PTM 具有挖掘数据背后的隐含知识的能力. PTM 的参数空间可看作是虚拟世界的一种表现形式. ChatGPT 已经给出了平行化训练的初步尝试, 通过人在回路的社会化大闭环, 将人类社会中的隐性知识蕴含于奖励模型中, 进而扩展到 PTM. 其优异的表现已充分表明, 真实世界与虚拟世界的交互模式会使 PTM 性能产生质的飞跃. 因此, 为充分克服上述难点, 必须

<sup>5</sup> <https://openai.com/blog/ai-and-compute>



进一步将 PTM 的训练和应用平行化, 通过预测智能充分挖掘潜在知识, 虚实交互、平行执行引导 PTM 的知识更新, 真正解决 PTM 的事实不一致问题. 目前, 已有部分研究工作对如何将平行智能与 PTM 结合的问题进行了探索<sup>[4-6, 44-48]</sup>.

### 3.4 超大决策预训练模型

目前, PTM 已经在自然语言、计算机视觉等领域取得了激动人心的成果, 但是在决策和控制等领域, 仅有少部分 PTM 尝试处理决策、控制任务, 且并未取得优异的成绩<sup>[21]</sup>. 尽管如此, 基于 DNN 的决策与控制算法早已成为实现人类水平决策与控制的基石<sup>[38, 41, 49-50]</sup>, 因此将 PTM 应用于决策和控制问题应当也具有巨大潜力. 我们认为目前 PTM 尚未在决策与控制方向展现出潜力的主要原因如下: 1) 决策与控制问题常见于动态系统, 其优化难度较大, DNN 参数量的增加并不能确保性能的提升; 2) PTM 的预训练高度依赖于大规模未标记的数据集, 决策与控制问题尚无这类数据集; 3) PTM 的预训练一般采用自监督的形式, 而对于决策与控制问题, 状态映射到动作随任务变化而大有不同, 或高度依赖特定领域专家知识, 很难以自监督的形式预训练. 因此, 我们认为应采用平行智能的方式实现决策与控制 PTM<sup>[29-31, 51-52]</sup>. 通过构造人工社会对物理社会进行扩展, 在人工社会中生成大量具有标记的数据集, 同时为决策和控制算法提供计算实验和验证平台并保证其安全性.

此外, 近期一款针对 Stable Diffusion 的条件控制插件 ControlNet 引起了大家的关注<sup>[53]</sup>, 其思路是: 通过给定人为额外的条件, 使模型产生更高质量、更符合要求的图片, 使生成图片质量可控. 这种 PTM 向定制化、专业化的改进, 为 PTM 提供了良好的周边生态环境, 为 PTM 应用打下了夯实基础. 另外, 如何使 PTM 从巨型化迈向小型化, 促进 PTM 的实际应用和部署也是未来研究的重点. 通过增加 DNN 参数量, PTM 取得了令人瞩目的成绩. 但也随着 DNN 参数量的增大, PTM 输出的实时性、可解释性、安全性等则更难以保证. 在一些实时性、安全性要求不高的场景, 例如自然对话、文本生成、自动摘要等, 可以直接应用. 但在一些需要实时优化的工业场景, 例如现场控制、状态预测以及故障预测报警等, 目前 PTM 的实时性和安全性难以胜任. 因此促进 PTM 从巨型化迈向小型化, 或者通过知识蒸馏等压缩模型, 在不损失 PTM 性能的情况下保证其实时性、安全性具有重大的研究意义.

## 4 超大预训练模型的应用场景

### 4.1 人工智能生成内容

互联网数字内容产业经历了早期的专业生成内容和元宇宙初期的用户生成内容. 当前, 在 PTM 的推动下, AIGC 核心技术快速发展: 在自然语言生成方面, 能够自动生成文章、摘要、新闻、广告、电子邮件等各种文字内容, 大大提高生产效率和工作效率; 在视觉生成方面, 能够自动生成有创意的图像、视频和 3D 场景, 成为艺术创作的有力工具<sup>[54]</sup>; 在音频生成方面, 能够自动生成音乐、语音、对话等内容, 例如虚拟主播、自动生成音乐等应用.

### 4.2 工业化超大预训练模型

基于 PTM 提升工业流程智能化程度具有巨大的潜力. 工业化 PTM 需应用到实际生产环境中, 并对其进行优化和改进, 以满足各种复杂的应用需求. 与现有 AIGC 不同, 由于工业场景对可靠性和安全性有极高的要求, 因此对 PTM 输出的可控性和可解释性要求极高. 因此, 我们认为工业化 PTM 的实现必然是以 Hanoi 平行智能<sup>[31]</sup>、场景工程<sup>[55]</sup>和 AI4Science<sup>[40]</sup>组合的方式实现的. 首先, 基于场景工程针对不同的应用场景, 设计出相应的 PTM, 再基于虚实交互的平行智能方式提供数据和可靠的计算实验和验证平台. 其次, 通过和 AI4Science 结合, 将客观物理知识融入 PTM 的设计与实现, 增加其可解释性以及安全性. 然后, 建立完善的评估和优化机制, 以对模型进行不断的监控和调整, 以适应不同应用场景和变化的需求. 最后, 利用知识蒸馏等确保 PTM 的快速部署和实用性.

### 4.3 数字人

“数字人”<sup>[56]</sup>是指数字化的虚拟人物, 可以基于真实人物形象和行为特征进行生成. 随着语音识别、NLP、计算机视觉等技术的发展, 特别是以 ChatGPT 为代表的 PTM 技术, 使得数字人物具有自然而丰富的表情, 可以像真人一样进行交互和对话. “数字人”的概念可进一步推广, 形成为企业服务的“数字助手”和“数字员工”, 具有更广阔的应用前景. “数字助手”通过 NLP 等人工智能技术为用户提供个性化的帮助和支持, 使人们更加高效地完成各种任务. “数字员工”指的是基于软件机器人、流程自动化软件和智能决策系统等构成的软件, 可以自动化地执行重复性、繁琐或危险的任务, 也可以像人类员工一样, 接受指令、自主决策、自我学习, 从而降低人工成本、提高生产效率. “数字员工”应用前景

广阔,例如:在制造业中,数字员工可负责产品、物料管理等;在零售业中,处理电子商务订单、自动化仓库管理、以及在线客服服务等。“数字员工”的另外一项价值在于帮助探索更好的企业管理模式,培训企业管理人员,探索更好的商业模式等。

如前所述,在自然人-数字人-机器人的三层社会化大闭环中,数字人扮演着连接自然人和机器人的桥梁。对上层,数字人是人机交互的直接参与方,对用户体验起着至关重要的作用。对下层,数字人承担控制策略的分发与监视,从而是任务成功执行的必要保证。对中层,数字人通过合作、对抗等,完成需求分解、资源协调、方案验证等功能,是人机系统管控目标到优化执行的关键环节。而以 ChatGPT 为代表的 PTM,则能够使得数字人更加高效、智能地完成上述社会化大闭环。一方面,ChatGPT 可直接纳入数字人的知识库,作为数字人的认知知识基础,为个性化、特色化的数字人研发提供构建起点。特别是在数字人与自然人直接交互的平行认知场景中,ChatGPT 出色的类人生成能力可赋予数字人强大的交互能力<sup>[57]</sup>。另一方面,ChatGPT 也可与数字人进行交互。受 RLHF 学习模式的限制,ChatGPT 的迭代更新需要人类经验的反馈,以确保其问答知识向符合用户需求的方向优化。而数字人作为自然人的虚拟映射,可代替后者提供指导信号,从而加速 ChatGPT 的学习过程。更进一步,异质平行人口(即数字人群体)驱动下的多样化指导信号,还能够帮助 ChatGPT 发现其知识的不足,提高多场景下的泛化性能。这在我们前期的研究中已初步得到验证<sup>[56, 58-59]</sup>。因此,ChatGPT 与数字人的结合必将带来金融、教育、法律、文学、音乐、影视创作等众多行业的革新与提升。

## 5 结论

为促进 PTM 研究的进一步深入,本文首先对 PTM 的发展进行简要梳理,分别从语言 PTM 和多模态 PTM 进行介绍。其次,详细分析和探讨 ChatGPT 的设计和实现,从 RL 切入分析 ChatGPT,并以平行智能的视角解读 ChatGPT,分析表明 ChatGPT 出色表现得益于其成功构建了社会化大闭环,是 Hanoi 框架下 PTM 设计与实现的一次成功示范。最后,对 PTM 发展趋势进行展望。

就在本文录用不久前,OpenAI 发布了多模态预训练大模型 GPT-4,除了像 ChatGPT 完成文本对话,GPT-4 还能够进行更复杂的图像理解任务。GPT-4 发布之后两天,微软宣布将 GPT-4 整合进自家 Office 办公套件,并推出人工智能助理 Copi-

lot(副驾驶员),用户通过自然语言描述需求,Copilot 即可完成相应功能,例如在 Word 中起草文章草稿,在 PowerPoint 中整合素材并制作美观的幻灯片,在 Excel 中完成数据分析与图表生成。Copilot 的发布将极大提高人类在知识创造方面的生产力。

尽管 ChatGPT 与 GPT-4 的成功让人们对人工智能的发展充满期待,但我们仍然需要严谨、务实地对待 ChatGPT 及相关技术。例如 ChatGPT 的技术仍不完善,经常会一本正经地说些不着边际的话。同时 ChatGPT 的出现也带来了许多问题,例如:算法偏见、著作权纠纷、行业垄断等等。这些问题不能忽视,应及时完善法律法规。我们应在社会化大闭环中充分发挥积极作用,促进人与社会和谐发展。

## References

- Wang F Y, Miao Q H, Li X, Wang X X, Lin Y L. What does ChatGPT say: The DAO from algorithmic intelligence to linguistic intelligence. *IEEE/CAA Journal of Automatica Sinica*, 2023, **10**(3): 575-579
- Han X, Zhang Z X, Ding N, Gu Y X, Liu X, Huo Y Q, et al. Pre-trained models: Past, present and future. *AI Open*, 2021, **2**: 225-250
- Bommasani R, Hudson D A, Adeli E, Altman R, Arora S, von Arx S, et al. On the opportunities and risks of foundation models [Online], available: <https://arxiv.org/abs/2108.07258>, August 16, 2021
- Lu J W, Wang X X, Cheng X, Yang J, Kwan O, Wang X. Parallel factories for smart industrial operations: From big AI models to field foundational models and scenarios engineering. *IEEE/CAA Journal of Automatica Sinica*, 2022, **9**(12): 2079-2086
- Wang X X, Cheng X, Lu J W, Kwan O, Li S X, Ping Z X. Metaverses-based parallel oil fields in CPSS: A framework and methodology. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, **53**(4): 2138-2147
- Yang J, Li S M, Wang X X, Lu J W, Wu H Y, Wang X. De-FACT in ManuVerse for parallel manufacturing: Foundation models and parallel workers in smart factories. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, **53**(4): 2188-2199
- Chen T, Kornblith S, Swersky K, Norouzi M, Hinton G. Big self-supervised models are strong semi-supervised learners. In: Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc., 2020. Article No. 1865
- Ramesh A, Dhariwal P, Nichol A, Chu C, Chen M. Hierarchical text-conditional image generation with CLIP latents [Online], available: <https://arxiv.org/abs/2204.06125>, April 13, 2022
- Rombach R, Blattmann A, Lorenz D, Esser P, Ommer B. High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE, 2022. 10674-10685
- Vaswani A, Shazeer N, Parmar N, Uszkoreit U, Jones L, Gomez A N, et al. Attention is all you need. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA: Curran Associates Inc., 2017. 6000-6010
- Peters M E, Neumann M, Iyyer M, Gardner M, Clark C, Lee K,

- et al. Deep contextualized word representations. In: Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. New Orleans, USA: Association for Computational Linguistics, 2018. 2227–2237
- 12 Devlin J, Chang M W, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis, USA: Association for Computational Linguistics, 2019. 4171–4186
- 13 Radford A, Narasimhan K, Salimans T, Sutskever I. Improving language understanding by generative pre-training [Online], available: [https://s3-us-west-2.amazonaws.com/openai-assets/research-covers/language-unsupervised/language\\_understanding\\_paper.pdf](https://s3-us-west-2.amazonaws.com/openai-assets/research-covers/language-unsupervised/language_understanding_paper.pdf), April 13, 2022
- 14 Radford A, Wu J, Child R, Luan D, Amodei D, Sutskever I. Language models are unsupervised multitask learners [Online], available: [https://cdn.openai.com/better-language-models/language\\_models\\_are\\_unsupervised\\_multitask\\_learners.pdf](https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf), April 13, 2022
- 15 Brown T B, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, et al. Language models are few-shot learners. In: Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc., 2020. Article No. 159
- 16 Raffel C, Shazeer N, Roberts A, Lee K, Narang S, Matena M, et al. Exploring the limits of transfer learning with a unified text-to-text transformer. *The Journal of Machine Learning Research*, 2020, **21**(1): Article No. 140
- 17 Zhong R Q, Lee K, Zhang Z, Klein D. Adapting language models for zero-shot learning by meta-tuning on dataset and prompt collections. In: Proceedings of the Findings of the Association for Computational Linguistics. Punta Cana, Dominican Republic: Association for Computational Linguistics, 2021. 2856–2878
- 18 Lu J S, Batra D, Parikh D, Lee S. ViLBERT: Pretraining task-agnostic visiolinguistic representations for vision-and-language tasks. In: Proceedings of the 33rd International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc., 2019. Article No. 2
- 19 Radford A, Kim J W, Hallacy C, Ramesh A, Goh G, Agarwal S, et al. Learning transferable visual models from natural language supervision. In: Proceedings of the 38th International Conference on Machine Learning (ICML). PMLR, 2021. 8748–8763
- 20 Alayrac J B, Donahue J, Luc P, Miech A, Barr I, Hasson Y, et al. Flamingo: A visual language model for few-shot learning. In: Proceedings of the 36th International Conference on Neural Information Processing Systems (NeurIPS). New Orleans, USA: 2022.
- 21 Reed S, Żołna K, Parisotto E, Colmenarejo S G, Novikov A, Barth-Maron G, et al. A generalist agent. *Transactions on Machine Learning Research*, 2022: 1–42
- 22 Zeng W, Ren X, Su T, Wang H, Liao Y, Wang Z, et al. PanGu- $\alpha$ : Large-scale autoregressive pretrained Chinese language models with auto-parallel computation [Online], available: <https://arxiv.org/abs/2104.12369>, April 26, 2021
- 23 Fei N Y, Lu Z W, Gao Y Z, Yang G X, Huo Y Q, Wen J Y, et al. Towards artificial general intelligence via a multimodal foundation model. *Nature Communications*, 2022, **13**(1): Article No. 3094
- 24 Lin J Y, Men R, Yang A, Zhou C, Ding M, Zhang Y C, et al. M6: A Chinese multimodal pretrainer [Online], available: <https://arxiv.org/abs/2103.00823>, May 29, 2021
- 25 Liu J, Zhu X X, Liu F, Guo L T, Zhao Z J, Sun M Z, et al. OPT: Omni-perception pre-trainer for cross-modal understanding and generation [Online], available: <https://arxiv.org/abs/2107.00249>, July 6, 2021
- 26 Zhang Z Y, Han X, Liu Z Y, Jiang X, Sun M S, Liu Q. ERNIE: Enhanced language representation with informative entities. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy: Association for Computational Linguistics, 2019. 1441–1451
- 27 Wang S H, Sun Y, Xiang Y, Wu Z H, Ding S Y, Gong W B, et al. ERNIE 3.0 Titan: Exploring larger-scale knowledge enhanced pre-training for language understanding and generation [Online], available: <https://arxiv.org/abs/2112.12731>, December 23, 2021
- 28 Hoffmann J, Borgeaud S, Mensch A, Buchatskaya E, Cai T, Rutherford E, et al. Training compute-optimal large language models [Online], available: <https://arxiv.org/abs/2203.15556>, March 29, 2022
- 29 Wang Fei-Yue. Parallel system methods for management and control of complex systems. *Control and Decision*, 2004, **19**(5): 485–489, 514  
(王飞跃. 平行系统方法与复杂系统的管理和控制. 控制与决策, 2004, **19**(5): 485–489, 514)
- 30 Wang Fei-Yue. Parallel control and digital twins: Control theory revisited and reshaped. *Chinese Journal of Intelligent Science and Technology*, 2020, **2**(3): 293–300  
(王飞跃. 平行控制与数字孪生: 经典控制理论的回顾与重铸. 智能科学与技术学报, 2020, **2**(3): 293–300)
- 31 Wang F Y. Parallel intelligence in metaverses: Welcome to Hanoi! *IEEE Intelligent Systems*, 2022, **37**(1): 16–20
- 32 Chen L, Zhang Y Q, Tian B, Ai Y F, Cao D P, Wang F Y. Parallel driving OS: A ubiquitous operating system for autonomous driving in CPSS. *IEEE Transactions on Intelligent Vehicles*, 2022, **7**(4): 886–895
- 33 Tian F Y, Li Z H, Wang F Y, Li L. Parallel learning-based steering control for autonomous driving. *IEEE Transactions on Intelligent Vehicles*, 2023, **8**(1): 379–389
- 34 Wang J G, Wang X, Shen T Y, Wang Y T, Li L, Tian Y L, et al. Parallel vision for long-tail regularization: Initial results from IVFC autonomous driving testing. *IEEE Transactions on Intelligent Vehicles*, 2022, **7**(2): 286–299
- 35 Wang K F, Gou C, Zheng N N, Reh J M, Wang F Y. Parallel vision for perception and understanding of complex scenes: Methods, framework, and perspectives. *Artificial Intelligence Review*, 2017, **48**(3): 299–329
- 36 Ouyang L, Wu J, Jiang X, Almeida D, Wainwright C L, Mishkin P, et al. Training language models to follow instructions with human feedback. In: Proceedings of the 36th International Conference on Neural Information Processing Systems (NeurIPS). New Orleans, USA: 2022.
- 37 Tian Yong-Lin, Wang Yu-Tong, Wang Jian-Gong, Wang Xiao, Wang Fei-Yue. Key problems and progress of vision Transformers: The state of the art and prospects. *Acta Automatica Sinica*, 2022, **48**(4): 957–979  
(田永林, 王雨桐, 王建功, 王晓, 王飞跃. 视觉 Transformer 研究的关键问题: 现状及展望. 自动化学报, 2022, **48**(4): 957–979)
- 38 Schulman J, Wolski F, Dhariwal P, Radford A, Klimov O. Proximal policy optimization algorithms [Online], available: <https://arxiv.org/abs/1707.06347>, August 28, 2017
- 39 Sutton R S, Barto A G. *Reinforcement Learning: An Introduction* (Second edition). Cambridge: MIT Press, 2018.
- 40 Lu Jing-Wei, Cheng Xiang, Wang Fei-Yue. Artificial intelligence and deep learning methods for solving differential equations: The state of the art and prospects. *Chinese Journal of Intelligent Science and Technology*, 2022, **4**(4): 461–476  
(卢经纬, 程相, 王飞跃. 求解微分方程的人工智能与深度学习方法: 现状及展望. 智能科学与技术学报, 2022, **4**(4): 461–476)
- 41 Mnih V, Kavukcuoglu K, Silver D, Rusu A A, Veness J, Bellemare M G, et al. Human-level control through deep reinforcement learning. *Nature*, 2015, **518**(7540): 529–533



- 42 Wang Fei-Yue. Parallel management: The DAO to smart ecological technology for complexity management intelligence. *Acta Automatica Sinica*, 2022, **48**(11): 2655–2669  
(王飞跃. 平行管理: 复杂性管理智能的生态科技与智慧管理之DAO. *自动化学报*, 2022, **48**(11): 2655–2669)
- 43 Li Li, Lin Yi-Lun, Cao Dong-Pu, Zheng Nan-Ning, Wang Fei-Yue. Parallel learning — A new framework for machine learning. *Acta Automatica Sinica*, 2017, **43**(1): 1–8  
(李力, 林懿伦, 曹东璞, 郑南宁, 王飞跃. 平行学习 —— 机器学习的一个新型理论框架. *自动化学报*, 2017, **43**(1): 1–8)
- 44 Wang X J, Kang M Z, Sun H Q, de Reffye P, Wang F Y. DeCASA in AgriVerse: Parallel agriculture for smart villages in Metaverses. *IEEE/CAA Journal of Automatica Sinica*, 2022, **9**(12): 2055–2062
- 45 Wang J G, Tian Y L, Wang Y T, Yang J, Wang X X, Wang S J, et al. A framework and operational procedures for metaverses-based industrial foundation models. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, **53**(4): 2037–2046
- 46 Li X, Tian Y L, Ye P J, Duan H B, Wang F Y. A novel scenarios engineering methodology for foundation models in metaverse. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, **53**(4): 2148–2159
- 47 Wang Y T, Wang J G, Cao Y S, Li S X, Kwan O. Integrated inspection on PCB manufacturing in cyber-physical-social systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, **53**(4): 2098–2106
- 48 Liu Y H, Shen Y, Tian Y L, Ai Y F, Tian B, Wu E, et al. RadarVerses in metaverses: A CPSI-based architecture for 6S radar systems in CPSS. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, **53**(4): 2128–2137
- 49 Lu J W, Wei Q L, Wang F Y. Parallel control for optimal tracking via adaptive dynamic programming. *IEEE/CAA Journal of Automatica Sinica*, 2020, **7**(6): 1662–1674
- 50 Wang Fei-Yue, Chen Jun-Long. *Intelligent Control Method and Application*. Beijing: Science and Technology of China Press, 2020.  
(王飞跃, 陈俊龙. 智能控制方法与应用. 北京: 中国科学技术出版社, 2020.)
- 51 Wang Fei-Yue. Parallel philosophy and intelligent science: From Leibniz's Monad to Blockchain's DAO. *Pattern Recognition and Artificial Intelligence*, 2020, **33**(12): 1055–1065  
(王飞跃. 平行哲学与智能科学: 从莱布尼茨的Monad到区块链之DAO. *模式识别与人工智能*, 2020, **33**(12): 1055–1065)
- 52 Lu J W, Wei Q L, Zhou T M, Wang Z Y, Wang F Y. Event-triggered near-optimal control for unknown discrete-time nonlinear systems using parallel control. *IEEE Transactions on Cybernetics*, 2023, **53**(3): 1890–1904
- 53 Zhang L M, Agrawala M. Adding conditional control to text-to-image diffusion models [Online], available: <https://arxiv.org/abs/2302.05543>, February 10, 2023
- 54 Guo C, Dou Y, Bai T X, Dai X Y, Wang C F, Wen Y. ArtVerse: A paradigm for parallel human-machine collaborative painting creation in metaverses. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, **53**(4): 2200–2208
- 55 Li X, Ye P J, Li J J, Liu Z M, Cao L B, Wang F Y. From features engineering to scenarios engineering for trustworthy AI: I&I, C&C, and V&V. *IEEE Intelligent Systems*, 2022, **37**(4): 18–26
- 56 Ye P J, Wang F Y. *Parallel Population and Parallel Human — A Cyber-Physical Social Approach*. Hoboken: Wiley-IEEE Press, 2023.
- 57 Ye P J, Wang X, Zheng W B, Wei Q L, Wang F Y. Parallel cognition: Hybrid intelligence for human-machine interaction and management. *Frontiers of Information Technology & Electronic Engineering*, 2022, **23**(12): 1765–1779
- 58 Wang Fei-Yue, Jiang Zheng-Hua, Dai Ru-Wei. Population stud-

ies and artificial societies: A discussion of artificial population systems and their applications. *Complex Systems and Complexity Science*, 2005, **2**(1): 1–9

(王飞跃, 蒋正华, 戴汝为. 人口问题与人工社会方法: 人工人口系统的设想与应用. *复杂系统与复杂性科学*, 2005, **2**(1): 1–9)

- 59 Ye P J, Wang F Y. Parallel population and parallel human — A cyber-physical social approach. *IEEE Intelligent Systems*, 2022, **37**(5): 19–27



**卢经纬** 青岛智能产业技术研究院副研究员. 2022 年获得中国科学院大学计算机应用技术博士学位. 主要研究方向为最优控制, 自适应动态规划, 深度强化学习和自动驾驶.

E-mail: [hujingwei11@gmail.com](mailto:hujingwei11@gmail.com)

(**LU Jing-Wei** Associate professor

at the Qingdao Academy of Intelligent Industries. He received his Ph.D. degree in computer application technology from University of Chinese Academy of Sciences. His research interest covers optimal control, adaptive dynamic programming, deep reinforcement learning, and autonomous driving.)



**郭超** 中国科学院自动化研究所复杂系统管理与控制国家重点实验室助理研究员. 主要研究方向为机器艺术创作, 智能机器人系统, 深度学习, 强化学习.

E-mail: [guochao2014@ia.ac.cn](mailto:guochao2014@ia.ac.cn)

(**GUO Chao** Assistant professor at

the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interest covers AI art, intelligent robotic systems, deep learning, and reinforcement learning.)



**戴星原** 中国科学院自动化研究所复杂系统管理与控制国家重点实验室助理研究员. 2022 年获得中国科学院大学控制理论与控制工程专业博士学位. 主要研究方向为人工智能, 强化学习, 智能交通系统.

E-mail: [xingyuan.dai@ia.ac.cn](mailto:xingyuan.dai@ia.ac.cn)

(**DAI Xing-Yuan** Assistant professor at the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences. He received his Ph.D. degree in control theory and control engineering from the University of Chinese Academy of Sciences. His research interest covers artificial intelligence, reinforcement learning, and intelligent transportation systems.)

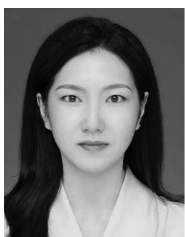


**缪青海** 中国科学院大学人工智能学院副教授. 2007 年获得中国科学院自动化研究所博士学位. 主要研究方向为智能系统, 机器学习, 计算机视觉.

E-mail: miaoqh@ucas.ac.cn

**(MIAO Qing-Hai** Associate professor at the School of Artificial In-

telligence, University of Chinese Academy of Sciences. He received his Ph.D. degree from the Institute of Automation, Chinese Academy of Sciences in 2007. His research interest covers intelligent systems, machine learning, and computer vision.)



**王兴霞** 中国科学院自动化研究所复杂系统管理与控制国家重点实验室博士研究生. 2021 年获得南开大学工学硕士学位. 主要研究方向为平行控制, 平行油田和多智能体系统.

E-mail: wangxingxia2022@ia.ac.cn

**(WANG Xing-Xia** Ph.D. candid-

ate at the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences. She received her master degree in engineering from Nankai University in 2021. Her research interest covers parallel control, parallel oilfields, and multi-agent systems.)



**杨 静** 中国科学院自动化研究所复杂系统管理与控制国家重点实验室博士研究生. 2020 年获得北京化工大学自动化学士学位. 主要研究方向为平行制造, 社会制造, 人工智能和社会物理信息系统.

E-mail: yangjing2020@ia.ac.cn

**(YANG Jing** Ph.D. candidate at the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences. She received her bachelor degree in automation from Beijing University of Chemical Technology in 2020. Her research interest covers parallel manufacturing, social manufacturing, artificial intelligence, and cyber-physical-social systems.)



**王飞跃** 中国科学院自动化研究所复杂系统管理与控制国家重点实验室研究员. 主要研究方向为智能系统, 复杂系统建模, 分析与控制. 本文通信作者.

E-mail: feiyue.wang@ia.ac.cn

**(WANG Fei-Yue** Professor at the

State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interest covers modeling, analysis, and control of intelligent systems and complex systems. Corresponding author of this paper.)