

AGI时代的电子及计算机工程师

周娜¹, 何铮², 何为民³

¹海口经济学院中芯依智网络学院, 海南 海口

²海南政法职业学院信息中心, 海南 海口

³海南久迪物联网科技有限公司, 海南 海口

收稿日期: 2024年7月4日; 录用日期: 2024年11月21日; 发布日期: 2024年11月29日

摘要

文章展望了AGI时代的特点及人才金字塔结构的分布, 重点分析了AGI时代电子及计算机工程师的行业发展趋势及特点, 并建言当今电子及计算机工程师如何应对AGI时代的来临。

关键词

通用人工智能AGI, 人才金字塔, 电子及计算机工程师

Electronic and Computer Engineers in AGI Era

Na Zhou¹, Zheng He², Weimin He³

¹ College of SMIC Network, Haikou Institute of Economics, Haikou Hainan

² Information Center, Hainan Vocational College of Political Science and Law, Haikou Hainan

³ Hainan Jiudi Internet of Things Technology Co., Ltd., Haikou Hainan

Received: Jul. 4th, 2024; accepted: Nov. 21st, 2024; published: Nov. 29th, 2024

Abstract

This paper looks forward to the characteristics of AGI era and the distribution of talent pyramid structure. This paper focuses on the development trend and characteristics of electronic and computer engineers in AGI era. It also suggests how electronic, and computer engineers should deal with the arrival of AGI era.

Keywords

General Artificial Intelligence AGI, Talent Pyramid, Electronic and Computer Engineer

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

自 1956 年 AI (Artificial Intelligence) 人工智能诞生到 2022 年, 它一直是属于电子、计算机及哲学领域前沿基础理论研究的范畴。期间, AI 只是艰难曲折地缓慢向上发展, 对社会各行业的影响都微乎其微。但 2022 年迎来一个重大转折, Open AI 生成式大模型 ChatGPT 发布, 成为一个以生成式大模型为核心的人工智能新时代来临的新起点。AI 也逐步发展为通用人工智能 AGI (Artificial General Intelligence), 它的标志是可以实现自我学习、改进和调整, 进而不需人为干预而解决任何问题, 使人 AI “工具” 开始有了 “生命”。

自此, AI 技术进入 AGI 一个按指数型发展新领域。至今 ChatGPT 已发展到多模态的 GPT4o 及谷歌的 Gemini 等。模型参数已从千亿级发展到万亿级; 所用算力从几十 PLOPS 扩展到几十 ELOPS。大模型已经学习了 5000~8000 万种人类有史以来创造的知识。大模型不仅能文字交互, 而且能语音交互聊天, 能看懂、生成图片, 输出一定程度上符合环境规律的视频。

2. AGI 时代的特点

2.1. AGI 可渗透应用到所有的行业中

AGI 是文本、语音、图像视频多模态交互的通用大模型, 它可渗透到当今世界的所有行业。当前, 大模型已在很多应用端爆发出它的巨大威力。人们都认识到大模型将改变世界; 世上所有行业都需要基于人工智能技术重新做一遍, 所有的人群和工作都可归纳为 “制造 AGI 工具的” 和 “使用 AGI 工具的”。虽然工作可取代, 但人类的做决策、情感、想象力和创造力是人工智能短期无法取代的。

(1) 经验性、重复性的工作将被 AGI 替代

AI 应用首先逐步替代的工作是一些大量经验型、重复流水性工作。如文案, 科技翻译、形象设计、码农等。但那些少数高端人脑创意性的工作则难以替代, 大量的人力劳动(如建筑行业)也无法替代。没有唯一标准答案的文学翻译, 好的美学作品, 系统规划师等还需要一些高端的人才。在医院, 急诊、各类手术科室的医生将加强, 但门诊、社区及全科大夫将大量使用 “AI 医生”, 互联网医疗将进化为无人的 AI 医疗。

(2) $1 + 1 > 2$

在 AGI 时代, AGI 工具的能力越来越强大。一个掌握了 AGI 工具人的工作效率将远远大于两个未掌握 AGI 工具人的工作效率。AI 翻译已 90% 取代了科技翻译。随着读片 AI 化, CT 及 MR 的工作效率将大步提升。一个掌握 AI 代码工具的嵌入式工程师的工作效率能顶上数个传统的嵌入式工程师。

(3) AGI 多模态生成式大模型技术的三大方向

A、本身发展研究: 类人神经算法进一步的研究, 加大算力, 减少能耗; 学完已有数据并生成数据。

B、智能体(Agent)研究: 如 AI PC、自主机器人、L5 无人驾驶等应用侧系统。

C、科学研究: 在各自科学研究中, 根据本专业建模, 建立本行业、企业、部门的大模型, 成为自身科学研究的工具, 成为人类科学家的好助手。

2.2. AGI 时代的人才金字塔的变化(如图 1)

(1) AI 时代人才金字塔体量变小

AI 时代, 掌握了 AGI 工具人的工作效率大幅提升, $1 + 1 > 2$ 。虽然增加了一些 “训练师”、“标记

师”等新行业，但老行业裁员已成新的大趋势。所裁人员向人力产业和第三产业转移。但总体上是使用 AI 工具的个人，能力越来越强大，这部分人可以养活全球人。福利社会将到来，高端人士贡献大，生活将会更好，每周工作 3 天、4 天已成可能。因为社会的稳定，低保已能满足一部分无业底层人的基本生活。

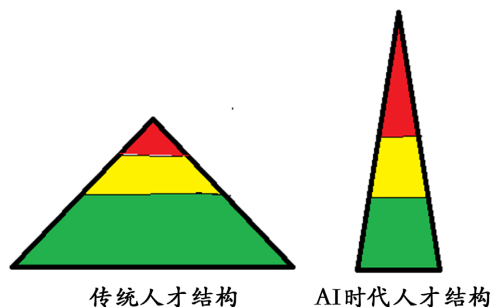


Figure 1. Talent pyramid of the AGI era
图 1. AGI 时代人材金字塔

(2) AI 时代人才金字塔更尖锐

在 AI 时代，人才高低端分离得更严重，人们可区分为“制造 AI 工具的人”和“使用 AI 工具的人”。在金字塔顶端，更需像奥特曼、黄仁勋、库克等这样的顶端大师，他们带领“制造 AI 工具的团队”，研究、推动 AGI 算法、算力的发展；生成多模态基础大模型及政府、行业、企业、个人的专属大模型。他们创造性思维带来的新工作远大于因 AI 消失的工作，它的体量会变大。在金字塔底端，“使用 AI 工具的人”因 AI 带来的高效率，消灭大量的凭经验及重复性的工作岗位。虽然它增加了训练师、标注师等工作岗位，但它总的体量会变小。

Table 1. Forecast of electronic and computer research and development team of small and micro enterprises for Internet of Things equipment

表 1. 物联网设备小微企业电子及计算机研发队伍预测

	系统规划师	软件开发工程师	测试工程师	结构工程师	总计
当前	3	12	4	1	20
AGI 时代	2	3	4	1	10

如表 1 所示，当前一个物联网设备小微企业，开发研究人员有 20 位(3 名系统规划师、12 名软件开发工程师，4 名测试工程师，1 名结构工程师)，到 AGI 时代可能只需 10 人(2 名系统规划师、3 名软件开发工程师，4 名测试工程师，1 名结构工程师)。其中创新思维，找到并提出解决方案的系统规划师仅减少 1 名，需手动非重复劳动的测试、结构工程师(非流水线上批量生产的)尚无法替代。而软件开发工程师由于上游提供大量标准电路及丰富的库函数，又有 AI 代写软件，其人数将大大压缩 3 分之一或更少。

所以 AGI 时代人才结构的金字塔会变窄、变高、变得更尖锐。

3. AGI 时代的电子及计算机工程师

作为 AGI 时代前沿的电子工程师和计算机工程师受 AGI 的影响最为直接。他们必须接受 AGI 强风暴雨的洗礼。

3.1. 硬件集成电路及微电子系统工程

(1) 迎对摩尔定律的失效

集成电路芯片制程从 5 nm 走到 3 nm, 生产成本翻了一番。摩尔定律“每 18~24 个月, 集成电路上晶体管面密度会增加一倍, 性能也将提升一倍(或同性能的成本减少一倍)”似乎已经走到了尾声, 1 nm 几乎成为芯片大厂追求的极限。应对摩尔定律的失效, 性能的提高, 不仅是算力、带宽、速度, 而且能耗也是越来越重要的考量因素。所以需要有根本性的突破, 研究发展类脑神经形态系统级芯片、类人的非硅系芯片, 光芯片[1]、量子芯片、DNA 计算芯片等。

“人脑是非常复杂庞大的神经网络系统, 总功耗仅为 20 瓦, 远小于现有的 AI 系统。”在算力比拼加速、能耗日益攀升的当下, 借鉴人脑的低功耗特性发展新型智能计算系统成为极具潜力的方向。采用人脑“神经形态动态计算”概念, 将人脑中的高抽象层次注意力机制应用于类脑芯片设计, 进一步挖掘了类脑芯片在性能和能效上的潜力。Speck 在一块芯片上集成了动态视觉传感器和神经形态芯片, 8 核具有极低的静息功耗。典型视觉场景任务功耗可低至 0.7 毫瓦, 为人工智能应用提供了高效、低延迟和低功耗的类脑智能解决方案[2]。

(2) 在摩尔定律的末端, 榨取它的剩余价值

在摩尔定律的尾声, 进一步提高性能, 就不能仅靠提高晶体管的面密度, 而是要推出新的晶体管架构、材料选择及先进制程、Chiplet 多核异构(CPU + NPU + GPU……)、光连接(TPU)、及多维连接等。

A、Chiplet 是一种芯片设计和集成的方法。它将一个大型 AI 芯片分解为多核异构多个独立的功能模块片段(称 IP 核或称 Chiplet)。每个 Chiplet 是已经过验证的、可以重复使用的具有某种确切功能的集成电路设计模块。各种 IP 核如图形处理单元 GPU (Graphics Processing Unit)、神经网络处理单元 NPU (Neural network Processing Unit)、视频处理单元 VPU (Video Processing Unit)、数字信号处理单元 DSP (Digital Signal Processor)、张量处理单元 TPU (Tensor Processing Unit)等。它使芯片设计更加模块化和灵活。不同的芯片片段可以独立设计和优化, 然后通过集成技术组合成一个完整的芯片。这种模块化的设计使得每个芯片片段可以在独立的制造工艺下进行生产(不必都使用同一种最高级制造工艺。这样可以降低制造成本, 并提高芯片的产量和良品率[3]。

B、多维集成芯片。当前计算机的计算单元和存储单元分离, 存储带宽制约了计算系统的有效带宽, 造成时延长、功耗高等问题。采用多维集成芯片, 将多个芯片堆叠在一起, 使存储与计算完全融合, 以新的高效运算架构进行二维和三维矩阵计算, 具有更大算力(1000TOPS 以上)、更高能效(超过 10~100 TOPS/W)、降本增效三大优势, 有效克服冯·诺依曼架构瓶颈, 实现计算能效的数量级提升。AMD 推出的 MI300 就采用了 3D 堆叠技术和 Chiplet 设计, 配备了 9 个基于 5 nm 制程的芯片组, 置于 4 个基于 6 nm 制程的芯片组之上。

C、互连就是电流在芯片中各个晶体管、存储器、处理单元和其他组件之间的连接方式。前期铜互连取代了铝互连。现在, 业界一直在寻找替代铜互连更优材料, 如碳纳米管(CNT)、单层石墨烯(SLG)、多层石墨烯(FLG)与钌(Ru)。台积电使用的石墨烯表现出高本征载流子迁移率和大载流能力, 具有高导热性和抗电迁移的竞争稳健性, 还可以制成原子级厚度, 有助于减轻厚度对 RC 延迟的影响。IBM 则使用钌。钌可以扩展到 1 纳米及以上节点不需要衬垫, 在导线的顶部通过减色图案化法形成的钌互连通孔, 形成连续的导线和自对准通孔。从而减少互连寄生电容, 助于实现更快、更低功耗的芯片[4]。

D、改变器件架构

采用堆叠式 CFET 场效应管架构。将晶体管组件垂直堆叠在一起, 而不是横向堆叠, 极大地增加了单个芯片上可以安装的晶体管数量。这种架构的集成密度进一步提升, 将 n 型和 p 型 MOS 元件堆叠在一起, 可以堆叠 8 个纳米片。

(3) AI PC [5]

AI 要为广大人群使用, 就必须将商用 PC 机集成人工智能技术, 演变为 AI PC。AI PC 必须满足五个

条件：配备个人智能体(AI Agent)；具备本地异构算力；具备本地化个人知识库；开放的 AI 应用生态以及设备级个人数据和隐私安全保护；可以离线运行。

要加大边缘芯片的算力，需在 AI PC 处理器芯片中加入 NPU (神经处理单元)等，采用 CPU + GPU + NPU 的多核异构的架构。以 TOPS 作为基础衡量指标。微软将自己的 AI-PC 命名为 Copilot + PC；处理器需拥有超过 40 TOPS 的算力。去年微软的首批 Copilot + PC 选择与高通合作，搭载了基于 Arm 架构的处理器。Arm 架构的功耗优势明显，更适合性能强大的 AI PC 落地普及。

AI PC 需要云端和应用侧混合运行。用户要求复杂，需要大型数据支撑，进行大型迭代，可依靠在云端运行的生成式 AI；但需快速反应、及有数据安全的要求时，则依靠自身的算力和应用侧端小型化的大模型运行。

(4)智能体芯片

智能体芯片和 AI 算力芯片一样，具有微处理单元、存储单元，可拥有自己的知识库和推理机，因而它能对自己的位置进行标识，自主地决定是否对外来信息作出响应或行为反应，而且具备与其它智能性通信的接口。它添加简单的外围部件就能构建成独立的智能体。它通常作为工业互联网的中间级或智能应用末端[6]。

3.2. 软件工程师

(1) 大模型算法 LLM(Large Language Model)研究

当前，LLM 算法研究基本都是源于类脑研究。文本生成、语音互聊、图像生成都已获得巨大进展，但无论在精细、速度、能耗等方面还欠缺不少，尤其是在多模态视频识别、视频生成方面还在摸索，在人工神经网络的可解释性、对齐和控制的领域仍处于起步阶段。亟待 MiniCPM-Llama3-V 2.5 等新模型的出现[7]。

(2) 推出算力运算平台

大模型所需的大算力需要上万块 AI 芯片堆叠运行。英伟达首推出了通用并行计算架构的 CUDA 算力平台，让全球超过 400 万开发人员依靠 CUDA 来构建 AI 及其他应用程序。CUDA 能够在运算基础不断增长的情况下，扩大生态系统，使成本不断下降。现在，随着计算膨胀和计算成本的提升，多种 AI 芯片的出现，推出各自(或共同)的 AI 芯片运算平台，对去垄断，多元化至关重要[8]。

(3) 跨行业，在各行业领域建模，用行业数据喂养，训练行业应用大模型

人类所有的行业都需要和值得用 AGI 重做一遍。这就需要跨界和行业交叉，在各个专业、行业、企业甚至个人根据自身的需求、条件及安全需求建模，利用开源系统，组建行业训练模型、标注数据队伍，建立各自的生成式大模型，应用于各自的工作、生活中。

(4) 为末端应用开发 API、SDK、APP 等

大厂不断推出了各种多模态生成式大模型，急于变现。非电子及计算机专业人士也急于大量应用 AGI。这就需要大厂为大模型的末端应用同步推出应用程序接口 API (Application Programming Interface)、软件开发工具包 SDK (Software Development Kit)及 APP 等二次开发应用端，以帮助二次开发商和电子及计算机人士能简便地安装应用在用户侧的手机及 PC 机上。

(5) 为应用端芯片开发库函数

AGI 时代，用户侧的电器都逐步演变成成为具有标识、状态、行为和接口智能体。这些智能体本身的功能越来越强大。一个多模态家居智能体加上各种前端传感器和后端执行机构可能就强似现在家庭的智能家居 + 智能健身养护 + 家居智能办公、学习 + ……等。

这种用户侧的智能体芯片都具有功能强大化、标准化、积木化、环境适应化、应用简单化的特点。

智能体芯片厂商都需要少量开发者为其研发完善的开源应用函数库，以提供给智能体开发者，也提供作为 AI 代码大模型的训练数据。这样采用智能体芯片开发智能体的工作效率就很高，所需的研究开发人员大大减少。

3.3. 应用侧电子及计算机工程师

快速发展的人工智能主体都依赖在云端的生成式大模型、大算力和大数据。但为了信息安全、反应快捷及消除拥堵，应用侧的“边缘 AI”像“边缘计算”一样也得到迅速地发展。

(1) 大模型的小型化

生成式 AI 在云端运行，但正在迅速演进至能在末端运行。如很多行业的垂直应用领域(如 L5 无人驾驶汽车、AI 医生)及个人 PC 用户，有时需要断网在用户侧末端运行，这就需要在用户侧末端能装入并运行百亿级以上的小型大模型。末端算力、存储有限，这就急需针对应用领域将大模型小型化，使其能装入智能体，单独在应用侧端运行。

(2) AI 工业物联网电子及计算机工程师

由于当今的物联网已转化为工业互联网的末端。物联网嵌入式工程师将会逐步演化为“工业物联网应用侧电子及计算机工程师”[9]。他们从事工业物联网末端智能体的开发、生产、维护。他们与传统物联网嵌入式工程师不同，有了新特点。

此时，物联网模块均成为智能体，其芯片功能强大、接口多样、协议标准，并备有推荐电路及配套的开源库函数。这种单体能力的提高，标准化，大规模生产的低成本化。使得小微物联网企业面临被兼并、淘汰、转行。这部分被裁员的工程师就只能向金子塔的两端转化。底层的工作越来越少。工程师们需更多地向软件发展，实现硬软结合，硬软交叉。也更多与应用专业交叉。要根据应用专业的要求搭载前端传感器，后端执行单元。要用应用专业的数据标记和训练，这就需要这部分工程师向应用专业跨界。

4. 当今的电子及计算机工程师的应对

4.1. 欢迎 AGI 的到来，迅速掌握出现的 AI 工具

人类历史已经证明，新技术都是不可抗拒的双刃剑，蒸汽机、电力、流水线直到核武器等都是如此。科技本身是无国界的，但使用科技的人是有政治观点的。技术的发展的阵痛之余，工作效率的提升，会使每个人的生活变得比以前更好。在这股 AI 技术浪潮中，与其害怕，倒不如积极拥抱，占据 AGI 的制高点，力所能及地顺风而上。不要囿于“模型有种种缺陷”，甚至“胡说八道”，而要看到这是必然趋势，成长中的短暂错误。学习它的利远大于弊。不想成为被逐渐淘汰的“不会使用 AI 工具的人”，那就赶快跑入掌握 AI 工具赛道中的第一梯队。

4.2. 在继承、否定、发展中成长[10]

在学习、掌握 AI 工具之时，要善于学习、继承，但不要迷信，要勇于怀疑、否定，勇于试错。科学技术就是在怀疑、否定之否定中发展的。让自己也在学习、怀疑、否定、创新探索中向金字塔顶端攀登。

4.3. 对电子及计算机教学进行大刀阔斧的改革

AGI 进入到指数发展的赛道，知识爆炸、更新淘汰得更快，部分前沿技术更新周期已大大缩短到以月计和年计，远小于课程设置、教材大纲审订、教材编写、出版的周期。所以这些权利应该下放给学校、专业、教研室及教师本人。允许教师有几分授课内容、方式裁定权。教师在教学中除自己不断学习外，

而且要指导学生打好基础，带领学生紧跟潮流、批判吸收。鼓励双学位和获取交叉专业学分。

5. 基金项目

2023 年海南省高等学校科学研究项目《基于大数据的海南特色分布式养老监护系统的研究》(项目编号: Hnky2023-50)。

参考文献

- [1] Xu, Z., Zhou, T., Ma, M., Deng, C., Dai, Q. and Fang, L. (2024) Large-Scale Photonic Chiplet Taichi Empowers 160-TOPS/W Artificial General Intelligence. *Science*, **384**, 202-209. <https://doi.org/10.1126/science.ad11203>
- [2] Yao, M., Richter, O., Zhao, G., Qiao, N., Xing, Y., Wang, D., *et al.* (2024) Spike-Based Dynamic Computing with Asynchronous Sensing-Computing Neuromorphic Chip. *Nature Communications*, **15**, Article No. 4464. <https://doi.org/10.1038/s41467-024-47811-6>
- [3] 周娜, 何铮, 何为民. 人工智能的生态树及算力研究[J]. 单片机与嵌入式系统应用, 2023, 23(9): 7-10.
- [4] 九林. 半导体工艺的极限 1 nm 之战[EB/OL]. <https://www.eefocus.com/article/1646589.html>, 2023-12-05.
- [5] CNMO 手机中国. AI 将助力 PC 迎来新一轮换机潮?事情恐怕没那么简单[EB/OL]. <https://baijiahao.baidu.com/s?id=1800147117699467870&wfr=spider&for=pc>, 2024-05-27.
- [6] The Blog of Bill Gates (2023) The Future of Agents AI Is about to Completely Change How You Use Computers and Upend the Software Industry. <https://www.gatesnotes.com/AI-agents>
- [7] 缪青海, 王兴霞, 杨静, 等. 从基础智能到通用智能: 基于大模型的 GenAI 和 AGI 之现状与展望[J]. 自动化学报, 2024, 50(4): 674-687.
- [8] 凤凰网. 打破英伟达 CUDA 霸权[EB/OL]. <http://ishare.ifeng.com/c/s/v002k46UWetBHM1S1hsUgvMUHdqWBr24cECNRVvXbPGNSOA>, 2024-03-27.
- [9] 工业互联网产业联盟. 工业互联网体系架构(版本 2.0) [EB/OL]. https://www.miit.gov.cn/cms_files/filemanager/1226211233/attach/20238/7b6171f454f94a5e9a14f2fd3b5f1c4c.pdf, 2020-04-23.
- [10] 徐光春. 马克思主义大辞典[M]. 武汉: 崇文书局, 2017.