

人工通用智能的世界

如果有一天，我们只需简单地向 AI 发出一个指令，就能让它创建一个类似 TikTok 或 YouTube 的应用程序，并且它还能提供给我们一个包含数百万行代码的完整项目，那将是非常令人惊叹的。

当 ChatGPT 在 2022 年 11 月底发布时，它能够执行诸如修复一段代码片段和回答各种问题的任务。它擅长解释神经网络等概念，并提供各种编程语言的 “Hello, World!” 示例。这就像拥有一个非常强大的搜索引擎，能够直接给出答案。

然后，它具备了搜索功能。我可以指示 ChatGPT 从一个网站检索链接列表，并创建一个关于这些链接的文档。

接着，ChatGPT 3.5、4、4o、o1-mini 和 o1 等版本相继发布。

现在，它能够处理诸如为网站添加暗黑模式的请求。它可以提供必要的代码和说明，以更新 HTML、CSS 或脚本，甚至建议添加暗黑模式切换功能。实现暗黑模式涉及更改 CSS，如果网站使用 Markdown，也需要进行相应的更新。

这就像 AI 能够实现整个功能或特性，而不仅仅是代码片段。

通过结合这些功能，我们可以创建一个应用程序。因此，有一天，如果我们告诉一个 AI 工具构建一个终端、浏览器、待办事项列表、任务应用、日历、代码协作工具或会议应用，它可能会提供整个项目代码。

然后，我们可以使任务变得更加复杂。例如，我们可以要求 AI 整合所有现有的 YouTube 代码，并使用 OpenAI、Claude 或 Deepseek 的 API 为 YouTube 添加 AI 功能。这可能包括添加智能助手、用 AI 驱动的翻译替换当前的翻译、用 AI 增强搜索能力，甚至策划专门的短视频，比如请求 YouTube 提供 100 个关于日本生活的搞笑短视频。

所以，这是一个应用程序。但更高级的任务呢，比如创建一个操作系统？我们可以告诉 AI 设计一个新的、完全开源的操作系统，具有全新的设计、基本应用程序、终端、命令行和调度程序，类似于 Oberon，使用数据结构在进程之间进行交互，而不是字符串。

接下来呢？我们可以要求 AI 设计最新的 Mac 并更新其操作系统。

然后呢？我们可以告诉 AI 设计和更新整个家庭，根据我们的活动、最新知识和需求定制所有电器产品，以创造更好的生活环境。

再然后呢？我们可以要求 AI 设计整个城市，根据市民的行为和最新知识量身定制，以改善他们的生活。

最后呢？我们可以告诉 AI 改善地球，利用所有可用的知识和信息来提升每个人的生活。

我正在为这篇文章想一个标题。让我们称之为 “人工通用智能的世界”。

根据 Deepseek 的说法，“人工通用智能（AGI）是指一种能够在广泛任务中理解、学习和应用知识，达到与人类智能相当水平的人工智能。”与专为特定任务（如面部识别、语言翻译或下棋）设计的狭义 AI 不同，AGI 可以执行人类能够完成的任何智力任务。

在考虑 AI 的未来时，有两个基本点需要掌握：算法和计算。AI 算法主要涉及微积分、反向传播、变换器、GPT 和多头潜在注意力。

在数字世界中，将会有从 X 到 Y 的映射，其中 X 可以是文本、图像、视频、音频、代码或任何字节数据。Y 也可以是这些中的任何一种。

计算机本身并不理解 AGI；它只是人类创造的一个定义，对机器来说并不重要。

AI 在物理世界中的应用将包括自动驾驶和机器人等领域。如果数字世界能够将 X 映射到 Y，物理世界也将效仿。例如，机器人可以将食材变成菜肴、搭建乐高、装饰房屋、铺设地板、安装空调和组装宜家家具。

已经有工业机器人在使用。日本著名的公司包括 FANUC、川崎重工和安川电机。

那么，为什么家庭中没有更多的机器人呢？消费级机器人需要多功能，能够执行多项任务。例如，一个烹饪机器人可能只能搅拌和炒食材，需要用户准备食材并事后清理。

未来，机器人将出现在家庭、商店、学校、办公室、电影院和旅游景点——基本上是目前人类工人所在的任何地方。

云端将有一个世界模型，一个非常大的模型，可能大约 100 拍字节大小。作为参考，1 拍字节是 1,024 太字节，1 太字节是 1,024 吉字节。Llama 3 70B 模型的一个版本文件大小为 21.1 吉字节。

世界上的机器人将需要咨询云端的世界模型以采取行动。100 毫秒甚至 1 秒的网络延迟是可以接受的，只要机器人能够有效地执行其任务。