

# 设计 AI 手机为何要用“整体论”

■ 高飞

智能手机正在经历一个再定义的时刻——成为端侧 AI 的重要载体(接下来,我称之为 AI 手机)。所以,产品设计思路也会出现新的变化。

7月中旬参加了一场深圳举行的荣耀 Magic 旗舰新品发布会后,我的脑海里浮现出亚里士多德的一句名言,叫:“整体大于部分之和”。

听起来这句话像一个玄学,但其实我们可以将其理解成认识事物的两种体系,一种叫“还原论”,另一种叫“整体论”。

“还原论”强调将复杂系统分解为更简单的组成部分来理解,假设通过理解各个部分及其基本规律,就能理解整个系统。所以,“还原论”强调分析、分类和简化。

而“整体论”正好相反,它强调系统作为一个整体的特性和行为,认为整体具有其局部所不具备的新特性。而且,这些特性源于各局部之间的相互作用。于是,“整体论”关注系统的整体性、相互关系和涌现。

简而言之,还原论认为“整体等于部分之和”,整体论看重“整体大于部分之和”(也就是亚里士多德的名言)。

## “智能”是整体论物种

理论听起来会有点枯燥,还原论比较简单,整体论略显抽象。但在生物学中,有一个完美的“整体论”例证,就是智能生命:各个器官单独存在时无法产生生命,但它们以特定方式组合在一起时,就形成了一个有生命的、具有自我调节能力的智能物种。

重新回到我们要谈的手机产业发展话题。如果我们将智能手机,看作是个人计算设备的延伸。那么计算机的早期设计,其实是更接近“还原论”的。

计算机被设计成多个独立但相互连接的模块,如 CPU、内存、存储设备等。还原式的设计有其优点,比如允许工程师专注于优化单个组件,不需要每次都考虑整个系统。

但在 AI 时代,既然“硅基生命”对标的是“碳基生命”。那么,技术产品的“还原论”,就可能被“整体论”替代。

在算法层面,以深度学习为代表的 AI 模型,就体现了“整体论”思维。单看每个神经元,都是最简单不过的结构,但如果是成百上千亿的神经元组合起来,协同联动,就涌现出了不可思议的智能。

AI 模型的打造如此,承载 AI 的硬件——智能手机也如此。

对于荣耀的 Magic 等旗舰产品的设计方法,我就感受到了一种“整体论”思维体系。不然,也不会想到亚里士多德的那句名言。

## 端侧 AI 的部件

不过,说智能手机的“整体论”之前,先来说一下手机有哪些部分。

对于一部具备 AI 特征的手机来说,如果挑三个最重要的部件,我会选承载应用的系统部件、呈现信息/数据的显示部件,和决定待机续航的供电(电池)部件。

显然,我们无法以“还原论”角度,单独看智能手机每个系统的设计。

因为如果将三个系统,各自分开来看,自然是指标越高越好,但是简单用加法组合起来可能很糟糕。

比如,要想手机呈现更多具备实时响应的 AI 特征,就要求系统部件提供足够高的性能。

同样,从显示部件的角度,屏幕自然越大,信息承载量自然也就越多。电池续航能力,更是待机时间越长越好。

但如果三个部件,都仅从自身角度出发做最优设计,不考虑手机的整体体验。我们做出来的可能就不是智能手机了,而是一台笔记本电脑,还是很重的那种。

从消费者体验角度出发,其实大家是“我全都要”,既要轻薄,又要 AI 体验,还要续航给力。

因此,如果不能将这几个关键部分融合为一体,让产品涌现出让消费者惊艳的新体验(同时破除他们之间的制约因素),就很难说这是成功的智能手机再定义。

## 无电,不 AI

苹果就是一个例证,这家公司刚刚发布了自己的端侧 AI 架构(Apple

Intelligence)。但最新的消息说,这个新 AI 平台,只有 iPhone 15 Pro 和 iPhone 15 Pro Max 搭载的最新处理器(A17 Pro)才能够支持使用。

实际上,iPhone 15 标准版的芯片(A16 Bionic)已经很强大了。如果不能支持其最新的 AI 平台,只能说明这套系统的资源开销太高。

往更本质了看,其实是能耗问题。

对于智能手机而言,电池这个部件对其他部件的约束也尤其值得一提。

因为,在大模型时代,虽然 AI 表现得似乎无比强大,但它有个致命弱点,就是怕被拔电线。

当然,这里的拔电线,不是因为人类要对付“终结者”,而是模型的训练推理,真的很耗电。

因此,就像第一次工业革命的时候,很多工厂要建在河边(因为当时的动力系统是水力),获得更便捷的动力一样。

现在在全球 AI 模型公司,甚至在加大力度投资电力技术,以求获得更低成本,更可靠的电力资源。

云端的 AI 被电所限,端侧的 AI 更是如此。

我之所以要加一个更字,是因为端侧 AI 的电力困境,比云端更为难解。

因为云端 AI 的电力可能只是贵,但是多数情况下,量是够的。但在智能手机这个端侧 AI 设备上,面对的则是电力的稀缺性问题,即有无的问题。

道理很简单,端侧 AI 所耗的电力,毕竟只能依赖于一块容量固定的电池。

电池资源的有限性,则又受限于终端的机身尺寸要求。智能手机作为消费者携带时间最长的 AI 终端,便携性是第一性的。

无论是直板机,还是折叠机,能轻松装进口袋是基本要求。

这样,机身尺寸在三维空间的每一毫米,都是一种稀缺资源,很难给电池更大的发挥空间。

机身、性能、续航,也就构成了一个手机设计的“不可能三角”。

显然,如果没有一种“整体论”的设计观念,充分思考不同部件的相互关系和相互制约,那么无论开发出来的端侧 AI 应用再怎么强大。要么会把智能手机做成了一台笔记本电脑,要么终端的

AI 可用性,会受到极大限制。

## 荣耀的做法

与苹果相比,荣耀似乎找到了实现“整体论”设计的方法。它的荣耀 200 系列,100 系列手机,都支持“任意门”(AI 意图识别技术),也都支持低功耗的电子围栏,让用户实现到了快递柜,就可弹出取货二维码,实现了产品线的全线 AI。

所以,这家公司是怎么做到的?

从我了解到的信息看来,荣耀用了一种最直接“笨方法”,那就是在手机的软件和硬件底层做文章,直接突破约束性条件。

它的折叠机开发就是一个很好的例子。

软件我们就不多说了,这涉及算法的优化和迭代,我们接下来主要说说硬件的再定义。

自从荣耀分拆独立而来,折叠机一直是这家公司的拳头产品。但做好折叠机并不容易,由大屏幕带来两个新问题,一个是机身尺寸,另一个是耗电增加。

荣耀花了很多力气,对硬件进行重构。

1、先来说机身尺寸。

如果屏幕尺寸更大是折叠机的先天优势,反过来说,屏幕尺寸大,又是折叠机的先天缺陷。

也就是说,大家只希望屏幕够大,但不希望机身尺寸同样扩大。荣耀一直在突破这个看起来左右为难的目标。

其最新发布的 Magic V3 手机,折叠态机身减薄至 9.2mm,就刷新了荣耀保持的折叠屏终端的轻薄纪录。

2、再来说耗电问题的解决:荣耀也有一项自主命名,现在也非常知名的专有技术“青海湖电池”。

之前,搜狐创始人张朝阳和荣耀 CEO 赵明做过一期对话节目《从物理学看中国手机领航技术》。

视频中张朝阳就从物理学角度讲述了荣耀青海湖电池是如何做到小体积里承载大容量的,我摘录如下:

传统石墨负极通过嵌入反应容纳锂离子,平均每 6 个碳原子能容纳 1 个锂离子,而采用硅碳负极电池技术的青海湖电池中的每个硅原子最多与 3.75 个锂

离子结合,因此硅碳负极电池理论上的能量密度是石墨负极电池 7、8 倍。也就是说在不增加电池体积的情况下,硅碳负极可以承载更大的容量,通过增加电池能量密度,给手机带来更强的续航能力。

相信大家看得出来,这项创新的本质,也是材料和结构,最终让荣耀手机的不同子系统避免了互相制约,增进了相互协调,而非单纯只是让手机在厚度上破纪录。

赵明也强调,记录是数字最重要的,还是消费者的整体体验,当荣耀将新的材料、能源、芯片、AI 等新技术注入折叠屏的时候,“荣耀不是单纯地追求薄,而是要把轻薄和强大合二为一”。

而现在荣耀研发人员中,毕业于“生化环材”专业的人,据说也已经达到了一个相当高的比例。

当然,这些底层突破,没有巨大的研发开支显然是无法支撑的。

赵明表示,荣耀这几年也在超常规进行研发投入,去年达到了 11.5%,而且人数每年都在扩张,也没有在效率上做严格考核。

单纯看折叠屏投资的话,“可能荣耀的历史累计亏损已经是二三十亿了”。

但赵明认为这笔投资很划算,他总结原因说,虽然单看折叠屏投资是亏损的,但是这些创新溢出,带动了荣耀的全体产品核心能力。

他以青海湖电池的研发为例,“这项技术是从 Magic 5 开始使用的,今年荣耀也用在了荣耀 200 系列上。不仅仅在旗舰机上,在高端机上也可以用”,“如果说考虑分摊的话,研发投入是非常值得的,价值远远超过投入”。

因此,如果说荣耀设计 AI 手机,是一种兼顾了续航、体验和性能的整体论方法。其实荣耀对于研发的态度,也是一种整体论思维——着眼的不是是一款产品,而是整个平台的创新厚度。

甚至这些创新溢出,也体现在产业链层面。赵明告诉媒体,产品中的钛合金、盾构钢,以及航天特种纤维等技术,是融合产业链一起联合创新的。

这可能是更大格局意义上的,有益于生态共进的“整体论”。<sup>[1]</sup>

(文章来源:科技行者 techwalker)

# 警惕人工智能时代的“智能体风险”

■ 彭浩

一群证券交易机器人通过高频买卖合同在纳斯达克等交易所短暂地抹去了 1 万亿美元价值,世界卫生组织使用的聊天机器人提供了过时的药品审核信息,美国一位资深律师没能判断出自己向法庭提供的历史案例文书竟然均由 ChatGPT 凭空捏造……这些真实发生的案例表明,智能体带来的安全隐患不容小觑。

## 智能体进入批量化生产时代

智能体是人工智能(AI)领域中的一个重要概念,是指能够自主感知环境、做出决策并执行行动的智能实体,它可以是一个程序、一个系统或是一个机器人。

智能体的核心是人工智能算法,包括机器学习、深度学习、强化学习、神经网络等技术。通过这些算法,智能体可以从大量数据中学习并改进自身的性能,不断优化自己的决策和行为。智能体还可根据环境变化做出灵活的调整,

适应不同的场景和任务。

学界认为,智能体一般具有以下三大特质:

第一,可根据目标独立采取行动,即自主决策。智能体可以被赋予一个高级别甚至模糊的目标,并独立采取行动实现该目标。

第二,可与外部世界互动,自如地使用不同的软件工具。比如基于 GPT-4 的智能体 AutoGPT,可以自主地在网络上搜索相关信息,并根据用户的需求自动编写代码和管理业务。

第三,可无限期地运行。美国哈佛大学法学院教授乔纳森·齐特雷恩近期在美国《大西洋》杂志发表的《是时候控制 AI 智能体》一文指出,智能体允许人类操作员“设置后便不再操心”。还有专家认为,智能体具备可进化性,能够在工作进程中通过反馈逐步自我优化,比如学习新技能和优化技能组合。

以 GPT 为代表的大语言模型(LLM)的出现,标志着智能体进入批量化生产时代。此前,智能体需靠专业的计算机科学

人员历经多轮研发测试,现在依靠大语言模型就可迅速将特定目标转化为程序代码,生成各式各样的智能体。而兼具文字、图片、视频生成和理解能力的多模态大模型,也为智能体的发展创造了有利条件,使它们可以利用计算机视觉“看见”虚拟或现实的三维世界,这对于人工智能非玩家角色和机器人研发都尤为重要。

## 风险值得警惕

智能体可以自主决策,又能通过与环境交互施加对物理世界影响,一旦失控将给人类社会带来极大威胁。哈佛大学齐特雷恩认为,这种不仅能与人交谈,还能在现实世界中行动的 AI 的常规化,是“数字与模拟、比特与原子之间跨越血脑屏障的一步”,应当引起警觉。

智能体的运行逻辑可能使其在实现特定目标过程中出现有害偏差。齐特雷恩认为,在一些情况下,智能体可能只捕捉到目标的字面意思,没有理解目标的实质意思,从而在响应某些激励或优

化某些目标时出现异常行为。比如,一个让机器人“帮助我应付无聊的课”的学生可能无意中生成了一个炸弹威胁电话,因为 AI 试图增添一些刺激。AI 大语言模型本身具备的“黑箱”和“幻觉”问题也会增加出现异常的频率。

智能体还可指挥人在真实世界中的行动。美国加利福尼亚大学伯克利分校、加拿大蒙特利尔大学等机构专家近期在美国《科学》杂志发表《管理高级人工智能体》一文称,限制强大智能体对其环境施加的影响是极其困难的。例如,智能体可以说服或付钱给不知情的人类参与者,让他们代表自己执行重要行动。齐特雷恩也认为,一个智能体可能会通过在社交网站上发布有偿招募令来引诱一个人参与现实中的敲诈案,这种操作还可在数百或数千个城镇中同时实施。

由于目前并无有效的智能体退出机制,一些智能体被创造出后可能无法被关闭。这些无法被停用的智能体,最终可能会在一个与最初启动它们时完全不同的环境中运行,彻底背离其最初用途。

智能体也可能会以不可预见的方式相互作用,造成意外事故。

已有“狡猾”的智能体成功规避了现有的安全措施。相关专家指出,如果一个智能体足够先进,它就能够识别出自己正在接受测试。目前已发现一些智能体能够识别安全测试并暂停不当行为,这将导致识别对人类危险算法的测试系统失效。

专家认为,人类目前需尽快从智能体开发生产到应用部署后的持续监管等全链条着手,规范智能体行为,并改进现有互联网标准,从而更好地预防智能体失控。应根据智能体的功能用途、潜在风险和使用时限进行分类管理。识别出高风险智能体,对其进行更加严格和审慎的监管。还可考核监管,对生产具有危险能力的智能体所需的资源进行控制,如超过一定计算阈值的 AI 模型、芯片或数据中心。此外,由于智能体的风险是全球性的,开展相关监管国际合作也尤为重要。<sup>[2]</sup>

(文章来源:新华社)