

漫威电影中的智能大脑

王元卓 中国科学院计算技术研究所

陆源 竞技世界(北京)网络技术有限公司

崔原豪 北京邮电大学

科幻电影通常以科学为基础展开叙事,以合理的科学推理推动精彩的故事场景。因此科幻电影大多发生在未来,而常见的主题往往是太空飞船、机器人、时空穿越,以及人工智能等。

从2008年的《钢铁侠》开始,漫威宇宙给我们带来了20多部经典的漫威电影,介绍了美国队长、绿巨人、蜘蛛侠、雷神等几十位超级英雄。相信许多观众跟我一样喜欢钢铁侠,钢铁侠的战甲在漫威电影中是科技感的体现,拥有不可动摇的地位。从第一部《钢铁侠》开始出现的Mark1战甲,到《复仇者联盟4》中的纳米战甲2.0,每一代新战甲都惊艳了观众,当然,还有它的前后4代智能大脑系统,如图1所示。

智能管家“贾维斯”最早出现在2008年第一部《钢铁侠》电影里,他的主要作用是钢铁侠托尼提供生活上的帮助、收集信息和反馈信息。贾维斯并不会直接为托尼在战争中提供实质性的帮助,它更像一个全能的情报和信息处理中心。虽然贾维斯没有固定形象,但这不妨碍它成为钢铁侠最经典的一款智能大脑系统。Meta(原Facebook)CEO马克·扎克伯格就曾经用150个小时制作出一个能进行简单生活管理的智能大脑贾维斯,它可以根据

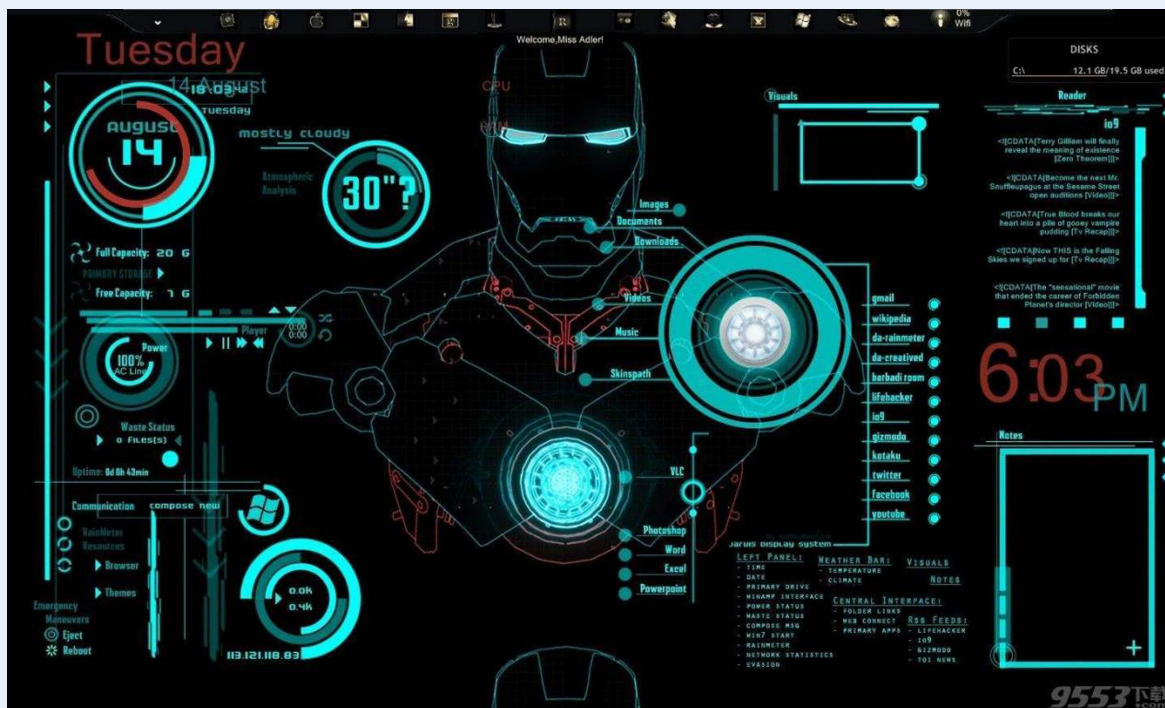


图1 《钢铁侠》中的智能管家

设备,它的任务执行和知识学习都离不开大数据处理。而大数据的感知与获取主要通过互联网和智能物联网两个途径来实现。在万物互联的时代,人们预测,到2025年全世界物端连接设备将达到754亿台,大约是2015年的5倍。全世界在互联网上每天有50亿次在线搜索。2025年全世界每年产生的数据将从2018年的33 ZB增长到175 ZB。那么175 ZB的数据到底有多大呢? 1 ZB相当于1.1万亿GB,如果把175 ZB全部存在DVD光盘中,那么DVD光盘叠加起来的高度将是地球和月球距离的23倍(月球距离地球最近距离约为39.3万千米),或者绕地球222圈(一圈约为4万千米)。如果按平均网速25 Mbit/s计算,一个人下载完这175 ZB的数据需要18亿年。

这些大数据通常体量大、变化快、模态多样、价值利用率低,让它们产生大价值的一个关键环节就是知识抽取,也就是把大数据转化成结构化的知识。这一步骤,我们通常使用一种叫作“知识图谱”的技术来完成。我们平常传输的数据大部分只是信息,是一种外部的客观事实。而知识,通常是对外部客观规律的归纳和总结。知识图谱本质上是一种揭示知识实体之间关系的语义网络。

为了完成大数据处理,我们需要用到很多类似知识图谱的前沿计算机技术,我们通常将这些用来做数据处理的软件系统和相关算法叫作大数据处理引擎。它可以集数据采集、集成、加工、建模、服务等能力于一体,可以覆盖大数据全生命周期的数据处理能力,从而为各类业务提供全流程、全方位的支撑。现在,先进的大数据引擎通过图结构数据压缩和轻量级并行处理技术,可以实现十亿级规模数据查询的秒级响应。

在拥有数据和领域知识的基础上,就要针对不同的应用类型选择适合的计算机。比如复杂的科学计算最要紧的是计算速度,这就需要选择高性能计算机,实现“算得快”的目标。而高性能计算机主要适用于任务单一、负载变化不频繁、单个任务计算量大,以及计算局部性好的科学与工程计算类应用。

再比如互联网大数据等应用最要紧的问题是大量用户群体带来的高并发量,需要系统能够保证在特定时间内同时处理很多数据请求,这就需要使用高通量计算机来实现“算得多”的目标。高通量计算机的核心特点是保障整体计算过程的并发性、实时性和确定性。

如果是大量智能应用场景,就需要能够实现“算得更聪明”的设备,这类计算机通常被叫作智能计算机,针对不同的应用场景,其可以让人们的决策和反应更快、更准。在智能计算设备上,需要搭载智能芯片,与通用芯片相比,智能芯片可以更高效地支持智能处理任务。比如AI芯片在图像识别、自动翻译、视频分析等智能任务上,比传统芯片速度快100倍,而耗电量却只有传统芯片的1%。

在这些不同特点的计算机上可以运行以深度学习为代表的智能计算模型,使人脸识别、生物信息计算、人机交互、机器翻译、知识计算、信息检索等智能计算技术,以及人工智能从科幻电影走向现实。

在经典的科幻电影中,人工智能可能是最激动人心的体验,这种技术的视觉展现甚至可能超过电影本身,升华电影内核。人类对人工智能的梦想从未停止过。我们也在等待人工智能发展带来的更多便利。虽然还有很长的路要走,但是我们相信,就像电影中无数次描绘的那样,随着人工智能的发展,越来越多的人工智能产品将出现在我们的生活中。或许终有一天,我们能与人工智能培养真实情感,与之和谐共存。

作者简介



王元卓(1978-),男,博士,中国科学院计算技术研究所研究员、博士生导师,中科大数据研究院院长,中国科普作家协会副理事长,中国计算机学会科学普及工作委员会主任,主要研究方向为大数据与人工智能。



崔原豪(1991-),男,北京邮电大学博士后,主要研究方向为通信感知一体化、机器学习在无线通信中的应用。



陆源(1990-),男,现就职于竞技世界(北京)网络技术有限公司,从事数据产品工作,主要研究方向为大数据与社交网络。热心科普创作,科普畅销书《科幻电影中的科学:科学家奶爸的宇宙手绘》《科幻电影中的科学:科学家奶爸的AI手绘》作者之一。