

评人工智能向何处去？

(兼谈国内外跟帖评论)

陆首群

2020.10.23

2019年8月8日，针对当时一个世界性的热门话题：人工智能如何走向新阶段？基础理论抓什么？我们感到有必要建立一个公开评论的平台，邀请或吸引国内外专家、草根（不拘一格）以跟帖方式参与讨论集思广益。到2020年6月26日（这时距一年还差一个多月），我们公布了来自国内外的535条跟帖，其中不乏真知灼见，真有点出乎意料！我们创办跟帖讨论平台有一个优势：几十年来我们在国内外结识了不少开源的、ICT的和AI的专家朋友（其中不少大师），以及朋友的朋友；人工智能讨论平台是以国内外专家评论（包括摘引的）为骨架，以发布跟帖方式起来的，这时我们可随时向他们请教或邀请他们发表跟帖参与讨论。

在已发表的535条跟帖中，从评论机器学习/深度学习那些不可解释的人工智能的初级阶段开始，而今天人工智能的繁荣正是基于机器学习/深度学习，说它们已近天花板（已经没有发展潜力）是不妥的。谈到人工智能的出路，归纳跟帖中指出的4条（包括研发其基础理论）：①打破机器学习的黑盒子研发可解释的人工智能，②基于异步脉冲神经网络的神经拟态计算系统，③从知识工程出发，依托大规模语义网络（知识图谱）的支持，破解认知智能解决方案，④脑机结合的理论 and 实践。目前①已有所突破，②已见亮点，③还差最后一公里，④国内外已有几十例试点。

最近，有关专家怀疑由人工智能不同学派提出的上述四个发展途径，他们认为人工智能不同学派互不相容，单打独斗，缺乏整体观，他们只是从不同侧面模拟人类心智（大脑），各自提出的“发展路径”均有片面性。他

们的建议是从改造、融合统一的人工智能发展范式出发，分别提出发展新一代人工智能的解决方案。我们将他们的解决方案收集到跟帖中来讨论。他们提出的方案尚属于构想，尚未见诸于实效，拭目以待！

必须指出，当今世界已进入“量子计算+人工智能+基因科学”新时代，作为新时代的系统工程，我们需要研究开源与新时代的关系。过去我们虽论证开源是人工智能与深度信息技术的基础（人工智能实行开源可克服其发展瓶颈，加快其研发速度，有利于解决维护升级，加快解决 BugFix, Patch），量子计算代表算力，今天量子计算机应该拥有像人工智能及人脑（或人类）一样的自学和思考能力，所以可认定人工智能、开源也应将是量子计算的技术基础之一，即开源是创建“新时代”的技术基础之一。在今后发布的跟帖中，我们也将收集反映“新时代”的风貌。人工智能、基因科学在以往发布的跟帖中已有所反映，量子计算也曾有几条跟帖，今后还将强化。

我们研究了针对“人工智能向何处去？”的跟帖评论要不要继续办下去？

回答是：既然很受欢迎，就应办下去！为此作为“人工智能国内外跟帖评论续集”的开场白告示诸位朋友，“续集”将以 552 条跟帖开始。

536. 由机器学习/深度学习（不可解释）算法技术支持的应用场景：

如语言识别、图像识别、有规则的围棋游戏（AlphaGo）、自动驾驶、自然语言处理、换脸技术、新材料、新药物等独领风骚的应用场景；

又如不耐药的新一代抗生素、英国第六代战斗机“暴风雨”设计、AlphaFold 预测蛋白质三维结构（向 AI 基因医疗领域挺进创造条件）、苹果的 face-id(使脸部识别达到支付级别)、加州理工学院全球最快相机（创纪录每秒捕捉 70 万亿点惊人速度）、美国加速器实验室加快加速器调谐过程（提高 5 倍，大大提高实验室工作效率）、英国研制可实时同步扫描心血管流成像技术（提高医疗诊断治疗效率）等应用场景均基于机器学习。

可查看 1-536 条中的跟帖。

537. 利用机器学习方法，确定导致结核病细菌生存所需蛋白质为靶标的化合物

MIT 利用一种包含不确定性的机器学习方法，确定了几个有希望以导致结核病的细菌生存所需的蛋白质为靶标的化合物。为克服算法在分析与训练数据相似性较低的数据时未能表现良好的问题，研究人员使用了一种称为高斯过程的技术，为算法所训练的数据分配不确定值，使得模型在分析训练数据时考虑预测的可靠性。例如，如果进入模型的数据可以预测特定分子与靶蛋白的结合强度以及这些预测的不确定性，则该模型可以使用该信息对以前从未见过的蛋白质与靶蛋白的相互作用进行预测。该模型还估计其自身预测的确定性。当分析新数据时，模型的预测可能对与训练数据相差较大的分子有较低的不确定性。研究人员可以使用这些信息来帮助他们确定要对哪些分子进行实验测试。此外，该算法只需要少量的训练数据，而

且一旦研究人员获得了额外的实验数据，他们就可以将其添加到模型中并对其进行再训练，从而进一步改善预测。目前，在该模型确定度最高的分子中，大约有 90% 被证实是真实命中，远高于现有用于药物筛选的机器学习模型的 30%-40% 的命中率。该方法还可以应用于许多其他类型的生物学研究。

摘自;MIT News

<https://news.mit.edu/2020/gaussian-machine-learning-tb-drug-1015>

538. 机器学习/深度学习是一个强大的数据分析工具，带动了今天人工智能的繁荣。但机器学习/深度学习是有缺陷的，它本质上是一项暗箱操作（黑盒子）技术或盲模型，其训练过程是不可解释、不可理解的，缺乏类似人脑的推理能力。在 527 条跟帖中，图灵奖得主、贝叶斯网之父 Judea Pearl 谈到，当前机器学习理论有局限性，完全以统计学或盲模型（黑盒子）的方式运行，所以不能成为强 AI 的基础。

在 1-535 条跟帖汇集中我们也能查到：ACM 图灵获奖得者、算法大师约翰·霍普克罗夫在 2019 年谈到，对深度学习这个黑盒子，人们知道它在学习，但不知它怎么学习，人类可能会在 5 年后大体读出深度学习的数学理论。

当下，打破机器学习黑盒子研发可解释的人工智能已成为一股热潮：

IBM 人工智能研究所 Vijay Arya 研发团队（19 人），为打破机器学习黑盒子研发可解释的人工智能，针对信用、打假、反篡改、公平等 4 个案例（课题），率先研发了可解释工具包 AIX360，目前已研发了 10 种算法。

马格德堡大学人工智能实验室 Sebastian Stober 团队也在研发可解释的人工智能技术，其研究项目的编号为为 CogxAI。

他们用分析深度学习人工神经网络的方法打破黑盒子，在认知神经科学启发下研发可解释、可理解的人工智能，使其学习、训练过程变得更加透明和容易理解。

539. **可信任人工智能（反欺诈、可解释、公平性）**

打破机器学习黑盒子研发可解释人工智能
IBM 开放技术副总裁 Todd Moore

（在《第十五届开源中国开源世界高峰论坛》线上会议上的报告）

大家好。我是 IBM 开放技术副总裁 Todd Moore。今天，我想和大家谈谈对我来说很重要的事。此话题我之前在 Linux 基金会的一个开源活动中也谈到过。那就是《可信的人工智能》、《负责的人工智能》。

当今世界正处于新冠肺炎大流行病的高危之中，我们面临着一场全球性健康危机，而要有效地拯救生命，这取决于数据，取决于我们如何收集这些数据，取决于我们如何有效地利用这些数据。同时，我们正处于一场围绕种族和平权的全球运动之中。这些事件相互交叉、影响。它们影响到我们如何利用人工智能、如何利用我们的模型、我们的数据，以及我们将如何从这里走向未来。

IBM 的核心价值观根植于权利平等。这些 IBM 长期以来的公司政策，可以追溯到 1953 年，托马斯·约翰·沃森那时所写的（IBM 公司）政策令第 4

号，确保 IBM 作为雇主将提供平等机会，确保 IBM 不会容忍因种族、肤色或宗教而产生的偏见。

今天，阿尔温德·克里希纳也已肩负这项使命。阿尔温德日前宣布 IBM 不再提供面容识别服务，因而切断了一条可能会因其使用而在社会中造成偏见的途径。他也将这一问题提升为一项公众对话，以探讨这项技术是否有意义，我们的社会应该如何应对。但是，人工智能的可信、透明不仅仅是面容识别，当我们收集的数据一旦投入使用，隐私问题就呈现出来。伦理问题关乎这些数据如何使用、关乎如何转化、关乎我们的社会如何应对：所有这些方面涉及的伦理问题都亟需解决。

IBM 为此提供了一系列工具。我们也将其中许多项目开源了。我们能够提供一个由一系列 Python 库组成的《稳健性工具包》，我们能够提供一个能在模型和数据中发现偏见的《公平工具包》。我们能够提供一个《可解释性工具包》，它可以让你能够快速检查你的潜在 AI 模型的黑匣子，从而把这个黑盒子变成一个你能给识别其行为方式的东西。这些都是帮助我们消除偏见必不可少的工具。

同时，我们也在开始一项我们称之为 AI FactSheets 的工作，这个项目是关于我们如何追究模型的责任。想想看，就如同食物产品上贴的标签（标签体系）花费了很长的时间才开发出来，但一旦有了这样的标签，你会从中得到相当多的信息，比如手里这罐汤，到底含有什么东西。这也是我们研发这个 AI FactSheets 项目希望达成的效果。

现在，这些项目已经取得了进展。我们参加了 Linux 基金会人工智能组织（LF AI）。我们希望这些项目都置于其开放治理模式之下，（LF AI 的）

技术顾问委员会（TAC）目前已经投票赞成孵化这些项目。大家将看到其中三个项目已经发布，并通过 LF AI 机构开源。我们非常感谢大家这样做。我们认为这是一个很好的场所能让其他人可加入进来公开影响和推动这些工作的进展。

我们认为这对社会很重要。LF AI 这个组织的团队已经有不少非常重要的公司和大学参与其中，此外还成立的《可信人工智能委员会》，我们希望大家也都能参加进来。因为他们一直在制定（人工智能）的基本原则，而我们将充分准照这些基本原则，使用由此而研发的 AI 技术所产生的数据和模型。这个工作组宣布了七项基本原则：即公平性、再现性、透明性、治理、隐私、安全和问责。这些是这个工作组正在研究的七项基本原则，与时俱进中。所以，现在就和我们一起来吧。

自从上次我们谈过这件事以来，就如我刚刚提到过的，AI FactSheets 已经提上议事日程，我们已经开始寻求其开源途径。AIF360 团队与 Scikit Learn 和 R Studio 展开合作，现在已经可以在（Learn 和 R Studio 中）使用了；Kubeflow MLOps 流水线也已经发布，我们可用这些流水线进行公平性和对抗性检测；Apache NiFi 也已经接纳 MLOps，现在我们也有了可用的（Apache NiFi）处理器。所有的项目也都取得了很大的进展，感谢一直以来为此做出贡献的人们，这些真的对我们的当今世界非常重要。作为小结，大家今天看到了不少东西，我们很高兴也能够提供非常有价值的（开源）社区提供支持，您如果有意愿，请加入这些项目或者加入 LF AI 组织的可信人工智能委员会，帮助我们确立发展方向。现在是至关重要的时刻，这个话题将帮助我们，作为世界性组织机构，作为开源贡献者，推动世界向前发展。非常感谢大家！

在今年 7 月 15 日召开的《第 15 届开源中国开源世界高峰论坛》（线上会议）取得了很大成功，海内外舆论反响热烈！

会上报告亮点很多！仅以那些环绕研发人工智能发展前沿的报告为例，如打破机器学习黑盒子发展可解释的人工智能（IBM 推荐十种破解黑盒子的工具或算法），系全球首创；又如英特尔研发基于异步脉冲的神经拟态计算系统，是当前人工智能的发展前沿；还有如基于知识工程，研发以大规模语义网络求解认知算法，阿里达摩院的报告虽然还差最后一公里（目前全世界研究水平均如此，尚未能突破），但还是对人工智能发展前沿的研发进行了有益的探索。

这里仅就基于异步脉冲神经网络的神经拟态计算系统研发动态作一简报。

异步脉冲神经网络（SNNs）是模拟人脑神经元连接、信息处理和运行方式的模型，以实行神经拟态计算或类脑计算，这也是人工智能的一条升级之路（从弱人工智能跃升为强人工智能）。

在传统计算（非人工智能）领域，乃至在弱人工智能领域，是采用计算机运作机制、传统数据处理通用算法，主要依靠硅基芯片微处理器，遵循冯·诺伊曼计算架构；在强人工智能领域，是采用模仿人脑的运作机制、神经拟态或类脑算法，将模仿人类大脑功能的神经拟态计算技术映射到神经拟态芯片（硬件）上，遵循神经拟态计算架构（或新型类脑计算架构）；英国曼彻斯特大学研发的 SpiNNaker，是一种新颖的大规模神经拟态计算+并行计算体系，其灵感来自人脑的基本结构和功能，因此也可归类于新型神经拟态计算或类脑计算架构，其中的并行处理可以认为，是为实现大规

模神经网络的仿真提供高性能大规模并行处理平台；在人类智能（智慧）领域，则依据人脑生物学的运作机制进行计算。

神经拟态计算或类脑计算是下一代人工智能的重要方向（或主流模式）。

当前计算技术发展遇到瓶颈：一是摩尔定律将失效，神经拟态芯片将取代传统的硅基微处理器芯片；二是冯·诺伊曼计算架构（读取、存储、处理数据能效低下）将被新的计算架构（类脑计算架构或神经拟态计算架构）所取代。

在进行学习、认知、处理等复杂计算时，人脑计算功耗最少，约 3-5W，类脑计算功耗约 350-500W，而标准计算机能耗约 50-200KW。据 Loihi 神经拟态计算研发团队报告：神经拟态计算比传统计算，速度快 1000 倍，能效高 10000 倍。神经拟态计算依靠神经元、突触进行信息编码和处理，神经元（计算）、突触（存储），计算存储不再分离，配置一体化，传输依靠脉冲（尖峰）神经网络，当尖峰运作时才进行信息传输，如此的传输机制有利于节能。

基于异步脉冲神经网络的神经拟态计算与传统计算运行方式比较表

	信息源	编码方式	传递方式	基于不同特征的计算方式
传统计算运行方式	数字信号 计算机程序	由数字源代码变换 0,1 的机器码进行编码	传统网络权重连接+激活方式，对机器码进行信息传递	传统计算机运行方式，符合冯·诺依曼硬件和软件计算架构
神经拟态计算运行方式	神经电脉冲信号和化学信号	采用稀疏脉冲时序编码机制	通过由仿脑自然神经元+突触组成的脉冲神经网络和运行方式模型，对神经电脉冲进行信息传递	神经拟态计算系统运行方式，打破冯·诺依曼硬件和软件计算架构

人脑或类脑计算模式与硅基计算机原理之间差别很大：

- ①人脑或类脑计算中，计算（神经元）和存储（突触）是一体化的，融合在一起的，训练和推理也是一体化的，在硅基计算中，计算（处理单元）和存储（存储单元）是分离的，训练和推论也是分离的；
- ②在人脑或类脑计算中，存在三维广泛连通性，在硅基技术上尚无法模拟三维连通，受限于二维连接；
- ③在人脑或类脑计算中，是基于脉冲的事件驱动型的随机计算，在硅基技术中主要为了构建确定性计算采用晶体管间布尔代数电路开关。
- ④在人脑中是生物学自主低能耗类型，在类脑计算中实行节能型，在标准计算机中能耗大。
- ⑤在人脑或类脑计算中，将模仿大脑的大规模并行通信方式（或结构），将数十亿信息同时发送到数千个不同的目的地，不会像传统通信模式那样通过标准网络从A点到B点发送大量信息来进行通信。

目前国内外研发异步脉冲神经网络与神经拟态计算（或类脑计算）系统案例如下：

①英特尔案例

2017年发布Loihi脉冲神经网络（SNN）芯片。

芯片制程14nm，管芯尺寸60mm，每颗芯片包含20多亿个晶体管。

每颗芯片128个核，每个核支持1000个神经元，每颗芯片支持13万个神经元、1.3亿个突触。

2019 年，英特尔将 Loihi 芯片集成为 PohoikiBeach:

64XLoihi=800 万个神经元，组成神经拟态网络。

2020 年 3 月 19 日，英特尔将 Loihi 芯片集成为 PohoikiSprings:

768XLoihi=1 亿个神经元，组成神经拟态网络（相当于小型哺乳动物大脑）。

应用场景

英特尔以开放姿态成立 INRC（神经拟态研究开源社区）以期在应用方面实行突破，做好应用落地。

目前其神经拟态计算切入点为：

① 基于神经拟态计算开展超算应用开发，

② 针对非结构化数据、实时性要求高的场景进行应用开发，

③ 针对多模拟、实时的场景进行引用开发

（如机器人、无人机等需持续学习或自适应学习的场景）

②浙江大学案例

似乎是模拟英特尔的案例

2015 年 12 月发表达尔文—1 脉冲神经网络芯片。

2019 年发表达尔文—2 脉冲神经网络芯片：每颗芯片 576 个核，每个核支持 256 个神经元，每颗芯片支持 15 万个神经元。

2020 年 9 月 1 日，浙大将达尔文—2 集成为：

792X15 万=1.2 亿个神经元（相当于小鼠大脑）。

应用场景正在起步

④ 国曼彻斯特大学案例

由曼彻斯特大学、南安普敦大学、剑桥大学、谢菲尔德大学、ARM、Silistix 公司、泰雷兹等共同研发，2016 年发布 SpiNNaker—1 脉冲神经网络（芯片）和类脑计算机，每颗芯片支持 1.6 万个神经元，曼彻斯特大学将 SpiNNaker 芯片集成为：30KX16K=4.6 亿神经元（SpiNNaker 的模拟规模相当于一只老鼠的大脑或十分之一个人脑。

541. 异步脉冲神经网络与神经拟态计算系统

神经拟态计算技术与非人工智能的传统计算技术（及具有人工智能加速技术的传统计算系统）的区别在哪里？

有人说：“目前的人工智能本质上还是计算机科学的一个分支（现在国际上将人工智能的论文都统计在计算机科学的前沿研究中）”，对这种模糊的说法，似有写这篇补充的必要。

在这里我们还是引用《评人工智能如何走向新阶段？》（国内外跟贴留言）第四集（386-486 条）中的 438 条来自国外的留言：

神经拟态计算技术（新一代 AI）PK 传统人工智能（深度学习 AI）加速技术

神经拟态计算技术与传统的人工智能加速技术相比，有哪些技术特点？两者有哪些相同或不同（包括所解决任务的不同）？

神经拟态计算技术与人工智能加速技术两者都是为了处理人工神经网络而设计的，同样是支持大规模运算任务的前沿计算技术。

传统设计的 AI 芯片，即人工智能加速器是一类专用于人工智能（如人工智能网络、机器视觉、机器学习等）硬件加速的微处理器或计算机系统。据英特尔神经拟态计算实验室主任 Mike Davies 谈，英特尔不打算做面向特定应用的专用芯片，而是希望能让神经拟态计算实现通用性。

神经拟态计算与传统人工智能加速器主要差异在哪里？

先从两者定义出发：

传统的 AI 芯片是一类专用于人工智能硬件加速的微处理器 (CPU、GPU) 或计算机系统。所谓神经拟态计算技术是将人类大脑功能映射到硬件上，即直接用硬件（神经拟态芯片）来模拟人类大脑结构（这种方法叫神经拟态计算）。

两者最大区别在：

神经拟态技术研究类脑芯片严格按人脑的机制，将训练和推理整合到一块芯片上，实现存储与计算融合，而基于冯-诺依曼计算架构设计的 AI 芯片，它并未严格按人脑机制，其训练与推理是分离的，存储与计算也是分离的。这样神经拟态计算技术变得更加智能、快速和高效，与传统处理器比速度快、能效高，神经拟态计算相较于传统架构的优势不在于吞吐量而在于超低延时和超低功耗。

所谓 Loihi 芯片拥有并行性和异步信号传输能力，在这里谈到并行计算，公平地说两者是一样的，但神经拟态计算自然拥有异步信号传输能力的优

势!它像人脑一样，根据环境的反馈来自动学习如何操作，随时间推理变得越来越智能、快速和高效。

人类大脑神经网络通过脉冲传递信息，根据脉冲时间调节突触强度或突触连接的权重，并把这些变化存储在突触连接处。在硬件层面上能不能复制大脑神经元组织、通信和学习方式，打造出一种更加智能的系统呢?这种颠覆传统计算架构的新型计算模式就是神经拟态计算。

以往国内外都在开发传统的专用 AI 芯片（人工智能加速器），近来神经拟态系统的研究在国外开始火红起来，英特尔、IBM、惠普 MIT、普渡大学、斯坦福大学、浙江大学等研究机构，均在紧跟开展这方面研究。

542. **神经拟态计算，探索智能互联时代的计算创新**

英特尔中国研究院院长 宋继强

（在《第十五届开源中国开源世界高峰论坛》线上会议的报告）

大家上午好，非常高兴来到这次开源中国开源世界的高峰论坛，跟大家分享英特尔在神经拟态计算方面的进展。我今天的报告主题是神经拟态计算探索智能互联时代的智能创新。

随着无线通信技术和人工智能技术的广泛应用，现在信息世界里的数据呈现了三大主要变化，首先是数据量，有指数级的爆发，最开始都是人工产生的数据，人往系统里输入数据，后来有了多媒体的数据。现在这个阶段，我们发现能够产生数据捕捉数据的设备越来越多。除了人使用的设备之外，

还有大量物联网设备产生的数据，还有人工智能的很多模型自己也有很多数据，包括参数和训练的结果，这些称为原数据。

这些数据量上升很快，形态跟原来的结构化数据有很大的不同，结构化数据指的是人为设定的数据结构，比较易于管理和存储，但是非结构化数据通常是来自很多自然界的的数据，有很多处理的变化。

同时，很多数据刚才讲的来自于物联网设备的，来自于前面实际应用场景的，需要在终端或者说边缘侧被采集，同时有实时处理再反馈到边缘侧的需求，也就是说我们对于数据实时处理的需求也在快速增长。

应对这样的变化，我们觉得未来的十年，处理这些数据需要的计算架构，就不会是单一的计算架构，而是需要好多种不同计算架构去混合处理。架构创新将会成为计算创新的关键驱动力，目前来看主流的计算架构有这么四种，图上表示了四种主流计算架构是标量的、矢量、矩阵和空间。标量指的一个一个去计算数据，有前后的顺序。矢量就是把一组数一起来做，就像把两个数组直接相加相乘。矩阵把一块数据同时做运算。空间说明这些数据本身有很强的稀疏性，它的处理需要采用特殊的方法加速。现在的人工智能领域的数据处理，至少是矢量级别的，很多是矩阵的，在某些应用场景需要空间数据结构来加速。

这些整合在一起，称为 XPU 超异构计算，X 代表有多种可能性。超异构是什么概念？传统异构领域，要么把不同计算的能力，不同计算架构整合到一个单芯片上去，那是传统的 SOC 的做法。或者把不同的芯片，比如说 CPU、GPU、IPGA 整合到一个板上，板级的整合，这是另外一种异构计算。这两种各有优劣势。

超异构计算希望把他们的优势集中起来，而把不足之处去掉。也就是说我们可以把已经做好的很多种不同架构的加速的芯片单元通过异构封装的方式组合到一个芯片里面，把高性能的芯片之间的传输通道建立起来，这样就避免了在板级异构的方式，牺牲了空间也牺牲了传输的效率，同时又可以把很多不同架构芯片整合在一起，小空间、低功耗的情况下提高计算的性能和能效比。

这个图是直接的例子，就是我们怎么样把多种不同的处理器放到一个大的芯片封装里面，利用上高速芯片间的 IO 还有封装领域的高级技术，这是非常重要的技术趋势。

继续往前看我们觉得架构方面需要更多的探索，现在常用的两种一种叫标准计算的思路，我们已经知道一个问题，同时我们也能够对这个问题进行很好的建模，我们可以用一个流程图的方式把怎么处理它的步骤写出来，根据它去编程，这是由人工选定怎么做，人工实现怎么把它实施到软件或硬件中去，我们称为标准计算模式。

第二种模式是现在神经网络给我们带来的，也就是说我们知道输入的数据是什么，我们也知道想产生的结果是什么，但是中间的计算模型我们不清楚怎么做，人工很难去描述它。比如说人脸识别，到底是怎么算的模型，不允许我们描述。但是我们有大量的数据，可以端到端的进行训练，有神经网络、神经网络模型可以作为框架进行训练，最后训练出一个符合特定任务解决的方式。这种是叫数据驱动的计算的架构和模式。

第三种未来要更多考虑怎么把现在已经有的优势整合起来，因为标准计算来自于人类已经有的知识，来自于一些推理。而神经网络利用的是数

据本身内涵的规律，计算的性能要求很高，对内存的使用要求很高，相应的功耗也会很高。如何能够达成优势互补，同时我们在最后使用的时候，又是能源友好，因为我们知道人类大脑只需要 20 瓦就能完成很多不同种类的事情，既可以做标准的逻辑推理运算，也可以做深度神经网络擅长的视觉识别、语音识别的运算。

我们要把它整合到一个架构里去，希望能够有一个能效比高的方案，同时又能够支持扩展。神经拟态计算是这方面的可行之道，我们来介绍一下它的进展。Loihi 神经拟态芯片，用了整合计算和存储架构来做的，不像传统计算一样，有 CPU，有内存，中间有数据传输通道，把计算单元和需要的存储空间紧密的整合在一起，形成非常多的小的核。

一个芯片里面就有 128 个小的核，每个核都是一样的，叫同构多核，每个核里可以做 1000 个神经元的模型，这个神经元就是在模拟人脑神经元的结构和数据处理和存储方式。每一个神经元又可以和另外的 1000 个神经元互相连接，也就是说构造人脑里面类似突触这样的概念。这样一个小的芯片就可以模拟人脑的 13 万个神经元，并且可以产生 1.3 亿个突触，已经是一个比较复杂的神经网络的突破。

它体积很小，能耗很低，需要使用的时候，只有工作那部分电路才耗电，其他电路可以保持非常低功耗的休眠状态，它是一个真正的绿色芯片，当世界越来越多的使用人工智能的时候，我们会需要这样高能效比，有自己的学习和省电能力的芯片。

我们也认为这是一种助力 AI 从 2.0 时代向 3.0 时代迈进的，为什么？因为 AI 的 2.0 时代更多利用数据驱动，做端到端的学习，去学习对识别，

对大规模计算领域，数据量比较大领域的好用的模型。AI3.0 时代，我们希望 AI 能力可解释，可以通过少量的数据去自己学习，不一定需要非常大量的数据，因为很多领域大数据是不一定容易获得的。同时，又希望它能够比较鲁棒，对恶意输入的数据，有抵抗能力。同时希望把多种不同的学习能力整合在一起，也就是说这么多种能力，希望在 3.0 时代获得。

Loihi 因为灵活的架构，可以支持多种事件驱动的学习模式。比如常规使用的无监督学习模式，这是在 2.0 时代深度学习已经非常常用的模式，一种是给出目标的具体的标签，另外一种是不用给出标签，主要做分类。

右边这两种是在未来各种各样的自主系统里面非常有用的，一种叫自监督学习，是关联不同的事件，去产生出新后的因果顺序，并且找出可能的关联，并且强化它。强化学习应用在自动驾驶、机器人领域或者交互式推荐这些方面，它是利用的观察到事件和作出的决策之间的关系，去反复做迭代，优化出一个最好的决策。这么多种不同的学习模式，在 Loihi 上面都可以编程支持，并且同时去学习。

Loihi 系统从 2017 年推出第一个单芯片，后面逐渐扩大它的规模，大家看这条线，从这几年来来看，我们已经能够用 768 个芯片集成在一个大的系统里面，这样可以提供接近 1 亿个神经元的规模，这样的规模已经接近于小的仓鼠哺乳动物的脑。有这样的规模未必现在有这个智能，这还要靠很多软件、很多算法去达到这样的能力，但是我们已经可以提供这样的规模，就是硬件的基础。

同时，这件事情本身又不是英特尔一家可以完全搞定的，还是需要广大的研究社区，还有很多企业参与，去寻找最好的应用领域和未来很多组合式

创新的方法。英特尔成立了一个研究社区，是把大家聚合起来，一起来去推动这个领域的发展，这个研究社区是面向全球的。

我们现在可以提供计算的平台和能力，不光是提供在设备上直接的使用能力还提供在云端使用这些服务和测试系统的能力。举一个使用的案例，这是低功耗使用的案例，最近我们也发表了一个合作，用一片 Loihi 芯片训练，让 Loihi 系统拥有了嗅觉。可以对十种不同的气味做识别，而且通过每个气味单一样本来训练，就可以让识别量达到 92%。传统深度学习的方法，至少要用 3000 倍以上的训练样本，才能达到这样的识别效率。同时，功耗也很低，Loihi 用的功耗非常低。

对于崭新的硬件架构，如何对它编程进行很好的能力释放是非常重要的，在结合之前我们讲的 GPU、硬件加速器，如果把它们统一做成一个异构系统，如何能让开发者或者说科研人员更方便的去使用它，做更多的创新，因为我们说未来不可能有人把所有的底层硬件都了解的很清楚。

对于英特尔来讲，这件事情非常重要，异构计算需要统一可控的编程模型，也需要软硬结合的方法去创新。我们针对不同硬件底层会提供直接的开发的性能库，包括底层的编程接口，但是在上层，对开发者来讲提供统一的 API，把这些能力释放出来，让在不同硬件之间去迁移能力变得非常简单和高效。这块就是我们称为通过 oneAPI 软硬融合，释放异构计算潜能的概念。希望大家来参与这方面的交流和合作，因为 oneAPI 是面向开源领域的。

最后总结一下，英特尔是通过创造改变世界的技术来造福地球上的每一个人，我们非常希望通过异构计算，通过开源的技术和平台，能够让新的技术快速覆盖到全球，产生造福人类的技术。谢谢大家。

543. 2018 年 11 月 SpinNaker 类脑计算机首次亮相，为实时大规模神经网络仿真提供高性能处理平台。

这是一种新颖的大规模神经拟态计算+并行计算的体系结构，打破冯-诺伊曼计算架构和常规超级计算规则，它采用适合于大型尖峰神经网络建模（或基于脉冲神经网络）进行信息传送，SpinNaker 可实时模拟大量生物神经元，其规模相当于一只老鼠的大脑，或千分之一个人脑。

它不同于传统计算机，它是一台由百万处理器驱动的计算机，更像人类大脑，采用神经拟态+平行处理以与人脑类似的方式传递数据，它不会通过标准网络从 A 点到 B 点发送大量信息来进行通信，相反，它模仿大脑的大规模神经拟态+并行通信体系结构，同时将数十亿的信息同时发送到数千个不同目的地，实现构建用于实时的大脑建模应用程序的 100 万个核心，利用该机器，研究人员将能够模拟 10 亿个简单的神经元或数百万个具有复杂结构和内部动力学的神经元。

SpinNaker 项目启动于 2006 年，由曼彻斯特大学、南安普敦大学、剑桥大学、谢菲尔德大学及 ARM、Silistix、泰雷兹等公司合作研发。

类脑计算机主要用于处理其擅长的人工智能任务，其神经拟态计算架构打破冯-诺伊曼计算架构，但不会替代冯-诺伊曼架构的传统计算机，两者互

补、融合可能会是未来的趋势。总的来说，在国际上类脑计算发展还处于初级阶段，未来走向成熟需要硬件、软件和算法的进步。

544. 2020 年 9 月 1 日浙江大学联合之江实验室发布：成功研制全球神经元规模最大的类脑计算机 DarwinMouse。

该类脑计算机含 792 颗达尔文—2 代类脑芯片（2019 年 8 月发布，15 万个神经元，相当于果蝇），该类脑计算机在对芯片集成处理后具有 1.2 亿个脉冲神经元和近千亿个神经突触，堪比小鼠大脑神经元规模，典型运行功耗仅 350-500 瓦，其运行机制与 SpinNaker 相同，但浙大声称完全是基于自主知识产权类脑芯片的类脑计算机。类脑计算机与传统计算机差别非常大，传统计算机的算法和软件无法在类脑计算机上运行。类脑计算机主要用于处理人工智能任务。

545. 2020 年 10 月 14 日清华大学计算机系张悠慧团队、精仪系施路平团队合作在《Nature》杂志上发文，提出类脑计算完备性（类脑计算即神经拟态计算）以及软硬件去耦合的类脑算法层次结构（首次提出）。

与通用计算机的“图灵完备性”概念与“冯诺依曼”体系结构相对应，本文题为《一种类脑计算系统层次结构》（Asystemhierarchyforbrain-inspiredcomputing）的论文，通过理论论证与原型实验证明，该类系统的硬件完备性与编译可行性，并扩展了类脑计算系统使之能支持通用计算。类脑计算处于起步阶段，国际上尚未形成公认的技术标准与方案，这一成

果填补了完备性理论与相应系统层次结构方面的空白，利于自主掌握新型计算机系统核心技术。

近年来类脑计算研究受到了越来越多的关注。类脑计算是借鉴生物神经系统信息处理模式和结构的计算理论、体系结构、芯片设计以及应用模型与算法的总称：在未来 10-20 年内，谁要引领世界经济，谁就必须在这个领域领先。

现在类脑计算系统方面的研究多聚焦于具体芯片、工具链、应用和算法的创新实现，而对系统基础性问题，例如计算完备性、系统层次结构等思考不足，导致软硬件紧耦合、应用范围不明确等一系列问题。但从现有通用计算机的发展历史与设计方法论来看，完善的计算完备性与软硬件去耦合的层次结构是计算系统蓬勃发展的计算理论与系统结构基础。

图灵完备性和冯诺依曼体系结构是通用计算机技术能够飞速发展并持续繁荣的关键因素——几乎所有的高级编程语言都是图灵完备的，冯诺伊曼架构通用处理器则可以通过图灵完备的指令集实现图灵完备性，这意味着编程语言编写的任何程序都可以转换为任意图灵完备处理器上的等价指令序列（即程序编译）。这样，由软件层、编译层、硬件层组成的计算机层次结构就能够确保应用软件、指令集、硬件设计在独立发展的同时相互兼容（即软硬件去耦合），为整个领域的繁荣发展打下了系统基础。

“完备性”可以回答系统能够完成什么、功能边界在哪里等问题。研究完备性，可以为软硬件系统的解耦合、划分不同研究领域的任务分工与接口提供理论基础。

546. 研发通用人工智能基础理论推出新一代人工智能

陆总，您寄来的《异步脉冲神经网络与类脑计算系统研发动态》收到了，非常感谢！

近年来，国内外人工智能研究高度活跃出现了许多新方法和新进展，令人欣喜。在人工智能发展的历史上有三大学派，他们鼎足三分，各自取得不少精彩成果，但三派争霸，无法实现统一。三大学派各自采取三种不同的研究路径（三种发展范式）：①是模拟人类大脑皮层生物神经网络结构的人工神经网络研究（也称结构主义研究路径），②是模拟人类逻辑思维功能的物理符号系统和专家系统研究（也称功能主义研究路径），③是模拟智能生物行为的感知动作系统研究（也称行为主义研究路径）。不同学派互相竞争，各自发展，均存在很大片面性。不过根据我五十多年研究所积累的知识来判断，这些新方法和新进展不可能取得人工智能的根本性突破，原因是他们都没有抓住人工智能学科范式（即其科学观和方法论）的有机整体这个龙头。

要想研发新一代通用人工智能，必须对学科范式及基础理论进行颠覆性变革，改革并融合三种不同的发展范式，并从统一的人工智能发展范式出发，正确把握和处理如下四个问题：即①纠正 AI 研究范式张冠李戴，②把握信息科学研究范式，③立足于研发通用 AI 基础理论，④推动中华文明优势伟大复归（于范式变革中）。谨向您作详细说明，并把我的一个 PPT 寄给您，供您仔细研究。关于新一代人工智能发展途径详见附件一。

请陆总批评指正！

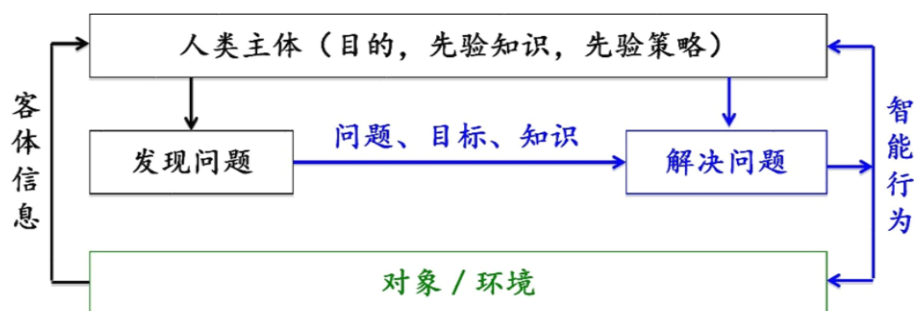
钟义信
北邮人工智能学院
2020.9.15

547. 附件一： 人工智能范式的颠覆性变革催生了“通用的人工智能基础理论”

本研究始于1965年，
至今已完成理论研究；
正转向原型开发阶段。

人类智慧 → 人类智能 → 人工智能

人类智慧：人类独具的卓越能力 — 为了实现生存与发展的目的而不断运用知识去发现问题、解决问题、发展自己。



人类智慧→人类智能→人工智能

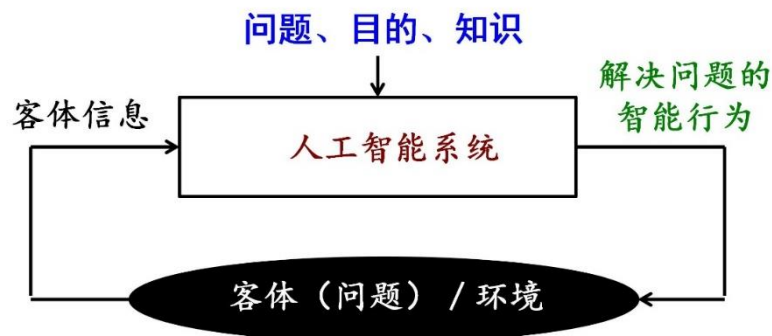
发现问题的能力：依据目的和知识，去寻求应当且能够解决的问题及目标的内隐性思辨能力，称为**隐性智慧**。

解决问题的能力：针对给定的问题、目标和知识，去寻求解决问题的智能策略的外显性操作能力，称为**显性智慧**。

人类智能专指人类解决问题的能力，即**显性智慧**。它是人类智慧的真子集。

人工智能：机器所能实现的人类智能（**显性智慧**）；所解决的问题、目标和知识由人类**隐性智慧**提供。

人工智能：开放、复杂、高级的信息系统

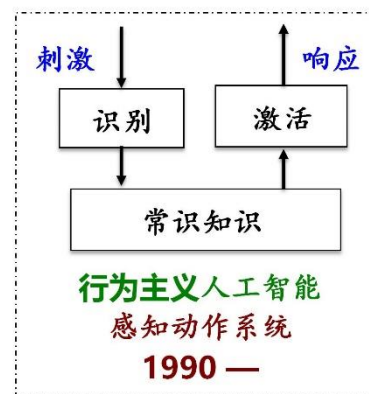
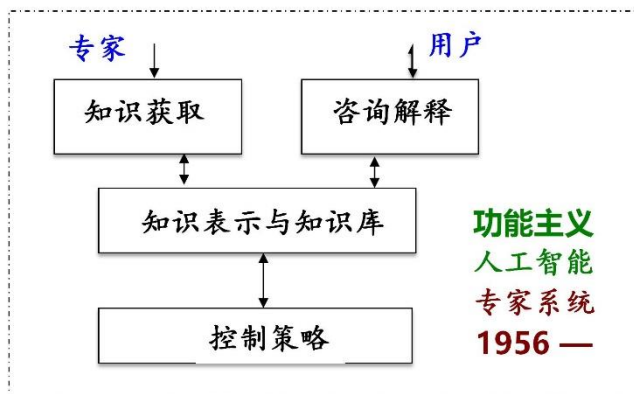
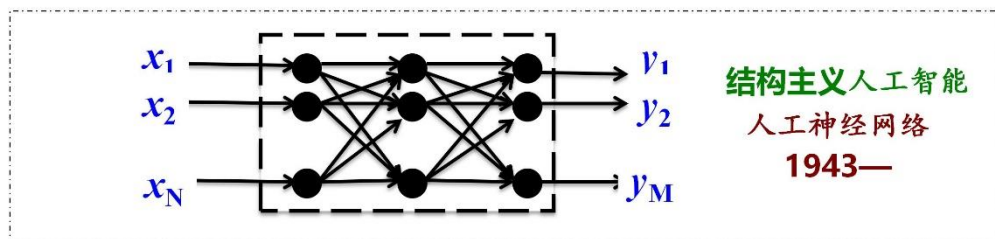


针对**人类给定**的问题、目的、知识等信息，通过**信息生态**生成**智能行为**，解决问题，达到目的。

二

人工智能的历史与现状

历史：结构、功能、行为的“鼎足三分”



现状（一）：鼎足三分，各有精彩

三大学派各自都取得了不少精彩成果。如：

结构主义方法：Pattern Recognition, Deep Network等

功能主义方法：Deep Blue, Watson, Alpha-Go等

行为主义方法：Sophia, Boston Dynamics等

现状（二）：整体被肢解，智能被掏空

现行人工智能研究存在许多致命性痼疾顽症：

（一）整体被肢解（分而治之）

- （1）三派争霸，无法实现统一
- （2）每个学派内部也不能通用

（二）智能被掏空（纯形式化）

- （3）理解能力低下
- （4）可解释性很差
- （5）需要大量样本
- （6）无意识，少情感

三

须在源头诊治 才能根除病患

学科范式：科学观与方法论的有机整体

学科的**范式**：学科的“**科学观**和**方法论**”的有机整体。

科学观在宏观上回答：这个学科是什么（认识）

方法论在宏观上回答：这个学科怎么做（实践）

科学观**方法论**两者：共同决定学科的研究范式。

学科不同，**范式**不同：物质科学，信息科学



学科成长规律：范式管控全程！

生长进程	进程名称	进程要素	要素解释
自下而上的摸索阶段	求索 (准备)	多方试探 总结提炼	常见：盲人摸象 任务：寻求 范式
自上而下的建构阶段	范式 (定义)	科学观	宏观上回答： 是什么
		方法论	宏观上回答： 怎么做
	框架 (定位)	全局模型	学科蓝图是什么？
		研究路径	研究路线怎么走？
	规格 (定格)	学术结构	交叉结构是什么？
		数理基础	数理方法怎么定？
理论 (定论)	基本概念	基本要素是什么？	
	基本原理	内在联系怎么做？	

惊人的重大发现：人工智能范式张冠李戴！

事项	范式要素之一：科学观	范式要素之二：方法论
传统物质科学	物质观 纯客非主，关注结构与功能 确定性演化，可分可合	机械还原方法论 形式化描述，比对式判断 分而治之的全局处置
现行人工智能	准物质观 纯客非主，关注结构与功能 接受可分性	真还原论 形式化描述，比对式判断 分而治之的全局处置
现代信息科学	信息观 主客互动，关注主体目的达成 不确定性贯彻全程	信息生态方法论 整体化描述，理解式判断 生态演化的全局处置

初级阶段“错冠”：学科体系大转变带来大阵痛
高级阶段“正冠”：学科范式大变革引领大发展

信息科学与物质科学相反相成，范式大不相同。信息学科初期的范式张冠李戴，是学科（时代）大变革带来的阵痛。

这是因为初期还没有提炼出信息科学的范式（滞后现象）；“盲人摸象”在所难免。

范式“张冠李戴”的人工智能研究注定不可能胜任人工智能高级阶段的全局研究：不可能担负建立统一理论的重任。

高级阶段自上而下建构：颠覆旧范式，确立新范式，创立新体系。这是**千年一遇的范式大变革！**

概念的释疑：范式是关键

类脑研究：并无新意：人工神经网络就是类脑的结构；专家系统就是类脑的功能。问题在范式！ ❌

计算机科学的分支：“算法、算力、数据是人工智能的三大要素”这类误解的思维源泉。问题也在范式！ ❌

强人工智能：不仅能够具有像人类一样的解决问题能力，而且可以具有像人类一样的发现问题的能力。 ❌

超强人工智能：发现和解决问题的能力都超过人类。 ❌

四

范式变革引领全局 学科体系全面创新

学科范式的颠覆性变革（1）：科学观 物质观 → 信息观

	科学观
现行 人工 智能 理论	准物质观 1, 原型对象是严格排除主观因素的物质客体; 研究的关注点是对象的结构。 2, 研究对象确定性演化。
通用 人工 智能 理论	信息观 1, 对象是主体驾驭的主客互动信息过程; 研究的关注点是主体的目标。 2, 研究对象不确定性演化。

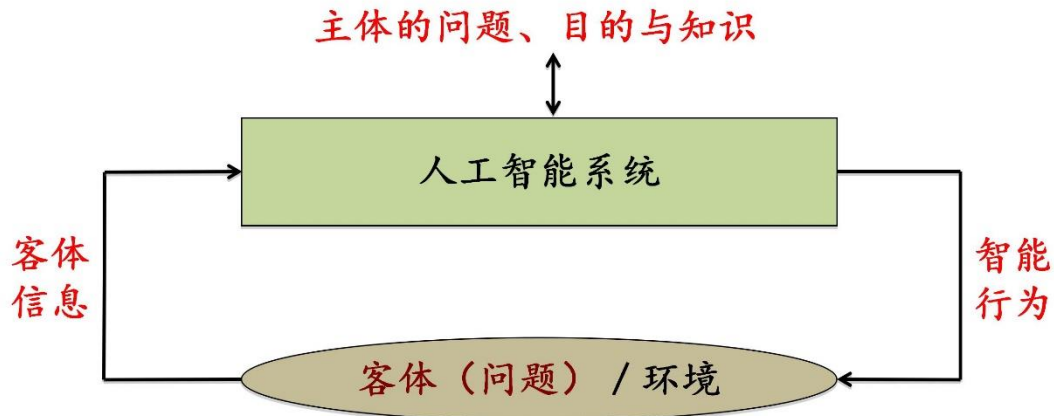
学科范式的颠覆性变革（2）：方法论 机械还原论 → 信息生态论

	方法论
现行人工智能理论	<p>机械还原论</p> <p>1, 关注对象的结构与功能, 采用纯粹形式化方法</p> <p>2, 对象确定性演化, 可以分而治之</p>
通用人工智能理论	<p>信息生态论</p> <p>1, 关注目标, 需用“形式—内容—价值”整体化方法</p> <p>2, 不确定性演化, 不可分割, 遵循生态演化。</p>

学科定位的颠覆性变革（1）：全局模型 脑模型 → 主客交互信息过程模型

	研究模型
现行人工智能理论	<p>脑模型</p> <p>1, 脑结构: 人工神经网络模型</p> <p>2, 脑功能: 物理符号系统 / 专家系统模型</p>
通用人工智能理论	<p>主体驾驭的主客互动的信息过程模型</p> <p>信息转换与智能创生: 见下页。</p>

主体驾驭的主客互动信息过程模型



人工智能：信息科学的完整篇章

学科定位的颠覆性变革（2）：研究路径 三分路径 → 整体路径

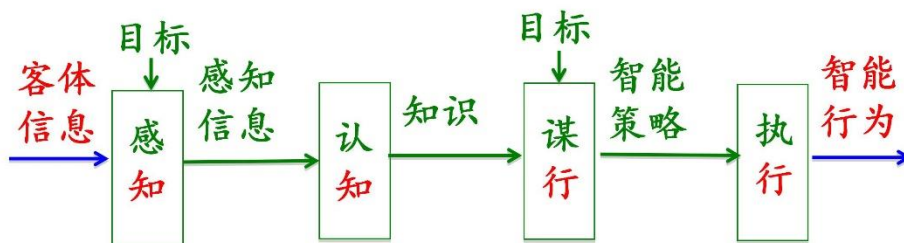
	研究路径
现行人工智能理论	<p>结构 / 功能 / 行为：研究路径鼎足三分</p> <ol style="list-style-type: none"> 1, 结构主义路径：人工神经网络研究路径 2, 功能主义路径：物理符号系统 / 专家系统研究路径 3, 行为主义路径：感知动作系统的研究路径
通用人工智能理论	<p>基于“普适性智能生成机制”的机制主义路径，实现“统一”！</p> <p>客体信息 → 感知信息 → 知识 → 智能策略 → 智能行为</p>

普适性智能生成机理的揭秘

智能生成机理由“智能生长动力学因素”协同决定：

- 全局动因—主客互动过程中主体追求生存发展目的
- 环境作用—客体信息（问题）的刺激
- 启动条件—客体信息被证明与目的相关，生成语义信息
- 起止标志—由语义信息起始，到智能行为生成为止
- 牵引力量—追求主体预设的目的
- 约束力量—遵守与问题相关的各种知识
- 检验准则—主体行为实效与目标之间的误差的满足
- 优化途径—误差反馈，学习新知，改善策略与行为
- 主体进化—若优化无效，则需主体修正目标（提升认识）

普适性智能生成机理：知行学原理



这是主体把客体信息加工成为智能行为的过程，即信息生态过程，本质是信息转换与智能创生原理。

学科定格的突破性变革（1）：学术结构

单一学科 → 交叉学科

	学术结构
现行人工智能理论	<p>单一学科</p> <p>计算机学科的分支</p>
通用人工智能理论	<p>交叉学科</p> <p>原型学科：神经科学，认知科学，人文学，社会学等</p> <p>本体学科：信息科学（信息转换与智能创生）</p> <p>基础学科：数学，逻辑学，哲学</p> <p>技术学科：微电子，微机械，新材料，新能源</p>

学科定格的突破性变革（2）：数理基础

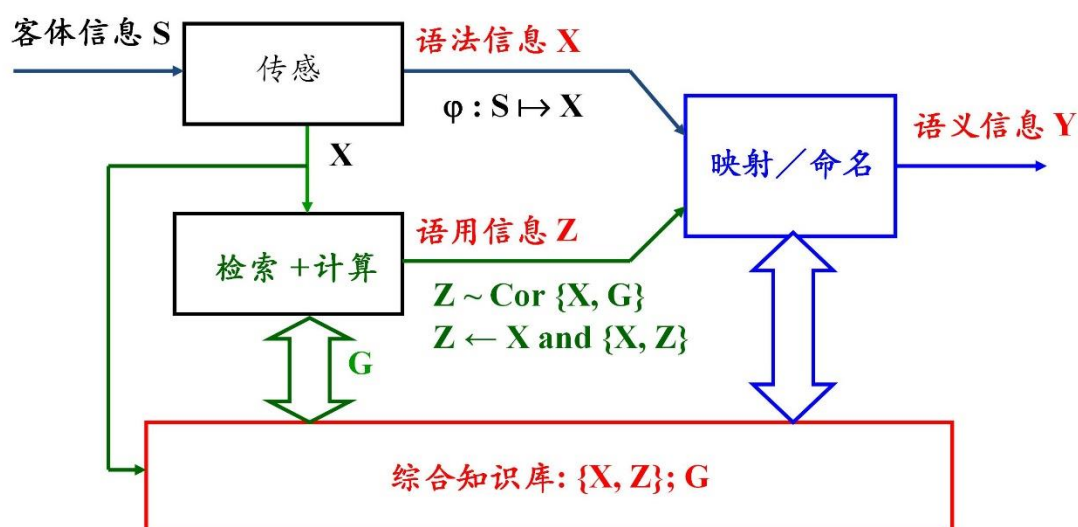
形式化碎片化 → 辩证化系统化

	数理基础
现行人工智能理论	<p>已有的数学与逻辑：</p> <p>数理（形式）逻辑；概率统计（形式）</p>
通用人工智能理论	<p>创新的理论：</p> <p>泛逻辑理论 — 面向通用人工智能的逻辑理论</p> <p>因素空间理论 — 面向通用人工智能的数学理论</p>

学科理论的突破性变革（1）：基本概念 形式化概念 → 内容性概念

	基本概念
现行人工智能理论	<p>形式化的概念：</p> <p>形式数据，形式知识，形式智能</p>
通用人工智能理论	<p>内容性的概念 — 奠定智能的“理解能力”：</p> <p>全信息：语义信息~（语法信息，语用信息）</p> <p>全知识：内容性知识~（形态性知识，价值性知识）</p> <p>全智能：基础意识、情感、理智的三位一体</p>

语义信息：定义与生成机制 — “形式科学”发展到“内容科学”的基础



语义信息：普遍的误解

有严格的定义、生成机制、表达式： $Y \leftarrow \lambda (X, Z)$

语义信息是感知信息的唯一合法代表。它表示了主体对客体的感性理解：客体是什么形态？有何效用？。这是人工智能系统理解能力的感性基础。

普遍的误解：

符号论（Semiotics）的误会
语义信息是词语的频率统计。
感知是传感！

学科理论的突破性变革（2）：基本原理 形式关系 → 信息转换

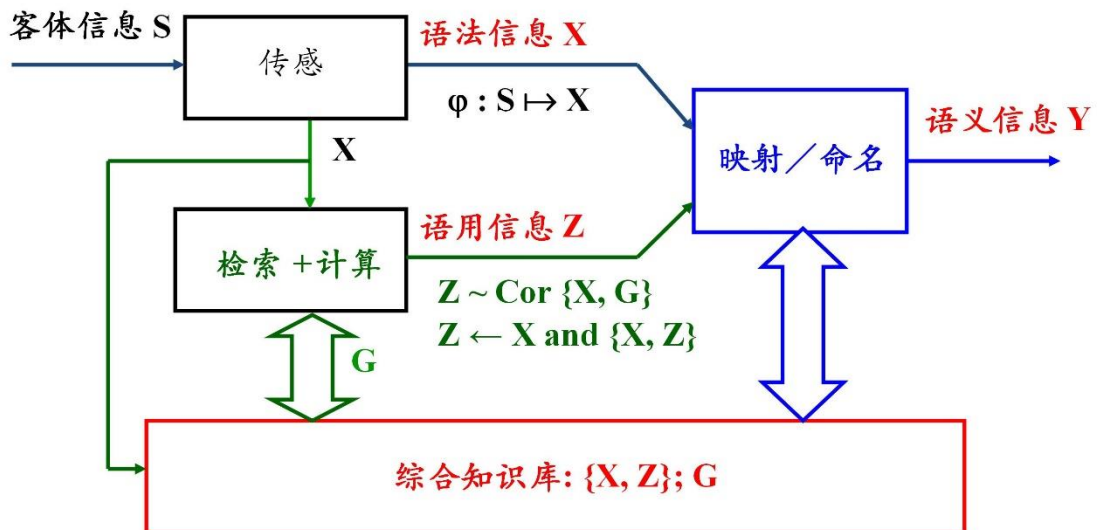
	基本原理
现行人工智能理论	尚未总结
通用人工智能理论	信息转换与智能创生定律： 客体信息—感知信息—知识—智能策略—智能行为—误差信息—学习新知—优化策略—不可消除的误差—优化目标

五

基本原理的落地

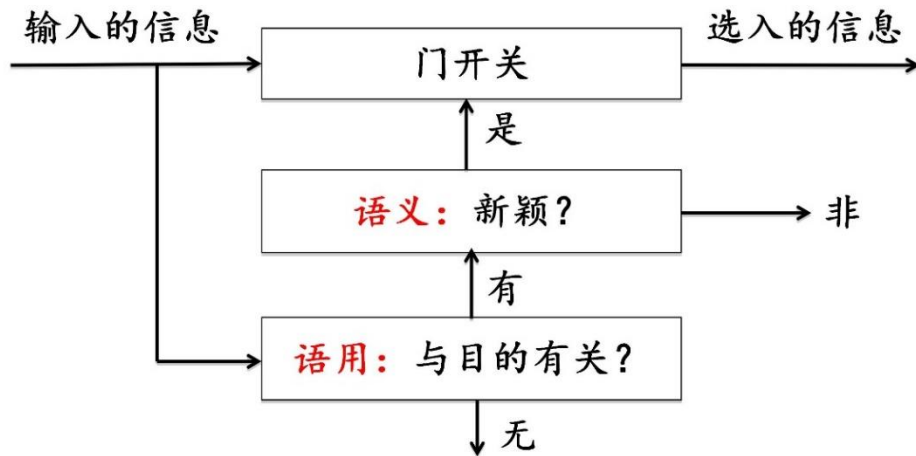
感知原理 (1)

客体信息 → 语义信息



感知原理 (2)

来者不拒 → 注意选择



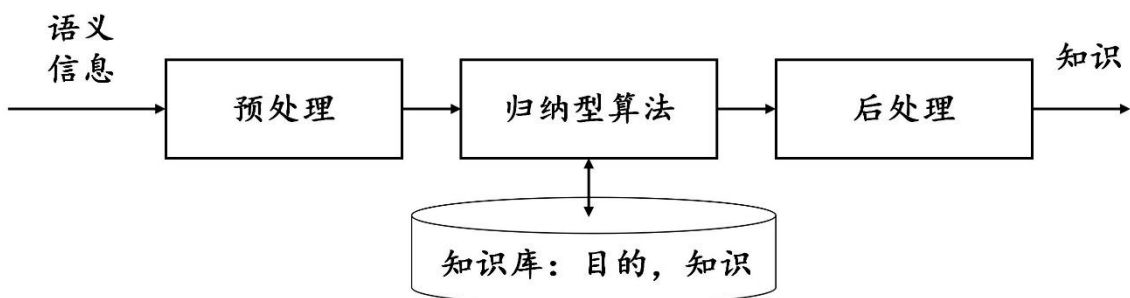
认知原理 (1)

语义信息 → 知识 (归纳)

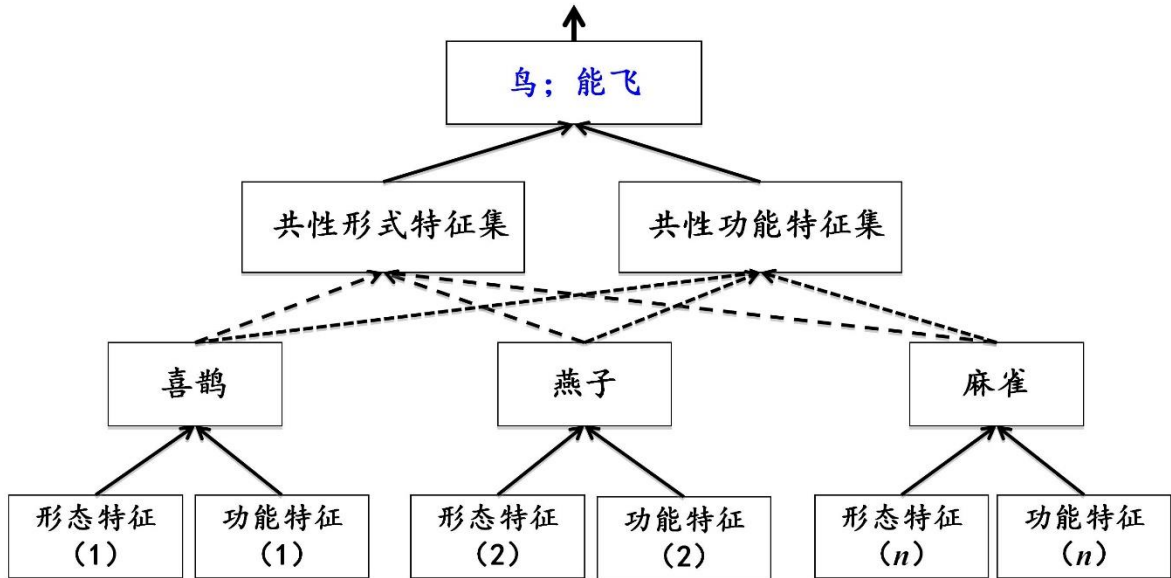
由“语义信息”提炼“知识”的主要算法是各种“归纳型算法”。

归纳型算法的真度与样本的典型性有关。这里体现了涌现机理。

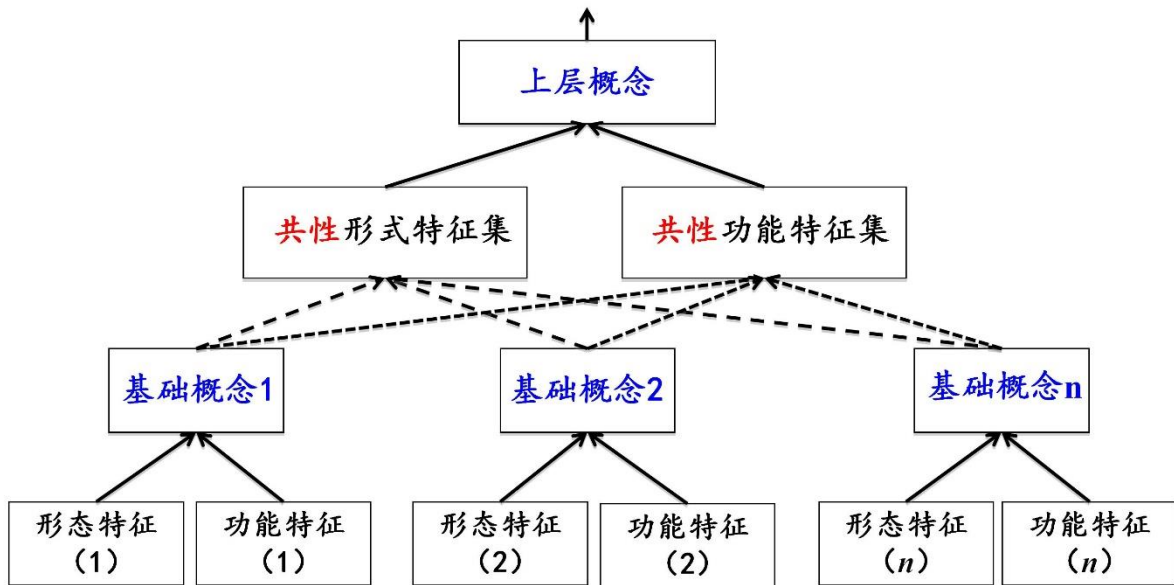
如：喜鹊会飞，燕子会飞，...，麻雀会飞 → 鸟能飞翔。



认知：由“语义信息”归纳学习“知识”



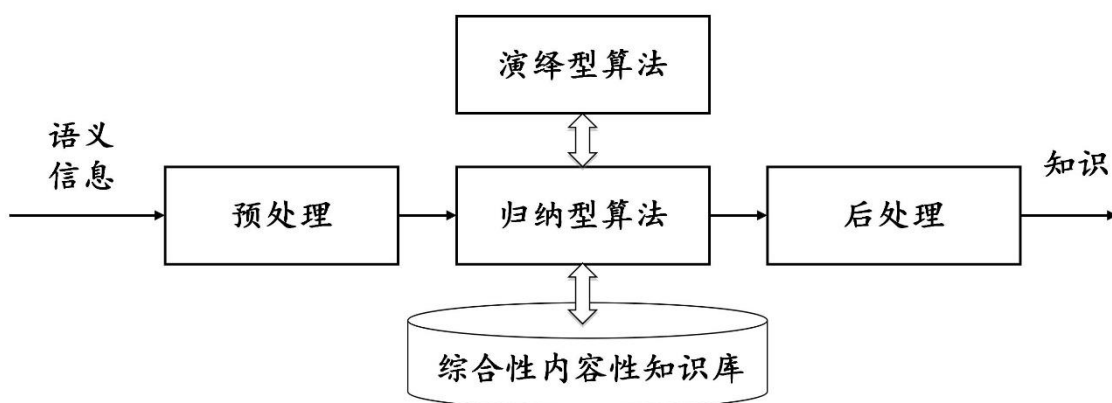
基于内容知识的知识库建构：
由低层概念到高层概念的逐级抽象



认知原理 (2)

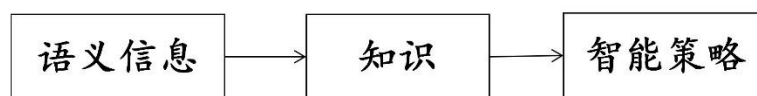
原知识 → 新知识 (演绎)

由“原知识”推理出“新知识”的主要算法是各种“演绎型算法”。
新知识与原知识之间，需要进行“同化”、“顺应”等扩展性处理。

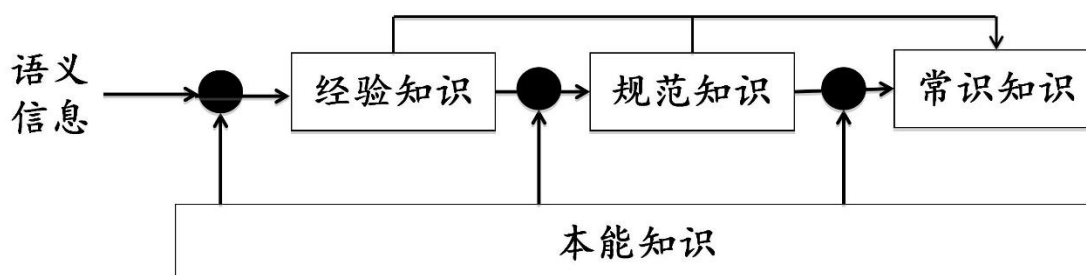


知识的内部和外部生态过程

外生态律:



内生态学:



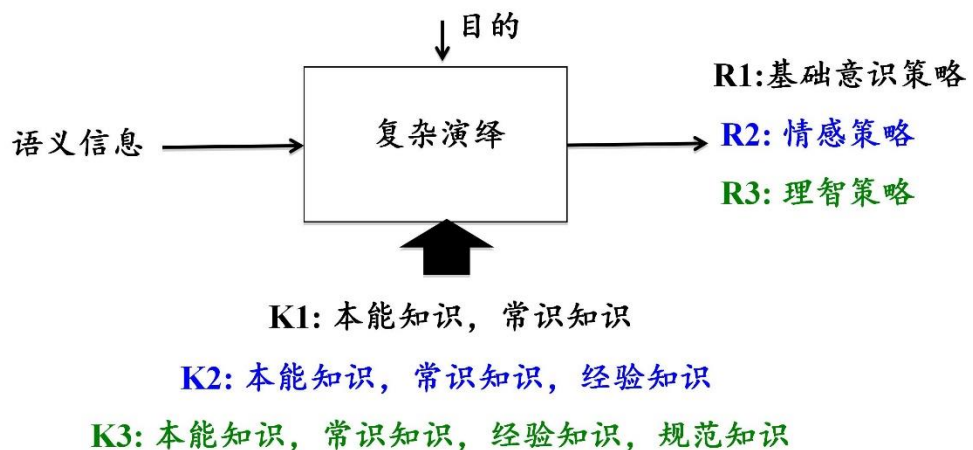
基础意识—情感—理智

基础意识是意识的基础部分，是在语义信息作用下主体根据自身的目的和基础知识（本能知识和常识知识）所产生的主观性反应。

情感是在语义信息作用下，主体根据自身的目的和情感性知识（本能知识，常识知识和经验知识）所产生的主观性反应。

理智是在语义信息作用下，主体根据自身的目的和理智性知识（本能知识，常识知识，经验知识和规范知识）所产生的主观性反应。

谋行原理 语义信息→理智策略



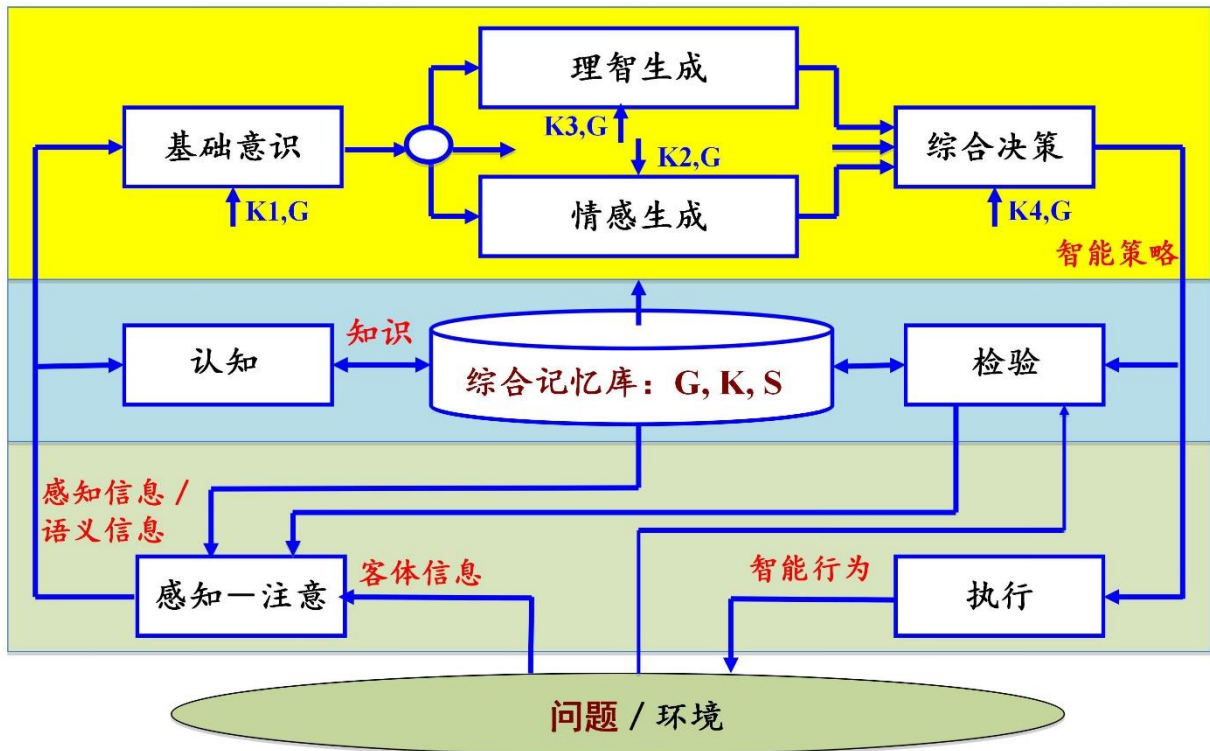
基础意识、情感、理智三者，是“意识”的有机成分。

执行原理 智能策略 → 智能行为



执行算法的作用就是普通的“控制技术”：把智能策略映射（转换）成为智能行为，反作用于控制对象。

基于普适性智能生成机理的 通用人工智能基础理论模型



六

通用人工智能理论的 首创性、优胜性、通用性

首创性（自顶向下）

	现有理论	新的理论
科学观	准物质观	信息观
方法论	机械还原论	信息生态论
全局模型	孤立的脑模型	主体客体互动的信息过程模型
研究路径	三分路径	整体路径
学术结构	单一学科	交叉学科群
数理基础	概率论，形式逻辑	面向智能：泛逻辑，因素空间
基本概念	形式化概念	理解性概念
基本原理	未总结	信息转换原理
基本结果	三个局部理论	通用人工智能基础理论

优胜性：全面消除了现行理论的痼疾顽症

颠覆“去主观性”，确立主体的主导性：实现系统的**目的性**

颠覆“分而治之”，创建信息生态方法论，实现**整体性**

突破“脑模型”，创建“主客互动模型”，保障模型**真实性**

突破“三分路径”，揭示“普适性生成机制”，实现**通用性**

突破“纯形式化”，创建了“语义信息论”，奠定**智能基础**

保证充分的“**理解能力**”

保证“**可解释性**”

提供“**小样本学习**”可能性

实现“**意识—情感—理智**”和谐生成

通用性：基本涵义

通用人工智能不是指某个人工智能系统的“巨无霸”。

通用人工智能是指：**智能的生成机理是普适性的**。

现实世界的“**人**”就是通用人工智能的原型。

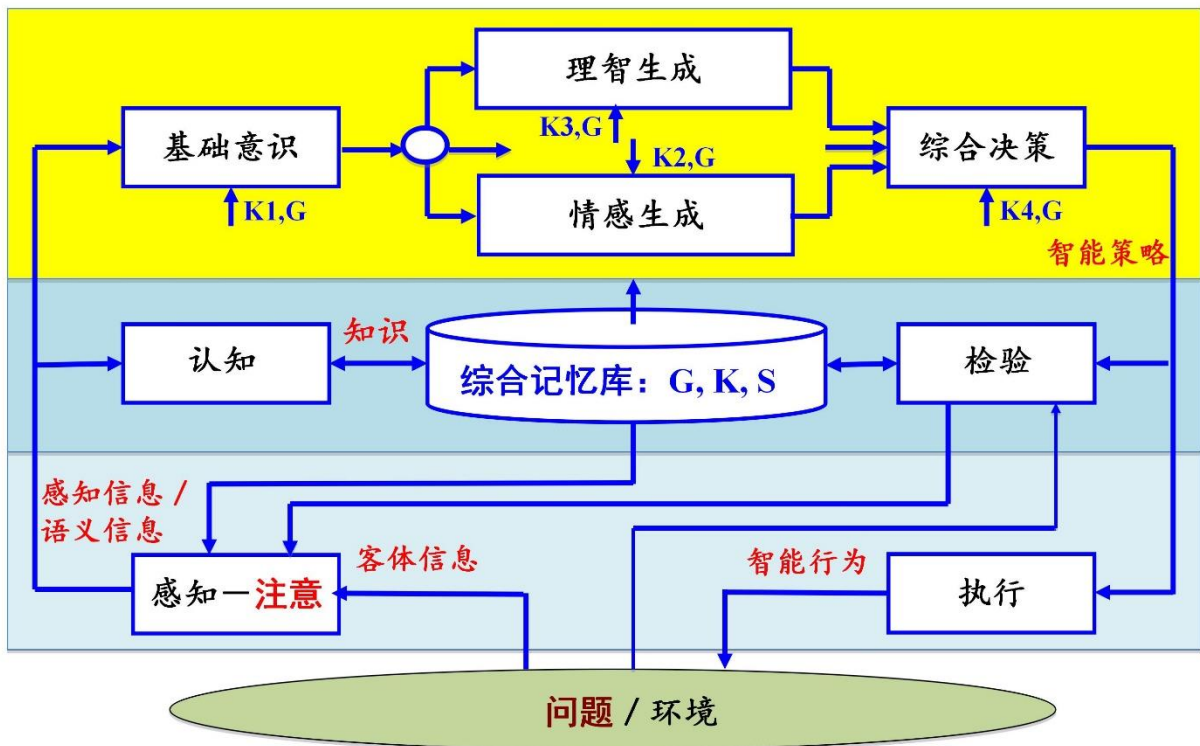
我们已经发现和实现“**普适性人工智能生成机理**”。

普适性人工智能生成机理就是**通用人工智能孵化器**。

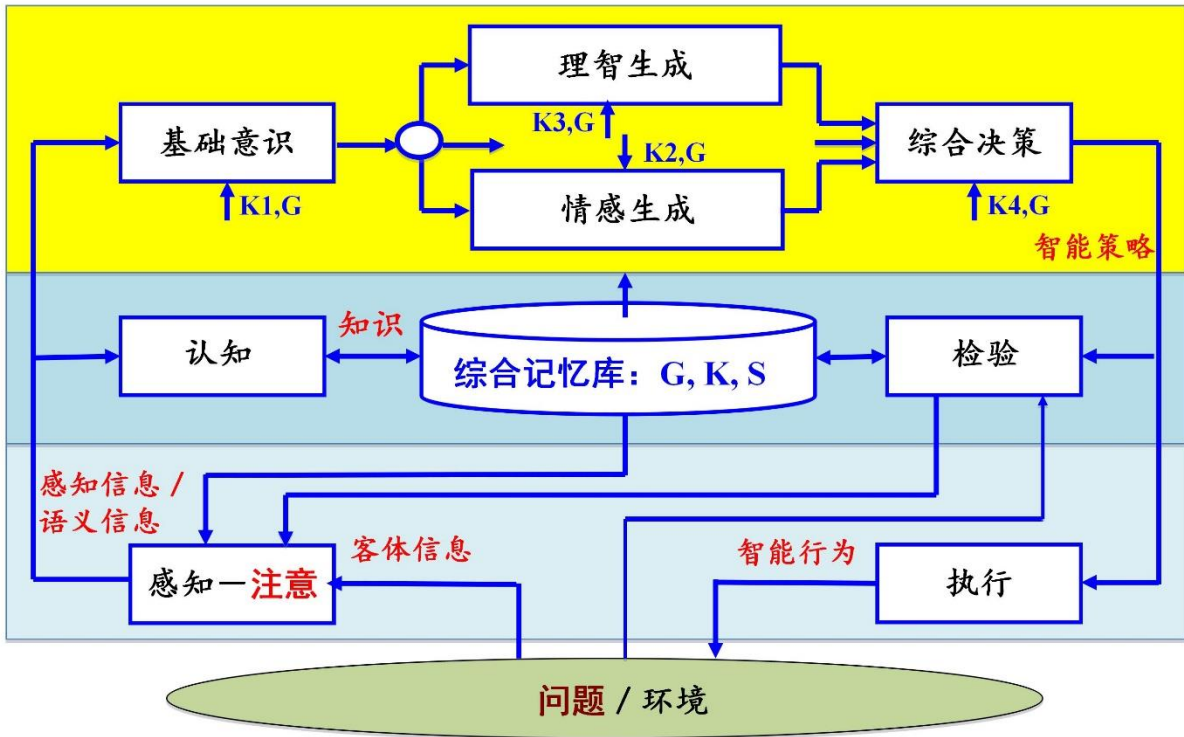
通用性的内部表现：实现AI的“三分归一”

机制主义方法	信息 \rightarrow 知识 \rightarrow 智能策略			实例
A型	信息	经验性知识	经验型策略	人工神经网络
B型	信息	规范性知识	规范性策略	物理符号系统
C型	信息	常识性知识	常识性策略	感知动作系统

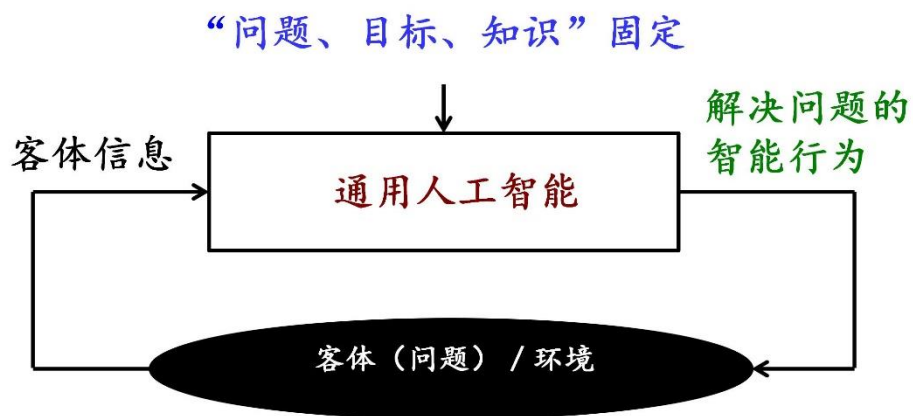
通用性 (1)：“大数据智能”的聪明型



通用性 (2) : “混合智能” 的基本型

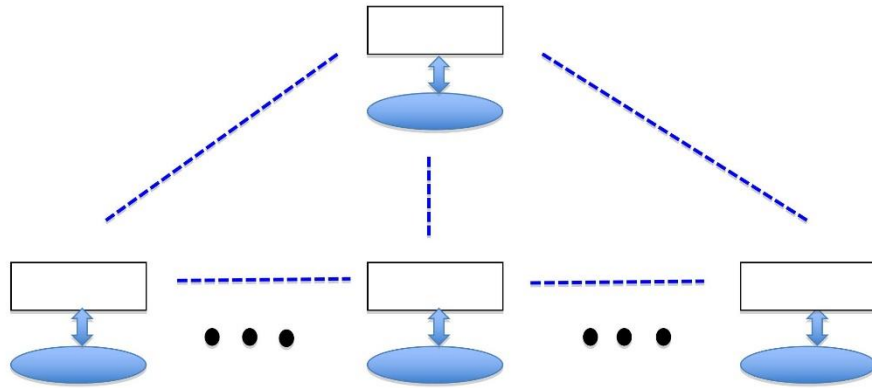


通用性 (3) : 是“自主智能”的母体

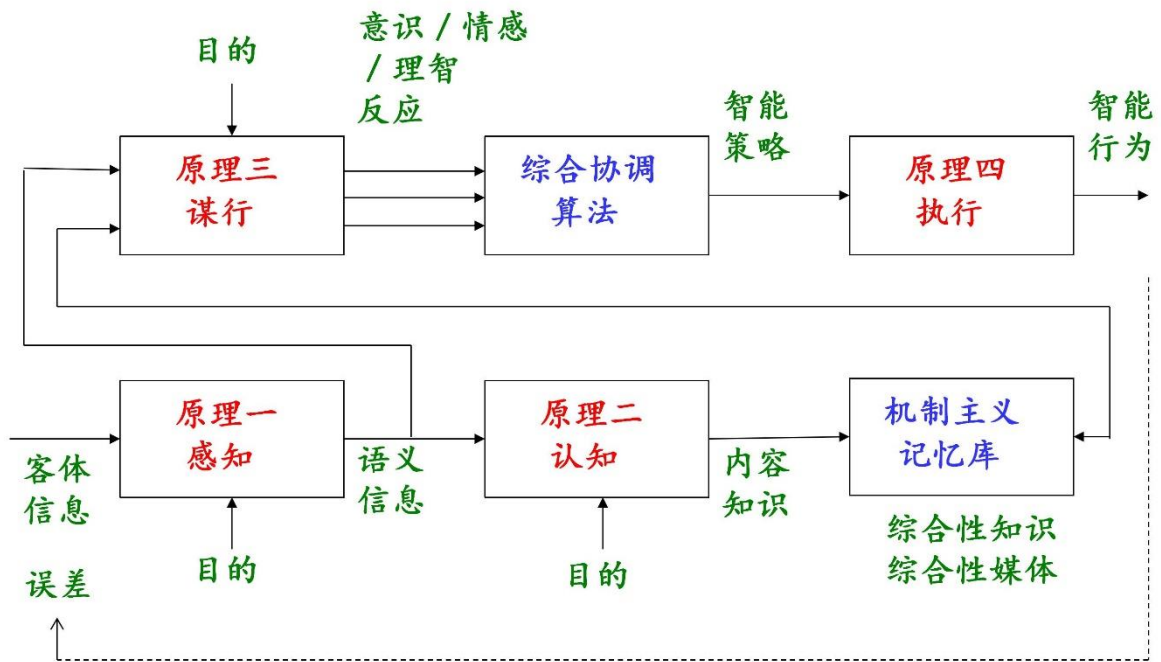


人工智能：机器所实现的人类“解决问题的能力”。

通用性 (4) : 是“群体智能”的基元



通用性 (5) : 是“跨媒体智能”的必由之路



通用性（6）：通用人工智能系统孵化平台



《洛神工程》设想：基于通用人工智能系统的万物互联，智能社会生产力的“社会生产工具网络体系”（待续）。

清华大学 AI 研究院 张钹、朱军、苏杭

摘录自“纪念《中国科学》创刊 70 周年专刊”

清华大学人工智能研究院张钹院士等人提出如何迈向第三代人工智能的理论及其发展途径。

他们认为，在人工智能 60 年发展中，一直存在相互竞争的范式：上世纪 80 年代，符号主义（第一代 AI）主导 AI 的发展，上世纪 90 年代至本世纪初，连接主义（第二代 AI）推动 AI 的发展，但这两种范式只是从不同侧面模拟人类的心智（大脑），具有各自片面性，不可能触及人类真正的智能。

他们提出的第三代人工智能的理念及其发展途径是：融合第一代知识驱动和第二代数据驱动的 AI 范式，把知识驱动和数据驱动结合起来，同时利用知识、数据、算法和算力等四要素，建立可解释和鲁棒的 AI 理论与方法，构造更强大的 AI，即发展安全、可信、可靠和可扩展的 AI 技术，这是发展第三代 AI 必经之路。

为了实现第三代 AI 的目标，他们采用“三空间融合”的模型，如能实现，这时机器就会像人类的行为一样。

549. 附件二

第三代人工智能单一空间模型、双空间模型、三空间模型见附件二：

第一代知识驱动的 AI，利用知识，算法和算力三个要素构造 AI，第二代数据驱动的 AI，利用数据、算法与算力三个要素构造 AI。由于第一、二代 AI 只是从一个侧面模拟人类的智能行为，因此存在各自的局限性。为了建立一个全面反映人类智能的 AI，需要建立鲁棒与可解释的 AI 理论与方法，发展安全、可信、

可靠与可扩展的 AI 技术，即第三代 AI。其发展的思路是，把第一代的知识驱动和第二代的数据驱动结合起来，通过同时利用知识、数据、算法和算力等四个要素，构造更强大的 AI。目前存在双空间模型与单一空间模型两个方案。

双空间模型

双空间模型如图 1 所示，它是一种类脑模型，符号空间模拟大脑的认知行为，亚符号（向量）空间模拟大脑的感知行为。这两层处理在大脑中是无缝融合的，如果能在计算机上实现这种融合，AI 就有可能达到与人类相似的智能，从根本上解决目前 AI 存在的不可解释和鲁棒性差的问题。

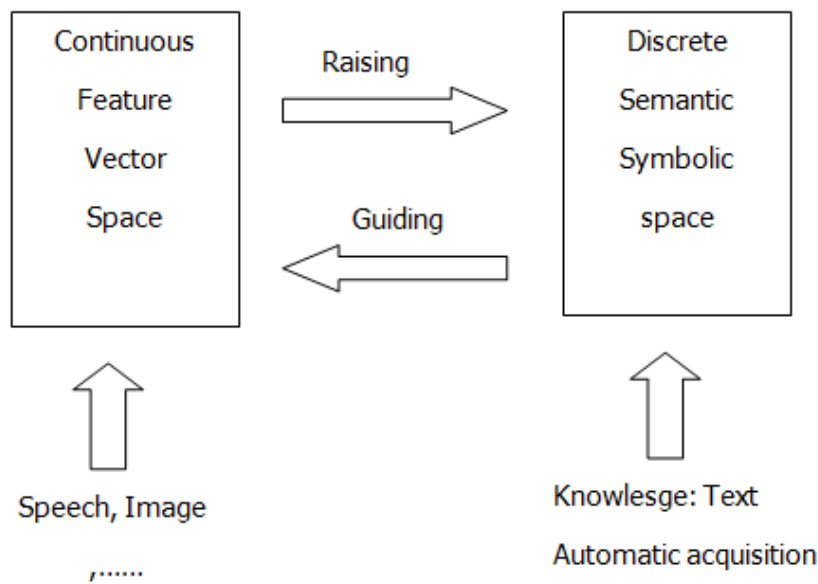


图 1 双空间模型

单一空间模型

单一空间模型是以深度学习为基础，将所有的处理都在亚符号（向量）空间，这显然是为了利用计算机的计算能力，提高处理速度。问题在于深度学习与大脑的学习机制不同，在许多方面表现不佳，如可解释和鲁棒性等，关键是要克服深度学习所带来的缺陷，如图 2 所示，下面讨论几个关键问题。

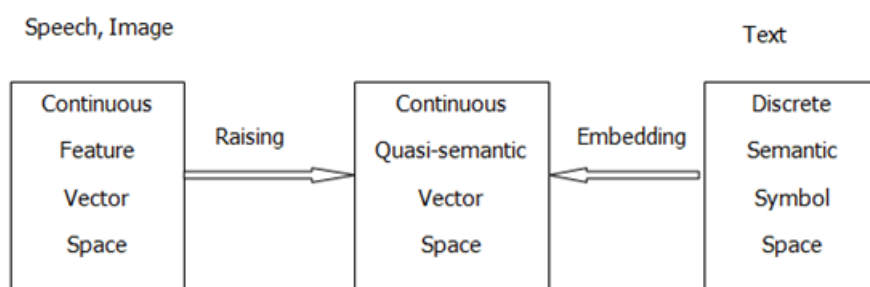


图2 单一空间模型

为了实现第三代 AI 目标，我们采用三空间融合模型，即融合双空间与单空间两种模型，如 3 所示。

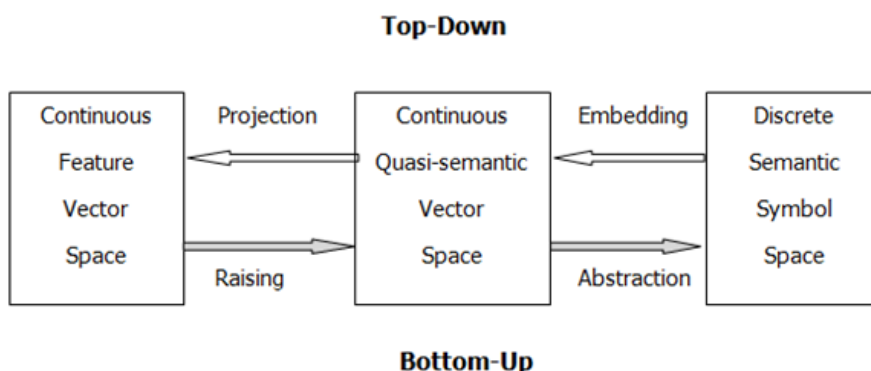


图3 三空间融合模型

550.

也谈新一代人工智能

忠实的跟帖者

读了钟义信、张钊等老师谈判新一代人工智能（第三代人工智能）的发文，眼前一亮！他们的观点不同于目前国内外多数学者，他们建议的人工智能发展路径不同于当今人工智能不同学派提出的发展路径，并认为后者均存在很大片面性，他们的共同点是要立足于人工智能融合的统一范式，从源头出发进行探索（虽然钟、张对范式的出发点以及发展途径也未必完全相同）。至今，钟、张

的建议尚属于构想，尚未见到他们推出具有实效的解决方案，我们还需拭目以待！

钟、张在提出“发展途径”时均十分关注创建语义网络作为通向新一代人工智能的关键环节，张老师谈到人类的分析和决策是利用自身包括常识在内的背景知识在语义空间当中完成，但是数据特征空间和人类的语义空间在结构和内涵上存在明显区别，我们赞成这样的论述。联想到当今国内外研究的大规模语义网络，对突破认知智能显得其支持力度不足，还差最后一公里！（对此，不少人在以往发表的人工智能跟帖中有所讨论）

期望钟、张等老师们在研究新一代人工智能的基础理论和实践中取得成功！

551. 你们创办的“人工智能国内外跟贴留言”平台，确实是一个巨大的宝藏！就以钟、张等教授谈到的新一代（或第三代）人工智能的构想而言，在 263、521 等条跟贴中，也谈到了同样的思路。其中最关键的是建设知识图谱（大规模语义网络）以往对大规模语义网络的研究还有待于进一步完善，以便能有有力的支持实现可解释的新一代人工智能。

其中 521 条跟贴提出了数据（第二代人工智能）与知识（第一代人工智能）双驱动，即将数据分类（结构化、非结构化、半结构化数据）→基于表结构的数据库→数据获取；将知识表示（基于符号的、向量的知识）→基于图结构的知识库→知识获取；再将数据、知识获取→知识融合→知识推理→知识图谱（大规模语义网络）→实现可理解、可解释的人工智能（或破解认知智能）。

但演进到最后一个环节，不管是钟、张建议的新一代人工智能构想，或者是不同学派的国内外 AI 专家各自研发的方案中，语义网络迄今均不够完善，不足以

支持其破解认知智能（或还达不到新一代人工智能），还差最后一公里！尚需进一步努力。

说三大学派（或两大学派）的学术观点有其片面性，但不要一概废弃它们已取得的学术成就，如能经过科学分析，吸收、改造、兼容，可能更好！

在 210 等条跟贴中谈到：机器最大的问题是缺乏常识，而建立常识库又非常困难，这个问题仍然阻碍机器对自然语言的理解。

鉴于对常识、专业知识、专家经验，机器很难识别，IBM Watson 提出具人（embodiment）概念，强调人工智能医疗研究人员必须与临床医生结合，在疾病诊断时要取得共识，以弥补研究人员常识之不足！但医疗实践证明，如此做法是非常粗糙的，难以成势。

在国内外跟贴留言中，大家还可以看到不少关于语义网络、常识等问题的研究和论述，但火候还差这么一点，期望继在这方面的研发中取得进展、突破和成功！