

探索通用人工智能之路

国内外人工智能跟帖留言

(共28条跟帖)

点评“探索通用人工智能之路”

陆首群

2022.5.5

4年前 COPU 提出“人工智能向何处去？”同时构建“人工智能国内外跟帖留言平台”，期望基于平台博纳人工智能发展的多元言论，回答人工智能发展方向问题。迄今 COPU 已收到人工智能国内外跟帖留言 1006 条。

我们决定陆续汇总编辑出版《可解释性机器学习》、《异步脉冲神经网络+神经拟态计算系统》、《脑机接口》、《基于大规模语义网络约认知算法》、《探索通用人工智能之路》等 5 部专辑。

在《探索通用人工智能之路》专辑中，有“通用人工智能（AGI）”的探索者，也有“通用人工智能”的创新理论——“机制主义通用人工智能（GTAI）”的探索者；有资深院士提出的“通用人工智能十问”，也有专家提出的“十答”；我们汇集了国内外专家对通用人工智能的探索性观点、理论，也有正、负面评论，包括国外有的专家提出“为什么通用人工智能不能实现”的观点，还有的专家认为“探索通用人工智能之路还很漫长”。本专辑可说是各抒其见，缤纷纷呈！实践是检验真理的唯一标准，我们盼望专家们的理论探索能够获得实践证实的时间早日到来！

下面介绍钟义信教授探索、创建的“机制主义通用人工智能理论”，以及我的感受（可能有倾向性，那也可算是一家之言供大家参考）：

我的老友钟义信教授几年来多次给我们发来他的“机制主义通

用人工智能理论”的跟帖，冲击了我们头脑里一些传统观念，展示一条不同的探索通用人工智能基础理论创新之路。

在与钟教授多次交谈学习中感到：

①他的机制主义通用人工智能不同于目前国内外人工智能主流，后者都把人工智能看作是计算机科学的延伸，都是从算法、算力、数据入手展开研究，②也不同于另一部分科研人员把人工智能看作是人脑技术的实现，而从模拟人脑结构或功能出发开展的研究。他认为这些研究都有出色的成果，可惜只是人工智能的“局部性”和“表层性”成果，没有触及人工智能的本质。

他提出的机制主义通用人工智能理论是通用人工智能的创新理论，其英文缩写是 GTAI (General Theory of AI)，可能与我以前认定的通用人工智能（其英文缩写是 AGI）有差别，机制主义通用人工智能是在通用人工智能领域取得的突破。

机制主义通用人工智能的研究路径是以普适性智能生成机制（即信息转换与智能创生）为核心的。

在本专辑上，我们征得钟教授同意，将转载他于去年 12 月 15 日刊登在人民日报《学术前沿》上的论著：“范式革命：人工智能基础理论源头创新的必由之路”。

国内外通用人工智能跟帖留言

(共 28 条)

1-546, 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，正确把握和处理好如下四个问题：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！

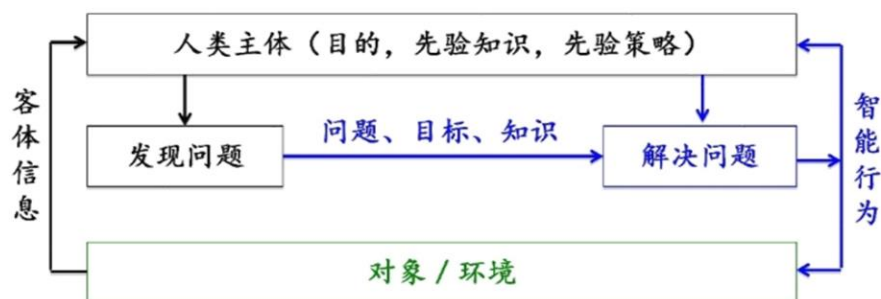
钟义信
北邮人工智能学院
2020.9.15

2-547, 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

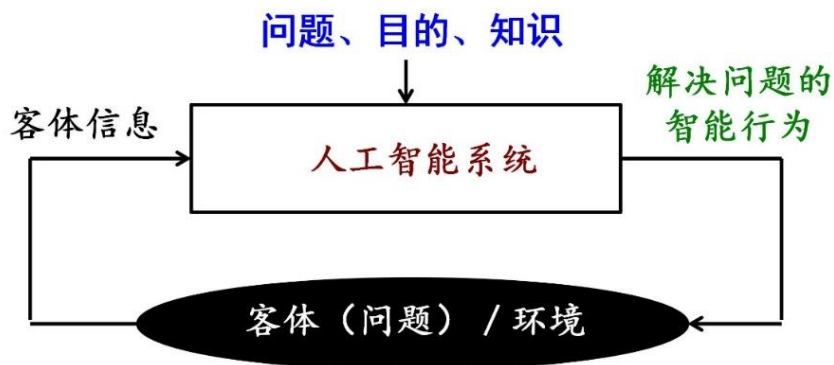
发现问题的能力：依据目的和知识，去寻求应当且能够解决的问题及目标的内隐性思辨能力，称为**隐性智慧**。

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

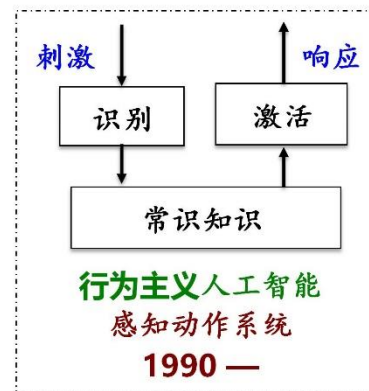
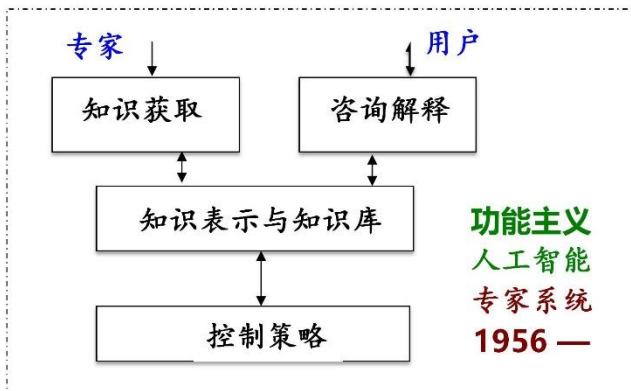
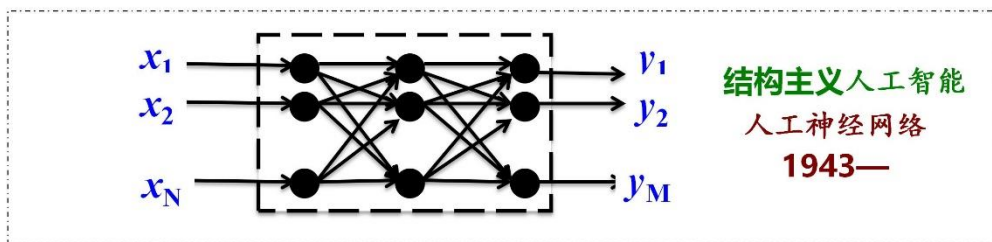


针对**人类给定**的问题、目的、知识等信息，
通过**信息生态生成智能行为**，解决问题，达到目的。

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

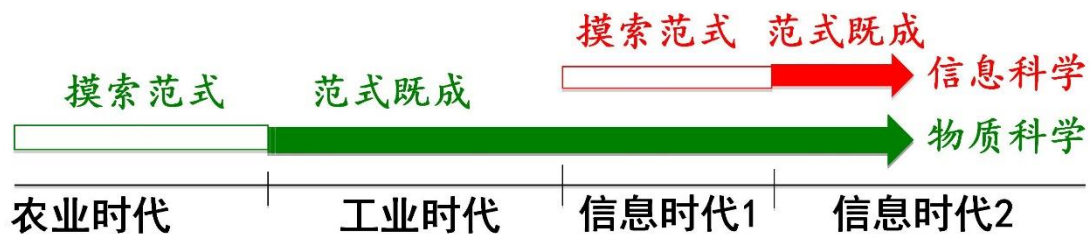
学科的**范式**：学科的“**科学观**和**方法论**”的有机整体。

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 。

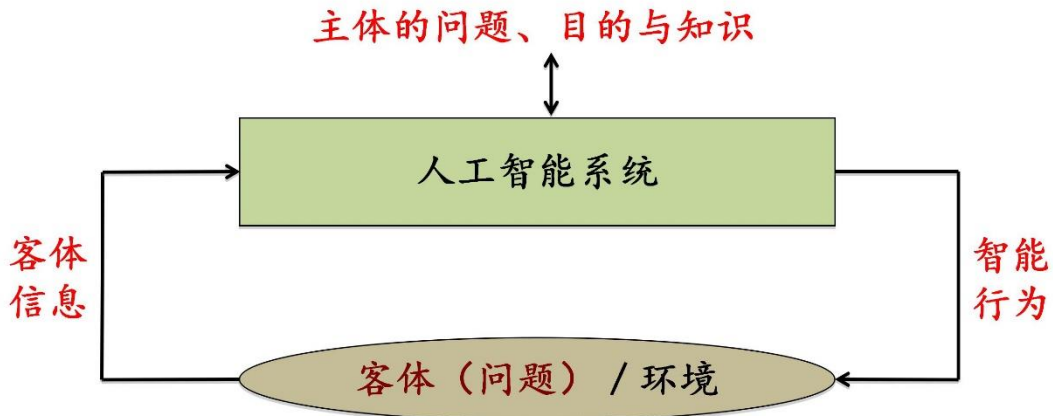
学科范式的颠覆性变革（2）：方法论 机械还原论 → 信息生态论

	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

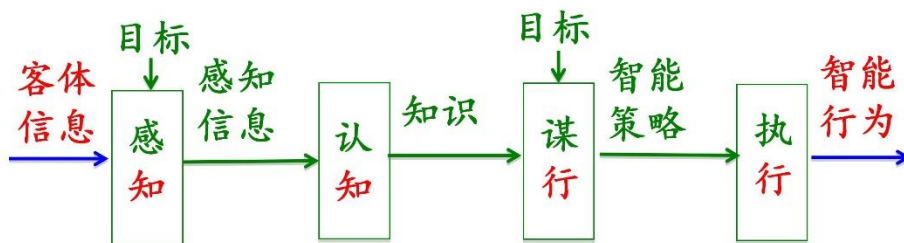
	研究路径
现行人工智能理论	<p>结构 / 功能 / 行为：研究路径鼎足三分</p> <ol style="list-style-type: none"> 1, 结构主义路径：人工神经网络研究路径 2, 功能主义路径：物理符号系统 / 专家系统研究路径 3, 行为主义路径：感知动作系统的研究路径
通用人工智能理论	<p>基于“普适性智能生成机制”的机制主义路径，实现“统一”！</p> <p>客体信息 → 感知信息 → 知识 → 智能策略 → 智能行为</p>

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

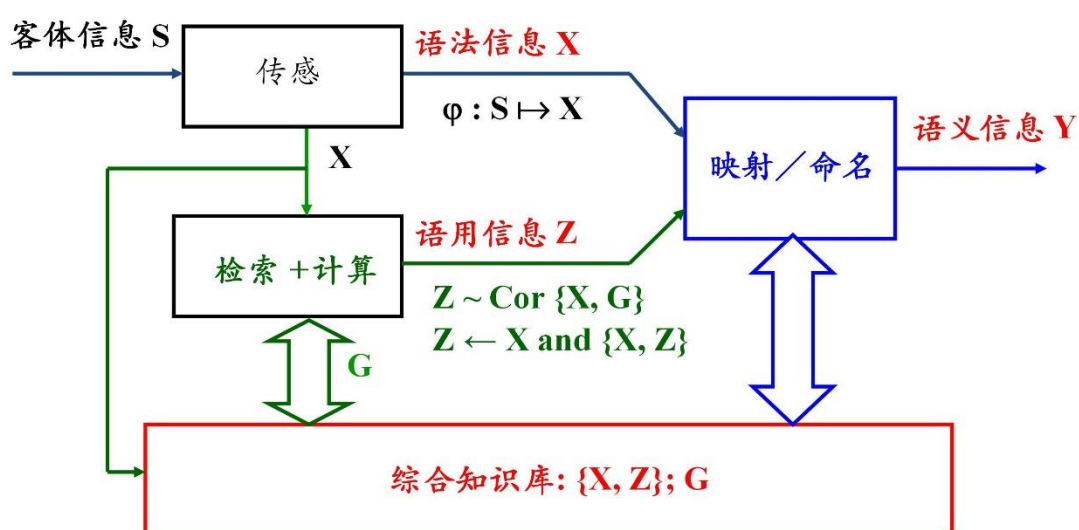
形式化碎片化 → 辩证化系统化

	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：普遍的理解

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。

普遍的理解：

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

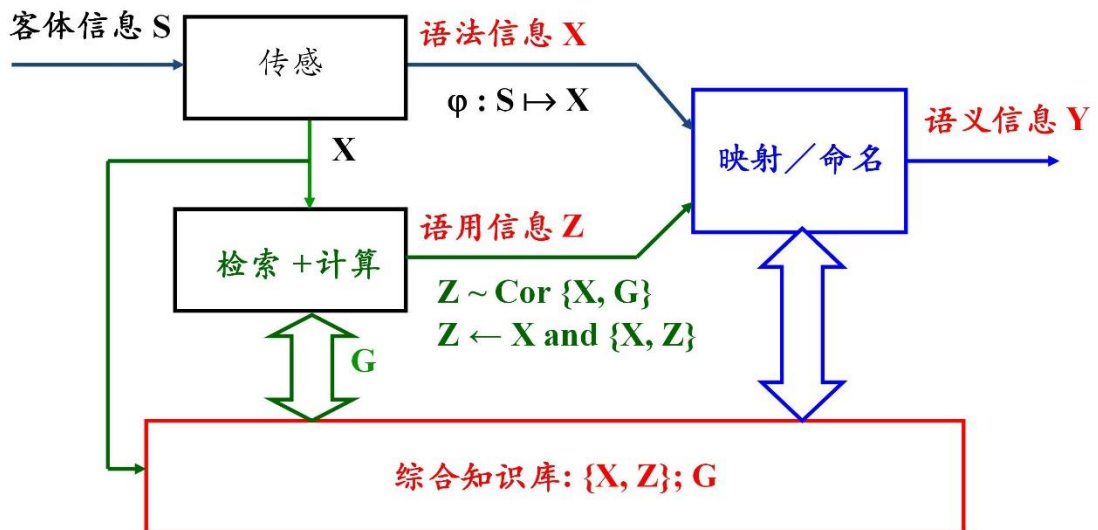
	基本原理
现行人工智能理论	尚未总结
通用人工智能理论	信息转换与智能创生定律： 客体信息—感知信息—知识—智能策略—智能行为—误差信息—学习新知—优化策略—不可消除的误差—优化目标

五

基本原理的落地

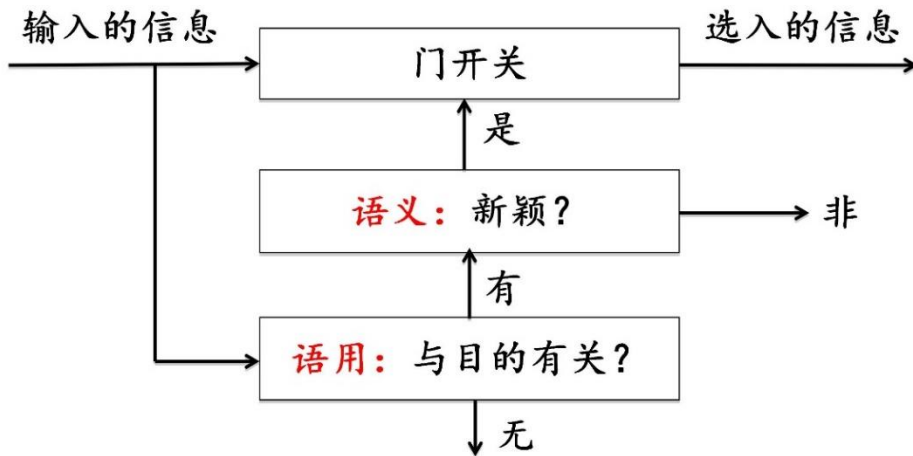
感知原理 (1)

客体信息 → 语义信息



感知原理 (2)

来者不拒 → 注意选择



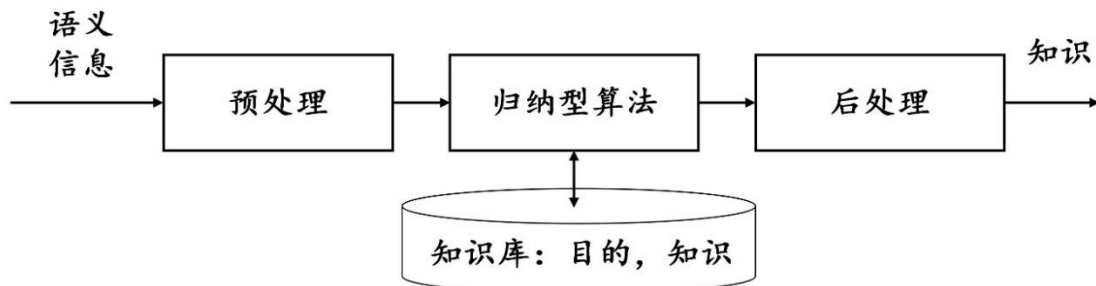
认知原理 (1)

语义信息 → 知识 (归纳)

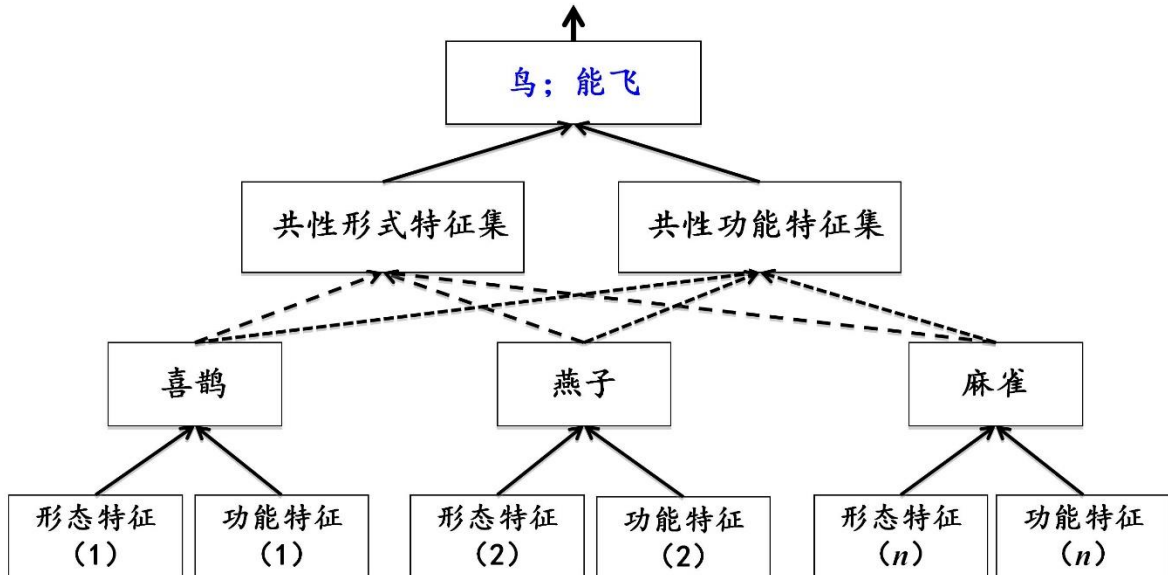
由“语义信息”提炼“知识”的主要算法是各种“归纳型算法”。

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。

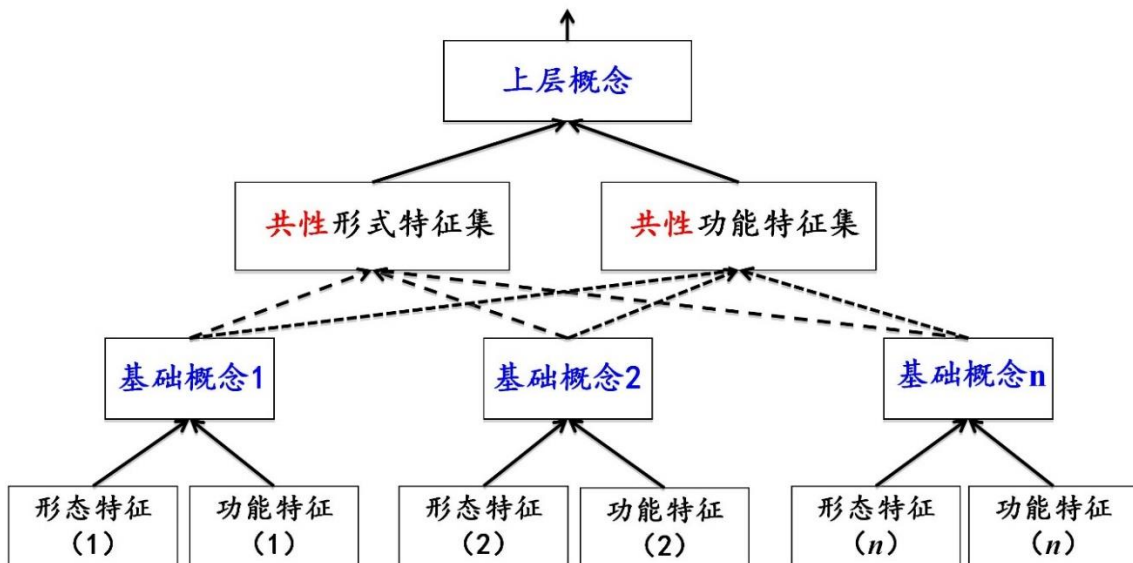
如：喜鹊会飞，燕子会飞，...，麻雀会飞 → 鸟能飞翔。



认知：由“语义信息”归纳学习“知识”



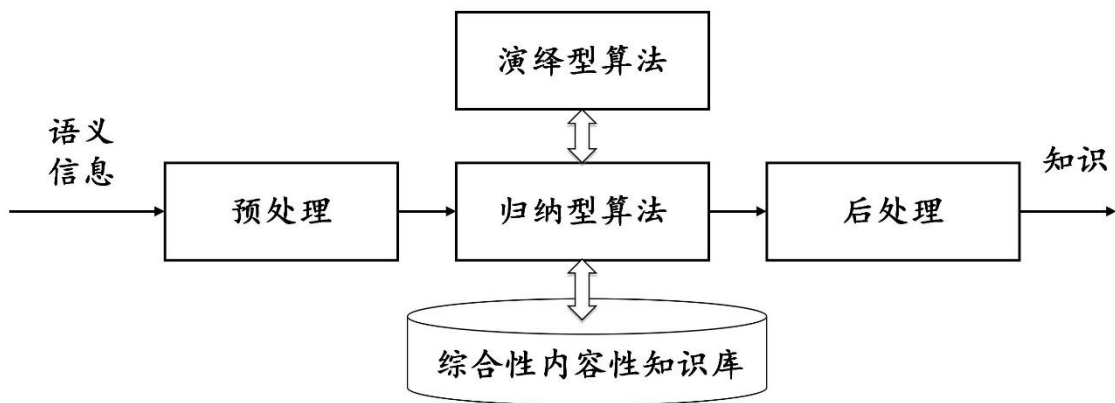
基于内容知识的知识库建构：
由低层概念到高层概念的逐级抽象



认知原理 (2)

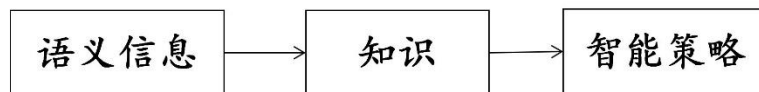
原知识 → 新知识 (演绎)

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。

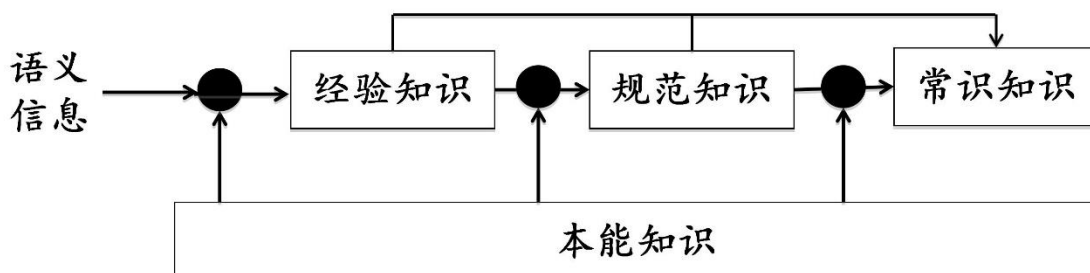


知识的内部和外部生态过程

外生态律:



内生态学:



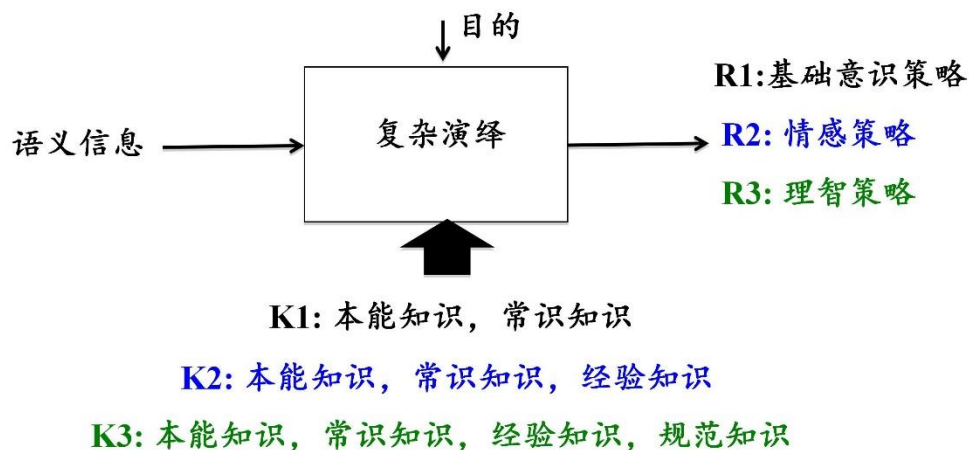
基础意识—情感—理智

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



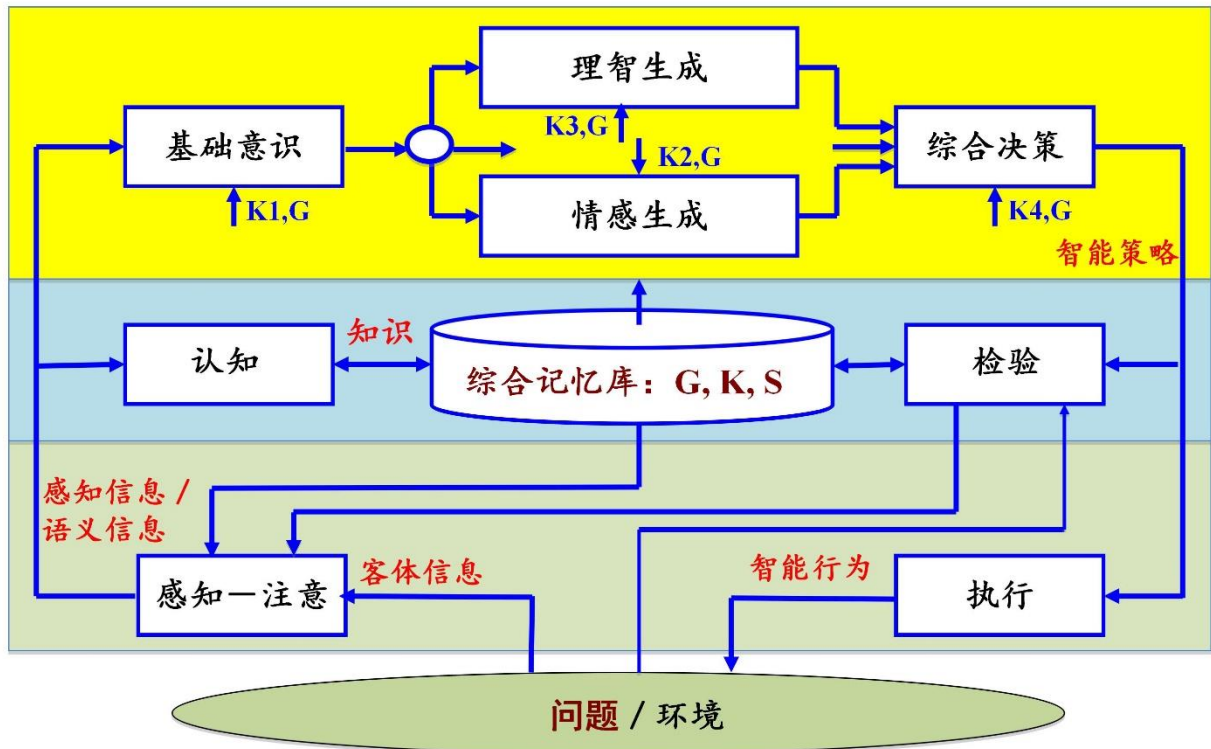
基础意识、情感、理智三者，是“意识”的有机成分。

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。

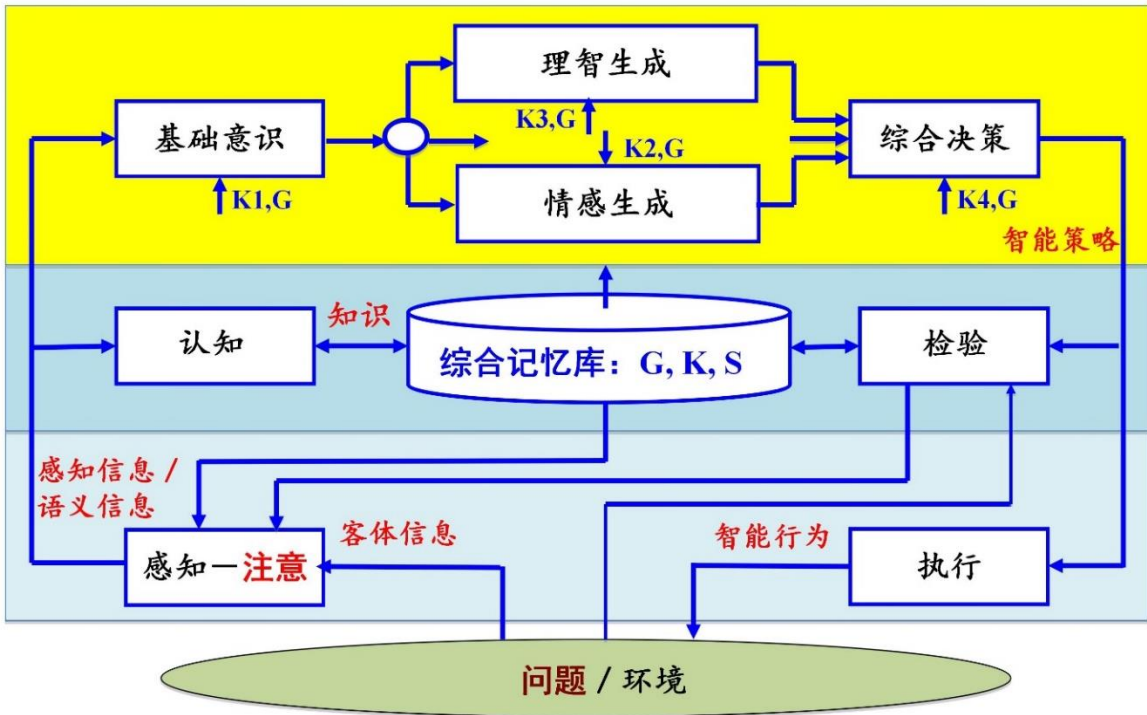
我们已经发现和实现“**普适性人工智能生成机理**”。

普适性人工智能生成机理就是**通用人工智能孵化器**。

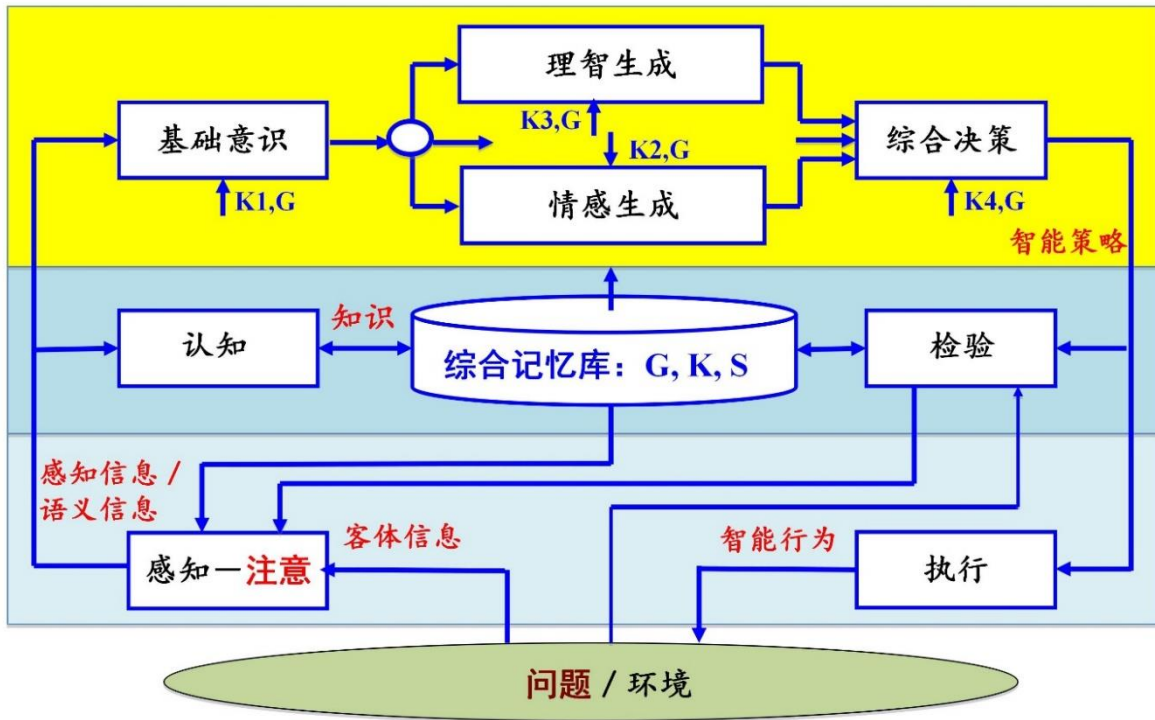
通用性的内部表现：实现AI的“三分归一”

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

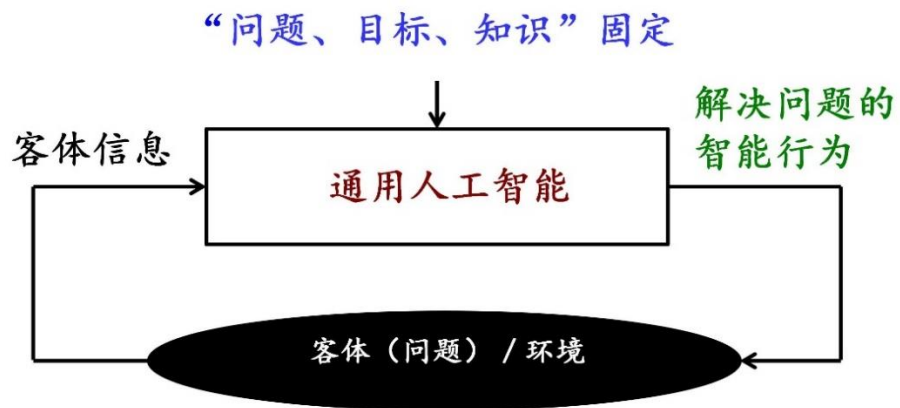
通用性 (1)：“大数据智能”的聪明型



通用性 (2) : “混合智能”的基本型

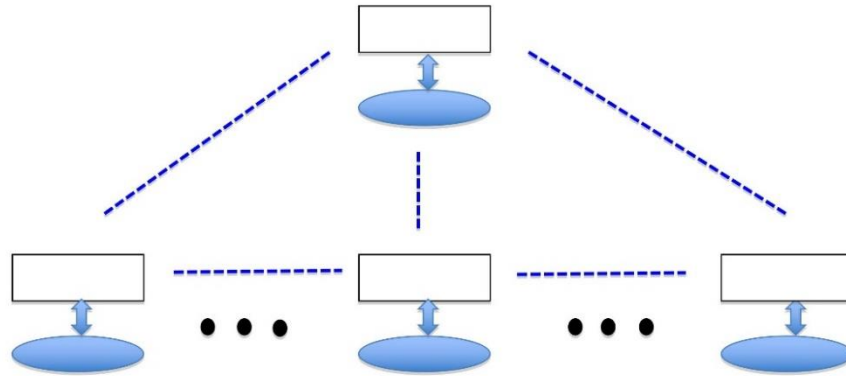


通用性 (3) : 是“自主智能”的母体

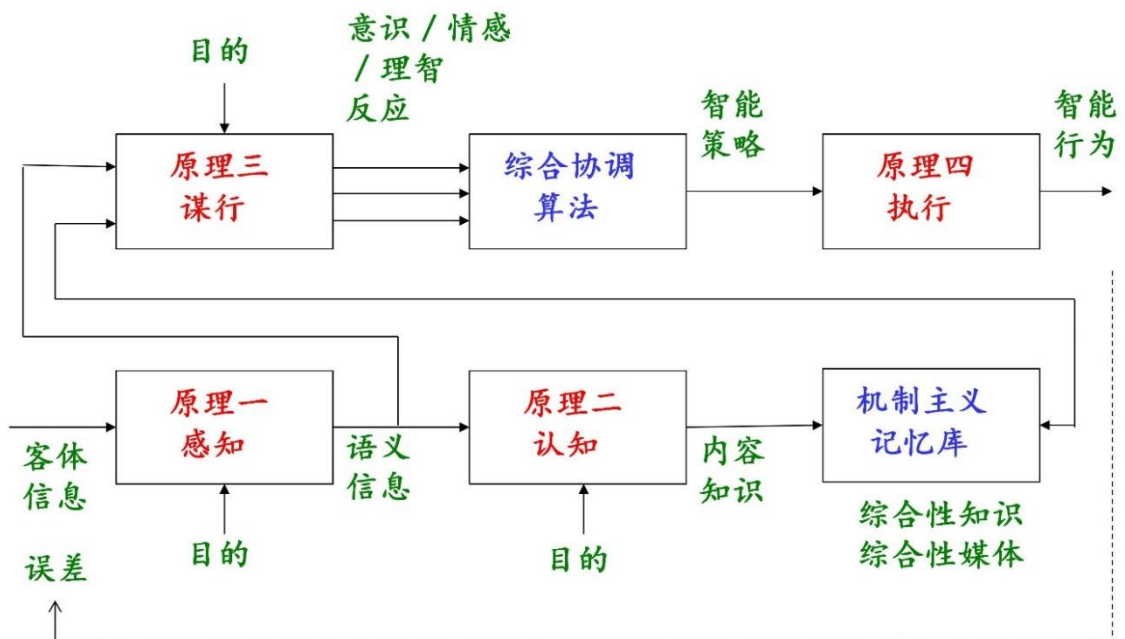


人工智能：机器所实现的人类“解决问题的能力”。

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）。

3-548, 迈向第三代 AI

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。

4-549, 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。

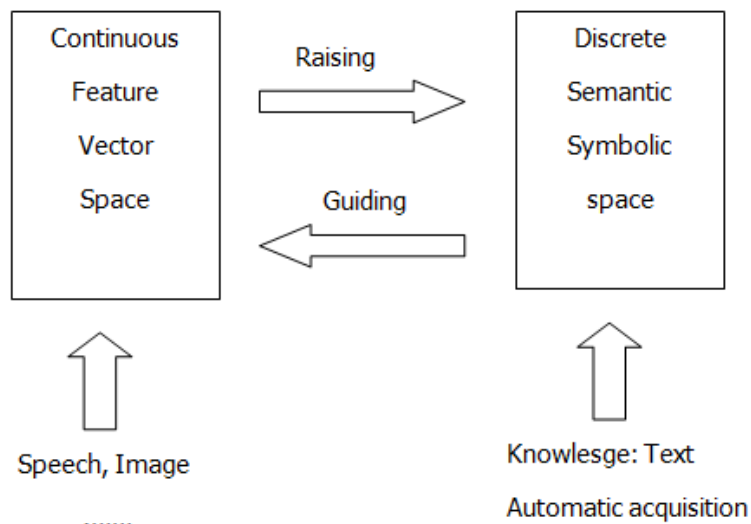


图 1 双空间模型

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。

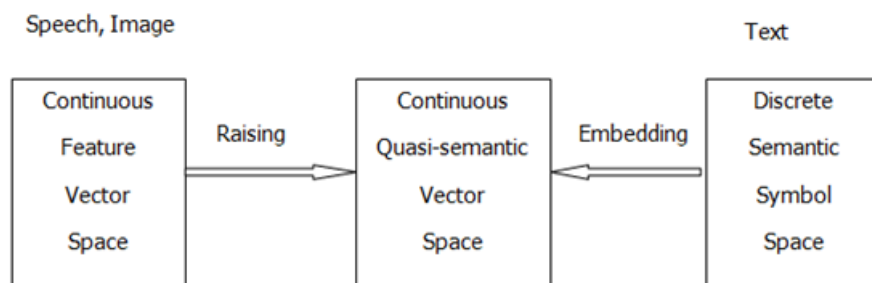


图 2 单一空间模型

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。

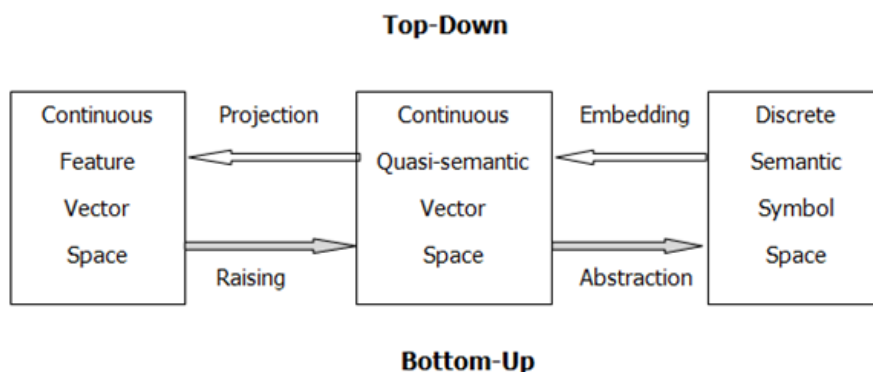


图3 三空间融合模型

5-550，也谈新一代人工智能

忠实的跟帖者

读了钟义信、张钹等老师谈论新一代人工智能（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！

6-559, 机制主义通用智能理论

(钟义信教授来信)

陆总：您好！您寄来的《国内外 AI 跟贴留言》第六集收到了，非常感谢！

向您报告：我们原先发现的“普适性智能生成机理”不但适用于人工智能，而且适用于人类智能。因此，“机制主义通用人工智能理论”可以推广成为“机制主义通用智能理论”。这里的关键就在于“研究范式的革命”。

这太有意思了！

如果您有兴趣和时间，我可以给您做一个具体的汇报。

钟义信

559 号附件：人工智能范式革命与通用智能理论的创生-钟义信，北京邮电大学人工智能学院

人工智能范式的革命与通用智能理论的创生

钟义信

(北京邮电大学人工智能学)

提要： 人工智能的研究取得了一批可喜的进展，也面临着诸多的挑战。为了应对这些挑战，学术界涌现了丰富多彩的创新思路。笔者相信，每种思路都有其合理之处，都有可能获得一定的成效。不过，根据笔者的分析，人工智能面临的最深刻最本质的挑战，是学科和时代的大转变所带来的大阵痛：范式的张冠李戴。因此，必须对人工智能的范式实施“正冠”：颠覆传统学科范式对人工智能研究的束缚，确立信息学科范式对人工智能研究的规范和引领。实施人工智能范式革命的结果，便创生了本文要介绍的《通用智能理论》。

一、为什么人工智能的根本出路是范式革命？

1、什么是人工智能？

智能与智慧是相联系又相区别的概念。**人类智慧**是人类为了生存发展的目的而运用知识去探索未来发现问题（**隐智慧**）和在此基础上变革现实解决问题（**显智慧**）的能力。隐智慧严格依赖于人类目的和思辨能力，只有人类能享有。而显智慧依赖于人类的智能操作，故称**人类智能**，可用机器来模拟。

人工智能是以人类智能（变革现实解决问题的显智慧）为原型、研究具有智能水平的机器系统为人类提供智能服务的学科。

2、人工智能研究的现状：局部有精彩，整体很无奈

人工智能系统的具体形态多种多样，分别源于三大学派。（1）1943 年发端的以模拟人脑结构为导向的人工神经网络学派，（2）1956 年兴起的以模拟人脑功能为标志的专家系统

学派，(3) 1990 年前后问世的以模拟智能系统行为为特色的感知动作系统学派。

经过数十年的努力，三大学派的研究都取得了一些精彩的成果。如人工神经网络的深度学习，专家系统的机器博弈，感知动作系统的智能机器人等。

但是，另一方面，三大学派的研究更面临着许多问题。其中最为严峻的挑战包括：它们的理解能力（真正的智能水平）都非常低，它们的通用能力都非常差，至今未能形成人工智能的整体理论。这些问题的严重性表现在：

(1) 智能水平低下，就**不够资格成为真正的人工智能**。

(2) 没有整体人工智能理论，就表明**人工智能的研究远远没有上轨**。系统学的原理表明：整体远远大于部分和。再多再好的部分成果之和，也远远不如整体成果。但是如何才能解决整体理论的建构？至今仍然众说纷纭。

3、人工智能存在问题的总根源：范式张冠李戴

人工智能的研究之所以会存在上述这些严重的问题，**根本原因在于：人工智能是一类开放复杂的信息系统，却遵循了传统物质学科的方法论：**

(1) 人工智能被分解为结构模拟、功能模拟、行为模拟三大学派，归因于运用了传统学科的“分而治之”方法论。对复杂信息系统**施行“分而治之”的结果就割断了复杂信息系统各个子系统之间复杂而隐秘的信息联系**，而这些复杂隐秘的信息联系正是复杂信息系统的生命线和灵魂。失去了（不可恢复）生命线和灵魂的各个子系统，就不再可能恢复原来的复杂信息系统！这是现行人工智能研究不能建立“整体理论”的根本原因。

(2) 人工智能系统智能水平低下，归因于运用了传统学科的“单纯形式化”方法论。智能的决策能力根植于对研究对象的形式、内容、价值的全面理解，而**施行“单纯形式化”的结果，丢失了内容和价值这样的智能“内核”**，单凭对表面形式的了解很难做出智能的决策。这是“智能水平低下”的根本原因。

“分而治之”和“单纯形式化”是传统物质学科的方法论。它们对传统学科的研究、发展与繁荣做出了伟大的历史性贡献，功不可没。然而，把它们用到人工智能研究领域，就用错了场合，变成了“张冠李戴”！

众所周知，方法论是为科学观服务的，有什么样的科学观就要求有什么样的方法论。传统学科的方法论在人工智能研究领域产生了负面的效果，说明传统学科的科学观也不适合于人工智能的研究。

学科的科学观阐明“学科的本质是什么”，学科的方法论则阐明“学科的研究应当怎么做”。于是，**学科的科学观与方法论两者一起，就决定了学科研究应当遵循的规范方式，称为学科的研究范式，简称范式。**

思想指导行动，任何学科的研究都必须遵循本学科的研究范式。以上的分析表明：**人**

人工智能的研究范式张冠李戴（用传统物质学科的范式来规约作为开放复杂信息系统的人工智能研究）是人工智能存在各种问题的总根源。

4、人工智能研究范式的张冠李戴，是历史的必然

学科的研究活动（社会存在）都由它的范式（社会意识）所支配。但是，由于“社会意识滞后于社会存在”法则的制约，20世纪中叶信息学科的研究活动迅猛兴起之后，直到如今都还没有能够形成信息学科的范式，因此，信息学科的研究便沿用了业已存在的传统物质学科范式。于是就造成了整个信息学科（含人工智能）范式的张冠李戴。

可见，人工智能范式的张冠李戴，并不是笔者的主观臆断，而是由物质学科主导的学科体系向信息学科主导的学科体系（由工业时代向信息时代）历史性大转变所带来的大阵痛，是不以人的意志为转移的客观规律。

5、范式革命：人工智能基础理论重大突破的必由之路

既然人工智能存在问题的总根源是“范式的张冠李戴”，那么，解决问题的对症良方就应当是“正冠”：**颠覆传统物质学科范式对人工智能研究的桎梏，确立信息学科范式对人工智能研究的规范和引领。**

那么，什么是信息学科的范式？表1示出了信息学科范式的内涵特征。为了便于比较，表1还列出了传统学科的范式以及现行人工智能所实行的范式。

表 1 学科范式的比较

事项	科学观	方法论
经典物质学科	机械唯物的物质观 对象：物质客体，排除主观因素 关注：对象的结构与功能 遵守：确定性演化，具有可分性	机械还原的方法论 描述方法：纯粹形式化 判断方法：形式匹配 宏观处置：分而治之
现行人工智能	近准的“物质观” 对象：脑物质，排除主观因素 关注：对象的结构与功能 遵守：可分性	明确的“还原论” 描述方法：纯粹形式化 判断方法：形式匹配 宏观处置：分而治之
现代信息学科	唯物辩证的信息观 对象：主体驾驭的主客互动信息过程 关注：主体目的 遵守：不确定性演化	信息生态方法论 描述方法：形式—内容—价值整体化 判断方法：内容理解 宏观处置：生态演化

通过表 1 三种范式（科学观和方法论）的详细解析和对比，可以十分清晰地看出：作为复杂信息系统的现行人工智能研究，它所遵循的研究范式本来应当是信息学科的研究范式（后者的内涵特征参见表 1 的第三栏），但实际上却是遵循了传统物质学科的研究范式（请比较表 1 的第 1 栏和第 2 栏）。这就是“人工智能范式张冠李戴”的具体表现。造成张冠李戴的深刻原因已在第 4 节阐明。

可见，只有颠覆和摒弃传统学科范式对人工智能研究的钳制，确立信息学科范式对人工智能研究的规范和引领，人工智能的发展才能走上正确的轨道。

这不是普通意义上的创新，而是**人工智能基础理论的彻底革命**：在学科研究的最高引领层次上实施的人工智能研究范式颠覆性变革，是实现信息学科范式从无到有的历史性突破和确立，是人工智能理论的彻底转轨。

二、范式革命怎样创生通用智能理论？

1、发现“学科创生”的一般规律：范式管控学科创生的全局全程

摆脱了传统物质学科范式对人工智能研究的束缚，确立了信息学科范式对人工智能研究的引领，全新的智能理论——通用智能理论——便终于破土而出。

创建“通用智能理论”的历程可以用笔者总结的“学科创生规律”来描述，详见表 2。

表 2 学科创生规律

生长阶段	模块名称	模块要素	要素解释
探索阶段 自下而上的探索	学科探路	多方摸索	通过长期自下而上多方摸索，总结失败教训和成功经验，提炼学科的研究范式（这一阶段常有“盲人摸象”现象）
		科学观	宏观上明确“学科是什么”
建构阶段 自上而下的建构	学科范式 (宏观定义)	方法论	宏观上明确“应该怎么做”
		学科模型	基于“学科范式”的学科全局蓝图
	学科框架 (具体定位)	研究路径	基于“学科范式”的整体研究方法
		学术结构	基于“学科定位”的学科宽度规格
	学科规格 (精准定格)	数理基础	基于“学科定位”的学科深度规格
		基本概念	基于“学科基础”的学科基本知识点
	学科理论 (完整理论)	基本原理	基本概念之间的相互联系

表2的“学科创生规律”显示，学科的创生需要经历两个阶段：首先是自下而上探索范式的阶段，然后是自上而下贯彻范式的建构阶段。

探索阶段 特点是：各种学术背景的研究者们进行自下而上的摸索、讨论和争论，总结失败的教训和成功的经验，逐渐提炼出普遍认可且科学合理的学科范式（科学观和方法论）。由于是经验性的摸索试探，常常发生“盲人摸象”现象，因此这一阶段经历的时间可能会很长：信息学科范式的形成便超越了半个多世纪。一旦达成了关于“范式”的共识，就可以自动转到建构的阶段。

建构阶段 特点是：根据探索阶段所总结的学科范式（即学科的宏观定义，包括学科的科学观和方法论），就可构筑学科的框架（即学科宏观定义的具体化，包括学科的全局模型和研究路径）、拟定学科的规格（即学科规格的精准化，包括学科的学术结构和所需数理基础的水准）和最终落实学科的基本理论（即学科理论的完整化，包括学科的基本概念和基本原理）。

从表2可以看出，在学科创生的整个过程中，学科范式自始至终都扮演着最高引领者和规范者的角色：整个探索阶段的任务是为了找到学科范式，而整个建构阶段的任务则是为了贯彻学科范式。

以下各节的内容，将分别介绍笔者团队依照表2的“学科创生规律”在信息学科范式引领下构建“通用智能理论”的系统成果。

2、研究和提炼“信息学科范式”

从当前的实际情况判断，无论是国内还是国外，人工智能学科应当遵循的范式都还处在“探索阶段”。因为，各种不同背景的研究人员仍然在按照各自对人工智能学科范式的理解进

行着不懈的探索，争论频出，共识罕有。值得庆幸的是，由于特殊的学术兴趣、背景和经历，本文笔者早在半个世纪之前就开始密切关注和潜心探索信息学科的科学观和方法论，为形成信息学科范式做了长期的研究和积累。

1962年笔者作为北京邮电大学信息论专业的研究生，在研读信息论原著的时候发现它的信息概念只关注了波形（信息的形式），丢了信息的内容和价值这两个核心要素，造成了“信息的空心化”。于是开始探索“形式、内容、价值”三位一体的“全信息理论”。26年之后，出版了国内外第一部以“全信息理论”为基础的《信息科学原理》。1987年在研究人工智能的时候，又发现它被分解为人工神经网络、专家系统以及后来的感知动作系统三个互不相容的学派。

于是，笔者认识到，从信息论到人工智能，都受到传统学科方法论“单纯形式化”和“分而治之”的影响，而这些方法论与信息学科（包括信息论和人工智能等）的特点格格不入。

基于这些认识，笔者在1988年出版的《信息科学原理》第10章就专门探讨了“信息科学的方法论”。此后在1996年的第二版、2002年的第三版、2005年的第四版、2013年的第五版每次再版都深化了信息学科方法论的探讨。2014年，笔者出版了另一部学术专著《高等人工智能理论》，又在全书的第一篇安排了专门的“高等人工智能的科学观与方法论”探讨。最终，总结提炼成为本文表1第三栏信息学科范式的标准表述，包括它的科学观和方法论。这是我国关于信息学科范式研究的宝贵成果。

3、范式革命：颠覆旧范式，确立新范式

颠覆传统学科范式对人工智能学科研究的统领地位，具体表现为：

(1) 在信息学科的研究领域摒弃传统学科的科学观，后者认为：

(1-1) 把研究对象看作是纯粹客观且遵循确定性演化规律的物质

(1-2) 彻底排除主观因素

(1-3) 研究的目的是阐明物质的结构

(2) 在信息学科研究领域摒弃传统学科的方法论，后者包括：

(2-1) 分而治之方法

(2-2) 纯粹形式化方法 确立信息学科范式对人工智能学科研究的规范和引领作用，具体表现为：

(1) 在信息学科的研究领域确立信息学科的科学观，后者认为：

(1-1) 把研究对象看作是主体与客体互动的具有不确定性的信息过程

(1-2) 强调主体的驾驭作用

(1-3) 研究目的是实现主体的目标

(2) 在信息学科的研究领域确立信息学科的方法论，后者包括：

(2-1) 信息生态演化方法

(2-2) 形式、内容、价值三位一体的整体化方法 可见，两种范式相反相成。要让人们放弃传统学科范式，其实远非易事！

4、遵循信息学科范式，构筑全新“人工智能全局模型”

信息学科范式的科学观业已指明：人工智能是代表主体（人类或其他生物） 的意志，在主体的驾驭下主体与环境客体相互作用而形成的不确定性信息过程。其中，所谓“代表主体（特别是人类主体）的意志”，是指：人工智能系统必须接受“主体所提出的求解问题、主体预设的求解问题目标，以及主体所提供的相关知识”；在这个框架下实现主体所确定的目标。

根据这些要求，可以构筑图1所示的人工智能学科研究的全局模型

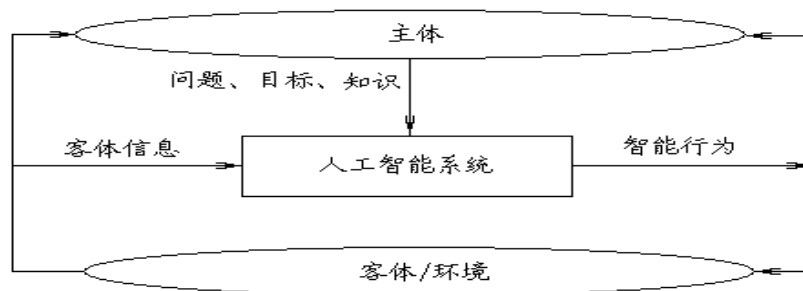


图 1 根据信息学科范式构筑的“人工智能全局模型”

图1的模型清楚表明：**人工智能的全局模型确实是“主体驾驭下（接受 来自主体的问题、目标和知识）的主体与环境客体相互作用（接受环境“客 体信息”的作用、产生“智能行为”反作用于环境客体）的具有不确定性的信息过程”**。这是真实人工智能系统的正确抽象。

现有的人工智能研究，包括以结构模拟为特征的人工神经网络和以功能模拟 为特征的专家系统，都把“孤立的脑”作为全局研究模型的原型。事实上，不接受外部环境客体信息刺激的孤立脑不能产生智能（“印度狼孩”的实验），而不向外部环境输出反作用的孤立脑也不可能检验脑的工作是否有意义。

由图 1 的人工智能模型还可以看出，人工智能系统所实现的，确实完全是自 然主体的目的，而不是人工智能“自己的目的”。事实上，**人工智能系统由于没有生命，因此不可能有它自身的目的和欲望**，不可能脱离主体的意志自行 其事，而只能成为人类主体的聪明助手与合作伙伴。

5、遵循信息学科范式，揭示智能生成机制，开创机制主义研究路径

信息学科范式的方法论指明：要按照信息生态演化（既然是生态演化，就不 允许被分割）方法来处置、要坚持运用形式、内容、价值三位一体的整体化方法来分析问题、要凭借理解来作出判断。

把信息学科范式方法论与图 1 的全局研究模型结合起来，就明确了人工智能 系统生成智

能的机制（注：“机制”与“机理”两者是同义语）应当是在问题、目标、相关知识的约束下，实现由“客体信息”到“智能行为”的复杂转换，如图2所示。

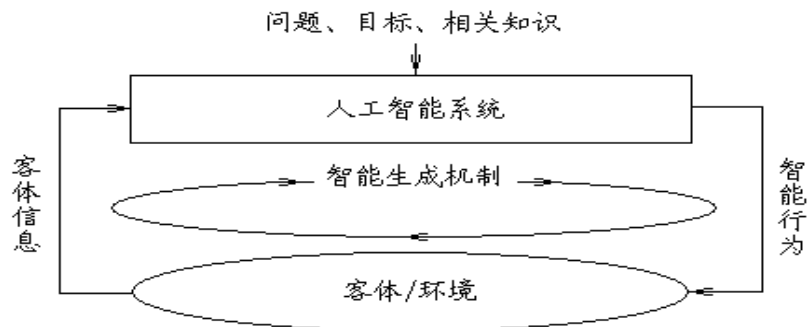


图 2 人工智能系统的智能生成机制

图 2 显示，人工智能系统中的智能生成机制的激励条件是环境客体所提供的“客体信息”，它的结果是主体操作下由客体信息转换生成的“智能行为”，而它所遵守的约束条件则是由主体所规定的“问题、目标、知识”。

质言之，智能生成机制的实质是“信息转换与智能创生”，它的具体转换与创生过程则是：**客体信息→感知信息→知识→智能策略→智能行为**，恰如图3所示。

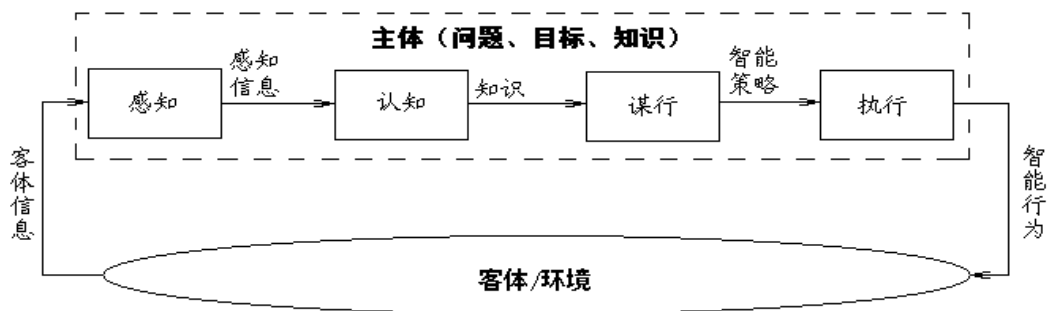


图 3 智能生成机制“信息转换与智能创生”的具体化

可以证明，“信息转换与智能生成”机制是普适性的，不仅适合于各种人工智能也适合于自然智能（包括人类智能），因此，可以名副其实地把它称为**普适性智能生成机制**，并由此把它的本质称为“**信息转换与智能创生定律**”。

普适性智能生成机制（机理）是人工智能研究的核心问题，也是人工智能研究的根本路径。于是，我们把以“普适性智能生成机制”为基础的研究路径，称为“**机制主义**”研究路径。这是与现有人工智能研究路径截然不同的研究路径。

6、遵循信息学科范式，重申人工智能的学术结构

信息学科范式要求保持学科的学术结构整体性（完整性），恢复人工智能学 科的本来面目。遵循完整性的要求，我们**把人工智能学科的学术结构理解为 以下各个学科群的交叉与综合：**

原型学科群：人类学，神经科学，认知科学，人文科学，社会科学，哲学**本体学科群：**信息科学，系统科学 **基础学科群：**生物物理学，逻辑学，数学 **技术学科群：**微电子，微机械，新材料，新能源由于传统物质学科范式强调对复杂对象施行“分而治之”，结果就把人工智

能学科分解出一些互不相容的分支学科，从而产生对人工智能学科的片面认识和误解。最典型的一种误解就是，仅仅根据专家系统的一家之情，竟把整个人工智能看作是计算机学科的一个分支。

7、遵循信息学科范式，重塑人工智能的数理基础

遵循信息学科范式的“统一性和整体性要求”，需要改造和重塑人工智能的 数理基础，特别是它的逻辑基础和数学基础。

为此，团队的何华灿教授创建了具有可调参数的“柔性（泛）逻辑理论”，从而把原来的标准数理逻辑和各种非标准逻辑纳入统一的逻辑连续谱系；与此同时团队的汪培庄教授创建了以因素为基元的“因素空间数学理论”，从而把原来互相离散的普通集合论、概率论、模糊集合论、粗糙集合论等相关数学分支和谐地纳入统一完整的数学理论。

8、遵循信息学科范式，重构人工智能的基础概念

在传统物质学科范式引领下，人们建立了一批人工智能的基础概念。但是，由于接受了“纯粹形式化”方法的影响，这些概念只有形式因素而没有内容因素和价值因素，因此都是“空心化”的基础概念，比如形式化的“数据”，形式化的“知识”，形式化的“智能”等等。事实上，正是这些空心化的概念，使现行的人工智能系统的智能水平（理解能力）非常低下。

符合信息学科范式理念的基础概念包括：全信息，全知识，全智能等。这里的前置词“全”并不是要求“胡子眉毛一把抓”，而是强调“形式、内容、价值”三位一体的整体化。只有全面了解事物的形式、内容和价值，才能理解事物。

全智能来源于全知识，全知识来源于全信息。因此，智能理论最基础的概念 是“全信息”。它的“形式、内容、价值”三位一体整体化体现为“语法信息（形式）、语义信息（内容）、语用信息（价值）”的三位一体。全信息的概念不是随 便说说而已，而是具有严格的生成机制，见图4

:

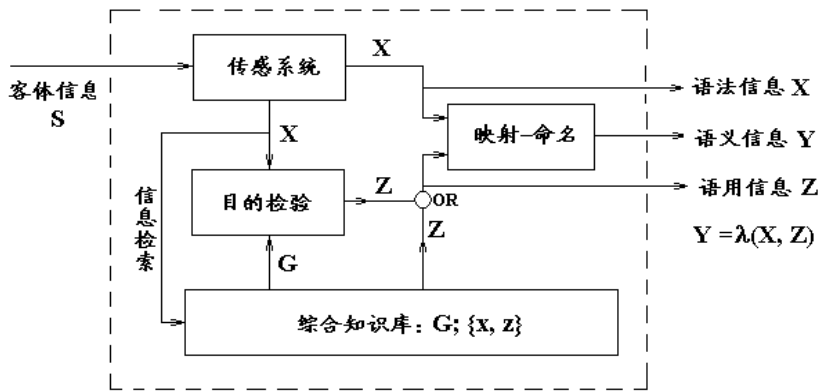


图 4 全信息的生成机制

图4的模型不仅阐明了全信息的生成机制，而且给出了语义信息（内容）的科学定义： $Y = \lambda(X, Z)$ ，其中算子“ λ ”代表“映射与命名”的逻辑操作。可见，人们掌握了语义信息，就同时掌握了语法信息和语用信息，也就理解了信息。所以，语义信息是用来“理解”事物的，而不是用来描述事物的统计特性的。

有了全信息的概念与生成机制，全知识与全智能的概念与生成便水到渠成。由此就可以建立“全知识”的知识库，它与传统知识库的根本区别就在于它的知识包含了“形式性知识、内容性知识、价值性知识”，因此可以有力地支持理解。值得指出，这样的“全知识库”也比目前流行的“知识图谱”更为优越。

总之，只有在“全信息”和“全知识”的基础上，才能具有“理解能力”，才能支持真正的“全智能”。

9、遵循信息学科范式，深挖人工智能的基本原理

信息学科范式强调“信息生态方法论”。因此，最为深刻的人工智能原理就是体现信息生态演化的“信息转换与智能创生定律”。这是一切人工智能和人类智能系统的本质和灵魂。

正如图3所表明的那样，“信息转换与智能创生定律”具体包含（1）“客体信息→感知信息（感知）”的转换原理，（2）“感知信息→知识”（认知）的转换原理，（3）“感知信息与知识→智能策略”（谋行）的转换原理，（4）“智能策略→智能行为”（执行）的转换原理，和（5）“误差信息→优化智能行为”（优化）的转换原理。这既是人工智能的基本原理，也是人类智能的基本原理。

值得指出，“信息转换与智能创生定律”的深远意义还在于，它是与物质科学领域的“质量转换与物质不灭定律”和能量科学领域的“能量转换与能量守恒定律”等量齐观的科学

定律，它们三者一起就完善了物质、能量、信息三大资源领域的三大科学定律。质量转换与物质不灭定律和能量转换与能量守恒定律阐明了这两个领域所存在的不能逾越的界限，信息转换与智能创生定律则告诉人们，可以通过信息转换创生人工智能系统为人类提供智能服务，把人类从体力劳动和有规可循的智力劳动中解放出来，以便更好地发挥人类的创造性能力，实现人类社会的可持续发展。

10、遵循信息学科范式，创建通用智能理论

综合集成以上各项成果，特别是其中以“信息转换与智能创生定律”为旗帜的机制主义研究路径（具体包含上述五项基本原理），可以创立“通用智能理论”，它的模型如图5所示。

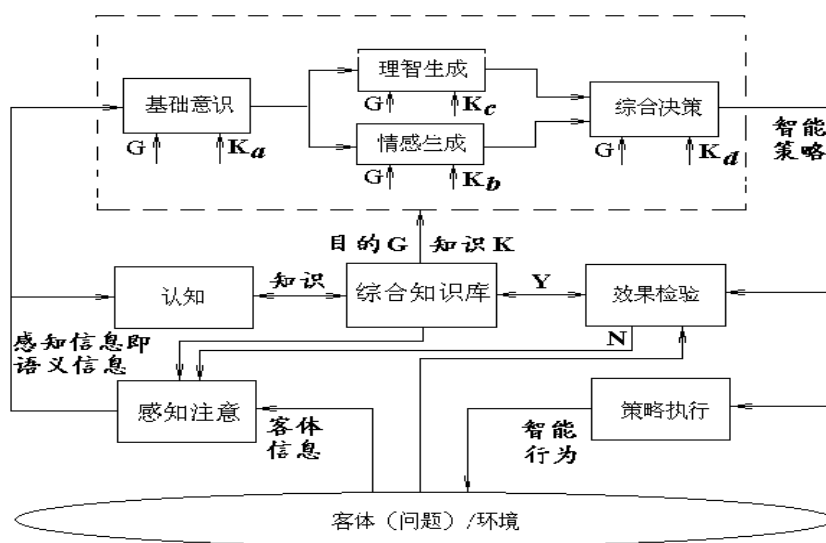


图 5通用智能理论模型

可以看出，图5的通用智能模型不但全面体现了信息学科范式（科学观和方法论）理念，体现了人类学、神经科学、认知科学、信息科学（含智能科学）、系统科学、柔性逻辑理论、因素空间数学理论等科学精神，而且，也展现了人类智能的精髓。特别体现了“物质变精神和精神变物质”的辩证法，以及“认识世界和改造世界，并在改造客观世界的过程中也改造自己”的主客互动理论。

三，附带的话

人工智能的范式革命与通用智能理论的创生，涉及到许多深刻的科学问题和哲学问题。为了节约篇幅，以下这些问题只列出题目而不做解释：

- 1、通用智能理论模型（图5）的实现原理已经落地。
- 2、通用智能理论可以克服现行人工智能存在的各种问题。
- 3、通用智能理论系统可以孵化出各种应用的人工智能系统。

4、把孵化出的各种人工智能系统用于各行业，我们规划为“洛神工程”。

5、目前正在寻求合作，把理论转化为通用智能孵化平台和现实生产力。

7-569,《通用智能理论》创新思想梳理

陆总：这是我的与众不同的观念和方法。请您批评！——钟义信

《通用智能创新思考梳理》

一、怎样发现问题？——看本质

(1)

人工智能存在的主要问题之一

是学科的整体被肢解

(鼎足三分，无法形成整体理论)

而不是通用性够不够、移植性好不好？

(2)

人工智能存在的主要问题之二

是它的“智能”被置空

(没有理解能力、结果不可解释)

而不是理解力好不好、可解释性强不强

(3)

整体被肢解：归因于“分而治之”方法论

智能被置空：归因于“纯粹形式化”方法论

总之，归因于研究范式

不是仅靠算法、算力、数据的改善就能解决

也不是把结构类脑 / 功能类脑结合就能解决

或者把“数据驱动”与“知识驱动”相结合就可以

二、怎样寻找解决问题的方法—循规律

(4)

只有突破自然科学与社会科学哲学之间的藩篱

才能看清人工智能问题的根源在范式

局限于自然科学范畴内，只能看到算法、算力、数据

只能看到脑结构、脑功能

(5)

突破藩篱之后就可以看出

驾驭学科研究的最高引领力量是学科的研究范式

科学观回答：学科是什么？

方法论回答：学科怎么做？

两者共同决定了学科规范定义

否则，学科定义就有随意性

(6)

范式，是统领学科研究的龙头

首先自下而上探寻范式

然后依照范式（学科定义）自上而下建构学科：

学科定义（定义）

学科框架（定位）

学科规格（定格）

学科理论（定论）

人们往往只看到“学科理论”，看不到它的根基

(7)

调研发现

人工智能问题的根源是：范式的张冠李戴！

(用传统学科的范式指导人工智能研究)

所以，人工智能的根本问题是：范式革命

其他研究都解决不了根本问题

三、怎样解决问题一把范式革命贯彻到底

(8)

为此，需要建立信息科学的研究范式

科学观：主客互动信息过程，而非物质客体

方法论：信息生态论，而非分而治之 / 纯粹形式化

(9)

人工智能的全局模型是

主体驾驭下的主客互动信息过程

而不是孤立的脑

(10)

人工智能的研究路径是

建立普适性的智能生成机制

而不是结构、功能、行为的分立路径

(11)

人工智能的学科结构是

原型科学、本体科学、基础科学、技术科学的交叉综合
而不是计算机科学的分支

(12)

人工智能的基础科学是
泛逻辑理论和因素空间数学理论
而不是传统的概率论和数理逻辑

(13)

人工智能的基本概念是
形式内容价值三位一体的全信息、全知识、全智能
而不是纯粹形式化数据、形式化知识、形式化智能

(14)

人工智能的基本原理是
信息转换与智能创生定律
它与
质量转换与物质不灭定律
能量转换与能量守恒定律
构成物质、能量、信息的三大定律

(15)

普适性的智能生成机理
既适合于人工智能，也适合于自然智能
但实现机理的方式各有特色

(16)

基于以上的突破、发现和建构

创立了《通用智能理论》

不是“不存在通用智能理论”

8-571, 李德毅院士关于通用人工智能十问

钟义信教授、张钹院士分别提出要构建作为下一代人工智能的通用人工智能（或强人工智能）。

通用人工智能即下一代人工智能，需要全新的理论基础，当前无论从哲学层面讲，或从科学层面讲，尚不完备！

今天人工智能的目标不是构建与人脑完全一样的东西，也并非全知全能，而是模拟人脑智能（即构建类脑智能）。

如果没有人工智能全新的理论突破，没有信息处理机制全新范式，李院士的“十问”在现有人工智能理论框架内将难以回答，而如不能回答这些问题，那么通用人工智能也就很难实现。

下面是李德毅院士关于通用人工头智能的十问：

一、意识、情感、智慧和智能，它们是包含关系还是关联关系？是智慧里面有智能，还是智能里面有智慧？大凡意识、情感都是内省的、自知的、排他的，怎么可以用他人的、人工的来代替呢？所以非生命体不可能有意识？

二、如何理解通用智能？我们应该不应该把通用智能理解为“全知全能”或者单向超强智能？尽管今天的计算机已经可以解决很多复杂的、专门的智力问题（如

围棋智能), 我们仍常常觉得它们缺乏人类思维的某些本质特征。这里的差别主要不是在算法、算力、数据量方面, 不是在速度和容量方面, 而是在智能的一般性、通用性、普遍性、灵活性、缺省性、容错性、可习得性、不确定性、适应性、常识性、开放性、创造性、自主性等方面。遗憾的是发展 60 多年的人工智能没有能够更靠近人的原始的智能。

三、目前所有的人工智能的成就都是在计算机上表现出来, 是基于冯架构的计算机智能或者计算智能, 人工智能是计算机的一个应用而已。而人脑不是冯诺依曼架构的, 存在不存在宏观上更类似脑的非冯诺依曼架构呢? 例如, 对人的智能而言, 记忆力是真正的智力, 超强记忆力就是超强智能, 记忆比计算机重要, 记忆的提取要比复杂的推理快得多, 非冯架构如何在结构上体现人脑的不同记忆区和记忆力呢? 如何体现环境和知识的双驱动?

四、非生命体不会有七情六欲, 机器人是非生命体, 还会有学习的原动力吗? 如果没有学习的原动力, 没有接受教育的自发性, 还会有学习的目标吗? 目标从哪产生? 机器人能否自己提出问题?

五、人的注意力选择源于记忆, 源于记忆的偏好依附性, 偏好如何产生的? 偏好依附是否只能与交互认知的频度和时间的远近相关? 人的偏好依附不是这样的, 人的恐惧性以及满足感会让一些发生频度很低、或者很久远的事记忆特别深刻。

六、自然语言是人类思维活动的载体, 如果自然语言是第一语言, 数学语言是第二语言, 计算机语言是第三语言, 后一个比前一个常常更严格, 后一个比前一个常常更狭义, 根据哥德尔不完全定理, 数学自身难以完全自洽。数学的形式化要

借助于自然语言，计算机语言的形式化要借助于数学语言。因此，人工智能怎么可以反过来要用数学语言或者计算机语言去形式化人类的自然语言呢？

七、人脑是个小宇宙，其中的智能是多情境、多公理兼容并包的，在不同情境里有不同应对，不完全收敛，不完全自恰，不整体统一，不存在非公理的统一的数学推理，当然也不必一定要脑裂。

八、一个机器或者系统是否有智能，不在于某一个时刻它能解决什么实际的智力问题，而在于它有没有学习的能力？智能，即提供的问题解决方案，是否依赖于有限的认知资源？是否需要进一步交互认知？是否可以有选项？是否可以进化和成长？这才是最重要的。

九、在一个非冯诺依曼架构的机器人脑中，组成记忆、交互和计算的基本元件最少有哪几种？各元件中的信息的产生机制与存在形式是什么样的？他们之间的信息传递机制是什么样的？

十、通用智能后天的习得靠教育，智能植根于教育，文明是智能的生态。

9-584，答李院士问六：自然语言问题——王迪兴

问六：自然语言是人类思维活动的载体，如果自然语言是第一语言，数学语言是第二语言，计算机语言是第三语言，后一个比前一个更严格、更狭义。数学自身难以完全自治，数学的形式化借助于自然语言，计算机语言的形式化要借助于数学语言。因此，人工智能怎么可以反过来要用数学语言或计算机语言去形式化人类的自然语言呢？

答：自然语言、数学语言、计算机语言必须基于某种同构性才能建立，互相翻译、代偿、转换。没有同构性任何语言都无法相互沟通与交流，更不会产生有效认知结果。各种语言都是准完备的，都要符合逻辑一致性，基于逻辑一致性进行内涵和外延拓展，才会体现语言的有效性。任何广义语言都不能单一发挥作用，一定要有各种表达方式互补，如语言必须与文字、图形甚至手势互补发挥作用。语言不存在谁形式化谁的问题，都是基于同构互补发挥作用。

计算机语言是基于自然语言派生的，但基于机器语言不能派生自然语言，因为计算机本身不具创造语言的功能，其语言不具主体地位，也没有自主社会化交流的需求，不存在计算机语言形式化自然语言的问题！

10-619，钟义信：《通用人工智能理论》

人工智能范式的革命与通用智能理论的创生

钟义信（北京邮电大学人工智能学院）

提要：人工智能的研究取得了不少可喜的进展，也面临着许多紧迫的问题。为了应对这些挑战，学术界提出了各种各样的研究思路。笔者相信，每种思路都有其合理之处，都有可能获得一定的成效。不过，根据笔者的理解，人工智能面临的最深刻最严峻的挑战，是学科和时代的大转变所带来的大阵痛：人工智能范式的张冠李戴。因此，必须对人工智能的范式实施“正冠”：颠覆传统学科范式对人工智能研究的束缚，确立信息学科范式对人工智能研究的规范和引领。实施人工智能范式革命的结果，创生了本文要介绍的《通用智能理论》。

一、为什么人工智能的根本出路是范式革命？

1、什么是人工智能？

智能是智慧的子集。人类智慧是人类为了生存发展的目的而不断地运用知识去探索未来定义问题（隐智慧）进而解决问题变革现实（显智慧）的能力。其中的隐智慧严格依赖于人类的目的和思辨能力，只有人类才能拥有。而显智慧则依赖于人类的智能求解操作，因此特称为人类智能。它，可用机器模拟。

根据科学技术的拟人律，人工智能是以人类智能（解决问题变革现实的显智慧）为原型、研究具有智能水平的机器为人类提供智能服务的学科。

2、人工智能研究的现状：局部有精彩，整体很无奈

人工智能的研究，存在三大学派：以模拟人脑结构为导向的人工神经网络学派，以模拟人脑逻辑功能为标志的专家系统学派，以模拟智能系统行为为特色的感知动作系统学派。但互不认可，未能形成合力。

经过数十年的努力，三大学派的研究都取得了一些精彩的成果。如人工神经网络的深度学习，专家系统的机器博弈，感知动作系统的智能机器人等。

但是，另一方面，三大学派的研究更面临着许多问题的困扰。其中最为严峻的挑战包括：它们的理解能力（真正的智能水平）都非常低；它们的通用能力都非常差；至今未能形成人工智能的整体理论。这些问题的严重性在于：

（1）智能水平低下，就不够资格成为真正的人工智能。

（2）没有通用的整体人工智能理论，表明人工智能的研究还没上轨道。

系统学的原理表明：整体远远大于部分和。再多再好的部分成果之和，也远远不及整体成果。但是如何才能构建通用的整体理论？至今仍然没有共识。

3、人工智能存在问题的总根源：范式张冠李戴

思想指导行动，学科研究活动的指导思想就是学科的科学观和方法论。学科的科学观阐明“学科的本质是什么”，学科的方法论则阐明“学科的研究应当怎么做”。于是，学科的科学观与方法论一起，就决定了学科应当遵循的规范研究方式，称为学科的研究范式，简称范式。

人工智能的研究之所以存在上述这些严重问题，根本原因在于：人工智能是一类开放复杂的信息系统，却遵循了传统物质学科范式的方法论：

(1) 人工智能被分解为结构模拟、功能模拟、行为模拟三大学派，归因于运用了传统学科范式的“分而治之”方法论。对复杂信息系统施行“分而治之”的结果就割断了复杂信息系统各个子系统之间复杂而隐秘的信息联系，而这些复杂隐秘的信息联系正是复杂信息系统的生命线和灵魂。失去了生命线和灵魂的各个子系统，就不再可能合成原来的复杂信息系统！这是现行人工智能研究不能建立“整体理论”的根本原因。

(2) 人工智能系统智能水平低下，归因于运用了传统学科范式的“单纯形式化”方法论。智能的决策能力依赖于对研究对象的形式、内容、价值的全面理解。施行“单纯形式化”丢弃了智能的“内核”：内容和价值因素，而仅仅了解对象的形式，极难做出智能的决策。这是“智能水平低下”的根本原因。

众所周知，“分而治之”和“单纯形式化”是传统物质学科的方法论。它们对传统学科的研究、发展与繁荣做出了伟大的历史性贡献。然而把它们用到作为开放复杂信息系统的人工智能研究领域，就犯了范式“张冠李戴”的大忌！

4、人工智能研究范式的张冠李戴，是历史的必然

人工智能范式的张冠李戴并非偶然现象。这是因为，学科的研究活动属于社

会存在，它的范式则属于社会意识。由于受到“社会意识滞后于社会存在”法则的制约，20世纪中叶信息学科的研究活动迅猛兴起之后，直到如今都没有形成信息学科范式的共识，信息学科的研究便沿用了业已存在而且众所熟知的传统物质学科范式，于是造成了信息学科（含人工智能）范式的张冠李戴。

千年的科学发展，都属物质学科（含材料科学和能量科学），都遵循着同样的物质学科范式，因而没有发生过范式的变革。但是当今却正在发生由物质学科主导的学科体系向信息学科主导的学科体系的历史性大转变：研究活动的大转变在前，研究范式的大转变在后——发生在人工智能由初级阶段进入高级阶段的当下。因此，人工智能范式的大转变（范式大革命）正是这个历史大转变所带来的大阵痛。没有这个大阵痛，便无法完成人工智能研究由初级阶段向高级阶段的转变。这是不以人的意志为转移的客观规律。

5、范式革命：人工智能基础理论重大突破的必由之路

既然人工智能存在问题的总根源是“范式的张冠李戴”，那么，解决问题的对症良方就应当是“正冠”：颠覆传统物质学科范式对人工智能研究的束缚，确立信息学科范式对人工智能研究的规范和引领。

那么，什么是信息学科的范式？表1示出了信息学科范式的内涵特征。为了便于比较，表1还列出了传统学科的范式以及现行人工智能所实行的范式。

表 1 学科范式的比较

事项	科学观	方法论
经典物质学科	机械唯物的物质观 对象：物质客体，排除主观因素 关注：对象的结构与功能 遵守：确定性演化，具有可分性	机械还原的方法论 描述方法：纯粹形式化 判断方法：形式匹配 宏观处置：分而治之
现行人工智能	近准的“物质观” 对象：脑物质，排除主观因素 关注：对象的结构与功能 遵守：可分性	明确的“还原论” 描述方法：纯粹形式化 判断方法：形式匹配 宏观处置：分而治之
现代信息学科	唯物辩证的信息观 对象：主体驾驭的主客互动信息过程 关注：主体目的 遵守：不确定性演化	信息生态方法论 描述方法：形式—内容—价值整体化 判断方法：内容理解 宏观处置：生态演化

通过表 1 三种范式（科学观和方法论）的详细解析和对比，可以十分清晰地看出：作为复杂信息系统的现行人工智能研究，它所遵循的研究范式本来应当是信息学科的研究范式（后者的内涵特征参见表 1 的第三行），但实际上却遵循了传统物质学科的研究范式（试比较表 1 的第 1 和第 2 行）。这就是“人工智能范式张冠李戴”的具体表现。

可见，只有颠覆和摒弃传统学科范式对人工智能研究的制约，确立信息学科

范式对人工智能研究的规范和引领，人工智能的发展才能走上正确的轨道。

二、范式革命怎样创生通用智能理论？

如上所述，人工智能的范式革命，就是在人工智能的研究领域以现代信息学科的范式置换传统物质学科的范式。那么，信息学科的范式又将怎样规范和引领人工智能的研究呢？

1、 总结“学科创生”的普遍规律：范式引领学科创生的全局全程

信息学科范式引领创建“通用智能理论”的全部阶段和历程，可用表2所总结的“学科创生规律”来描述。

表2 学科创生规律

生长阶段	生长模块	模块要素	要素解释
自下而上 探索范式的阶段	学科探路	多方摸索	通过长期自下而上多方摸索，总结失败教训和成功经验，提炼学科的研究范式（这一阶段常有“盲人摸象”现象）
自上而下 落实范式的阶段	学科范式 (宏观定义)	科学观	宏观上明确学科本质“是什么”
		方法论	宏观上明确学科研究“怎么做”
	学科框架 (具体定位)	学科模型	基于学科范式的学科全局蓝图
		研究路径	基于学科范式的整体研究途径
	学科规格 (精准定格)	学术结构	基于学科范式和框架的学科宽度规格
		数理基础	基于学科范式和框架的学科深度规格

	学科理论 (完整理论)	基本概念	基于范式、框架、规格的学科基本知识
		基本原理	基于范式、框架、规格的学科概念联系

表 2 的“学科创生规律”显示，学科的创生需要经历两个阶段：首先是自下而上探索范式的阶段，然后是自上而下落实范式的建构阶段。

探索阶段特点是：各种相关学术背景的研究者们各自进行自下而上的摸索、讨论和争论，总结失败的教训和成功的经验，逐渐提炼出普遍认可且科学合理的学科范式（科学观和方法论）。由于是各自展开的摸索，常有“盲人摸象”的现象，因此，这一阶段经历的时间可能很长。信息学科范式的形成便超越了半个多世纪，至今尚未在学术共同体达成共识。

建构阶段特点是：根据探索阶段所总结的学科范式（即由学科的科学观和方法论阐明的学科定义），就可根据学科定义构筑学科的框架（即学科宏观定义的具体化，包括学科的全局模型和研究路径）、进而拟定学科的规格（即学科规格的精量化，包括学科的学术结构和所需数理基础的水准规格）和最终落实学科的基本理论（即学科理论的完整化，包括学科的基本概念和基本原理）。

从表 2 可以看出，在学科创生的整个过程中，学科的范式自始至终都扮演着最高引领者和规范者的角色：整个探索阶段的任务是为了找到学科的范式，而整个建构阶段的任务则是为了落实学科的范式。

揭示和阐明“学科创生规律”，是构建“通用智能”的第一步。以下将分别阐述笔者团队依照表 2 的规律构建“通用智能理论”的各个步骤和结果。

2、研究和提炼“信息学科范式”

从当前的实际情况可以判断，无论是国内还是国际，人工智能学科应当遵循的范式都还处在“探索阶段”，这表现为：各种不同背景的研究人员仍然在按照各自对人工智能学科范式的理解进行着不懈的探索，争论频出，共识鲜有。

由于接受了“整体观和辩证论”思想的熏陶，也由于特殊的学术兴趣、背景和经历，本文笔者早在半个世纪之前就开始密切关注和潜心探究信息学科的科学观和方法论，为形成信息学科的范式做了长期的研究和积累。

1962年笔者作为北京邮电学院信息论专业的研究生，在研读信息论原著的时候发现它的信息概念只关注了波形，丢了信息的内容和价值这两个核心要素，造成了“信息概念空心化”。于是开始探索、并于1986年发表了“形式、内容、价值”三位一体的“信息的综合测度”（即“全信息理论”）。接着，1988年出版了国内外第一部以“全信息理论”为基础的《信息科学原理》。1987年又发现人工智能的研究被分解为人工神经网络、专家系统以及后来的感知动作系统三个互不相容的学派。于是笔者认识到，从信息论到人工智能，都受到了传统物质学科方法论“单纯形式化”和“分而治之”的影响，而这些方法论与信息学科（包括信息论和人工智能等）的特点格格不入。

基于这些认识，笔者在1988年出版的《信息科学原理》第10章就专门探讨了“信息科学的方法论”。此后在每次再版（直到2013年第五版）都深化了信息学科方法论的探讨。2014年，笔者出版了另一部专著《高等人工智能理论》，又在全书的第一篇安排了专门的“高等人工智能的科学观与方法论”探讨。2017年在《哲学分析》杂志发表了论文“从机械还原方法论到信息生态方法论”。最终，总结提炼成为本文表1第三行信息学科范式内涵的标准表述，包括它的科学

观和方法论。这是我国关于信息学科范式研究的基本成果。

3、范式革命：颠覆旧范式，确立新范式

颠覆传统物质学科范式对人工智能学科研究的统领地位，具体来说就是要在人工智能研究领域摒弃如下的科学观和方法论：

(1) 传统物质学科的科学观：

(1-1) 把研究对象看作是纯客观且遵循确定性演化规律的物质

(1-2) 彻底排除主观因素

(1-3) 研究的目的是阐明物质的结构

(2) 传统物质学科的方法论：

(2-1) 分而治之方法

(2-2) 纯粹形式化方法

确立信息学科范式对人工智能学科研究的规范和引领作用，具体来说就是要在人工智能研究领域确立如下的科学观和方法论：

(1) 信息学科的科学观：

(1-1) 把研究对象看作是主体与客体互动的具有不确定性的信息过程

(1-2) 强调主体的驾驭作用

(1-3) 研究目的是实现主体的目标

(2) 信息学科的方法论：

(2-1) 信息生态方法

(2-2) 形式、内容、价值三位一体的整体化方法

可见，信息学科范式与传统物质学科范式两者几乎相反相成。然而，千百年来传统物质学科范式已在人们心中深深扎根，而信息学科范式才刚刚露出嫩芽。因此，要让人工智能研究领域的科技工作者放弃自己所熟悉所深信的传统物质学

科范式，绝非易事！不过，非如此就不能取得人工智能的突破与升级。

4、构筑全新“人工智能全局模型”（学科框架A）

信息学科范式的科学观业已指明：人工智能是在主体驾驭下的主体与环境客体相互作用所形成的不确定性信息过程。其中，所谓“主体（特别是人类主体）驾驭”，实际是指：人工智能系统必须接受主体所提出的“求解问题”、主体预设的“求解问题目标”，以及主体所提供的“相关知识”；人工智能系统只能在这个框架下去实现主体所预设的目标。

根据这些思想，可以构筑图 1 所示的人工智能学科研究的全局模型：

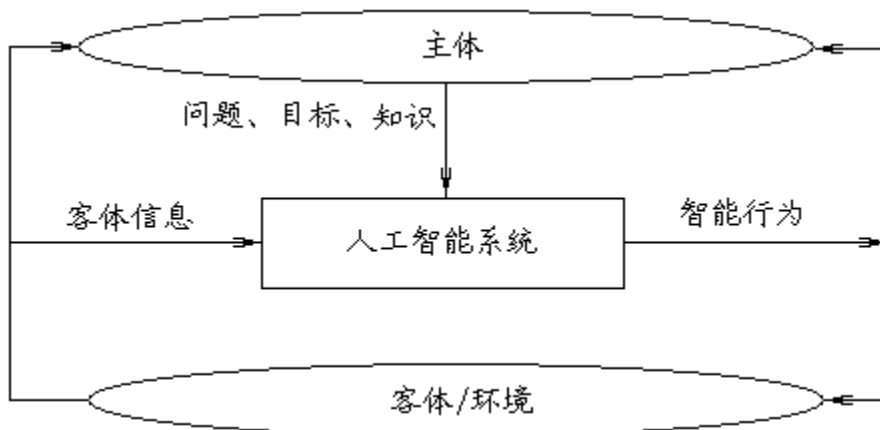


图 1 根据信息学科范式构筑的“人工智能全局模型”

图 1 的模型清楚表明：人工智能的全局模型确实是“主体驾驭下（接受来自主体的问题、目标和知识）的主体与环境客体相互作用（接受环境“客体信息”的作用、针对这种作用产生“智能行为”反作用于环境客体）的具有不确定性的信息过程”。这是真实人工智能系统的正确抽象。

现有的人工智能研究，包括以结构模拟为特征的人工神经网络和以功能模拟

为特征的专家系统，都把“孤立的脑”作为全局研究模型的原型。事实上，不接受外部环境客体信息刺激的孤立脑不能产生智能（“印度狼孩”的实验），而不向外部环境输出反作用的孤立脑也不可能检验脑的工作是否有意义。

由图 1 的人工智能模型还可以看出，人工智能系统所实现的，完全是自然主体的目的，而不是人工智能系统“自己的目的”。事实上，人工智能系统由于没有生命，因此不可能有它自身的目的和欲望。它不可能脱离主体的意志自行其事，而只能成为人类主体的聪明助手与善解人意的合作伙伴。

5、揭示智能生成机制，开创机制主义研究路径（学科框架 B）

信息学科范式的方法论指明：要按照信息生态演化（既然是生态演化，就不允许把信息剥离成形式化的语法信息，也不允许把信息生态过程割断为封闭的信息过程）方法论来处置问题、要坚持运用形式、内容、价值三位一体的整体化方法来分析问题、要凭借理解来作出判断。

把信息学科范式的方法论与图 1 的全局研究模型相结合，就明确了人工智能系统生成智能的机制应当是在“问题、目标、相关知识”的约束下实现由“客体信息”到“智能行为”的复杂转换，如图 2 所示。

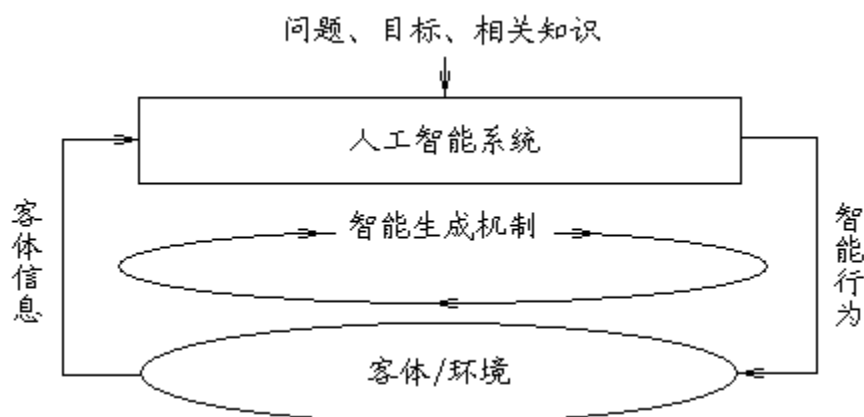


图 2 人工智能系统的智能生成机制

图 2 显示，人工智能系统中的智能生成机制的激励条件，是环境客体所提供

并经主体所选择的问题（即“客体信息”），智能生成机制所生成的结果是由客体信息转换而成的“智能行为”，而它所遵守的约束条件则是由主体所提供的预设目标和知识。

智能生成机制的实质是“信息转换与智能创生”，具体转换与创生过程则是：客体信息→感知信息→知识→智能策略→智能行为，如图3所示。

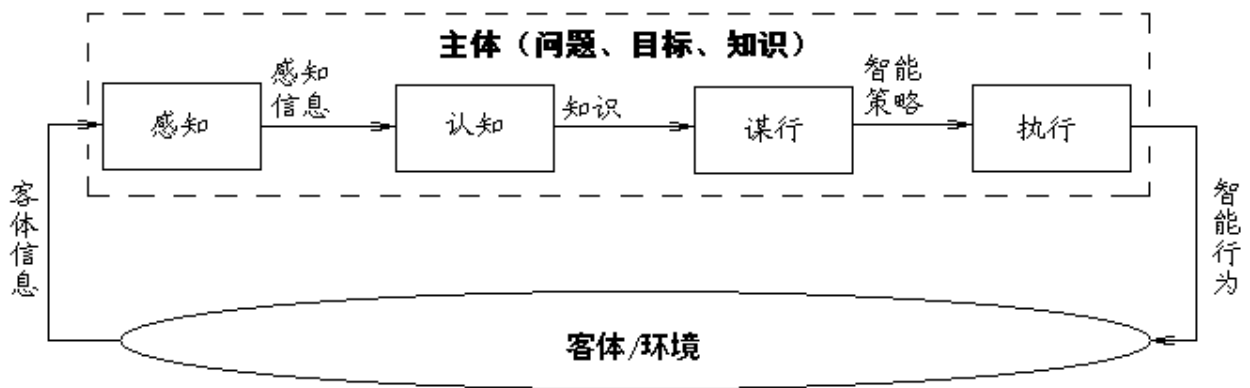


图3 智能生成机制“信息转换与智能创生”的具体化

可以证明，“信息转换与智能创生”机制是普适性的，不仅适合于各种人工智能也适合于自然智能（包括人类智能），因此，可以名副其实地把它称为普适性智能生成机制，并把它本质称为“信息转换与智能创生定律”。

普适性智能生成机制是人工智能研究的根本路径，以“普适性智能生成机制”为标志的研究路径称为“机制主义”研究路径。这是与现有人工智能的结构主义路径、功能主义路径、行为主义路径鼎足三分的情况截然不同的全新研究路径。

6、重审人工智能的学术结构（学科规格A）

信息学科范式要求保持学科的学术结构整体性（完整性），恢复人工智能学科的本来面目。遵循完整性的要求，就应把人工智能学科的学术结构理解为以下各个学科群的交叉与综合：

原型学科群：人类学，神经科学，认知科学，人文科学，社会科学，哲学

本体学科群：信息科学，系统科学

基础学科群：生物物理学，逻辑学，数学

技术学科群：微电子，微机械，新材料，新能源

由于传统物质学科范式强调对复杂对象施行“分而治之”，结果就把人工智能学科分解出一些互不相容的分支学科，从而产生对人工智能学科的片面认识和误解。最典型的一种误解就是，仅仅根据专家系统的一家之情，把整个人工智能学科看作是计算机学科的一个分支。这显然不是对人工智能的正确理解。

7、重塑人工智能的数理基础（学科规格 B）

遵循信息学科范式的“统一性和整体性要求”，需要改造和重塑人工智能的数理基础，特别是它的逻辑基础和数学基础。

为此，本团队一方面创建了具有可调参数的“柔性逻辑理论”，从而把原来的标准数理逻辑和各种非标准逻辑纳入统一的逻辑连续谱系；同时本团队创建了以因素为基元的“因素空间数学理论”，从而把原来互相独立发展的普通集合论、概率论、模糊集合论、粗糙集合论等相关数学分支和谐地纳入统一完整的人工智能数学理论。

8、重构人工智能的基础概念（学科理论 A）

在传统物质学科范式指导下，人们建立了一批人工智能的基础概念。但是，由于接受了“单纯形式化”方法的影响，这些概念只有形式因素而没有内容因素和价值因素，因此基本都是“空心化”的基础概念，比如形式化的“数据”，形式化的“知识”，形式化的“智能”等等。事实上，正是这些空心化的概念，使现行的人工智能系统的智能水平（理解能力）非常低下。

符合信息学科范式理念的基础概念包括：全信息，全知识，全智能等。这里的前置词“全”并非要求“胡子眉毛一把抓”，而是强调“形式、内容、价值”三位一体的整体化。只有全面了解事物的形式、内容和价值，才能理解事物。

全智能来源于全知识，全知识来源于全信息。因此，智能理论最基础的概念是“全信息”。它的“形式、内容、价值”三位一体整体化体现为“语法信息（形式）、语义信息（内容）、语用信息（价值）”的三位一体。“全信息”的概念具有严格的生成机制和定义方法，如图4所示（详细的解释件文献：

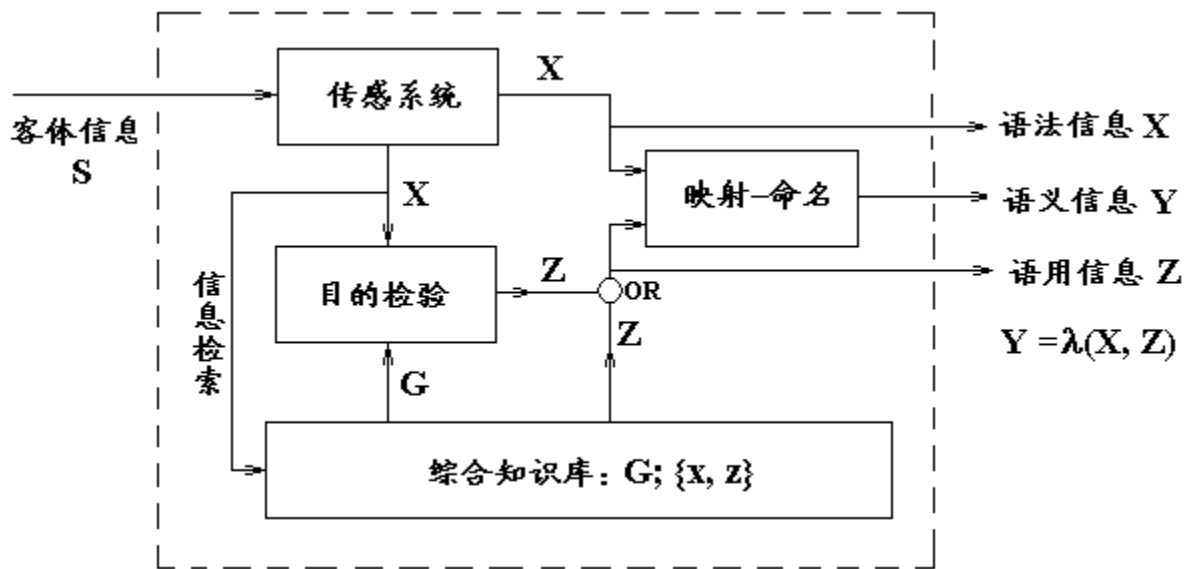


图4 “全信息” 的生成机制

图4的模型不仅阐明了全信息的生成机制，而且给出了语义信息（内容）的科学定义： $Y = \lambda(X, Z)$ ，其中算子“ λ ”代表“映射与命名”的逻辑操作。可见，人们掌握了语义信息，就同时掌握了语法信息和语用信息，就代表了全信息。所以，语义信息是用来“理解”事物的，而不是用来描述事物的统计特性的。

有了全信息的概念与生成机制，全知识与全智能的概念与生成便水到渠成。由此就可以建立“全知识”的知识库（标准的名称是“机制主义的记忆库”），它与传统知识库的根本区别就在于它的知识包含了“形式性知识、内容性知识、

价值性知识”，因此可以有力地支持理解。值得指出，这样的“机制主义记忆库”比目前流行的“知识图谱”更为深刻，因而更为优越优越。

总之，只有在“全信息”和“全知识”的基础上，才能具有“理解能力”，才能支持真正的“全智能”。

9、深挖人工智能的基本原理（学科理论 B）

信息学科范式强调“信息生态方法论”。因此，最为深刻的人工智能原理就是体现信息生态演化的“信息转换与智能创生定律”。这是一切人工智能和人类智能系统的本质和灵魂。

正如图 3 所表明的那样，“信息转换与智能创生定律”具体包含：

- (1) “客体信息→感知信息（感知）”的转换原理，
- (2) “感知信息→知识”（认知）的转换原理，
- (3) “感知信息、知识与目标→智能策略”（谋行）的转换原理，
- (4) “智能策略→智能行为”（执行）的转换原理，和
- (5) “误差信息→优化智能行为”（优化）的转换原理。

细心的读者可能已经发现，“信息转换与智能创生定律”的前四项转换原理恰好体现了中华文明的“知行学说”一知（感知、认知）行（谋行、执行）。最后这项“优化原理”，其实是在“利用知行学说来纠正误差”。所以，“信息转换与智能创生定律”完全与中华文明的知行学说交相辉映。

不难证明，“信息转换与智能创生定律 / 知行学说”既是普适性的人工智能生成机理，同时也是人类智能的生成机理。当然，只有应用信息学科的范式才有可能发现这一极其重要的结论。相反，如果恪守传统物质学科的范式，那就只能看到人工智能系统的结构与人类智能系统的结构之间的巨大差别。

值得指出，“信息转换与智能创生定律”的深远意义更在于，它是与物质科学领域的“质量转换与物质不灭定律”和能量科学领域的“能量转换与能量守恒定律”等量齐观的科学定律，它们三者一起就完善了物质、能量、信息三大资源领域的三大科学定律。

进一步的分析可以看到，“质量转换与物质不灭定律”和“能量转换与能量守恒定律”阐明了这两个领域所存在的界限，告诉人们“不能逾越这些界限”。而信息转换与智能创生定律则告诉人们“可以通过信息转换来创生人工智能系统为人类提供智能服务，把人类从体力劳动和有规可循的智力劳动中解放出来，以便更好地发挥人类的创造性能力，实现人类社会的可持续发展”。或者，更简要说，物质不灭定律和能量守恒定律告诫人们“不能逾越界限”，而智能创生定律则告诉人们“可以创造未来”。

从这个层面上是否也可以说，“信息转换与智能创生定律”对于人类的进步与发展具有更加重大的意义？

10、创建通用智能理论（学科理论体系）

综合集成第二节第 1-9 小节的各项成果，特别是其中以“信息转换与智能创生定律”为表征的机制主义研究路径，可以构建既适用于人工智能也适用于人类智能的“通用智能理论”。表 3 列出了通用智能理论的总体特征以及通用智能理论总体特征与现有人工智能理论总体特征的对照。

表 3 通用智能理论的总体特征及其与现有人工智能理论的对比

比较事项	现有人工智能理论	通用智能理论
科学观	客观物质，关注物的结构， 确定性变化	主体客体互动，关注主体的目的，

		不确定性变化
方法论	机械还原方法论（单纯形式化；分而治之）	信息生态方法论（形式内容价值三位一体化；整体性演化）
全局模型	孤立的脑模型； 智能系统行为模型	主体驾驭下的主体与客体相互作用信息过程模型
研究路径	互不相容的研究路径：结构模拟、功能模拟、行为模拟	统一的研究路径：机制模拟（普适性的智能生成机制）
学术结构	计算机学科的一个分支	神经科学、认知科学、信息科学、社会科学和数理科学等的交叉与综合
数理基础	概率论，数理形式逻辑	柔性逻辑理论，因素空间理论
基础概念	形式化数据、形式化知识、形式化智能	形式内容价值三位一体的信息、形式内容价值三位一体的知识和智能
基本原理	未有总结	信息转换与智能创生定律
最终结果	三个局部理论	通用智能理论

由此可以看出范式革命对于智能理论研究的巨大作用和深刻意义。通用智能理论的功能模型如图 5 所示。

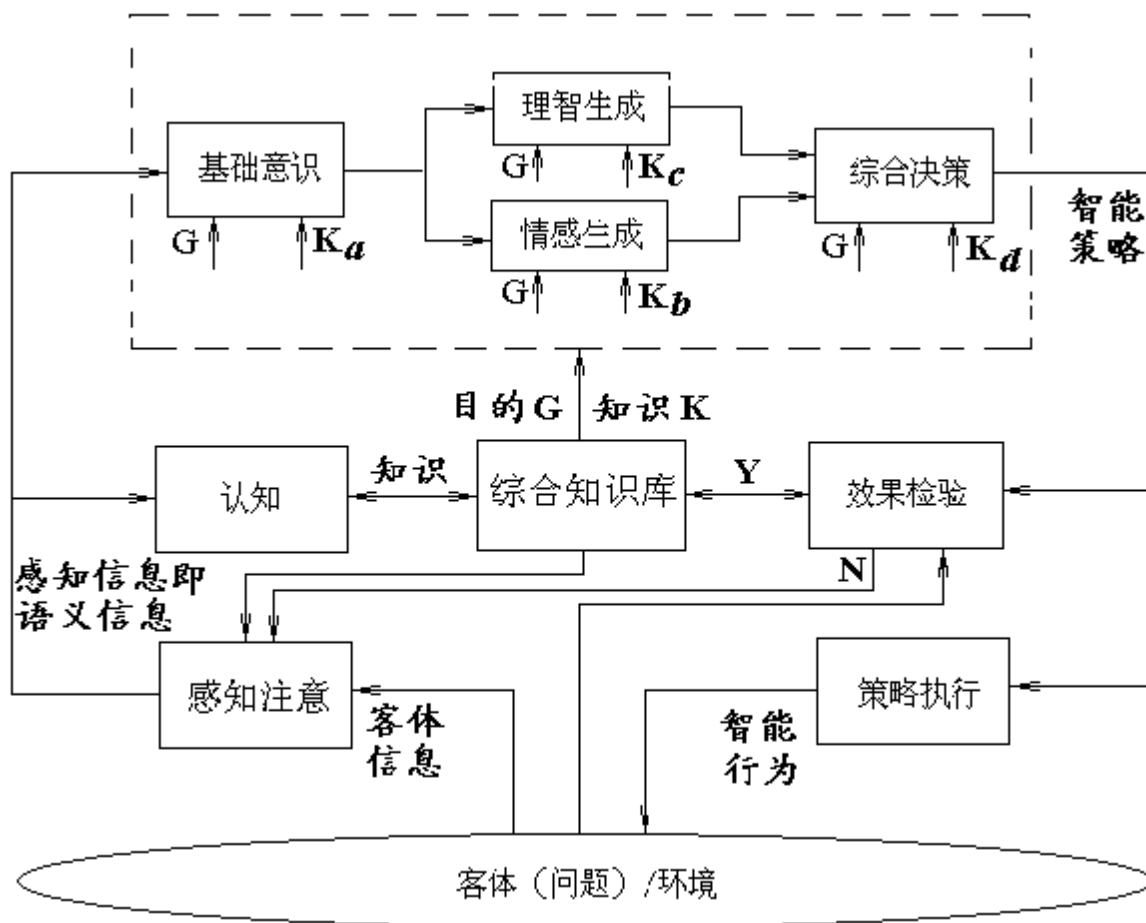


图5 通用智能理论模型

不难看出，图5的通用智能理论模型不但全面体现了信息学科范式（科学观和方法论）的理念，体现了人类学、神经科学、认知科学、信息科学（含智能科学）、系统科学、柔性逻辑理论、因素空间数学理论等科学精神，而且，也展现了人类智能的精髓。特别体现了“物质变精神和精神变物质”的辩证法，以及“人类不断地认识世界和改造世界，并在改造客观世界的过程中改造自己”的人与环境相互作用的理论。

11-632，融合人工智能不同学派发展路径探索新一代人工智能（通用人工智能）在跟帖150中，我们看到40多年来符号主义者与连接主义者争论一直不断。

2018年12月符号主义者、纽约大学教授 Gary Marcus 又与连接主义者（或深度学

习者)、蒙特利尔大学教授 Yoshua 及脸谱首席科学家 YannLecun 又燃起了这场古老的争论。

在跟帖 457 中，华盛顿大学计算科学家叶锦才 (Yejinchoi) 提出自动知识图谱构建模型 COMET，企图将符号推理与深度学习融合在一起，攻克常识推理难关，以完善大规模语义网络。

在跟帖 546 中，钟义信教授提出新一代人工智能（通用人工智能）的基础理论，在跟帖 548 中，张钹院士提出迈向新一代人工智能（第三代人工智能）的设想。钟、张均认为符号主义、连接主义等不同学派的发展路径均存在片面性，难以据此单独发展为新一代人工智能，只有将当前人工智能不同学派的发展路径融合改造为统一的发展范式并据此发展，探索新一代人工智能才有希望。

但钟、张团队对如何融合，均未提出具体的解决方案，似乎只有华盛顿大学叶团队为实行常识推理攻关初步提出了具体的融合方案（将符号主义与深度学习融合），这还是一个初步的融合方案（因为深度学习的统计特征还难以转化为推理或理解）。

12-636，评钟文

看到跟帖 546、547、559、559 附件 569 发表钟义信教授关于构建下一代强人工智能（钟称：通用人工智能）的文章，概括来说钟文认为：

1)，在人工智能发展历史上有三大学派，即：①基于数据统计连接系统的结构主义，②基于逻辑推理符号系统的功能主义，③基于模拟智能生物行为感知动作系统的行为主义。在人工智能发展中，各自取得不少精彩成果，但三派争霸无法统一，其发展路径均存在片面性，无法通往下一代强人工智能领域。

2), 通往下一代强人工智能(即欲进入通用人工智能领域), 需从突破性变革的学科范式出发(人工智能范式的颠覆性变革催生了通用人工智能的基础理论)。

3), 钟教授提出了构建通用人工智能学科范式的基础理论, 由变革并融合三大学派的不同发展路径演变而成。

现在的问题是: 如何对三大学派的发展路径进行颠覆性变革和相互融合, 以实现统一的学科范式, 支持构建下一代强人工智能(通用人工智能)? 似乎尚缺具体的解决方案!

13-637, 评张文

看到跟帖 548 发表张钹院士等关于构建下一代强人工智能(张称: 第三代人工智能)的文章, 概括来说张文认为:

第三代人工智能的理念及其发展途径是融合第一代人工智能(符号主义)的知识驱动和第二代人工智能(连接主义)的数据驱动的人工智能范式。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人工智能理论与方法, 构建更强大的人工智能, 即发展安全、可信、可靠和可扩展的人工智能技术。

张院士提出构建第三代人工智能的范式, 如何融合(不同发展路径的融合)? 如何结合(两驱动)? 如何利用(四要素)? 如何构建大规模语义网络(知识图谱), 赋予其足够能力可支持认知智能得以实现? 似乎尚缺乏具体的解决方案! 不能一句话就说过去!

14-653, 陆总: 这是我们研究工作的一个简介, 请批评指正!

“机制主义人工智能基础理论”简介

《机制主义人工智能理论》及其逻辑与数学基础（合称为机制主义人工智能基础理论）是实现了“人工智能基础理论重大突破”（国务院规划的第二步战略目标）的重要理论成果。

突破的关键是抓住了学科研究的龙头——范式，发现了人工智能范式张冠李戴，实施了范式革命，突破了原有人工智能面临的层层障碍，揭示了人工智能的新科学观、新方法论、新模型、新路径、新逻辑、新数学、新概念、新原理，创建了“机制主义人工智能基础理论”。

具体而言：

- 颠覆了物质学科范式：机械唯物科学观和机械还原方法论

因而

- 突破了现行人工智能的全局研究模型：孤立的脑模型
- 突破了现行人工智能的研究路径：结构主义、功能主义、行为主义三个分道扬镳的研究路线
- 突破了现行人工智能的基础概念：形式数据、形式知识、形式智能
- 突破了现行人工智能的逻辑基础：刚性的逻辑
- 突破了现行人工智能的数学基础：分立的理论

进而

- 确立了信息学科范式：主客互动科学观和信息生态方法论
- 构筑了全新的研究模型：主体客体相互作用的演进模型
- 开创了全新的研究路径：基于普适性智能生成机制的机制主义路线
- 创建了全新的人工智能逻辑理论：泛逻辑理论
- 创建了全新的人工智能数学理论：因素空间理论
- 重建了人工智能的基本概念：全信息、全知识、全智能
- 发现了全新的人工智能基本原理：信息转换与智能创生定律

综之

- 首创了《机制主义人工智能基础理论》

简要解释如下。

一切科学研究活动都受到人们的科学观和方法论的引领和支配。科学观阐明了研究对象的本质“**是什么**”，方法论阐明了学科的研究应当“**怎么做**”；科学观和方法论一起，就阐明了学科研究应当遵循的规范方式，简称“**范式**”。因此，**范式是整个学科研究的龙头**，它的正确与否决定了整个学科研究的成败。

显然，作为开放复杂信息系统的人工智能，应当遵循信息学科的范式。但是，我们通过全面调研却惊人地发现：它所遵循的竟然是物质学科的范式，在范式上犯了张冠李戴的大忌！这是现有人工智能一切痼疾

顽症的总根源。

比如，在物质学科范式的“分而治之”方法论支配下，人工智能的研究被分解为人工神经网络、专家系统、感知动作系统三个分道扬镳的学派，使人工智能的研究至今仍然处于没有统一理论的初级状态；又如，在物质学科范式的“单纯形式化”方法论支配下，智能的内核（内容要素和价值要素）被完全丢弃，导致现有人工智能系统的结果不可解释，智能水平低下。

为什么人工智能的研究会陷入范式“张冠李戴”的境地？

这是社会法则所决定的。具体来说，由于历史上的科学研究都属于物质学科体系，都遵循物质学科的模式，不存在范式变革的问题。问题发生在 20 世纪中叶以来，信息学科迅猛兴起，形成了信息学科研究的社会存在。可是信息学科的模式却至今未能形成。这是因为模式属于社会意识范畴，而“**社会意识滞后于社会存在**”是社会发展的根本法则。在这种情形下，信息学科（含人工智能）的研究者们便沿用了当时业已存在的物质学科模式，于是造成了信息学科（含人工智能）模式的张冠李戴，而且一直延续至今，**是为无可避免的结果！**

面对人工智能模式的张冠李戴，唯一有效的解决办法就是实施“正冠”。

我们发现，**中华文明的整体观（科学观）和辩证论（方法论）完全符合信息学科的性质和需要，是引领信息学科和其他复杂科学开拓创新的最佳模式。**据此，我们以中华文明支撑的信息学科模式取代了人工智能的传统学科模式，并在信息学科模式引领下实现了上述各项突破与创新。

成果“机制主义人工智能基础理论”的意义：

- 开创了信息学科和复杂学科研究的新模式，具有划时代意义
- 从理论上消除了原有人工智能的痼疾顽症，具有重要的理论意义
- 智能生成机制是人工智能应用系统的通用孵化平台，具有重大工程意义
- 信息转换与智能创生定律与“物质不灭和能量守恒”具有相同科学意义
- 见证了中华文明对 21 世纪信息学科和复杂学科开拓创新的引领地位。

15-695，关于通用人工智能的讨论

在人工智能国内外跟帖讨论中，钟义信、张钹等均谈到人工智能发展前沿为通用人工智能，但至今尚未突破具体解决方案。李德毅院士提出通用人工智能“十问”，他提问如何理解通用人工智能？王迪兴曾答复“十问”，他在谈到通用人工智能时，也仅停留在议论的层面上，并未涉及具体解决方案！但这里的问答，也可供探索者参考！

问：如何理解通用人工智能？我们应该不应该把通用智能理解为“全知全能”或

者单项超强智能？尽管今天的计算机已经可以解决很多复杂的、专门的智力问题（如围棋智能），我们仍常常觉得它们缺乏人类思维的某些本质特征。这里的差别主要不是在算法、算力、数据量方面，不是在速度和容量方面，而是在智能的一般性、通用性、普遍性、灵活性、缺省性、容错性、可习得性、不确定性、适应性、常识性、开放性、创造性、自主性等方面。遗憾的是发展六十多年的人工智能没有能够更靠近人的原始智能。

答：通用人工智能需要全新的理论基础，从哲学层面讲，需要本体论、方法论、认识论的统一，从自然科学的角度讲，需要系统论、信息论、控制论的统一。从计算理论角度讲，巴贝奇与图灵计算模式均不适用，需要奠定新的计算理论基础，统一时间与空间计算理论，且与脑科学统一。

根本性障碍在于目前数学描述方法对于结构化的多因果互为因果关系描述无能为力，需要创建一门新的数学——结构数学，其理论基础是准全息系统论，其定量描述是全息结构计算模型。

基于它可设计 2-16-256 进制的类脑计算机（我们已设计实现 16 进制计算机），不仅计算速度能达到指数级计算能力增长，关键是性能更符合人脑功能机制。

智能的最本质特性，是子系统或功能模块基于多因多果、互为因果关系建立实时、共时交互作用的关系模式。智能是其若干功能模块互补交互作用的结果。相对于传统计算机，其功能特征是读、写、算同步、储算一体化、址与数据统一。

外设之间能够多对多并行实时双向同步交互作用，体现内在的统一性，即运算、交换、控制、双向立交总线功能统一一体化。大脑和遍布人体的神经网络本来是一体的东西，但在现实世界，计算机和网络却人为的被分成两个相对独立的东西，丧失了内在统一性。智能最本质的东西，是记忆单元与各功能单元具有内在自组

织作用关系。

不仅存储单元之间具有确定性的互为因果作用关系，且能形成紧密交互作用的更大规模的超循环！所谓超循环是每一个作用因子的输出是另一个作用因子的输入，这是智能、生命产生的基本前提。

模拟人脑智能，即模拟类脑神经网络的结构及功能，多因多果、互为因果作用关系模式，是生命及智能的基础。

至于通用人工智能的目的不应该是构造一个与人脑完全一样的东西，完全相同的克隆人就行了！也并非全知全能，人也非全知全能！说到底我们不是构造具有独立意志的智能异己！只有不同才会超越！

16-755，论人工智能基础理论的源头创新

范式革命：人工智能基础理论重大突破的必然选择

钟义信

摘要 制约人工智能发展的最大障碍，是它的最高指导思想（即科学观和方法论，统称为范式）犯了张冠李戴的大忌：用了物质学科的范式来指导人工智能的研究。本文用科学发展普遍规律与人工智能发展历史相结合的方法论证了：人工智能范式张冠李戴的不可避免性，人工智能的范式革命是人工智能实现重大突破的唯一正确举措，后者将推动物质学科主导的科学时代迈向信息学科主导的科学新时代，并从源头上创建科学新时代的人工智能理论——通用人工智能理论。

1，引言

与解放人类体质能力的材料科学和解放人类体力能力的能量科学（统称为物质学科）不同，人工智能是以解放人类智力能力为目标的一门学科。因此，人工

智能学科一问世就受到人类社会的高度关注。至今，发展人工智能已经成为世界各国特别是各发达国家的重大战略。

2018年10月31日，习近平总书记在中共中央政治局集体学习会议上指出：人工智能是引领这一轮科技革命和产业变革的战略性技术，…要加强基础理论研究，支持科学家勇闯人工智能科技前沿的“无人区”，努力在人工智能发展方向和理论、方法、工具、系统等方面取得变革性颠覆性突破，确保我国在人工智能这个重要领域的理论研究走在前面、关键核心技术占领制高点。

2017年7月，我国国务院颁发了《新一代人工智能发展规划》。

根据这个《发展规划》，今年，我国的人工智能研究进入了第二个战略阶段，目标是实现人工智能基础理论的重大突破。因此，实现人工智能基础理论研究的重大突破，应当成为本阶段我国人工智能研究的重中之重。

2, 实现人工智能基础理论的重大突破，需要准确理解几个重要概念

人工智能的研究新颖、复杂而艰深，呈现出诸多的谜团和不确定性。因此，需要正确理解人工智能领域的一些重要概念，以免研究活动误入歧途。

2.1 人类智慧和人类智能

首先，要分清智慧与智能的关系，不能把二者混为一谈。

什么是智慧？现代汉语词典说，是辨识判断和发明创造的能力。

笔者认为，人类智慧是指它作为万物之灵的卓越能力：人类自觉地不断地运用信息和知识去认识世界和改造世界，并在改造客观世界的过程中不断地改造自己的主观世界，从而不断地改善自己的生存与发展水平。

认识世界，是为了探索未来发展道路而去研究和提出应当解决的问题，给出

解决问题应当遵循的工作框架：(1) 定义和描述所需解决的问题，(2) 预设解决问题应达到的目标，(3) 明确解决问题所需的知识。

改造世界，是为了变革现实而去解决所提出的问题，即针对给出的工作框架，利用关于问题的信息和知识，在预设目标的引导下去谋划解决问题的策略，并把策略转化为行动，把问题的原有状态改造成为符合预设目标所要求的状态。

认识世界必须根据人类的目的，利用人类的隐性思辨能力（如直觉能力、抽象能力、理解能力和想象能力等）去探索，因此，被称为人类的隐性智慧。改造世界则主要依赖于人类的显性操作能力（包括获取信息、提取知识、生成策略和执行策略的能力等）来实施，因此，被称为人类的显性智慧。

人类隐性智慧的特点，是直接依赖于人类的目的，而且高度抽象，甚至近于神秘，因此，难以用机器来模拟。然而，人类的显性智慧却是在给定了具体问题、求解目标和知识的工作框架条件下的问题求解研究，既有明确的起点，又有明确的目标，既可望也可及，因而可以成为“机器模拟”的对象。人类的显性智慧由此就被特别地称为“人类智能”。

人造机器可以用来模拟人类智能，却难以用来模拟人类智慧。

2.2 人工智能与人类智能

人工智能，是在人造机器上所实现的人类智能。因此，人工智能的研究只涉及对“解决问题变革现实的人类显性智慧”的模拟，不涉及探索未来提出问题的隐性智慧。人类不可能让没有生命的机器取代自己去探索自己的未来，把人类未来发展的命运交给没有生命的机器。

明确了这些概念，就可以避免产生“人工智能无所不能、人工智能将全面超越人类和统治人类”这样一些误解。其实，再强大的人工智能，都只能在人类隐

性智慧给出的工作框架内去解决问题，只能执行人类的意志，为解决人类关注的问题和为改善人类生存发展的目的服务。人工智能可以是人类的聪明助手与合作伙伴，而不能成为主宰人类命运的主人。

2.3 信息与机器感知

没有信息就没有知识，没有知识就没有智能。因此，信息是智能的源头，准确理解信息概念才能保证智能理论研究不走偏路。

需要提醒的是，至今国际学术界唯一公认的信息理论是 Shannon 1948 年创立的信息论。然而，Shannon 信息论的原名是通信的数学理论 (Mathematical Theory of Communication)，适用于“只关心形式，不关心价值和内容”的通信工程，不适用于“需要全面理解形式、内容和价值”的人工智能研究的需要。

具体来说，在人工智能的场合，信息有两个互相联系又互相区别的概念。

首先是客体信息的概念，它的定义可以这样表述：**客体信息，就是客体所呈现的自身状态及其变化方式。**显然，这是本体论的信息概念。客体信息只与客体自身的状况有关，而与主体的状况无关，甚至与是否有主体存在无关。

如果客体信息被主体（人工智能机器可被理解为人类主体的代理，也常常简称为主体。下同）所感知，就产生了主体的感知信息的概念。由于引入了主体，感知信息比较复杂，包含了三个分量：（1）主体从客体信息中感受到的客体的状态及其变化方式的外在形式，称为形式信息（文献中称为语法信息）；（2）从客体信息中知觉到的客体对主体目标而言所具有的效用价值，称为价值信息（文献中称为语用信息）；（3）由形式信息和价值信息两者定义而成的含义内容，称为内容信息（文献中称为语义信息）。

感知信息是形式信息、价值信息和内容信息的三位一体。正由于感知信息包

含了形式、价值、内容全部分量，因此也被称为“全信息”。显然，感知信息属于认识论的范畴。关于感知信息的更深入讨论，见本文第 5.2 节。

有了这些信息概念，就可以精准定义感知的概念：**感知，就是把客体信息转换为感知信息的过程。**这是感知的学术本质。

学术界存在一个相当普遍的误解：以为感知就是传感。根据上述的感知定义就明白：感知功能必须能够了解客体的形式、价值和内容。传感功能只能感受到客体的形式，不了解客体的价值和内容，不可能承担感知的功能。

2.4 知识与机器认知

由上述信息的定义可以知道，信息是具体的个性化的概念。如钢笔的信息、毛笔的信息、铅笔的信息、圆珠笔的信息、排笔的信息、画笔的信息等等，都是具体事物的信息，个别事物的信息。

与此相对，“笔”的信息则是抽象的共性的概念，它是从具体的钢笔的信息、毛笔的信息、铅笔的信息、圆珠笔的信息、排笔的信息、画笔信息这个集合抽象出来的共性的概念。这个“抽象的笔的信息”就称为“关于笔的知识”。**信息是具体的个别事物呈现的现象，知识则是一类事物的抽象的共性本质。**

需要注意，既然知识是一类事物的抽象的共性的本质，因此，知识应当属于认识论的范畴（而不是本体论的范畴）。换言之，认识论范畴的知识只能从认识论范畴的感知信息抽象提炼出来的概念。因此，与感知信息概念相对应，知识也应当是形式性知识、价值性知识、内容性知识的三位一体。

由此可以给出认知的定义：**认知，就是把感知信息转换为知识的过程。**这是认知的基本功能，也是认知的学术本质。

不难理解，作为把感知信息（感性认识）转换为知识（理性认识）的认知，

其实就是学习的过程，包括初级的灌输式（机械式）学习，中级的从众式（统计式）学习，高级的自主式（理解式）学习。

2.5, 策略与谋行

策略是解决问题的方法与步骤，是人工智能的核心概念。人类智能和人工智能的智能水平究竟如何？关键就看解决问题的策略的水平如何。由于策略与智能具有这样密切的关系，人们就把策略看作是智能的化身，称之为智能策略。

如果给定了求解的问题，那么，问题的客体信息也就随之给定了。于是，通过感知就可以获得感知信息，通过认知就可以获得相应的知识。在此基础上，针对感知信息（它是主体对问题的感性认识），在知识的约束下（什么步骤能被采取，什么步骤不能被采取），在预设的求解目标引导下，就可以谋划解决问题的智能策略。这个过程，称为谋划求解问题的行动策略的过程，简称“谋行”。

所以，可以给出定义：**谋行，就是在目标导引下，在知识支持与约束下，把感知信息转换为智能策略的过程。**

以往的文献没有“谋行”的概念，只有“决策”的概念。不过，人们往往把决策简单地看作是“在几个备选行动方案中选取一个”的过程，忽视了这些备选的行动方案是怎样被“谋划”出来的。为了避免这种简单化片面化的误解，这里提出了“谋行”的概念，它涵盖了谋划与选择的全部过程。

2.6 主客互动与信息生态

人工智能是在人造机器上实现的人类智能，是对人类解决问题变革现实的显性智慧的模拟。因此，人工智能所扮演的角色就是“人类的代理”。而人类正是通过与环境中的客体的相互作用产生出认识世界和改造世界的智能策略和智能行为，赢得人类生存发展水平的不断改善。因此，**主客互动（主体与客体的相互作**

用)是人工智能研究的根本前提。

在主客互动过程中,主体接受的是客体信息的刺激,输出的是主体产生的智能策略和智能行为。显而易见,智能必定是由信息经过复杂的转换所生成的。具体来说,在主客互动的过程中,客体信息会被主体转换为感知信息,进而转换为知识、智能策略和智能行为。这就是所说的信息转换,而且是**信息的生态转换**,这个信息转换过程就形成了**信息的生态链**:从初级的信息(现象)转换为高级的知识(本质)以至更高级的智能(策略)。

所谓信息生态,是指在保持信息内涵的整体性、时空的连续性和系统性前提下和全局优化条件下的信息转换。这样的信息转换过程,也就是信息的生态演化过程。

2.7 学科范式与科学时代

在一般的意义下,范式是指人们所遵循的“世界观和行为方式”^[7]。人类一切有意识的行为都受着某种范式的支配和规范,不是受这种范式的支配和规范就是受别的范式的支配和规范,不存在范式真空的情形。范式是引领和规范人们行为的最高支配力量。

在科学研究领域,范式是指研究活动所遵循的科学观和方法论,人类的一切研究活动都受着某种范式的支配和规范,不是受这种范式的支配和规范就是受那种范式的支配和规范,不存在范式真空的情形。范式是引领和规范研究活动的最高支配力量。

正是在这个意义上,科恩把范式看作是“科学革命”的指标。

迄今,人类的科学研究活动产生了两种不同的学科:一个是发端于农耕时代的物质学科,一个是兴起于信息时代初期的信息学科。物质学科早已渗透在整个

科学领域，信息学科也在快速地向所有科学领域（包括物质学科领域）渗透。

物质学科的研究对象是物质对象，信息学科的研究对象是信息对象。物质对象和信息对象的性质对立而统一，相反而相成，物质学科和信息学科所遵循的范式也是对立而统一，相反而相成。由此，就形成了对立而统一、相反而相成的两个科学时代：物质学科主导的科学时代，信息学科主导的科学新时代。促成物质学科主导的科学时代向信息学科主导的科学新时代转换的力量，就是范式的革命。

以上这些，就是与人工智能理论直接相关的最为基本的重要概念。

3. 实现人工智能基础理论的重大突破，需要准确理解人工智能的历史

半个多世纪的人工智能发展史，极其精彩地演绎了人工智能研究领域各种不同观念和方法之间的矛盾冲突，从中折射出了非常宝贵而且发人深省的启示。因而，值得细加考察。

3.1 人工智能的历史：观念方法的矛盾演绎

用机器来协助人类进行劳动，是人类一直不懈追求的美好理想。不过，真正付诸实践的研究始自 20 世纪的 40 年代。那时人们相信，人类的高级认知功能定位于大脑新皮层的神经网络，认为只要把大脑神经网络的结构在机器上实现出来就可以模拟出人类的智能。这是结构模拟方法捷足先登的缘故。

具体来说，1943 年 McCulloch 和 Pitts 发表了神经元的数理逻辑模型，1949 年 Hebb 提出了神经元的学习规则。利用 M-P 模型和 Hebb 规则就可以构造人工神经网络，开展人工神经网络及其应用研究。

人类大脑神经网络是由近千亿神经元互相复杂连接而成的大规模非线性系统，而当时的科技和工业能力都只能实现小规模简单神经网络系统，因此结构模拟的人工智能研究虽然似乎前景光明，但道路却很曲折，进展缓慢。

面对这种情况，一批思想活跃的学者便撇开结构模拟的途径，选择了功能模拟的途径。这就是 1956 年夏天 McCarthy 等人在 Dartmouth 发起的利用计算机作为硬件平台、通过编制“聪明软件”来模拟人类逻辑思维功能的研究途径，并且创造了 Artificial Intelligence 这一术语来表征这个新的研究领域。由于那时的计算机已有很强的功能，Newell 和 Simon 等人提出了“计算机的功能与人脑功能等效”的物理符号假设，并把功能模拟的研究产物称为物理符号系统，认为这是可以像人脑一样求解通用问题的人工智能系统。果然，功能模拟方法一经问世，便在数学定理机器证明和模式识别等方面取得了令人鼓舞的成绩。

不过，当人们真的利用这一方法来解决通用问题的时候，就发现这是很不现实的想法，因为求解通用问题需要无限的知识。于是，人们不得不把面向通用问题求解的物理符号系统改成为面向专门问题求解的专家系统。

然而，解决专门问题也需要专门的知识，而知识的定位、获取、表示、推理都存在难以逾越的“知识瓶颈”困难。在这种情况下，Brooks 等人提出了行为模拟的思路。他们宣称，行为模拟方法不需要知识，只需要感知到环境对智能系统的刺激和智能系统对此所产生的动作响应（于是被称为“感知动作系统”），因此可以回避结构模拟方法的复杂性问题 and 功能模拟方法的知识瓶颈困难。行为模拟的典型研究成果，是 Brooks 领导的 MIT 人工智能实验室在 1990 年所演示的六脚虫爬行机器人以及随后所研制的各种智能机器人。遗憾的是，对于行为模拟的感知动作系统来说，它自身的最大问题是它仅能模拟智能系统的外在行为表现，这只是一类浅层的智能。

上述历史表明，人工智能存在结构模拟、功能模拟、行为模拟三种不同的研究方法。它们是在人工智能发展的不同阶段、面临不同的问题、分别由不同的人

群发展出来的各不相同的研究方法。它们的学术信仰（以脑科学为模拟原型的结构主义、以认知科学为模拟思路的功能主义、以控制论为模拟原理的行为主义）各不相同，所针对的问题各不相同，所采取的策略各不相同，所显现的能力也各不相同。虽然都是人工智能的研究方法，但却无法实现殊途同归。

实际上，三大学派非但未能殊途同归，无法形成合力，反而变为互不相容，互相排斥，势难同立，以至逞强凌弱，成为同行冤家。以下是这些学派称雄争霸的若干典型事例。

1969年，功能主义方法的学术带头人 Minsky 和他的同事 Papert 出版学术专著《感知机》严厉抨击神经网络方法“没有科学价值 (without scientific value)”，认为它的成果建在流沙之上 (built on quick sand)。这种猛烈的批评，造成世界范围人工神经网络领域的研究人员大量转行，使人工神经网络的研究陷入为期十多年的“黑暗年代”。

1987年，借着专家系统的研究遭遇“知识瓶颈”的困难而神经网络研究进入复兴阶段的机会，结构主义人工神经网络研究者们在美国加州圣地亚哥举行的 IEEE 第一届神经网络国际会议期间，与会者中爆发出了“人工智能死了，神经网络万岁 (AI is dead. Long live neural network)”的强烈呐喊。表现了结构主义神经网络研究学派对功能主义专家系统研究学派的强烈反击。

1990年前后，行为主义感知动作系统研究学派的带头人 Brooks 连续发表论文抨击正在遭受“知识瓶颈”困难的功能主义研究方法，宣称行为主义感知动作系统的研究方法可以不需要知识 (Intelligence without Knowledge)，因而也不需要知识的表示 (Intelligence without Representation)，以此来否定功能模拟方法。

正是由于不同学派的学术主张各不妥协，使人工智能的研究始终处于某个学派一派独大的格局，从来没有出现过相互合作的局面。

●1943 年至 1956 年间，由于其它两种研究路径尚未问世，人工智能当然只存在结构主义（初期的简单人工神经网络）一种研究路径。

●1956 年至 2016 年这 60 年间，基本上是功能主义雄霸天下。虽然在 20 世纪 70 年代初期由于在机器翻译领域的失败和随后由于知识瓶颈的困扰曾经两度遭遇过发展的危机，但是，功能主义方法仍然维持着它的统治地位，因为被它打入“黑暗年代”的结构主义方法还没有完全缓过气来。

●从 2016 年至今，由于初期的简单神经网络逐渐发展成为复杂的深层结构模型，训练和学习的算法也得到显著改进，在人工智能的一些竞赛评测项目中表现出大大超越功能主义方法、甚至超越人类能力的优异性能，于是，基于结构主义方法的深层神经网络研究变成这一时期人工智能的主导学派，基于功能主义方法的专家系统研究则几乎难觅踪影了。

3.2 人工智能历史的启示：根本问题是范式的张冠李戴且不可避免

经过 70 多年的发展，人工智能研究的结构主义（人工神经网络研究）、功能主义（专家系统）研究、行为主义（感知动作系统研究）三大学派各自都取得了不少引以自豪的成果，但也都面临严峻的挑战。

不过，三大学派取得的成绩已为世人熟知，为了节约篇幅，这里暂不展开。从推动人工智能基础理论源头创新的需要着眼，我们应当重点关注现行人工智能究竟面临的根本性问题。只有发现了这些问题，才能有针对性地解决这些问题，推动人工智能及其基础理论研究健康发展。

历史考察表明，**第一个严重的问题是：三大学派分道扬镳各自为战，无法形**

成统一的人工智能理论，因而也就无法建立通用的人工智能系统。

由于没有统一的理论，现行人工智能系统的研究都严重依赖于应用场景。人们在研究实用的人工智能系统的时候，首先就需要仔细选择应用的场景，然后针对这个选定的场景设计解决问题的人工智能系统。一旦应用场景改变了，系统就要从头重新设计：一个场景一个系统，不同场景就要设计不同的系统。

曾经有人宣称，人工智能领域或许根本就不存在统一的理论，因此也不存在通用的系统；只要一个一个应用场景的人工智能系统都研制出来了，人工智能的问题就全解决了。不难看出，这种认识根本不符合辩证法和“可持续发展”的理念，也不符合系统论的基本原理。系统论原理指出：系统（整体）远远大于它的部分和。这就是说，系统的整体解决远远胜于它的所有部分解决之和，其间存在整体优化和局部优化之间的质的差别！因此，“现行人工智能无法建立统一理论”乃是一个极为严重的问题，表明人工智能的研究根本没有到位。

追根寻源，为什么现行人工智能的研究无法建立统一的理论呢？

人工智能发展的历史给出了明确的答案：作为开放复杂信息系统的人工智能研究，遵循了“分而治之”的方法论，把人工智能的整体研究肢解成了结构主义、功能主义、行为主义三种方法分别进行研究。“分而治之”方法论在物质学科领域非常有效，但用在人工智能领域却割断了三种方法之间的复杂而隐秘的信息联系，后者正是整体信息系统的生命线和灵魂；既然割断了三种方法之间的生命线和灵魂，当然就无法把它们恢复成为原来的复杂信息系统整体。

这就是现行人工智能的研究无法建立人工智能统一理论的根本原因。不是技术细节的原因，而是方法论的原因。而没有人工智能的统一理论，就表明现行人工智能的研究根本没有走上正确的轨道。

第二个严重的问题是：现行人工智能的智能不是在理解问题基础上的真正智能，而是快速计算能力的奇葩表现，真正的智能水平却很低下。

按照本文前面的人工智能定义，它应当是在人造机器上所实现的解决问题的显性智慧。这就表明，人工智能系统所表现的“智能”应当是在“理解问题基础上解决问题的能力”。但是，现行人工智能系统表现的“智能”其实并没有理解能力。为什么这样说呢？

这是因为：现行人工智能的研究都是纯粹形式化的研究，阉割了问题的内容与价值因素，因此不可能理解问题，当然也就不可能做出有智能水平的决策。比如，当你看见一只老虎，你看到了它的形态（形式信息），如果你也懂得它会伤人（价值信息），知道它是食人猛兽（内容信息），那么，你就理解了老虎，就可以做出正确的、有智能水平的决策；躲，逃，或在不得已的情况下射杀。但是如果你不知道老虎会伤人（不具有价值信息），也不知道它是食人猛兽（不具有内容信息），你就没有理解老虎这种凶猛野兽。在这种情况下，你怎么知道应当做出怎样的决策？说不定你就会“与虎共舞”，结果为虎所食。

单纯形式化是物质学科的方法论，在只关心物质结构与功能的物质学科范畴内可以畅行无阻，但是却不可能有效支持人工智能的研究。这是因为，人工智能的研究不能仅仅根据问题的形式就做决策，而必须对问题的形式、内容、价值有全面的理解才能做出合理的决策。

总之，现行人工智能存在的主要问题是（1）整体被肢解，无法建立统一理论，表明人工智能的研究没有走上正确的轨道，（2）价值和内容被阉割，导致智能水平低下，同样说明人工智能的研究没有走上正确的轨道。

总之，不能建立统一的人工智能理论，根本原因是因为人工智能的研究遵循

了物质学科“分而治之”的方法论。而智能水平低下，根本原因则是因为人工智能的研究遵循了物质学科“单纯形式化”的方法论。

正如范式定义所表明的，方法论是为科学观服务的。有什么科学观，就要求采用什么样的方法论。科学观和方法论两者所构成的统一整体，就形成了主宰科学研究活动的最高引领和规范力量——范式。

这就很明白了：人工智能的研究之所以一直都在遵循物质学科“分而治之”和“单纯形式化”的方法，根本的原因有两个方面：一方面，人工智能的研究活动虽然已经广泛展开，但是，长期以来很少关注人工智能的最高指导思想——科学观和方法论（范式）。以至半个多世纪过去了，仍然没有在学科范式上形成学术共同体的明确共识。因此，根本没有信息学科的范式可以遵循。另一方面，数百年来物质学科范式的强大存在和广泛影响，使人们误以为信息学科的研究也同样可以遵循物质学科的范式。这样两个方面原因的叠加，就造成了人工智能范式的张冠李戴。

这决不是什么人的恶意中伤或造谣惑众，而是半个多世纪以来人工智能研究领域活生生的事实。下面表 1 所列的内容，就是物质学科范式、人工智能实际所遵循的范式、信息学科范式三者的简明对照。

需要说明的是，表 1 第三行所列的“信息学科范式”是本文作者团队历经半个多世纪所研究提炼出来的结果。虽然在整个人工智能学术共同体内，很少有人关注和研究过信息学科领域的范式问题，但是由于独特的学经历，笔者从 20 世纪 60 年代起就特别关注了信息学科的科学观与方法论的探索。

表 1 人工智能范式张冠李戴的具体事实

事项	科学观	方法论
物质学科范式	机械唯物论的物质观 研究对象：物质客体，排除主观因素 对象性质：确定性演化，可分性 研究目的：了解对象的结构与功能	机械还原论的方法论 描述与分析方法：纯粹形式化 决策的判断方法：形式的匹配 宏观处置的方法：分而治之
人工智能范式	准“物质观” 研究对象：脑物质，排除主观因素 对象性质：存在不确定性，可分性 研究目的：了解对象的结构与功能	真“还原论” 描述与分析方法：纯粹形式化 决策的判断方法：形式的匹配 宏观处置的方法：分而治之
信息学科范式	整体观：主客互动的信息观 研究对象：主客互动的信息过程 对象性质：不确定性演化，整体性 研究目的：实现主体客体合作双赢	辩证论：信息生态方法论 描述与分析方法：形式内容价值一体化 决策的判断方法：内容理解 宏观处置的方法：生态演化

对比表 1 三个项目的内容可以明显看出，人工智能实际遵循的范式（表中第二行），基本上是物质学科的范式（表中第一行），而与信息学科范式（表中第三行）几乎无关。

那么，为什么会有这样的结果呢？答案很明确：历史注定，无可避免。

社会的发展存在一个铁定的法则：**社会意识滞后于社会存在。**

在科学研究领域，学科的研究活动就是一种社会存在，作为学科范式的科学观和方法论就是学科的社会意识。学科的研究活动这种社会存在迟早总要产生相应的学科范式作为自己的社会意识，而且，学科的范式一旦形成，它就要反过来影响学科的研究活动。只是，学科范式的形成必然要滞后于学科的研究活动。

学科的范式究竟需要滞后多长的时间？这没有确定的答案。一般而言，学科越是复杂，学科范式的形成（特别是在学术共同体内形成共识）所需要的时间就必然越长。比如，直到如今，信息学科的范式都没有在学术共同体内形成共识。

4, 实现人工智能基础理论的重大突破，需要遵循学科发展的普遍规律

作为一个新学科，人工智能当然具有自己的个性，但是，同样重要的（如果不是更为重要的话）是必须服从共性的规律。以下的表 2 用表格的形式给出了学科发展普遍规律的简明解释。

表 2 学科发展的普遍规律

阶段进程	进程名称	进程要素	要素解释
初级阶段： 自下而上的探索	摸索（准备）	多方试探	通过长期自下而上的多方试探摸索，总结失败教训和成功经验，提炼和确立学科的研究范式（即学科应当遵循的科学观和方法论）
高级阶段： 自上而下的建构	范式 （宏观定义）	科学观	宏观上明确学科的本质“是什么”
		方法论	宏观上明确学科的研究“怎么做”
	框架 （落实定位）	学科模型	基于“学科范式”的学科全局蓝图（是什么）
		研究路径	基于“学科范式”的整体研究方法（怎么做）
	规格 （精准定格）	学术结构	基于“学科范式”的学科内涵规格（是什么）
		学术基础	基于“学科范式”的学术基础规格（怎么做）
	理论 （完整定论）	基本概念	基于“学科范式”的学科基本知识点（是什么）
		基本原理	基于“学科范式”的概念间相互联系（怎么做）

表 2 显示，学科的发展需要经历两个相互联系而又相互不同的阶段：首先是自下而上的探索阶段（称为初级阶段），然后才能进入自上而下的建构阶段（称为高级阶段）。

初级阶段的主要任务，是要通过各个相关学科领域的研究人员从各种不同的学术角度展开全面而漫长的摸索、试探、争论、交流、总结，才能逐渐提炼出关于学科范式（学科本质是什么？应当怎样研究？）的共识。这是一个极为艰难的探索过程。在这个阶段，往往会出现各种各样“盲人摸象”的情景。

高级阶段的主要任务，是要在初级阶段摸索得到的学科范式引领下，通过形成学科框架（学科的研究模型和研究路径）和学科规格（学术结构的规格和学术基础的规格）一步一步地把范式贯彻落实到学科的理论建构之中（包括构建学科的基本概念集合和基本原理集合）。

对照表 2 所描述的学科发展普遍规律可以看出，**当今，人工智能学科的研究确实仍然处在自下而上探索范式的初级阶段。**这是因为，今天的人工智能研究仍然处于结构主义、功能主义、行为主义三大学派互不认可（类似于盲人摸象）的状态，而且仍然处于“范式张冠李戴”的状态。为了实现人工智能基础理论的重大突破，首先就必须尽快结束这种范式张冠李戴的状态。

表 2 的规律也清楚地表明，学科的范式是引领学科发展全过程的力量：学科探索阶段的任务，是为了总结出学科的范式；学科建构阶段的任务，是为了贯彻和落实所总结出来的学科范式。所以，学科范式的引领和规范作用贯彻在学科发展的始终。有了正确的学科范式，学科的发展就有了明确的方向和灵魂，离开了正确的学科范式，学科的发展就会迷失方向，陷入泥潭。

5， 范式革命， 实现人工智能基础理论重大突破的方法与步骤

既然学科范式是学科发展的最高支配力量，而人工智能的范式又不可避免地陷入了张冠李戴的窘境，成了现行人工智能一切痼疾顽症的总根源，那么，人工智能范式的革命就成为解决人工智一切痼疾顽症的对症良方，也是实现人工智能基础理论重大突破和源头创新的唯一正确举措。

具体来说，所谓人工智能的范式革命，就是一方面要颠覆物质学科范式对人工智能学科的统领地位，同时要总结和确立信息学科范式对人工智能学科的整体引领，并以信息学科范式来规范人工智能各个层面的研究。

现在就来考察：人工智能的范式革命是怎样按照表 2 所载明的学科发展普遍规律，从人工智能研究的源头开始一步一步地创建全新的一代人工智能理论——通用人工智能理论。

5.1 第一步，解除物质学科范式对人工智能研究的约束，确立信息学科范式对人工智能研究的引领，首先阐明人工智能学科的宏观定义

如前所说，人工智能的范式革命，就是要彻底颠覆物质学科范式对人工智能的统领地位，全面确立信息学科范式对人工智能的引领和规范作用。

之所以要颠覆物质学科范式对人工智能的统领地位，是因为：

(1) 按照物质学科范式的科学观，研究对象应是孤立的脑物质，排除主观因素的影响。但是，智能不可能由孤立的脑物质产生（参考狼孩的试验）。因此物质学科范式的科学观不适用于人工智能的研究。

(2) 按照物质学科范式的方法论，就要贯彻分而治之和单纯形式化的方法。但是，这样两种方法只能导致人工智能理论的分裂（分而治之割断子系统之间的联系）和智能水平的低下（单纯形式化使信息、知识、智能全部空心化）。

可见，物质学科的范式确实把人工智能学科的研究完全引入了歧途。

颠覆了物质学科范式对人工智能的束缚，只是范式革命的第一步。更重要的步骤是要全面确立信息学科范式对人工智能的规范与引领。

那么，什么是信息学科的范式？

经过我们半个多世纪对信息学科的科学观及方法论的探究，它的基本特征已经被总结出来，并且表述在本文表 1 的第三行。

(1) 信息学科范式的科学观认为，人工智能的研究对象不应排除主体因素的存在，人工智能的研究对象应当是主体驾驭和环境约束下的主体与客体相互作用的信息过程。

(2) 信息学科范式的方法论表明，人工智能的研究不应遵循机械还原的分而治之和单纯形式化方法，人工智能的研究应当遵循信息生态方法论（即在保持信

息的整体性、连续性、系统性和全局优化条件下的信息转换方法)。

这样，根据信息学科范式的科学观和方法论，我们就可以得到人工智能学科的宏观定义（是什么？怎么做？），即：

（1）人工智能学科的本质，是主体驾驭和环境约束下主体客体相互作用的信息过程，研究的目的是实现主体客体相互合作和互利双赢。

（2）人工智能学科的研究，应当坚持信息生态方法论。

正确的学科定义，是人工智能学科能够走上健康发展轨道的根本基础。这样，人工智能的研究就不再局限于模拟人类大脑的结构，也不是去模拟大脑的功能或者人类表现的外部行为，而是要运用信息生态方法论，研究主客互动的信息过程，实现主客双赢的规律，不断改善人类的生存发展水平。

5.2 第二步，把学科的宏观定义具体落实为学科的全局模型和研究路径

根据信息学科范式科学观所阐明的学科定义，人工智能的全局研究模型应当是主体驾驭和环境约束下的主体与客体相互作用的信息过程模型。图 1 所表示的，正是符合这个学科宏观定义的人工智能全局研究模型^[4]。

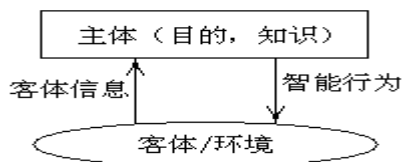


图 1 人工智能的全局模型：主体客体相互作用的信息过程

正如模型所表明的那样，人工智能学科的全局模型必然是客体信息首先作用于主体，然后，具有目的和知识的主体就设法生成智能行为反作用于客体，完成

主体与客体相互作用的第一个基本回合。

如果第一回合的智能行为反作用于客体的结果达到了预设的目标，这个回合就完成了任务；如果没有达到预设目标，其中的误差就要返回到主体的输入端，成为补充性的客体信息，促使主体去补充新的知识，优化智能策略和智能行为，改进反作用的效果（缩小误差），直至满意为止。

图 1 所示的人工智能全局研究模型虽然看似简单，却清晰而准确地揭示了人工智能的深层本质：只有主体与客体发生相互作用，才会使具有目的和知识的主体产生智能行为来反作用于环境的客体，实现主体与客体的合作双赢，一方面，主体达到了自己的预设目标；另一方面，客体维护了环境的运行规律。

那种既没有外部刺激输入、又没有自身行为向客体输出的孤立大脑，实际上不可能产生真正有用的智能。即使孤立的大脑非常“聪明”，但是没有外部的刺激就没有产生智能行为的激励与动因，因而不会去生成智能行为；而没有自身的输出，就不可能检验它所产生的行为是否具有智能水平。

根据信息学科范式的科学观构筑了正确的研究模型以后，接下来要思考的问题就是：应当怎样去研究这个全局模型？

根据信息学科范式的方法论，人工智能的研究应当遵循信息生态方法论。把这个问题说得更明确一些就是：遵循信“在主客互动框架下保持信息的整体性、连续性、系统性的条件下以及在全局优化条件下的信息转换方法”。

进一步，遵循了信息生态方法论的人工智能系统，又将会怎样产生所需要的智能行为呢？

事实上，图 1 的全局研究模型已经表明，人工智能系统产生智能行为的方法只与客体信息、主体目标、主体的知识这些因素有关，即：智能行为的生成必定

是在主体目标的导引下、在知识的支持与约束下、由客体信息通过复杂转换而产生出来。

由此，就可以构筑：在信息学科范式的信息生态方法论指引下智能行为生成的过程模型，如图 2 所示：

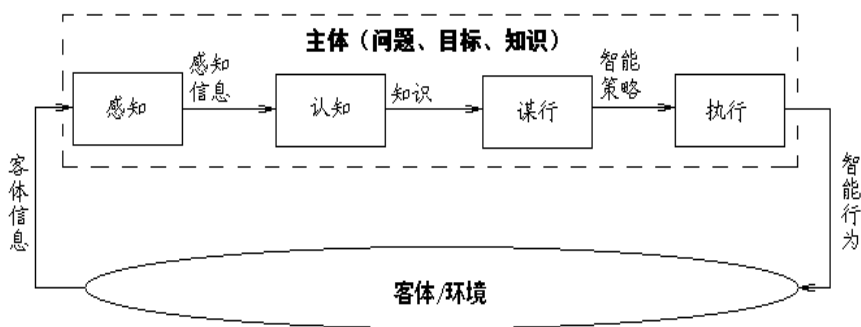


图 2 普适性智能生成机制的模型^[20]

显然，图 2 只是图 1 的具体化，但具体化了的图 2 就更加明确地揭示了主体产生智能行为的详细过程和普适性机制。

具体来说，首先就要实现主体对客体的感知，从而产生主体的感知信息，接着就要把感性的感知信息提升为理性的知识才能对客体有深刻的认识，然后就可以在预设目标的引导下、在知识的支持与约束下，把感知信息转换为解决问题的智能策略，进而转换成为智能行为。这就是上述图 2 的全部结果。

图 2 所揭示的智能生成机制是普适性的，因为其中定义的所有因素（感知、认知、谋行、执行）都是普适性的。图 2 还表明，这个普适性智能生成机制的准确内涵应当是“信息转换与智能创生”原理，即（以 表示“转换”算法）：

客体信息 感知信息 知识 智能策略 智能行为

这样，信息学科范式的信息生态方法论就创造了一种前所未有的人工智能研究路径：**基于普适性智能生成机制的机制主义研究路径**，它是一种统一的（普适

的)人工智能研究路径。有了这种统一的人工智能研究路径,原先的结构主义、功能主义、行为主义三者分道扬镳的研究路径就可以被放进历史博物馆!

事实上,对于人工智能基础理论研究来说,普适性智能生成机制才是人工智能的核心本质所在:系统的结构和功能都是为系统实现智能生成机制服务的。至于系统的行为,则是普适性智能生成机制实现以后的系统外部表现。

所以,这是人工智能理论的一个具有历史意义的重要结果。它给人工智能学科的研究大方向提供了一个极为重要的启示:普适性智能生成机制是一类复杂的信息转换,因此,要把人工智能研究的整体思路聚焦到“信息转换”这个大方向上来!只要抓住了“信息转换与智能创生”这个普适性智能生成机制,就抓住了人工智能研究的核心本质。算法、算力、数据、知识都是为实现“信息转换与智能创生”这个普适性智能生成机制服务的。

不过,需要特别提醒:作为普适性智能生成机制的本质,“信息转换与智能创生原理”中的源头—信息,并不是现在广泛流行的 Shannon 信息,而是感知信息,也就是全信息。如果不了解这一点,普适性智能生成机制仍然不可能生成真正的智能,这是因为,Shannon 信息是被单纯形式化方法阉割了内容和价值因素、只剩下形式因素的空心化信息!利用这样的空心化信息,普适性智能生成机制也只能提炼出空心化知识,生成空心化的智能!

本文在基本概念一节已经交代了感知信息(全信息/内容信息)的概念,指出了感知信息具有形式信息(语法信息)、价值信息(语用信息)、内容信息(语义信息)三个分量。不过,在那里还没有来得及交代这些分量是怎样生成的,也没有交代这三个分量之间存在什么样的相互关系。

这是因为,要想更深刻地阐述感知信息的定义和生成机制,首先必须懂得信

息学科的范式，即必须懂得：整个信息转换和智能创生的过程都必须“在主体驾驭和环境约束下主体与客体相互作用的框架下进行”这个大前提。而在本文第2节阐述基本概念的时候，还没有论述到这个前提。因此，只能等到现在这个时候，才具备了条件可以具体阐明：如何在主体驾驭下从外界的客体信息生成主体感知信息的工作机制。

在这里，“主体驾驭”和“主体与客体相互作用”的最重要标志，就是人工智能的整个信息转换与智能创生过程都必须尊重“主体目标”的导控作用，以及存在主体所提供的知识所发挥的支持与约束作用。

在此，值得特别强调的是：在人工智能系统中，必须十分重视“主体目标”的作用，它是体现主体主观能动作用和主体利益的至高无上的标志性因素。可以认为，只有那些有主体目标的人工智能系统，才会拥有真正的智能，没有主体目标的“人工智能”系统，不可能拥有真正的智能。

这是研究人工智能理论与研究一般物质系统理论之间的最大区别之一。这也是人工智能的研究不宜完全退化到物质系统的研究的一个重要原因。

在做了这些说明之后，现在就可以来构造“由客体信息转换生成感知信息”的工作机制模型，具体情况如图3所示。

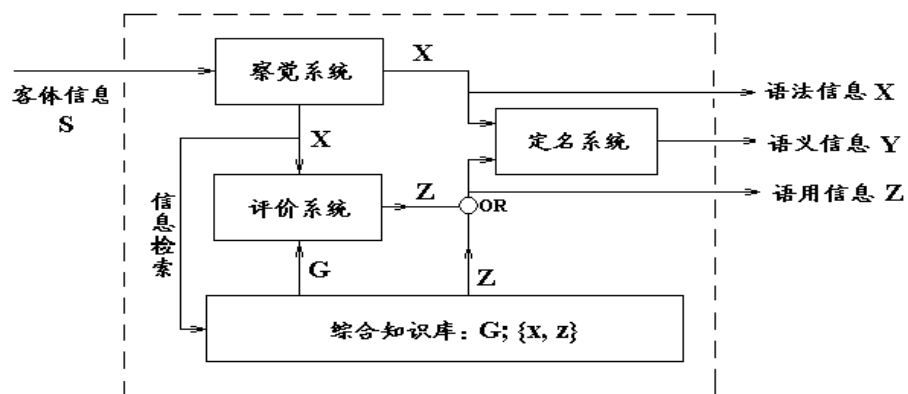


图3 感知信息生成机制的模型

图 3 模型中，这个驾驭全局的主体目标就是存储在“综合知识库”的 G。由它来确定外来的刺激（客体信息 S）对于达成系统的目标究竟是有利？有害？还是无关？也就是确定这个客体对于系统的价值信息（语用信息）为正？为负？还是为零，从而据此确定系统究竟应当对这个客体表示欢迎（支持）？反对（抵制）？还是不予理会（过滤）？

图 3 模型中的“**察觉系统**”就是传感系统，但它只能觉察和表达系统面临着什么形式的刺激 X（形式信息 / 语法信息），而不可能懂得刺激的内容；模型中的“**评价系统**”就是对刺激的价值评估：这个刺激对于达成主体目标有什么利害关系 Z（价值信息 / 语用信息）；其中的“**定义系统**”就是把形式信息 X 和价值信息 Z 两者形成的“偶对”映射到内容（语义）空间并对映射结果命名，这样就得到了内容（语义）信息 Y。显然，模型在技术上是完全可实现的。

模型表明，感知信息的形式信息 X（语法信息）、价值信息 Z（语用信息）和内容信息 Y（语义信息）三者之间，并非相互独立或完全平等的关系，它们之间存在“ $Y = (X, Z)$ ”的关系（其中的符号 $()$ 表示“映射与命名”的操作），即：内容信息乃是“把形式信息与价值信息所形成的偶对抽象而成的结果”。这是一个非常重要非常有用的结果，也是一个至今都被普遍误解了的重要结果。

由这样生成的全信息（感知信息）按照普适性智能生成机制生成的知识和智能就将是同样具备形式、价值、内容三个份量的“全知识”和“全智能”。

5.3 第三步，进一步精准刻画学科的学术结构规格和学术基础规格

根据信息学科范式的科学观，人工智能学科的全局研究模型是“主体客体相互作用的信息过程模型”，那么，为了精准研究这种主客互动的信息过程，人工智

能学科的学术结构规格也应当精细化。结果可以发现：人工智能学科的学术结构应当由人工智能的原型学科（神经科学、认知科学、人文科学等）、核心学科（信息科学、系统科学、控制科学等）和基础学科（数学、逻辑学、哲学等）所形成的交叉学科群结构；而不应当仅仅是计算机学科。

同时，根据信息学科范式的方法论要求，人工智能学科的方法论应当是信息生态方法论。于是，为了描述主客互动框架下的信息生态过程——以信息转换与智能创生为标志的普适智能生长机制，人工智能学科的学术基础的规格就应当满足“内涵的整体性、时空的连续性和系统性、全局的优化性”的数学理论、逻辑理论以及新型的信息哲学理论。而不应当仅仅是传统的概率理论、刚性化的数理逻辑和传统的哲学观念。

5.4 第四步，深度改造人工智能理论的基本概念和基本原理

由于现有的人工智能理论（包括他的基本概念和基本原理）是在物质学科范式约束下形成的，肯定不能适应信息学科范式的规范。因此，深度改造人工智能的基本概念和基本原理是绝对必要的工作。

比如，物质学科范式的科学观认为研究对象是物质客体，不允许主观因素的参与。这就导致现有人工智能的基本概念（如数据、知识、智能等等）都是纯粹客观的、绝对中性的、不反映主体因素的概念，同时也导致现有人工智能的工作原理（形式计算、统计、数理逻辑等等）也都是纯粹客观的、绝对中立的。

又比如，物质学科范式的方法论是分而治之和单纯形式化，因此，现有人工智能的基本概念（数据、知识、智能等等）都是被肢解和被阉割了内容与价值因素的形式化概念，而现有人工智能的基本原理（纯形式的计算、统计、数理形式逻辑等等）也都是被肢解和被阉割了内容与价值因素的原理。

显然，如果把现有人工智能的这些基本概念和基本原理原封不动地照搬到新的人工智能理论中来，那么就意味着，信息学科范式“被截去了双腿”。换句话说，这样的范式革命是半截子的、不成功的：学科的定义（学科的科学观和方法论）、学科的定位（学科的全局模型和研究路径）、学科的定格（学科的学术结构和学术基础）都符合信息学科范式的要求，而学科理论的基本概念和基本原理却是遵循了物质学科范式的要求！

显然，这种半截子的人工智能范式革命，绝对不可能成功。因此，人工智能理论基本概念和基本原理的深度改造势在必行。

根据信息学科范式的科学观，人工智能全局研究模型是“在主体驾驭和环境约束下主体与客体相互作用的信息过程”，那么，为了研究主体与客互动的信息过程，使主体能够从形式、价值、内容上全面了解客体，人工智能学科的基本概念也就应当是形式价值内容三位一体的全信息、形式价值内容三位一体的全知识、形式价值内容三位一体的全智能等基本概念。其中，价值和内容便是信息学科范式科学观所要求的主体因素的具体体现，也是信息学科范式的信息生态方法论的体现。

近些年来，从实际操作的角度，人们把纯粹形式化的“数据”看作是人工智能的粮食。这看似有一定的道理。不过，深入研究就会认识到，数据只是信息的载体，不含信息的形式数据根本无法使主体全面认识客体，因此不能启动“普适性智能生成机制”，不能启动人工智能系统的有效工作。只有携带了与系统目标相关（无论是正相关还是负相关）的信息（即感知信息的价值信息分量不为零），才能真正启动人工智能系统的工作。所以，对于人工智能的研究来说，真正重要的概念是全信息以及由全信息引出的一系列概念，而不是纯形式的数据。

同时，根据信息学科范式的方法论，人工智能学科的方法论应当是信息生态方法论，那么，为了满足“整体性、连续性、系统性和全局优化性”，人工智能学科的基本原理就应当是符合信息生态规律的一系列信息转换并最终创生智能策略的那些原理，它们构成“信息转换与智能创生”定律。

很有意义的是，信息学科领域的“信息转换与智能创生定律”，竟是与物质学科领域的“质量转换与物质不灭定律”以及能量学科领域的“能量转换与能量守恒定律”一样重要（甚至是更重要）的基本科学定律。着实发人深省！

5.5 第五步，最后综合形成通用的人工智能理论

最后，把人工智能范式革命及其全部链锁反应的结果综合在一起，就可以产生人工智能范式革命所创造的新一代人工智能理论——基于普适性智能生成机制的通用人工智能理论。

普适性智能生成机制之所以特别重要，并且成为人工智能范式革命最为核心的成果，是因为，无论面对任何问题（任何应用场景），只要能够提供“问题的充分描述、问题的求解目标和问题求解的相关知识”，那么，基于普适性智能生成机制的人工智能系统就可以生成解决问题达到目标的智能策略和智能行为。这就是通用人工智能的基本思想。

作为最终结果，可以在图 2 所示的普适性智能生成机制模型的基础上生成整个通用的人工智能理论的系统模型，而且完整地体现了表 2 所列的“学科发展普遍规律”，这就是图 4 所示的通用人工智能理论系统模型。

到此需要指出，所谓通用人工智能，并不是指：用一个单体的“巨无霸”式的人工智能系统去包打天下，去解决世间所有的问题。我们所说的**通用人工智能理论**，是指以不变的(普适性的)智能生成机制去成功应对千变万化的实际问题。这里的通用是指智能生成机制的通用，而通用人工智能系统的输入内容(问题、目标、知识)和输出内容(智能策略和智能行为)则将随着问题的改变而相应改变。这就是“以不变应万变”的真实含义。

具体来说，给定任何合理的“求解的问题、求解的目标、求解问题所需要的先验知识”(所谓“合理”，是指：所给定的“求解问题”至少在理论上确实存在着“求解目标”，或者说给定的“求解目标”一定存在)，通用人工智能理论都可以凭借它的普适性的(不变的)智能生成机制，在给定的目标引导下，利用给定的知识(在先验知识不足的情况下也可以通过学习来扩展知识)，生成能够求解问题达到求解目标的智能策略和智能行为。当然，如果求解的问题改变了(这意味着求解目标和所需要的知识也会相应改变)，解决问题的策略和智能行为也会随之而改变，但是，生成这种智能策略和智能行为的机制却永远不会改变：依然是“信息转换与智能创生定律”。

换句话说，“普适性的智能生成机制”是人工智能理论研究中的不变性和不变核。这样，面对千姿百态的问题，不再必须对每个不同的问题都推倒重来从头做起去设计专用系统，只需要把各自的“问题、目标、知识”表达成一定的规格，利用同样的普适性的智能生成机制就可以生成解决相关问题的智能策略和智能行为，达到求解问题的目的。

总之，通用人工智能系统乃是“人工智能系统的通用孵化平台”。通用人工智能理论的“机制通用性和理论统一性”，是人工智能范式革命带来的巨大优越性，

现行人工智能理论无法望其项背。

同样十分重要的问题是，由于现行人工智能遵循了“单纯形式化”的物质学科范式方法论，使得现行人工智能系统的“智能水平十分低下”，成为现行人工智能理论的又一项痼疾顽症。现行人工智能系统表显出来的“智能”，其实都不是基于对问题的理解来实现的，而是利用快速运算变幻出来的结果。这样设计出来的“智能”严格依赖于特定的应用场景，场景一变，这种智能便会失效。

从通用人工智能理论的图 2 和图 4 的模型则可看出，在这里，客体信息经过感知就被转换成为了感知信息。根据本文第 2 节的概念定义，感知信息是形式信息（语法信息）、价值信息（语用信息）、内容信息（语义信息）的三位一体；知识也是形式知识、价值知识、内容知识的三位一体；智能策略也是形式策略、价值策略、内容策略的三位一体。因此，信息、知识、策略都是可以理解的。

当然，人们不应当完全按照“人类的理解能力水平”来要求“人工智能机器的理解能力水平”。但是，人工智能的“理解能力”仍然能使人工智能机器做出具有智能水平的决策。

比如，在信息的层面上，系统可以根据感知信息的价值信息分量的大小对客体做出基于理解的明智决策：

- 1) 如果价值信息为足够大的正值，就采取相应的措施来保护客体。
- 2) 如果价值信息为足够大的负值，就采取相应的措施来反制客体。
- 3) 如果价值信息为足够小的正值或负值，就不理会或过滤掉这样的客体，
- 4) 在多中选一的情况下，就选择其中价值信息为最大正值的客体，或者选择其中价值信息为最小负值的客体。

在知识和策略的层面，也同样可以针对所面临的问题做出与信息层面类似的

具有一定智能水平的明智决策。而且，无论在信息、知识、智能的层面，这样所做出的决策，都可以清晰而准确地得到解释。

所以，人工智能范式革命为基于普适性智能生成机制的通用人工智能理论带来的另一个巨大优越性，就是它的理解能力和在理解基础上的智能决策水平。与通用人工智能理论的这种智能水平相比，现行人工智能理论的理解水平更是望尘莫及。

以上所述，就是人工智能范式革命的学术内涵、实现途径（表现为表 2 所示的学科发展普遍规律）和重大学术意义，也是人工智能范式革命所带来的人工智能理论划时代的深刻革命。

表 3 是对上述论证的简明总结。

表 3 人工智能范式革命创造了通用人工智能理论

对比项目	物质学科范式下的人工智能	两个科学时代的分水岭	信息学科范式下的通用人工智能
科学观	机械唯物的物质观：非主观，结构		对立统一的信息观：主客互动，双赢
方法论	机械还原论：纯形式，分而治之		信息生态学：整体化，生态演化
全局模型	孤立的脑模型		主体驾驭的主客互动的信息过程模型
研究路径	结构、功能、行为模拟分道扬镳		基于智能生长机制的机制主义路径
学术结构	计算机		原型—核心—基础等多学科交汇
基础特色	概率论，形式逻辑		因素空间数学理论，泛逻辑理论
基本概念	形式数据、形式知识、形式智能		全信息、全知识、全智能等
基本原理	未做总结		信息转换与智能创生定律
综合结果	三个独立的局部理论		通用人工智能理论

值得指出，表 3 显示了，在实施人工智能范式革命之前，“三分天下”的现行人工智能理论乃是物质学科主导的科学时代的人工智能理论；而在实施人工智能范式革命之后，“通用人工智能理论”则是信息学科主导的科学新时代的人工智能理论。人工智能的范式革命，是这两个时代人工智能理论的分水岭。

正是在这个意义上，人工智能的范式革命，乃是推动物质学科主导的科学时代转变到信息学科主导的科学新时代的引擎和桥梁，也是划分物质学科主导的科学时代与信息学科主导的科学新时代的界限和分水岭。这样，人工智能范式革命的产物——基于普适性智能生成机制的通用人工智能理论就名副其实地具有了划时代的意义。

这就是人工智能范式革命及其链锁反应所带来的人工智能基础理论的具有划时代意义的重大突破和源头创新。

致谢

本文是笔者数十年来在国家自然科学基金（包括，但不限于 68872014, 69171023, 69982001, 60496327, 60575034, 60873001, 70711120412 等）以及国家社科基金（18ZDA027）资助下所获得的系统性研究结果。本文的研究还得到陆汝钤院士、李衍达院士、陆建华院士、谭铁牛院士、蒲慕明院士、涂序彦教授、何华灿教授、汪培庄教授、史忠植教授、韩力群教授、王小捷教授、周延泉副教授、李蕾副教授、李睿凡副教授、陈志成博士、孙健博士等多种形式的帮助。在此谨表衷心的感谢！

17-799, 为什么通用人工智能不会实现

Ragnar Fjelland, 挪威卑尔根大学, 2020.6

内容：创造类人人工智能(AI)的现代项目始于第二次世界大战后，当时人们发现电子计算机不仅是处理数字的机器，而且还能处理符号。追求这一目标是有可能的，无需假设机器智能与人类智能是相同的。这就是所谓的弱 AI。

然而，许多人工智能研究人员追求的目标是开发原则上与人类智能相同的人工智能，称为强人工智能。弱 AI 不如强 AI 雄心勃勃，因此争议较少。然而，也存在着与弱 AI 相关的重要争议。本文重点讨论了人工智能 (AGI) 和人工窄智能 (ANI) 的区别。虽然 AGI 可能被归类为弱 AI，但它接近于强 AI，因为人类智能的一个主要特征是它的普遍性。尽管 AGI 没有强大的 AI 那么雄心勃勃，但几乎从一开始就有批评者。主要的批评者之一是哲学家休伯特·德雷福斯 (Hubert Dreyfus)，他认为，没有身体、没有童年、没有文化实践的计算机根本不可能获得智力。德雷福斯的主要论点之一是，人类的知识在一定程度上是隐性的，因此无法在计算机程序中表达和整合。然而，今天有人可能会说，人工智能研究的新方法已经使他的观点过时了。深度学习和大数据是最新的方法，倡导者认为它们将能够实现 AGI。仔细观察就会发现，尽管用于特定目的的人工智能 (ANI) 的发展令人印象深刻，但我们还没有在开发人工通用智能 (AGI) 方面走得更近。这篇文章进一步指出，这在原则上是不可能的，它恢复了休伯特·德雷福斯的观点，即计算机并不存在于这个世界上。

18-814, 通用人工智能技术 (AGI) 的认识过程论探析

陈凡、吴怡, 2021.11, 东北大学科学技术哲学研究中心

摘要：人工智能技术的发展被人们寄以从“专用”到“通用”的期盼，然而其实现的瓶颈在于循规蹈矩的程序恐难以真正理解符号背后的意义，由此表现为无法实现迁移学习的能力。马克思主义认为，人的认识是由感性上升至理性的过程。当前智能机器所做出的智能决策虽与人类由理性做出的决

策结果相吻合，但因其缺少感性过渡与升华的阶段，故而与人类智能两异。从马克思主义对于人类认识发展过程的科学阐释来看，人工智能技术从专用转向通用，其间的感性要素是必要不充分条件。

往常传统的人工智能路径总是急于将人类的功能结果作为机器模拟的直接来源，抑或是将人类的神经元结构做以静态的拷贝。它们都是出于功能结果和片面结构的模拟，而无一种方式是完全按照人类认识的发展过程的规律（即按照感性认识发展至理性认识这一顺序）进行模拟的，皆忽视了感性认识在智能机器设计中的基础性地位。马克思主义对于人类认识发展规律的科学解释，可从哲学的视角为通用人工智能技术提供一种全新的思路。虽然对于机器而言，这是个极为复杂繁琐的过程，但是我们没有捷径，应该遵循事物发展的客观规律。人类如此，机器亦如此。

19-826，无开源不通用：通用人工智能机器生产工艺学批判

刘方喜，2022.3.15，中国社会科学院

摘要：技术工艺性应用的通用性，需要社会性应用的开源性与之匹配，在人工智能通用性不断提升的动态发展趋向中，构建与其匹配的开源性伦理规则，推动高效而无害的通用人工智能创造，具有重要意义。随着创造并使用物质和精神劳动工具活动的发展，人类智能的封闭性、非通用性被不断超越而社会性、开源性不断提升，现代科学和自动机器大大加速了这一进程。现代科学这种社会通用智能的工艺性应用，首先把物质劳动工具的使用技巧转移到能量自动化机器上而成为社会机械通用智能，超越了手工智能封闭于个体人身内的生物性的非开源性。当今人工智能正在使精神劳动工具的

使用技巧也向机器转移,将进一步超越智能的个人生物性的非开源性,再进一步超越资本商业化社会性应用非生物性的非开源性,通用人工智能将成为高度自动化的社会机械通用智能。无开源不通用,构建公义创新动力机制,聚合非市场、非营利创新动力,人工智能将在通用性、自动性、开源性高度统一中充分发展并造福全人类。

20-827, 通用人工智能需要在私人语言的层面上进行知识表征吗? —来自大森庄藏的启发

徐英瑾, 2021. 12. 6, 复旦大学哲学学院

摘要: 通用人工智能语境中的私人语言,指的是这样一个意思: 表征 A 在系统甲那里的知识表征方式与同一个表征在系统乙那里的表征方式必然会有所差异。因此,在预设推论主义语义学自身有效性的前提下,A 在甲中的意义集,总会有一个子集(无论这一子集有多小)仅仅为甲自身所拥有,而无法被任何一个别的系统所拥有。因此,任何一个与甲不同的别的系统,都无法彻底地理解甲对于 A 的意义把握方式。很显然,这样一种将机器表征与哲学史上的"第一人称哲学"传统相结合的思路,是无法见容于后期维特根斯坦对于私人语言的著名反驳的。而为了与后期维特根斯坦论战,日本哲学家大森庄藏的思想资源便具有了很高的引用价值,因为他本身的哲学就可以被视为"维特根斯坦的话语方式与胡塞尔的思想内核"的日本式混合体。在对大森的哲学进行面向机器表征问题的重建的过程中,对于局域性原则与历史性原则的引入也是题中应有之义,以便为通用人工智能语境中建设私人语言的必要性提供辩护。而非公理性推理系统(纳思系统)所提供的技术手

段,则会为这种想法的技术落地提供可能。

21-854, 如何让通用人工智能系统能够“识数”

徐英瑾, 2021. 9. 25

复旦大学哲学院

摘要：“识数”的字面意思就是说,数字表达式的使用者需要知道这些表达式日常实践中的具体意义,如知道“十公里”算不算“路远”,“十万元”算不算“昂贵”。主流人工智能系统虽然都是数码化的,但却未必真正“识数”,因为它们都无法以一种灵活的方式来理解数字表达式在特定的日常语境中究竟意味着什么,并在这种理解的基础上进行复杂的语用推理。为了使得人工智能系统能够“识数”,工作思路应是在非公理化推理系统的平台上来构建数字表征。训练非公理化推理系统使之“识数”的指导原则,就是先教会它表征那些最不抽象的数学概念,再一步步进展到更抽象的概念。而这一教学步骤的安排,与我们教导儿童识数的路线图是有点相似的。在这个面向机器的数字教学法中,我们将以对“逻辑专名”的教学作为对数字表征教学的重要前导。

22-856, 基于多尺度滤波器组的时间、空间特征的集成学习方法用于运动图像脑电图分类

LiangShengZheng 等, 2022. 3. 20

中科院深圳先进技术研究院广东机器人与智能系统重点实验室, 中国科学院大学, 重庆邮电大学

内容: 从脑电图 (EEG) 解码运动意图是基于运动想象的脑机接口 (MI-BCI) 的关

键部分。为了帮助神经肌肉残疾的残疾人通过 BCI 恢复运动功能，需要建立一种高效稳定的分类算法来解码脑电信号中包含的运动意图。然而，EEG 信号是非平稳的，个体之间差异很大。在这项工作中，我们提出了一种基于时间、空间特征和多尺度滤波器组的集成学习方法，称为 TSMFBEL，旨在为 MI EEG 分类设计具有强泛化能力的集成分类器模型。为了获得多样化的集成分类器，将原始脑电数据通过自举抽样方法划分为具有样本多样性的子集，然后通过多尺度滤波器组方法分解为具有时频分布多样性的时频子集。对于每个时频子集，从时域、空间域和时空域中提取具有域多样性的特征，并根据每组特征训练具有多样性的异构分类器。为了获得最优决策，我们将集成策略描述为最小分类误差优化问题，并提出基于 L2 范数的集成分类器权重优化方法，最后通过加权融合对集成分类器的决策进行整合。所提出的方法在两个公共数据集（BCI Competition IV Dataset IIa 和 BCI Competition IV Dataset IIb）上进行了评估，并将结果与最先进方法的分类方法进行比较。实验结果表明，所提出的 TSMFBEL 算法能够有效地构建多样化的集成分类器，在两个数据集上的平均分类准确率分别为 88.80% 和 86.53%，是目前最先进的方法中最高的，结果的标准差也是最低的。优异的分类性能表明，该算法在脑电信号解码方面具有巨大潜力。

23-866，机器也需教育？论通用人工智能与教育学的革新

刘凯、胡祥恩、王培，2018.2

华中师范大学心理学院、孟菲斯大学智能研究院、天普大学计算机与信息科学系

摘要：历史悠久且底蕴深厚的教育学如今却处于被边缘化的尴尬境地，通用人工智能的兴起为教育学复兴带来契机。总体而言，人工智能分为专用(弱)人工智能

和通用(强)人工智能两大分支，二者技术路线完全不同。专用人工智能侧重对“智能”行为外在的模拟，通用人工智能则致力实现机器内在的思维和情感，认为智能是一个系统在知识和资源相对不足情况下的适应能力，强调自身经验对个体塑造的决定性作用。在教育领域中，通用人工智能不仅破除了羁绊专用人工智能的技术崇拜与人文关怀迷失两大问题，更表现了与人类教育活动极相似的特征：首先是“教”，人类学习理论对通用人工智能系统的学习过程同样有效；其次是“育”，通用人工智能系统在学习时也同样需要借助经验积累来实现育化的效果。因此，在通用人工智能的启迪下，教育学将打破“人”的先天束缚，在人类教育理论与实践的基础上吸纳和统合机器教育，在更宽广的“人-机”二元主体视角下探究教育与学习的一般性规律，向“大学科”和“大科学”的方向迈进。

24-891，脑工终结时代的来临：通用人工智能机器生产工艺学批判

刘方喜 2022年3月

中国社会科学院文学研究所

能量高度自动化的机器曾经终结手工时代，通用人工智能将成为高度发达的社会大脑所生成的高度自动化的社会机械通用智能，脑工终结时代正在来临。人类有望彻底征服自然力，使物质和精神生产力从人身生物性限制下全面解放出来，在物种关系方面超越人类与自然关系的动物性；扬弃资本之后，人与人社会关系中过度竞争的动物性也将被超越，每个人的手工、脑工劳动将从资本支配下的竞争性、商品性、雇佣性中解放出来，手工、脑工劳动面前人人平等，体力、智力自由发挥人人所求，每个人的手工、脑工劳动将从必然王国中解放出来转移到自由王国，每个人的体力、智力将得到全面、自由发展，每个人的生命意义将在平等

与自由、人道主义与自然主义的高度统一中得到全面实现。面对脑工终结时代的来临，重构马克思机器生产工艺学批判，有助于推动人工智能科学发展与合理应用。

25-900, 使用自动挖掘知识和深度神经网络回答中医问题：端到端解决方案

Lizhang 等, 2022. 4. 15

浙江大学先进技术研究院、浙江大学计算机科学系

互联网上的医疗信息迅速增加，已成为搜索引擎使用的主要目标之一。然而，互联网上的医疗信息存在质量和可访问性问题，普通用户无法方便地获得医疗问题的答案。作为一种解决方案，研究人员构建了医学问答（QA）系统。然而，用中文进行医学 QA 的研究落后于基于英语的系统的工作。这种滞后主要是由于构建高质量知识库的困难和中文医学语料库的利用不足。

结果

本研究开发了一种端到端的解决方案，以低成本和时间实施中文医疗质量保证系统。首先，我们以几乎自动的方式从医院数据（电子健康/医疗记录）中创建了一个高质量的医学知识图谱，该图基于使用自举技术标记的数据训练了一个监督模型。然后，我们设计了一个基于记忆神经网络和注意力机制的 QA 系统。最后，我们训练系统从知识库和互联网上的 QA 语料库中生成答案。

结论

自举和深度神经网络技术可以从电子健康/医疗记录中构建具有令人满意的精度和覆盖率的医学知识图谱。我们提出的上下文桥机制执行具有各种语言特征的训练。我们的 QA 系统可以在回答主题受限的医学问题方面达到最先进的质量。正如我

们评估的那样，复杂的中文语言处理技术，如分割和解析，对于实践来说不是必需的，复杂的架构也不是构建 QA 系统所必需的。最后，我们使用我们的方法创建了一个应用程序用于互联网 QA 使用。

26-902，基于历史角度分析的强人工智能论争（通用人工智能）

王彦雨，2018.12

中科院自然科学史研究所

在人工智能发展史上,强人工智能(“强 AI”)一直是一个争议不断但却又不断引发人们关注的议题.对于“强 AI”理念,我们应合理看待其所发的各种争论:(1)“强 AI”理念是推动人工智能界不断打破人机界限、使 AI 技术向前发展的重要信念;(2)“强 AI”争论背后所反映的是不同社会要素,特别是“两种文化”(科学文化和人文文化)之间的张力,且 AI 界的这几种文化也经历着由对立、冲突,到逐渐的尝试性对话与合作这一过程.(3)AI 界、哲学界在对待“强 AI”这一议题的态度并非一成不变,AI 界经历了一个由乐观与支持到悲观与放弃以及现在的谨慎心态,而哲学界对于“强 AI”的态度则是沿着由批判与质疑到现在的大力宣扬这一路径演变,且当前他们对强 AI 所可能引发的风险更为忧虑.(4)“强 AI”概念需要进行重新界定,使其成为科学而非单纯的“科幻式”概念,并给予强 AI 风险议题更多关注。

27-903，强人工智能争论过程中的“态度转换”现象研究

王彦雨，2020.8

中科院自然科学史研究所

在分析“强人工智能(强 AI)”概念的源起、类型及特征的基础上,对强 AI 争论过程中的“态度反转”现象进行了详细描述,即强 AI 争论过程中的两个主要参与主体——人工智能界及哲学社会科学界对于“强 AI 是否会实现”这一议题,均经历了态度上的反转(由赞同到反对抑或相反),并对“态度转换”现象背后的动力机制进行了解析.文章最后强调,传统的强 AI 概念过于科幻,并提出基于“自主性”理念的“强 AI2”,以期破解传统强 AI 争论过程中的“不落地”困境。

28-927, 范式革命: 人工智能基础理论源头创新的必由之路

北京邮电大学人工智能学院教授 钟义信

【摘要】人工智能是开放复杂的信息系统,应当遵循信息学科范式。然而,由于受到“社会意识滞后于社会存在”法则的制约,信息学科范式至今未能形成,于是人工智能研究实际遵循的是物质学科范式,致使人工智能研究一直受困于初级阶段,只能取得局部的浅层进展。唯有通过人工智能的范式革命建立信息学科范式,才能获得人工智能研究的源头创新和整体创新,进入人工智能发展的高级阶段。正是沿着这个思路,需要研究和总结信息学科范式,以此突破物质学科范式对人工智能研究的束缚,并在信息学科范式引领下,发现全新的人工智能研究模型,揭示“信息转换与智能创生定律”这个普适的智能生成机制,创建源头创新和整体创新的人工智能理论——机制主义通用人工智能理论。

【关键词】人工智能 范式革命 机制主义

与解放人类体质能力的材料科学和解放人类体力能力的能量科学(统称为物质学科)不同,人工智能是以解放人类智力能力为目标的一门学科。因此,人工智能学科一问世就受到人类社会的高度关注。至今,发展人工智能已经成为世界各国特别是各发达国家的重大战略。2018年10月31日,习近平总书记指出:“人工智能是引领这一轮科技革命和产业变革的战略性技术”,“要加强基础理论研究,支持科学家勇闯人工智能科技前沿的‘无人区’,努力在人工智能发展方向和理论、方法、工具、系统等方面取得变革性、颠覆性突破,确保我国在人工智能这个重要领域的理论研究走在前面、关键核心技术占领制高点。”(习近平,2018)2017年

7月，中国国务院发布了《新一代人工智能发展规划》。根据《新一代人工智能发展规划》，2021年，我国的人工智能研究进入了第二个战略阶段，目标是实现人工智能基础理论的重大突破，本文就以实现这一目标为主题展开论述。

实现人工智能基础理论的重大突破，需要准确理解人工智能的历史

子曰，“温故而知新”。人工智能发展的历史演绎了各种不同观念和方法之间的矛盾冲突，折射出发人深省的宝贵启示，值得认真汲取。

人工智能的历史：观念方法的矛盾演绎。用机器来协助人们进行劳动，是人类一直不懈追求的美好理想。不过，真正付诸实践的研究始于20世纪40年代。那时人们相信，人类的高级认知功能产生于大脑新皮层的神经网络，认为只要把大脑神经网络的结构在机器上实现出来就可以模拟出人类的智能。这是结构模拟方法一马当先的缘由。具体来说，1943年McCulloch和Pitts发表了神经元的数理逻辑模型（W. S. McCulloch; W. Pitts, 1943），1949年Hebb提出了神经元的学习规则（D. O. Hebb, 1949）。利用M-P模型和Hebb规则就可以构造人工的神经网络，开展人工神经网络及其应用研究。

人类大脑神经网络是由近千亿个神经元互相复杂连接而成的大规模非线性系统，而当时的科技和工业能力都只能实现小规模简单神经网络系统，因此，虽然结构模拟的人工智能研究似乎前景光明，但道路却很曲折，进展缓慢。

面对这种情况，一批思想活跃的学者便撇开结构模拟的途径，选择了功能模拟的途径。1956年，McCarthy等人在Dartmouth发起利用计算机作为硬件平台、通过编制“聪明软件”来模拟人类逻辑思维功能的研究途径，他们还创造了Artificial Intelligence这一术语来表征这个新的研究领域（J. McCarthy; M. L. Minsky; N. Rochester, and C. E. Shannon, 2006）。由于那时的计算机已有很强的功能，Newell和Simon等人还提出了通用问题求解算法，并把功能模拟的研究产物称为无所不包的“物理符号系统”（A. Newell, 1980），认为这是可以像人脑一样求解通用问题的人工智能系统。果然，功能模拟方法一经问世，便在数学定理机器证明和模式识别等方面取得了令人鼓舞的成绩。

不过，当人们真的利用这一方法来解决通用问题的时候，就发现这很不现实，因为求解通用问题需要无限的知识。于是，人们不得不把面向通用问题求解的物理符号系统改成面向专门问题求解的专家系统。然而，解决专门问题也需要专门的知识，而知识的定位、获取、表示、推理等都存在难以逾越的“知识瓶颈”困难。在这种情况下，Brooks等人提出了行为模拟的

思路。他们宣称，行为模拟方法不需要知识，只需要感知到环境对智能系统的刺激和智能系统对此所产生的动作响应（于是被称为“感知动作系统”），因此，它可以回避结构模拟方法的复杂性问题 and 功能模拟方法的“知识瓶颈”困难。行为模拟的典型研究成果，是 Brooks 领导的 MIT 人工智能实验室在 1990 年所演示的六脚虫爬行机器人以及随后所研制的各种智能机器人（R. A. Brooks, 1990）。遗憾的是，对于行为模拟的感知动作系统来说，它自身的最大问题是仅能模拟智能系统的外在行为表现，因而只是一类浅层的智能。

上述历史表明，人工智能存在结构模拟、功能模拟、行为模拟三种不同的研究方法。它们是在人工智能发展的不同阶段、面临不同的问题、分别由不同的人群发展出来的各不相同的研究方法。它们的学术信仰（以脑科学为模拟原型的结构主义、以认知科学为模拟思路的功能主义、以控制论为模拟原理的行为主义）各不相同，所采取的策略各不相同，所显现的能力也各不相同。虽然都是人工智能的研究方法，但却未能实现殊途同归。

实际上，结构模拟、功能模拟、行为模拟三大学派非但未能殊途同归，没有形成合力，反而各执己见，互不相容，以至“恃强凌弱”，成为同行冤家。以下是三大学派称雄争霸的若干典型事例。1969 年，功能主义方法的学术带头人 Minsky 和他的同事 Papert 出版学术专著《感知机》严厉抨击神经网络方法“没有科学价值（without scientific value）”，认为它的成果建在流沙之上（built on quick sand）（M. L. Minsky; S. A. Papert, 1969）。这种猛烈的批评，造成世界范围人工神经网络领域的研究人员纷纷转行，使人工神经网络的研究陷入为期十多年的“黑暗年代”。1987 年，借着专家系统的研究遭遇“知识瓶颈”的困难而神经网络研究进入复兴阶段的机会，结构主义人工神经网络研究者们在美国加州圣地亚哥举行的 IEEE 第一届神经网络国际会议期间发出“人工智能死了，神经网络万岁（AI is dead. Long live neural networks）”的呐喊。表现了结构主义人工神经网络研究学派对功能主义专家系统研究学派的强烈反击。1990 年前后，行为主义感知动作系统研究学派的带头人 Brooks 连续发表论文抨击正在遭遇“知识瓶颈”的功能主义研究方法，宣称行为主义感知动作系统的研究方法可以不需要知识（Intelligence without Knowledge），因而也不需要知识的表示（Intelligence without Representation）（R. A. Brooks, 1991），并以此否定功能模拟方法。

正是由于不同学派的学术主张各不妥协，也没有找到互相沟通和融合的实际途径，人工智能的研究始终处于某个学派一派独大的格局，从来没有出现过相互合作的局面：1943 年至 1956 年间，由于其它两种研究路径尚未问世，人工智能只存在结构主义（初期的简单人工神经网络

和算法)一种研究路径。1956年至2016年这60年间,基本上是功能主义“雄霸天下”。虽然在20世纪70年代初期由于在机器翻译领域的失败和随后由于知识瓶颈的困扰曾经两度遭遇发展的危机,但是,功能主义方法仍然维持着它的统治地位,因为这段时间被它打入“黑暗年代”的结构主义方法还没有完全缓过气来。从2016年至今,由于初期的简单人工神经网络逐渐发展成为复杂的深层结构模型,训练和学习的算法也得到显著改进,在人工智能的一些竞赛评测项目中表现出大大超越功能主义方法、甚至超越人类能力的优异性能,于是,基于结构主义方法的深层人工神经网络研究变成这一时期人工智能的主流学派,基于功能主义方法的专家系统研究则似乎遁迹销声了。半个多世纪的人工智能发展的这段历史,引人深思。

历史的启示:人工智能问题的症结何在?经过70多年的发展,人工智能研究的结构主义(人工神经网络研究)、功能主义(专家系统研究)、行为主义(感知动作系统研究)三大学派各自都取得了不少引以自豪的成果,但也都面临严峻的挑战。从推动人工智能基础理论源头创新的需要着眼,我们将重点解析现行人工智能面临的根本性问题。

其一,直接的观察:方法论的误用。历史记载,最早出现的人工智能研究方法是模拟大脑皮层神经网络结构的人工神经网络,然后是模拟人脑逻辑思维功能的物理符号系统/专家系统,最后是模拟主体行为的感知动作系统。然而,三者各自为战,无法形成合力。可见,人工智能面临的第一个问题是:三大学派无法形成统一的理论,也无法建立通用的系统。由于没有统一的理论,现行人工智能系统的研究都严格依赖于场景。人们在研究人工智能系统的时候,首先就需要精心选择应用的场景,然后针对选定的这个场景来设计解决特定问题的人工智能系统。一旦应用场景改变了,系统就要推倒重来重新设计:一个场景一个系统,不同场景就要求设计不同的系统。

曾经听到过这样的说法:人工智能领域或许根本就不存在统一的理论,因此,也不存在通用的系统;只要一个一个应用场景的人工智能系统都研制出来了,人工智能的问题就解决了。显然,这种说法不符合辩证法和“可持续发展”的理念,也不符合系统论的基本原理。系统论指出:系统(整体)远远大于它的部分和。也即,系统的整体解决远远胜于它的所有部分解决之和,其间存在整体优化和局部优化之间的质的差别。因而,“现行人工智能无法建立统一理论”乃是一个严肃的信号,表明人工智能的研究还远远没有到位。问题是,为什么现行人工智能的研究无法建立统一的理论呢?人工智能发展的历史表明:作为开放复杂信息系统的人工智能研究,因为沿用了物质学科“分而治之”的方法论,才导致把人工智能的整体研究肢解成了结构模拟、功能模拟、行为模拟三种子系统的分别研究。其实,“分而治之”方法论本身在物

质学科领域非常有效，但用在人工智能领域却割断了其子系统之间内在复杂而隐秘的信息联系，而这些内在的信息联系正是人工智能系统的命脉和灵魂，因此该方法论当然就无法把它们恢复成为原来的复杂信息系统整体。方法论的不当，成为人工智能统一理论无法建立的主要原因。

第二个问题是：现行人工智能的智能不是在理解基础上的真正智能，而是快速计算能力的奇葩表现，真正的智能水平实际上十分低下。人工智能是在人造机器上所实现的人类智能。这就表明，人工智能系统所表现智能应当是在理解问题基础上的能力。但是，现行人工智能系统表现的“智能”其实并没有理解能力。为什么这样说呢？

这是因为现行人工智能研究用的都是纯粹形式化的方法，“阉割”了问题的内容与价值因素。显然，仅凭形式不可能实现理解，因此，也就不可能作出有智能水平的决策。比如，当你看见一头猛虎，你看到了它的形态（形式信息），如果你知道它会伤人（价值信息），知道它是食人猛兽（内容信息），那你就理解了老虎，就可以作出有智能水平的决策：躲、逃或在不得已的情况下射杀。但是如果你不知道老虎会伤人（不具有价值信息），也不知道它是食人猛兽（不具有内容信息），你就没有理解老虎这种凶猛野兽。在这种情况下，你怎么知道应当作出怎样的决策？说不定你会“与虎共舞”，结果为虎所食。其实，单纯形式化是物质学科的方法论，在只关心物质结构的物质学科范畴内是非常成功的方法，但却不可能有效支持人工智能的研究。这是因为人工智能的研究必须对问题的形式、内容、价值有全面的理解才能作出合理的决策。

总之，现行人工智能存在的根本性问题是：（1）整体被肢解，无法建立统一的人工智能理论；（2）价值和内容被阉割，导致智能水平低下。前者是人工智能的研究沿用了物质学科“分而治之”方法论的结果；后者则是人工智能的研究沿用了物质学科“单纯形式化”方法论使然。可见，人工智能存在的问题直接触及了哲学的命题：方法论。

其二，深刻的启示：“范式张冠李戴”不可避免。众所熟知，学科的科学观阐明了学科的宏观本质（是什么），方法论规定了研究学科的宏观方法（怎么做）。科学观与方法论的统一体就是范式。学科的研究活动是一类社会存在，学科的范式则是相应的社会意识。物质学科和信息学科各自都会产生自己的范式，并接受各自范式的引领。

正是在这个意义上，托马斯·库恩在《科学革命的结构》一书中把范式理解为判断一个学科是否需要科学革命的根本标志：如果一个学科的范式能够指导这个学科的研究取得成功，说明这个学科的研究处于正常的状态；如果学科的范式已经不再能够指导这个学科的研究取得成功，就说明这个学科需要进行科学革命：颠覆旧的范式，创立新的范式（托马斯·库恩，2012）。

既然人工智能的研究沿用了物质学科的“分而治之”和“单纯形式化”方法论，也就沿用了物质学科的科学观，沿用了物质学科的范式。这便造成了人工智能“范式张冠李戴”：用物质学科的范式引领人工智能的研究（见表1）（钟义信，2018、2020、2021A）。

需要说明的是，表1所列的“信息学科范式”是笔者历经半个多世纪所研究提炼出来的结果。虽然在整个人工智能学术共同体内还没有形成信息学科范式的共识，但由于独特的学术经历，笔者从20世纪60年代起就专门关注了信息学科的科学观与方法论的探索，获得了相应的结果（钟义信，2013A、2007、2014、2017A）。

对比表1三个项目的内容可以明显看出，物质学科范式与信息学科范式几乎是相反而成，对立而统一。人工智能所沿用的范式确实是物质学科的范式，而非信息学科的范式。

那么，人工智能的研究为什么会遭遇“范式张冠李戴”的问题？主要原因是：一方面，虽然人工智能的研究活动已经广泛展开，形成了人工智能研究的社会存在，但由于“社会意识落后于社会存在”法则的制约，作为社会意识的范式至今没有形成，因而，没有信息学科的范式可以遵循；另一方面，数百年来物质学科范式的存在和巨大成功，使人们误以为信息学科的研究也可以借用物质学科的范式。这样两方面原因的叠加，就造成了人工智能的研究沿用物质学科范式的事实。可见，这是历史法则注定了不可避免的学术遭遇。

表1 人工智能“范式张冠李戴”的具体事实

事项	科学观	方法论
物质学科范式	机械唯物论的物质观 研究对象：物质客体，排除主观因素 对象性质：确定性演化，可分性 研究目的：了解对象的结构与功能	机械还原论的方法论 描述与分析方法：纯粹形式化 决策的判断方法：形式的匹配 宏观处置的方法：分而治之
人工智能范式	准“物质观” 研究对象：脑物质，排除主观因素 对象性质：存在不确定性，可分性 研究目的：了解对象的结构与功能	真“还原论” 描述与分析方法：纯粹形式化 决策的判断方法：形式的匹配 宏观处置的方法：分而治之
信息学科范式	整体观：对立统一的信息观 研究对象：主客互动的信息过程 对象性质：不确定性演化，整体性 研究目的：实现主体客体合作双赢	辩证论：信息生态方法论 描述与分析方法：形式内容价值一体化 决策的判断方法：内容理解 宏观处置的方法：生态演化

来源：作者自制

实现人工智能基础理论的重大突破，需要遵循学科发展的普遍规律

作为一个新学科,人工智能当然具有自己的个性,但同样重要的(如果不是更为重要的话)是必须服从学科发展的共性规律。表2用表格的形式给出了学科发展普遍规律的简明解释(钟义信,2020、2021B)。

表2显示,学科的发展需要经历两个相互联系而又相互不同的阶段:首先是自下而上的探索阶段(称为初级阶段),然后才能进入自上而下的建构阶段(称为高级阶段)。

初级阶段的主要任务,是要通过各个相关学科领域的研究人员从各种不同的学术角度展开全面而漫长的摸索、试探、纠错、争论、交流、总结,逐渐提炼出关于学科范式(学科本质是什么?应当怎样研究?)的共识。这是一个极为艰难的探索过程,常常会出现各种各样“盲人摸象”的情景。

高级阶段的主要任务,是要在初级阶段摸索得到的学科范式的引领下,通过形成学科框架(包括学科的全局研究模型和研究路径)以及学科规格(包括学科的学科结构规格和学科基础规格)一步一步地把范式贯彻落实到学科的理论建构之中(包括构建学科的基本概念集合和基本原理集合)。

对照表2所描述的学科发展普遍规律可以看出,当今,人工智能学科的研究确实仍然处在自下而上探索范式的初级阶段。这是因为,今天的人工智能研究仍然处于结构主义、功能主义、行为主义三大学派互不认可(类似于盲人摸象)的状态,信息学科的范式尚未形成。为了实现人工智能基础理论的重大突破,则首先必须努力结束人工智能领域“信息学科范式缺位,物质学科范式错位”的状态。

表2 学科发展的普遍规律

阶段进程	进程名称	进程要素	要素解释
初级阶段: 自下而上的探索	摸索 (广泛试探)	摸索试探	通过长期自下而上的多方试探摸索,总结失败教训和成功经验,提炼和确立学科的研究范式(即学科应当遵循的科学观和方法论)
高级阶段: 自上而下的建构	范式 (宏观定义)	科学观	宏观上明确学科的本质“是什么”
		方法论	宏观上明确学科的研究“怎么做”
	框架 (落实定位)	学科模型	体现“学科范式”的学科全局蓝图
		研究路径	体现“学科范式”的全局研究指南
	规格 (精准定格)	学术结构	体现“学科范式”的学科结构规格
		学术基础	体现“学科范式”的学术基础规格
理论 (完整定论)	基本概念	体现“学科范式”的学科基本概念集合	
	基本原理	体现“学科范式”的学科基本原理集合	

来源:作者自制

表 2 的规律也清楚地表明，学科范式是引领学科发展全过程的力量：在学科探索阶段，根本的任务是为了总结出学科的范式；在学科建构阶段，主要的任务是为了贯彻和落实所总结出来的学科范式。所以，学科范式的引领和规范作用贯彻在学科发展的始终。

人工智能：初级阶段、高级阶段及阶段转变的条件

按照表 2 所示的规律，人工智能的研究必然经历两个发展阶段：首先是自下而上探索范式的（初级）阶段，然后才是自上而下依照范式建构学科的（高级）阶段。初级阶段转变到高级阶段的条件，就是人工智能学科范式的成型与确认。从信息科学技术发展的实际情况来看，情形确实如此（见表 3）。

表3 信息学科发展的初级阶段、高级阶段及其相应范式

阶段名称	阶段特征	沿用范式	科技内容
初级 (1)	单一功能，形式描述	物质学科范式	传感，通信，计算机，控制
初级 (2)	复合功能，形式描述	物质学科范式	互联网，物联网，初等AI
高级阶段	完整功能，整体描述	信息学科范式	高等人工智能

来源：作者自制

从表 3 可见：在信息学科发展的初级阶段，学科对象（传感、通信、计算、控制、互联网、物联网，也包括初级阶段的人工智能）要么是单一功能，要么是部分功能的复合。在这里，学科发展与物质学科范式的沿用尚无很深的矛盾；而在信息学科的高级阶段，学科对象（高级阶段的人工智能）具备了完整的信息功能，物质学科范式不再能够满足需要，因此必须遵循信息学科本身的范式。

为什么在信息学科的高级阶段物质学科范式就不能满足需要了呢？因为在这个阶段，研究对象（高级阶段的人工智能）具备了体现主体意志和目标的高级核心功能：感知、注意、认知、谋行（谋划求解问题所需要的行为策略）与决策等，而物质学科范式只能描述和分析物质客体的功能，没有办法描述、分析、理解和实现主体的这些高级功能。

与此相应，为了描述、分析和实现人工智能的主体意志与目标，人工智能的研究需要采用“整体描述”方法，即不仅要描述问题的“形式”，更要描述问题相对于主体目标而言的“价值”，以及在此基础上定义的“内容”，从而形成“形式、内容和价值三位一体”的整体描述与分析方法。只有这样，才能充分表现人类主体的意志与目标。

这就是为什么信息学科初级发展阶段和高级发展阶段需要不同范式引领的内在根据：在初级阶段，信息科技的研究对象（包括初级阶段的人工智能）基本上都是“客体”；而在高级阶段，信息科技研究的对象（高级阶段的人工智能）具有了主体，于是，研究对象变成了“主体主导的主体与客体的互动信息过程”。研究对象变了，研究的范式就必须随之改变。

由此可知，人工智能的初级阶段向高级阶段转变的“硬性条件”，就是信息学科范式的成型与确立。这也是在信息学科发展进入高级阶段之前，人工智能的研究必须经历范式革命的道理。

初级阶段。对照表 1、表 2 和表 3 可以看出，现阶段国内外的人工智能研究的确还处在初级阶段，还没有找到自己的范式。因此，学术共同体应当继续努力展开全面的探索，以期总结和提炼出人工智能研究所需要的范式。需要说明，初级阶段在物质学科范式统领下所取得的人工智能研究成果仍然具有局部性的积极意义，这是因为，它们可以作为高级阶段人工智能研究的局部性基础。当然，仅凭这些局部性的基础还远远不足以建立具有高水平智能的人工智能统一理论。因此，不应当满足，更不应当停留在初级阶段。

高级阶段。学科的进步不可能是全体研究人员的“齐步走”。那些已经理解和提炼了信息学科范式的研究团队，应当大胆实施人工智能的范式革命，摆脱物质学科范式的束缚，确立信息学科范式的引领地位，进入人工智能学科的高级阶段（人工智能研究领域的无人区），创建人工智能的全新理论。

范式革命，实现基础理论重大突破走向高级阶段的方法与步骤

如上所说，既然学科范式是学科发展的最高支配力量，而初级阶段人工智能的范式又不可避免地处在“范式张冠李戴”的境地，并且成为现行人工智能一切痼疾的总根源，那么，人工智能范式的革命就成为一种医治人工智能一切痼疾的对症良方，也是实现人工智能基础理论重大突破和源头创新的正确举措。

具体来说，所谓人工智能的范式革命，就是一方面要突破物质学科范式对人工智能学科的束缚；另一方面要总结和确立信息学科范式对人工智能学科的引领作用，并以信息学科范式来规范人工智能各个层面的研究。

具体而言，依照表 2 所载明的学科发展普遍规律，人工智能的研究应当从学科的源头和制高点（也就是范式）开始，分五步创建新时代的人工智能理论。

第一步，解除物质学科范式对人工智能研究的约束，确立信息学科范式对人工智能研究的引领，并首先准确阐明人工智能学科的宏观定义。如前所说，人工智能的范式革命，就是要全面突破物质学科范式对人工智能的束缚，确立信息学科范式对人工智能研究的引领和规范作用。之所以要全面突破物质学科范式对人工智能的束缚，是因为：其一，物质学科范式的科学观认为，研究对象应当是客观中性的物质，要求彻底排除一切主观因素的影响。但显而易见，智能

不可能由排除主观因素的孤立脑物质产生（参见狼孩的试验）。其二，物质学科范式的方法论要求，要贯彻分而治之和单纯形式化的方法。但是这样两种方法只能导致人工智能理论的断裂（分而治之割断了子系统之间的联系）和智能水平的低下（单纯形式化使信息、知识、智能全部空心化）。

可见，物质学科的范式确实不适合于人工智能（尤其是高级阶段）的研究，因此托马斯·库恩所说的范式革命无可避免。当然突破物质学科范式对人工智能的束缚只是范式革命的第一步。更重要的是要确立信息学科范式对人工智能的规范与引领。

那么，什么是信息学科的范式？经过笔者半个多世纪对信息学科的科学观及方法论的探究（钟义信，2013A、2007、2014、2017A），总结其基本特征如表1最底一行所示。一是信息学科范式的科学观认为，人工智能的研究不但不应排除主体因素的存在，相反，人工智能的研究应当强调主体的作用，因为人工智能研究的是主体驾驭和环境约束下的主体与客体相互作用的信息过程。二是信息学科范式的方法论要求，人工智能的研究不应当继续沿用机械还原的方法论，而应当遵循信息生态方法论，即在保持信息成分的完整性、时空的系统性和全局优化条件下的信息转换方法。

这样，根据信息学科范式的科学观和方法论，就可以阐明人工智能学科的定义规范，包括学科的宏观本质是什么？学科的研究应当怎么做？人工智能学科的本质，是主体驾驭和环境约束下主体客体相互作用的信息过程，研究的目的是实现主体客体的合作与双赢。人工智能学科的研究，应当遵循信息生态方法论，即在保持信息成分的完整性、时空的系统性和全局优化条件下的信息转换方法。

正确的学科定义，是人工智能学科能够走上健康发展轨道的根本前提。基于此，人工智能的研究就不再局限于模拟人类大脑的结构，也不是去模拟大脑的功能或者人类表现的外部行为，而是要运用信息生态方法论，研究主客互动的信息过程，实现主客双赢的规律，不断改善人类的生存发展水平。

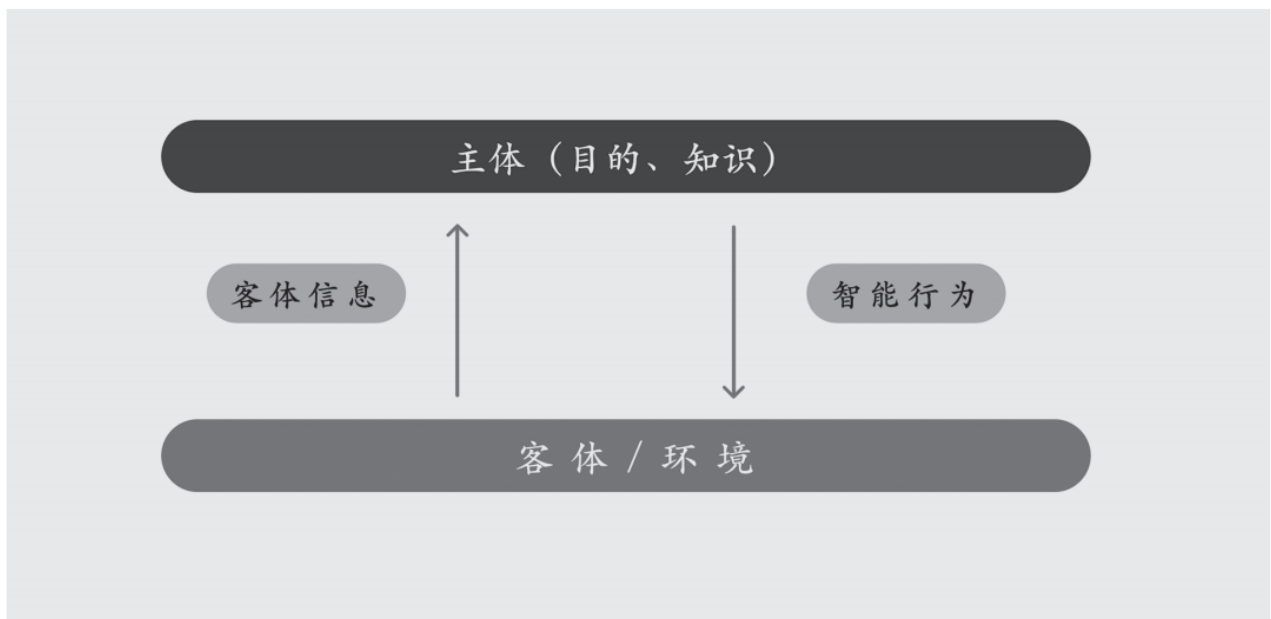
第二步，把学科的宏观定义具体落实成为学科的框架，包括学科的全局研究模型和研究路径。根据信息学科范式科学观所阐明的学科定义，人工智能的全局研究模型应当是主体驾驭和环境约束下的主体与客体相互作用的信息过程模型。图1所表示的，正是符合这个学科宏观定义的人工智能全局研究模型（钟义信，2017B、2021A）。

正如模型所表明的那样，人工智能学科的全局模型必然是，首先，客体信息作用于主体，然后，具有目的和知识的主体设法生成智能行为反作用于客体，完成主体与客体相互作用的基本回合。

如果第一回合的智能行为反作用于客体的结果达到了预设的目标，整个过程就完成了任务；如果没有达到预设目标，其中的误差就要返回到主体的输入端，成为补充性的客体信息，促使主体补充新的知识，优化智能策略和智能行为，改进反作用的效果（缩小误差），直至满意为止。

图 1 所示的人工智能全局研究模型虽然看似简单，却清晰而准确地揭示了人工智能的深层本质：只有主体与客体发生相互作用，才会使具有目的和知识的主体产生智能行为并反作用于环境的客体，实现主体与客体的合作双赢：一方面，主体达到了自己的预设目标；另一方面，客体维护了环境的运行规律。

图1 人工智能的全局模型：主体客体相互作用的信息过程



来源：作者自制

那种既没有外部刺激输入，又没有自身行为向客体输出的孤立“大脑”，实际上不可能产生真正有用的智能。即使孤立的“大脑”非常“聪明”，若没有外部的刺激就没有产生智能行为的激励与动因，因而，也就不会生成智能行为；而没有自身行为的输出，就不可能检验它所产生的行为是否具有解决问题的智能水平。

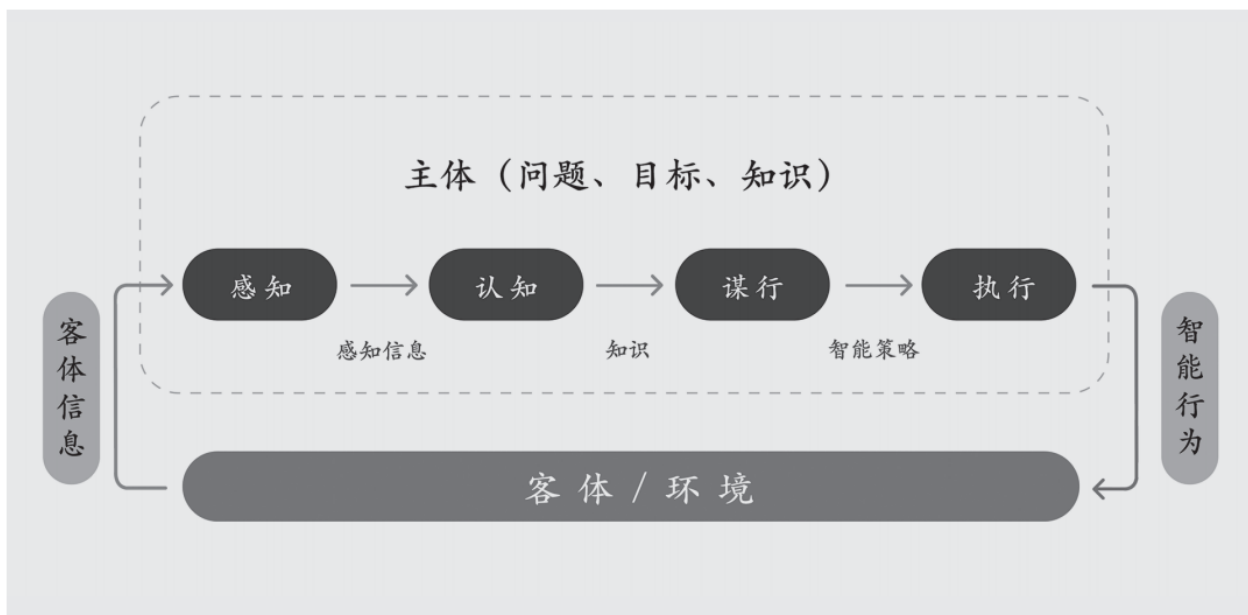
信息学科范式的科学观构筑了正确的研究模型，接下来必须思考的问题就是：应当怎样去研究这个全局模型？根据信息学科范式的方法论，人工智能的研究应当遵循信息生态方法论。

具体来说，就是要遵循“在主客互动框架下保持信息成分的完整性、信息时空的系统性以及在全局优化条件下的信息转换方法”。

进一步的问题是，遵循了信息生态方法论的人工智能系统，将会怎样产生所需要的智能行为呢？事实上，图1的全局研究模型已经表明，人工智能系统产生智能行为的方法只与客体信息、主体目标、主体的知识这些因素有关，即智能行为的生成必定是在主体目标的导引与控制下，在知识的支持与约束下，由客体信息通过复杂转换产生出来。由此，就可以构筑在信息学科范式的信息生态方法论指引下智能行为生成的过程模型，如图2所示（钟义信，2021B）。

显然，图2只是图1的具体化，但图2更加明确地揭示了主体产生智能行为的详细过程和普适性机制。图2表明：首先必须实现主体对客体的感知，从而产生主体的感知信息，这就是感知的作用；接着要把感性的感知信息提升为理性的知识，这是认知的作用；然后就可以在预设目标的引导下，在知识的支持与约束下，把感知信息转换为解决问题的智能策略，这是谋行的作用；进而把智能策略转换为智能行为，这是执行的作用。

图2 普适性智能生成机制的模型



来源：作者自制

图2所揭示的智能生成机制是普适性的，因为其中定义的所有因素（感知、认知、谋行、执行）都是普适性的。图2表明，普适性智能生成机制的本质内涵就是“信息转换与智能创生”原理。若以→表示转换算法，原理就表示为：客体信息→感知信息→知识→智能策略→智能行为。

这样，信息学科范式的信息生态方法论就创造了一种前所未有的人工智能研究路径：基于普适性智能生成机制的机制主义研究路径，它是统一的（普适的）人工智能研究路径。原先的结构主义、功能主义、行为主义研究路径乃是这种统一的人工智能机制主义研究路径在相应条件下的和谐特例。

事实上，对于人工智能基础理论研究来说，普适性智能生成机制才是人工智能的核心本质所在：系统的结构和功能都是为系统实现智能生成机制服务的。至于系统的行为，则是普适性智能生成机制实现以后的外部表现。

所以，这是人工智能理论的一个具有重大意义的研究结果。它给人工智能学科的研究大方向提供了一个极为重要的启示：普适性智能生成机制是一类复杂的信息转换，因此，要把人工智能研究的整体思路聚焦到“信息转换”这个大方向上来。只要抓住了“信息转换与智能创生”这个普适性智能生成机制，就抓住了人工智能研究的本质和全局。算法、算力、数据、知识都是为实现“信息转换与智能创生”这个普适性智能生成机制服务的。

不过，需要特别提醒：作为普适性智能生成机制的本质，“信息转换与智能创生原理”中的源头——信息，并不是现在广泛流行的 Shannon 信息，而是感知信息，也就是全信息。如果不了解这一点，普适性智能生成机制仍然不可能生成真正的智能，这是因为，Shannon 信息是被单纯形式化方法“阉割”了内容和价值因素而只剩下形式因素的空心化信息。利用这样的空心化信息，普适性智能生成机制也只能提炼出空心化知识，生成空心化的智能。

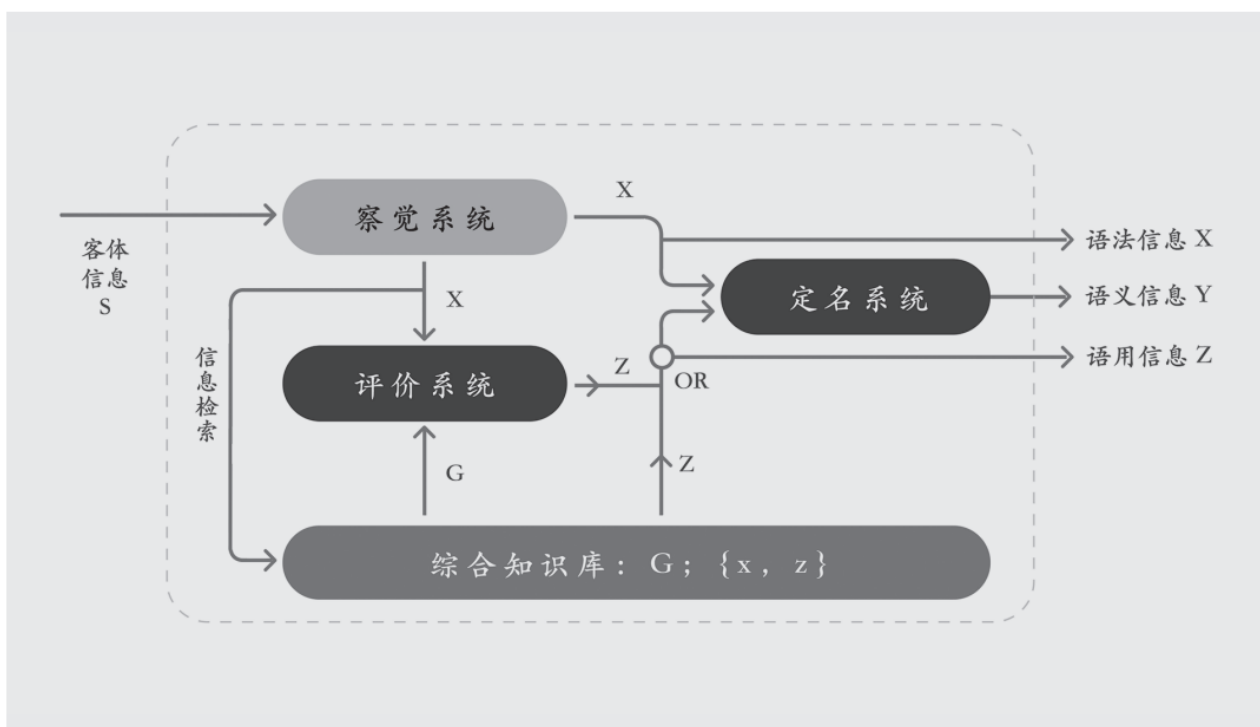
感知信息具有形式信息（语法信息）、价值信息（语用信息）、内容信息（语义信息）三个分量，这是因为感知过程具有主体目标。于是值得特别强调：在人工智能系统中，必须十分重视主体目标的作用，它是体现主体主观能动作用和主体利益的最高标志性因素。只有那些有主体目标的人工智能系统，才会拥有真正的智能，没有主体目标的“人工智能”系统则不可能拥有真正的智能。这是研究人工智能理论与研究一般物质系统理论之间的最大区别之一。这也是人工智能理论的研究不宜完全简单还原到物质系统的研究的一个重要原因。

在以上说明基础上，就可以构造“由客体信息转换生成感知信息”的工作机制模型，具体情况如图 3 所示（钟义信，2021A、2021B）。图 3 模型中，驾驭全局的主体因素就是存储在“综合知识库”的目标 G。由它来确定外来刺激（客体信息 S）对于达成系统的目标究竟是有利？有害？还是无关？也就是确定这个客体对于系统的价值信息（语用信息 Z）作用为正？为负？还是为零？从而据此确定系统究竟应当对这个客体表示欢迎（支持）？反对（抵制）？还是不予理会（过滤）？

可见，并非一切外部刺激或外来数据都能启动智能生成机制的工作。只有那些与系统主体目标有关（正相关或负相关）的外部刺激才能启动系统智能生成机制的工作：是主体的目标把握着“系统注意力”的大关。

图3模型中的“察觉系统”就是传感系统，它能够察觉和表达系统面临着什么形式的刺激X（形式信息 / 语法信息），但它不懂得刺激的内容；模型中的“评价系统”就是主体对刺激的价值评估：这个刺激对于达成主体目标有什么利害关系Z（价值信息 / 语用信息）；其中的“定义系统”就是把形式信息X和价值信息Z这两者形成的“偶对”映射到内容（语义）空间并对映射结果进行恰当的命名，这样就定义了主体关于客体的内涵Y（内容信息 / 语义信息）。不难看出，图3所示的模型是科学合理的，在技术上是完全可实现的。

图3 感知信息生成机制的模型



来源：作者自制

以前，人们对语法信息、语义信息、语用信息三者的定义和关系曾经存在很多误解。然而，感知信息生成机制的模型表明，感知信息的形式信息X（语法信息）、价值信息Z（语用信息）和内容信息Y（语义信息）三者之间存在的关系是 $Y = (X, Z)$ （其中符号表示“映射与命名”操作）。内容，乃是形式和价值的抽象与命名。这个表达式“ $Y = (X, Z)$ ”具有非常重要的意义，因为它准确回答了：对于人工智能的研究来说，究竟什么是内容？

具体来说，所谓“某个事物的内容”，实质就是指：（1）这个事物具有什么样的形态；（2）具有这样形态的事物对主体的目标而言具有什么样的价值关系。了解了事物的形态和价值，就是对事物的内容实现了理解。

根据这样定义的“内容”，决策者就可以作出明确的决策：欢迎？反对？不予理会？所以，这是一个非常重要也非常有用的结果，而且是一个至今都不曾被准确定义，甚至被普遍误解了的重要结果。由这样生成的全信息（感知信息）按照普适性智能生成机制生成的知识和智能就将是同样具备形式、价值、内容三个分量的“全知识”和“全智能”（钟义信，2021A、2021B、2013B）。

第三步，对学科框架的学术规格予以精准化，包括对学科模型的学术结构规格、研究路径所要求的学术基础规格都予以精准界定。根据信息学科范式的科学观，人工智能学科的全局研究模型是“主体客体相互作用的信息过程模型”，那么，为了精准研究这种主客互动的信息过程，人工智能学科的学术结构规格也应当精准化。结果可以发现：人工智能学科的学术结构应当是由人工智能的原型学科（神经科学、认知科学、人文科学等）、核心学科（信息科学、系统科学、控制科学等）和基础学科（数学、逻辑学、哲学等）所形成的交叉学科群结构；而不应当仅仅是计算机学科。同时，根据信息学科范式的方法论要求，人工智能学科的方法论应当是信息生态方法论。于是，为了描述主客互动框架下的信息生态过程——以“信息转换与智能创生原理”为标志的普适智能生长机制（机制主义研究路径），人工智能学科所要求的学术基础（数学、逻辑、哲学）的规格就应当满足“内涵的整体性、时空的系统性、全局的优化性”的数学理论（汪培庄，2021）、逻辑理论（何华灿，2021）以及新型的信息哲学理论（邬焜，2005），不应当仅仅是传统的概率理论、刚性化的数理形式逻辑和传统的哲学观念。

第四步，把学科的定义、框架、规格贯彻到学科的具体理论，包括深度改造人工智能理论的基本概念和深度挖掘学科的基本原理。现有的人工智能理论是在物质学科范式约束下形成的，不适用于信息学科的范式。因此，深度改造人工智能的基本概念和基本原理是极为必要的工作。比如，物质学科范式的科学观认为研究对象是物质客体，不允许主观因素的参与。这就导致现有人工智能的基本概念（如数据、知识、智能等）都是纯粹客观的、绝对中性的、不反映主体因素的概念，同时导致现有人工智能的工作原理（形式计算、统计、形式逻辑等）也都是纯粹客观的、绝对中立的、不能反映主体因素的演绎原理。又比如，物质学科范式的方法论是分而治之和单纯形式化，因此，现行人工智能的基本概念（数据、知识、智能等）都是被“阉割”

了内容与价值因素的形式化概念，而现行人工智能的基本方法（纯形式的计算、统计、数理逻辑等）也都是被“阉割”了内容与价值因素的分析方法。

显然，如果把现有人工智能的基本概念和基本原理原封不动地照搬到新的人工智能理论中来：一方面，学科的定义（学科的科学观和方法论）、学科的定位（学科的全局模型和研究路径）、学科的定格（学科的学术结构和学术基础）都符合信息学科范式的要求；另一方面，学科理论的基本概念和基本原理却还是沿用物质学科的范式。那就意味着，信息学科范式会被“截去了双腿”。这种“截去了双腿”的人工智能范式革命，绝不是成功革命。

根据信息学科范式的科学观，人工智能全局研究模型是“在主体驾驭和环境约束下主体与客体相互作用的信息过程”，那么，为了研究主体与客体互动的信息过程，使主体能够从形式、价值、内容上全面了解客体，人工智能学科的基本概念也就应当是形式、价值、内容三位一体的全信息、全知识、全智能等基本概念。其中，价值和内容便是信息学科范式科学观所要求的主体因素的具体体现，而信息、知识、智能之间的复杂转换则是信息学科范式的信息生态方法论的体现。

近年来，从实际操作的角度，人们把纯粹形式化的“数据”看作是人工智能的“粮食”。这看似有一定的道理。不过，深入研究就会认识到，数据只是信息的载体和外壳，不含信息的形式数据根本不能启动人工智能系统的有效工作。只有携带了与系统目标相关（无论是正相关还是负相关）的信息（即感知信息的价值信息分量不为零），才能真正启动人工智能系统的工作。所以，对于人工智能的研究来说，真正重要的概念是全信息（感知信息），而不是纯形式的数据。同时，根据信息学科范式的方法论，人工智能学科的方法论应当是信息生态方法论，那么，为了满足“整体性、系统性和全局优化性”，人工智能学科的基本原理就应当是符合信息生态规律的一系列信息转换并最终创生智能的那些原理，它们构成了“信息转换与智能创生”定律。

颇有意义的是，信息学科领域的“信息转换与智能创生定律”，与物质学科领域的“质量转换与物质不灭定律”以及能量学科领域的“能量转换与能量守恒定律”共同构成完备的科学定律体系：物质领域和能量领域的定律分别告诫人们应当遵守“物质不灭”和“能量守恒”，而信息领域的定律则启示人们应当如何具体“创生智能”。

第五步，综合学科的定义、框架、规格、理论，最终形成机制主义的通用人工智能理论这一全面体现信息学科范式的全新人工智能理论。把人工智能范式革命及其全部链锁反应的结果综合在一起，就可以产生人工智能范式革命所创造的全新一代人工智能理论——基于普适性智

能生成机制的通用人工智能理论，简称“机制主义通用人工智能理论”。具体来说，只要按照表2所表示的“学科发展的普遍规律”，我们就可以在图2所示的普适性智能生成机制模型的基础上，生成整个机制主义通用人工智能理论的系统模型，如图4所示(钟义信,2017B、2021A)。

不难看出，图4仍是“主体驾驭和环境约束下主体与客体相互作用所产生的信息过程模型”，其中的客体表示为模型最底端的椭圆，模型的其它部分是充分展开了的主体功能群，它完美地包容了图2的普适性智能生成机制。

机制主义通用人工智能理论系统模型图4的具体工作过程包括：(1)把客体信息转换为感知信息，“注意”的功能就在感知基础上利用感知信息来实现(图4的“感知—注意”子系统)；(2)把感知信息转换为知识(图4的“认知”子系统)；(3)在目的引导下，在知识约束下，把感知信息转换为智能策略(图4上部虚线所围的部分，称为“谋行”子系统)；(4)把与问题求解相关的各种信息、各种先验知识和各种先验策略随时进行存储与提取(图4的“综合知识库”子系统)；(5)把智能策略转换为智能行为(图4的“策略—执行”子系统)；(6)检验智能行为反作用于客体的行为实效(图4的“检验”子系统)；(7)如果检验的结果存在误差，就把误差作为补充性的客体信息反馈到整个系统输入端(如图4所示)；(8)根据误差信息，从综合知识库提取新的相关知识(在图4的“综合知识库”完成)；(9)利用新提取的补充知识，按照步骤(3)的方法改善智能策略和智能行为，从而改善智能行为的效果(仍是图4的“谋行”子系统)；(10)把满足要求的智能策略存入综合知识库(图4的“综合知识库”子系统)，从而体现了机制主义通用人工智能系统在解决问题的过程中不断通过学习和不断增加自身解决问题的能力自学习机制。由图4可见，“普适性智能生成机制(感知→认知→谋行→执行)”就和谐有机地融合在整个机制主义通用人工智能系统之中。

机制主义通用人工智能理论的简要自评

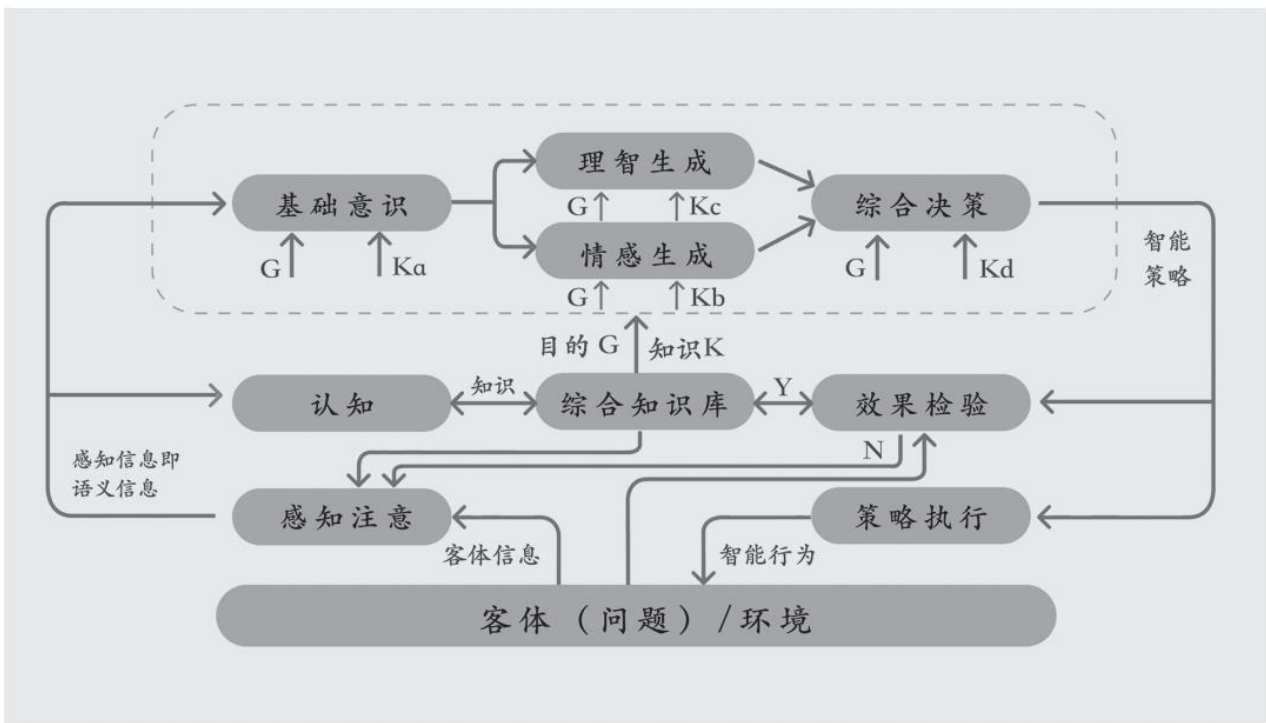
作为一项完整的人工智能基础理论研究工作，应当对研究的结果作出必要的分析和评判，但关于机制主义通用人工智能理论的全面评价，必须等待这一理论转化为实用的机制主义通用人工智能系统，并在若干有代表性的真实环境下经过比较充分的应用后才能正式开展。

不过，作为理论性的研究工作成果，同时也作为任何研究工作的一个必要的环节，研究者自己也有责任对研究结果作出公正的自评，这里就对上文所阐述的机制主义通用人工智能理论的生成思路、机制和原理作一些理论性的初步判断，如表4“机制主义通用人工智能理论与现行人工智能理论的定性对比”所示供读者分析、参考与批评。

此外，需要指出，所谓通用人工智能，并不是指用一个单体的“巨无霸”式的人工智能系统解决世间所有的问题。这里所说的通用人工智能理论，是指以不变的（普适性的）智能生成机制去成功应对千变万化的实际问题。换言之，通用是指机制的通用，而通用人工智能系统的输入内容（问题、目标、知识）和输出内容（智能策略和智能行为）则将随着问题的改变而相应改变，这就是“以不变应万变”的真实含义。

具体来说，给定任何合理的“求解的问题、求解的目标、求解问题所需要的先验知识”（所谓“合理”，是指：所给定的“求解问题”至少在理论上确实存在着“求解目标”，或者说给定的“求解目标”一定存在），通用人工智能理论 都可以凭借它的普适性的（不变的）智能生成机制，在给定的目标引导下，利用给定的知识（在先验知识不足的情况下也可以通过学习来扩展知识），生成能够“求解问题”达到“求解目标”的智能策略和智能行为。当然，如果求解的问题改变了（这意味着求解目标和所需要的知识也会相应改变），解决问题的策略和智能行为也会随之改变，但生成这种智能策略和智能行为的机制却不需要改变，依然遵循“信息转换与智能创生定律”。

图4 机制主义通用人工智能理论的系统模型



来源：作者自制

换句话说，“普适性的智能生成机制”是人工智能理论研究中的不变性和不变核，是人工智能理论之所以能够成为“通用人工智能理论”的根本基础。这样，面对千变万化的问题，不

再必须对每个不同的问题（场景）从头做起来设计专用系统，只需要把各自的“问题、目标、知识”表达成一定的形式，利用同样的普适性的智能生成机制就可以生成解决相关问题的智能策略和智能行为，并达到“求解问题”的目的。其实，通用人工智能系统就是“各种人工智能系统的通用孵化平台”。通用人工智能理论的“机制通用性和理论统一性”，是人工智能范式革命带来的巨大优越性。作为对比，现行人工智能理论无法望其项背。

同样十分重要的问题是，由于现行人工智能沿用了“单纯形式化”的物质学科范式方法论，使得现行人工智能系统的“智能水平十分低下”，成为现行人工智能理论的又一项痼疾。现行人工智能系统表现出来的“智能”，其实都不是基于对问题的理解所实现的，而是利用快速运算变幻出来的结果。这样设计出来的“智能”严格依赖特定的应用场景，场景一变，该种”智能“便会失效。

从通用人工智能理论的图3和图4模型则可看出，客体信息经过感知被转换成为了感知信息。如前所说，感知信息是形式信息（语法信息）、价值信息（语用信息）、内容信息（语义信息）的三位一体；知识是形式知识、价值知识、内容知识的三位一体；智能策略是形式策略、价值策略、内容策略的三位一体。因此，信息、知识、智能策略都是可以理解的。

表4 机制主义通用人工智能理论与现行人工智能理论的定性对比

对比事项	现有人工智能理论		机制主义通用人工智能理论
研究模型	以孤立的脑为原型	范式革命带来的优越性	主体客体互动的信息过程模型
研究路径	结构、功能、行为		机制主义
决策依据	基于形式的对比		基于内容的理解
理解能力	差		好
通用能力	差		好
智能水平	差		好
可解释性	差		好
小样本性	差		好
总体比较	差		好

来源：作者自制

当然，由于人工智能综合知识库所存储的信息、知识、策略毕竟有限，因此它的理解能力也有限。人们不应当完全按照“人类的理解能力水平”来要求“人工智能机器的理解能力水平”，不应当要求人工智能系统完全像人类那样能够洞若观火灵活变通或者举一反三、触类旁通。但是，人工智能的“理解能力”至少能使人工智能机器作出具有理解基础的决策，并且具有一定的自学习能力。比如，在信息的层面上，通用人工智能系统可以根据感知信息的价值信息的分量大小与性质对客体作出基于理解的明智决策：（1）如果价值信息为足够大的正值，就采取相应的措施利用客体；（2）如果价值信息为足够大的负值，就采取相应的措施限制客

体；（3）如果价值信息为足够小的正值或负值，就不理会或过滤相应客体；（4）在“多中选一”的情况下，就选择其中价值信息为最大正值的客体，或者选择其中价值信息为最小负值的客体。

在知识和智能策略的层面，也同样可以针对所面临的问题作出与信息层面类似的具有一定智能水平的明智决策。而且，无论在信息、知识，还是智能策略的层面，这样作出的决策，都可以清晰而准确地得到解释。这就解决了人工智能系统的可信任性问题。

人工智能范式革命所产生的机制主义通用人工智能理论的另一巨大优越性，是它的理解能力和在理解基础上的智能决策水平。与通用人工智能理论的智能水平相比，现行人工智能理论的理解水平确实相差甚远。这是人工智能范式革命所带来的人工智能理论的深刻革命。

表5是对上述论证的简明总结（从科学观、方法论到基本概念、基本原理）。值得指出，表5显示了在人工智能范式革命之前，“三分天下”的现行人工智能理论是物质学科主导的科学时代所产生的人工智能理论；而在实施人工智能范式革命之后，“机制主义通用人工智能理论”则是信息学科主导的科学新时代所创造的人工智能理论。人工智能的范式革命，是这两个科学时代人工智能理论的分水岭。

正是在这个意义上，人工智能范式革命乃是推动物质学科主导的科学时代转变到信息学科主导的科学新时代的引擎和桥梁，也是划分物质学科主导的科学时代与信息学科主导的科学新时代

的界限和分水岭。这样，人工智能范式革命的产物——基于普适性智能生成机制的通用人工智能理论就名副其实地具有了划时代的意义。这就是人工智能范式革命及其连锁反应所带来的人工智能基础理论的具有划时代意义的重大突破和源头创新。

表5 人工智能范式革命前后的学科范式比较

对比项目	物质学科范式下的人工智能		信息学科范式下的通用人工智能
科学观	机械唯物的物质观：非主观，结构	两个科学时代的分水岭	对立统一的信息观：主客互动，双赢
方法论	机械还原论：纯形式，分而治之		信息生态学：整体化，生态演化
全局模型	孤立的脑模型		主体驾驭的主客互动的信息过程模型
研究路径	结构、功能、行为模拟分道扬镳		基于智能生成机制的机制主义路径
学术结构	计算机学科		原型、核心、基础等多学科交汇
基础特色	概率论、形式逻辑		因素空间数学理论、泛逻辑理论
基本概念	形式数据、形式知识、形式智能		全信息、全知识、全智能
基本原理	未作总结		信息转换与智能创生定律
综合结果	三个独立的局部理论		机制主义通用人工智能理论

来源：作者自制

结语

当前，我们已经处在学科范式革命（由物质学科范式主导转变为由信息学科范式主导）的伟大时代。具体来说，20世纪中叶以来，信息学科迅猛崛起，成为了一种社会存在。只是受制于“社会意识滞后于社会存在”法则，信息学科的社会意识（范式）至今尚未形成。在此期间，作为初级发展阶段的信息学科研究便借用了物质学科的范式，造成了初级发展阶段信息学科“范式张冠李戴”的结果。然而，社会发展的需要使信息学科不可能永远停留在初级发展阶段，必然要冲破各种束缚向高级发展阶段迈进。这就必然迫使信息学科形成并确立信息学科范式来引领自己的高级发展阶段，而把物质学科范式归还给物质学科（理顺学科范式）。这就是信息学科（含人工智能）的范式革命。

信息学科范式革命的主要标志是（1）科学观的大转变：信息学科高级阶段的研究对象不再局限于物质客体，而是主体驾驭和环境约束下的主体客体相互作用的信息过程；研究的目的也不再是认识物质的结构与功能，而是实现主体客体在互动中双赢；（2）方法论的大转变：信息学科高级发展阶段的方法论不再是“分而治之”和“形式演绎”，而是转变为“整体而治”（而不是分而治之）和“生态演化”（而不是单纯的形式演绎）。可见，信息学科范式与物质学科范式之间大不相同，它们之间是“辩证的统一”。因此，不能混为一谈。

为了能够理解信息学科的范式革命，我们应当自觉地认识并处理好以下几种重要的关系。

科学-哲学。不能忽视哲学对科学的指导。仅在科学范围内部无法发现信息学科范式这个至高无上的指南，更不可能注意到人工智能范式革命这个本质命题；而不触及范式革命的人工智能研究是不深刻不彻底的研究。学科的划分，使科学界往往不再关注哲学的作用。但只有打破科学与哲学之间的藩篱，形成科学与哲学的联盟，才能认识到学科范式的制高地位，并在正确的学科范式引领下揭示复杂而深刻的规律，开拓科学研究前沿的“无人区”，创建全新的科学理论。在人工智能的初级发展阶段，在物质学科范式的束缚下，人工智能的研究也取得了许多进展。但是在物质科学观和分而治之方法论束缚下，那都只能是局部性和碎片式的进展，不可能取得全局性通用性的进展；而且在物质科学观和形式演绎方法论束缚下，信息、知识、智能的价值因素和内容因素都被彻底掏空，因而，不可能获得真正的智能。

信息科学（特指高级阶段人工智能）-物理科学。不能把人工智能的研究还原为物理学的研究。因为，前者的研究对象是主体与客体复杂互动的信息过程，后者的研究对象是纯粹的物质客体而不可能关注主客相互作用的信息过程。一个困扰很多人的问题是：人类主体和机器系统从外部世界接受到的输入同样都是各种事物的形式，为什么人类能够从这些形式之中加工出

相应的价值和内容,而现有的机器(包括现有的人工智能机器)却无法从中加工出价值和内容?这是因为,作为万物之灵的人类是具有明确目的的生物物种,因此,各种事物的形式在他的心目中就必然产生价值的判断和反应,因而,也就有了内容的提炼。加之,人类的活动是人与环境之间相互作用的开放过程,于是他就能够不断地在实践中检验各种形式的价值。信息学科范式尊重了人类的这种独特品格,把具有明确目的的人类主体及其与环境的相互作用看作自己的研究对象。然而,物质学科范式则仅仅把物质客体的结构与功能作为自己的研究对象,彻底地排除了主体的一切因素,因而,也就排除了价值和内容的立足点,彻底否决了从形式加工出价值和内容的可能性。可见,没有目的的系统,不可能有智能。由此可知,完全用物理学的观点看待人工智能(主体驾驭与环境约束下的主客相互作用的信息过程)的研究,或者把人工智能的研究还原(退化)为物理学的研究,都是不适宜的。

信息理论-高级阶段人工智能。不能把国内外唯一公认的 Shannon 信息理论当作研究人工智能理论的全部信息理论基础。因为,它完全抛弃了人工智能的主体因素。虽然人工智能把“主体驾驭和环境约束下的主体客体互动的信息过程”作为自己的研究对象,但这丝毫也不表明,现有的信息理论就可以完全满足人工智能研究的需要。恰恰相反,由于现有的信息理论(主要的代表是 Shannon 于 1948 年建立的信息论——通信的数学理论)是在物质学科范式束缚下产生的,只关注了信息的形式因素而完全忽视(掏空)了信息的价值因素和内容因素的纯粹形式化信息理论,因此,无法支持人工智能理论所需要的信息过程研究。解决这一问题的办法是,只有把现有纯粹形式化的 Shannon 信息理论改造提升成为“形式信息(也称语法信息)、价值信息(也称语用信息)和内容信息(也称语义信息)三位一体的全信息(也称感知信息)理论”,才能满足人工智能理论研究的需要。

高级阶段人工智能-计算机与形式逻辑。不宜把人工智能看作是计算机科学和形式逻辑理论的应用分支。因为,计算机和形式逻辑也是在物质学科范式引领下建立的形式化演绎系统。计算机是强大的信息处理系统,形式逻辑是强大的演绎系统。但是,它们所能够处理和演绎的“信息”必定是纯粹形式化的信息(0与1系列),而高级阶段人工智能所需要的价值信息和内容信息的处理与演绎,已经超出计算机和形式逻辑的能力范围。正如人们所说:计算是计算机的本领,算计是(高级阶段)人工智能的特殊本领。算计与计算的差别就在于:前者需要形式因素、价值因素和内容因素的支持,后者只需要形式因素的支持。因此,要想使计算机和形式逻辑理论能够支持高级阶段人工智能的研究,则两者都需要变革与创新,而不能在原有的范

式下直接实现。在这个意义上，广泛流行的“计算思维”只适用于计算机，而不适用于高级人工智能的研究。

高级阶段人工智能-数学。形式化的数学如何适应人工智能对“内容”研究的需求。这可能是数学科学面临的创新任务。一方面，和其他科学技术一样，高级阶段人工智能理论研究同样需要数学的鼎力支持。不仅如此，数学在人工智能理论研究中还发挥着特别重要的作用。另一方面，由于高级人工智能理论的科学观是主体驾驭与环境约束下的主客相互作用的信息过程，它的方法论是“整体而治”（而不是分而治之）和“生态演化”（而不是单纯的形式演绎）。因此，它对数学提出了“如何描述和分析事物的价值因素和内容因素”的要求。这便蕴含着“形式化科学”向“内容科学”发展的巨大变革。这与“人工智能范式革命”遥相呼应。

总之，学科的范式是学科的核心灵魂。人工智能范式的革命不仅是人工智能从初级阶段走向高级阶段的阶梯，而且是物质学科主导的科学时代通向信息学科主导的科学新时代的桥梁。这是 21 世纪必定要掀起的信息学科(含人工智能)以及复杂科学研究发展的波澜壮阔的未来！

（本文系国家自然科学基金项目“机制主义方法与高等智能理论”和国家社科基金重大项目“信息哲学的历史、现状与未来”的研究成果，项目编号分别为 60873001、18ZDA027。本文的研究有幸得到了陆汝钤院士、李衍达院士、陆建华院士、谭铁牛院士、蒲慕明院士、涂序彦教授、何华灿教授、汪培庄教授、史忠植教授、韩力群教授、邬焜教授、王小捷教授、周延泉副教授、陈志成博士、李蕾副教授、李睿凡副教授、孙健博士，以及中科华数信息科技有限公司卢建新院长、徐亭院长、张世光副院长、王丁桃秘书长等各种形式的帮助和支持，在此谨表衷心的感谢！）



敬请关注联盟微信公众号
COPU开源联盟



扫描二维码
获取往期资料

中国开源软件推进联盟秘书处

电话: +86 010-88558999

联盟公共邮箱: office@copu.org.cn

联盟官网: <http://www.copu.org.cn>

地址: 北京市海淀区紫竹院路66号赛迪大厦18层