

预测加工:走出两代认知科学的历史僵局

董 达 陈 巍

【摘 要】表征-计算观与具身行动观对认知的本质几乎做了截然相反的强调。近年来,预测加工理论的发展为统一两代认知科学提供了契机。预测加工是层级预测加工与主动预测加工这两大理论部件的合称,前一部件主要继承了第一代认知科学中的层级计算加工进路,后一部件则发扬了第二代认知科学中与行动有关的理论,这两大理论部件被视为同一个统一整合理论的两个不同方面。在当代,预测加工被认为有望成为未来认知科学的新范式。

【关 键 词】预测加工;层级预测编码;自由能原理;主动推断

【作者简介】董达,绍兴文理学院大脑、心智与教育研究中心(绍兴 312000),浙江大学哲学系,浙江大学语言与认知研究中心(杭州 310007);陈巍(通讯作者),绍兴文理学院大脑、心智与教育研究中心,E-mail:anti-monist@163.com(绍兴 312000)。

【原文出处】《心理科学》(沪),2022.1.235~241

【基金项目】本研究得到国家社科基金青年项目(20CZX015)、浙江省哲学社会科学规划新兴交叉学科重大扶持项目(21XXJC05ZD)、中国博士后科学基金特别资助(2019T120530)和中国博士后科学基金面上资助(2019M652134)的资助。

1 引言

两代认知科学之间以往数十年的斗争不断割裂了两派理论的历史渊源。以 Marr 与 Gibson 之间的理论关联为典型,在当代,Marr(1982)的遗著《视觉:人类视觉信息的表征和加工的计算研究》(以下简称《视觉》)被认为是第一代认知科学或标准认知科学的里程碑著作,而 Gibson 生态知觉理论被奉为以 4E 认知(具身认知、生成认知、延展认知、嵌入认知)为标志的第二代认知科学的重要理论源头之一(Shapiro,2019)。不过在《视觉》一书中,Marr 认为正是 Gibson(1966)提出的视觉理论(较之其他理论)最接近 Marr 本人的理念。站在 Marr 的立场上,他认为 Gibson 理论存在两点重要不足:第一,Gibson 未能充分认识到对环境光阵不变量(invariants)的视觉检测归根到底是信息加工问题;第二,Gibson 被看(seeing)这一行为的直接性误导,从而低估了三维知觉对象认知加工的困难程度。最终,Marr 提出了其三层次视觉计算理论框架。可以公允地说,

尽管二人存在基本理念上的轻重取舍,Gibson 对 Marr 存在显而易见的重要启发。

进入 21 世纪,两代认知科学之间的斗争似乎陷入了无可挽回的历史僵局。第一代认知科学提倡表征概念,表征被界定为对外界对象的直接的、被动的重现(re-presentation)。第一代认知科学的理论内核可概括为:表征-计算观(李恒威,黄华新,2006)。第二代认知科学将重心转移到认知主体与其周围环境的具身动力耦合,猛烈抨击基于表征的自上而下加工,转而强调基于身体和行动的自下而上加工。第二代认知科学的理论内核可概括为:具身行动观(Varela et al.,1991)。Gibson 生态学进路三原则被激进具身认知主义(Chemero,2009)等 4E 认知派别擎举为“正确的”反表征-反计算纲领:知觉是直接的(不涉及表征和计算),由行动引导,由动允性(affordances)构成。如此,Gibson 派别与 Marr 派别彻底对立。

如上,将 Gibson 与 Marr 之间的理论渊源彻底

割裂并推向对立面是两派斗争的一个历史缩影,但却并非是合乎全部历史事实的理性选择。心智的本性是什么?表征-计算观与具身行动观各自为阵,似乎成为了非此即彼的选择。

近年来,一些学者尝试整合两派理论资源,以期提出一种统一的理论框架,以打破两代认知科学的历史僵局。这一框架简称为“预测加工”(predictive processing),它表面上直接发展自1980年代以来视听觉的层级预测编码(hierarchical predictive coding)(吕雪靖,侯欣,2019;Rao & Ballard,1999),之后由神经科学家Friston(2005,2010)等人实质性地改造为由行动引导的(action-oriented)理论部件,至此理论框架初步定型。预测加工实际上是“层级预测加工”与“主动(active)预测加工”这两大理论部件的合称(Clark,2013)。层级预测加工主要继承了第一代认知科学中的层级计算加工进路(如Marr多层次视觉计算理论、层级神经网络等),主动预测加工则发扬了第二代认知科学中与行动有关的进路(如Gibson动允性理论、生成认知等)。这两大理论部件被视为同一个统一整合理论的两个不同方面。

本文认为,预测加工承担了融合表征-计算观与具身行动观这两大分裂派别的迫切时代任务(在本文中以Marr与Gibson的理论融合尝试为线索),在Hohwy(2013)、Clark(2016)等认知科学哲学家的着力推动下,被认为有望成为未来认知科学的新范式(Metzinger & Wiese,2017)。

2 层级预测加工

2.1 自由能原理:适应的最优解

认知科学的预测加工思潮体现了两代认知科学纲领的融合态势。这意味着,原本在一方被刻意忽视的理论部件得到重新发掘和重视。在表征-计算观中,Marr进路的基本理念包括!(视)知觉以世界对信息加工路径的自然约束(constraints)为前提,其目标是从可视的外部环境中获得内部表征,神经机制问题并非根本问题(而是背后的算法问题),视觉计算是多层级加工的(Warren,2012)。其中,Marr(1982)概用“约束”概念代指全局环境对信息加工的实际约束。在第一代认知科学中,约束作为无关信息加工过程本身的前置概念往往被忽

略了。

Teufel和Fletcher(2020)重新调用了Marr的约束概念,并在预测加工背景下赋予其更大的理论权重。他们认为,神经网络中存在两种不同的预测形式:自下而上的(情景独立)约束与自上而下的(情景依赖)期望(expectation)。约束式预测指神经生物学机制独立于认知加工经验内容的、先验的既定路径。该预测形式作为遗传和形态发生的结果无法基于感知输入得到后验修正。期望式预测指认知系统依赖已有经验对感知输入的自上而下加工。该预测形式由脑的抽象贝叶斯模型决定,可根据感知输入得到及时修正。这两种预测加工方式实际上同时发生作用。这样,Teufel和Fletcher把Marr处原本笼统的前置约束概念具身化,赋予其明确的神经生物学机制内涵。

约束概念的再发现有助于弥合两派之间的裂隙。预测加工的约束形式重新肯定了具身的神经生物学机制对认知的重要性。在第二代认知科学中,Maturana和Varela(1980)开辟的自创生式(auto-poietic)生成认知始终强调生命与心智之间的深刻连续性。Thompson(2007)指出认知活动的基本特征更适合被理解为生命自组织特征的更丰富形式,认知活动首先是生命活动,不能被简单还原为脑中的神经过程或者抽象的信息加工过程。如果约束式预测被接受为预测加工的基本形式,那么在一般系统论的意义上,脑与认知系统的这一可预测性(predictivity)是环境中的生命作为适应系统的必然结果,由其本身的系统属性决定。

Friston的自由能原理(free energy principle)以系统论和控制论中的良好调节器定理(good regulator theorem)等为理论基础,尝试给出适应问题的最优解(Teufel & Fletcher,2020)。根据良好调节器定理,每一个足够好的调节器(regulator)必定是其所调节系统的一个模型(Conant & Ashby,1970)。Friston进而称,每一个足够好的“脑”是关于认知主体所适应周围环境的一个足够好的模型(Friston & Buzsáki,2016)。简言之,就环境对脑(调节器)的每一刺激(S)而言,脑至少存在一个正确适应性反应(R)与之对应,当刺激-反应趋近于一一对应关系($s_1 - r_1, s_2 - r_2, s_3 - r_3, \dots$)

时,两者之间存在理想的同构(isomorphism)映射关系,于是脑被视为其所适应环境的一个同构模型。反之,当脑对某一环境刺激无法给出正确适应性反应时,严重时可能导致认知主体因无法适应环境而死亡。

自由能原理将适应性(adaptivity)视为脑与认知系统以至一般自组织系统理论建构的最基本原则(Friston,2010)。层级预测加工的整体目标是使认知个体尽可能适应环境,维持自身。Kirchhoff(2018)认为自由能原理不仅与自创生(以及相应生成认知版本)相兼容,而且出于理论建构的统一性和简洁性要优于后者。

2.2 上行预测-下行预测误差权衡

在预测加工中,最优化适应问题被Friston借助贝叶斯统计学转化为惊异(surprise)最小化或预测误差(prediction errors)最小化问题(Friston,2005)。设观察为 o ,环境变量为 s ,于是,先验概率 $P(s)$,观察概率 $P(o)$,后验概率 $P(s|o)$,似然 $P(o|s)$ 联合分布概率(即生成模型) $P(o,s)$ 。因为观察概率实际衡量了模型(脑)预测外部环境变量的能力,所以又被称为“模型证据”(model evidence)。总之,模型证据概率越大,即证明模型建构得越好(趋近于与周围环境同构),模型的预测能力越强。由贝叶斯定理,模型证据: $P(O) = P(s)P(O|S)/P(s|o)$ 。

Friston 尝试引入新术语“惊异”,尽量同时从数学上和常识上统一对模型证据概率值的表述。“惊异”被规定为模型证据的负对数值,即: $-\log P(O)$ (等于 $-\log \sum_s P(o,s)$); 它的量化等价概念为“惊异度”(surprisal), 可对应信息论中的自信息(self-information)概念(Buckley et al.,2017)。“惊异”概念可被应用于对预测误差的日常生活理解。例如,当预测误差出现时,认知主体会对超出预期的发生事件感到惊讶,“我对我的房间中出现一条蛇感到惊讶!”。总之,模型证据最大化=惊异最小化。惊异的上界被规定为“自由能”。于是,惊异 $\leq$ 自由能, $-\log \sum_s P(o,s) \leq -\sum_s \log P(o,s)$ 。自由能原理的基本宗旨是设法增大模型证据或减小惊异,这一问题最终变成尽可能使惊异的上界自由能最小化。这即是“自由能原理”这一名称的由来(Friston,

2005)。Friston 等(2012)甚至将自由能原理中蕴涵的适应性原则上升为指导生存和生活的律令(Imperative):避免惊异,你会活得更长。

在层级神经网络中,因为模型证据或惊异的真实值难以求得,所以通常用变分推断(variational inference)等近似推断算法替代精确推断,即通过构造可知的简单近似分布($Q(s)$)去逼近未知的复杂真实分布($P(s)$)(Goodfellow et al.,2016)。最终,自由能最小化问题被转化为变分自由能(variational free energy)最小化问题。

层级预测加工的基本加工方式为:上行预测-下行预测误差权衡(ascending prediction - descending prediction errors tradeoff)(Friston,2019)。如图1,认知主体通过身体与外部世界发生交互作用(“S/A”指“感知/行动”),例如通过眼和手从外部世界中采集信息。预测信号(又称为预期信号)(P_i)指模型关于世界若干真实状态的自上而下的预测,预测误差信号(E_j)指感知输入(S)与模型关于该输入的预测的差值。

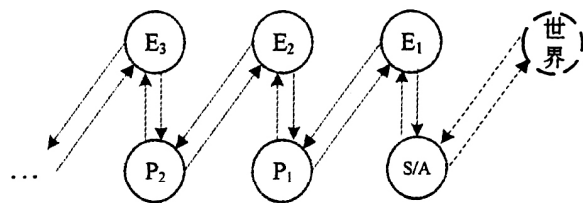


图1 上行预测-下行预测误差相互作用

基于适应性原则,上行预测与下行预测误差之间始终存在着权衡。一般而言,上行预测信号是抑制性的(inhibitory),下行预测误差信号是兴奋性的(excitatory)。Yon 等(2019)指出,预测误差与认知主体关于外部世界状态的确定程度(先验信念)有关。当认知主体对未来状态较为确定,例如较为“自信”或“骄傲”时,下行预测误差信号在更大程度上被上行预测信号抑制,例如,“我完全不相信我的房间中存在一条蛇!”,此时,即使房间中的确有一条骇人的蛇,认知主体关于真实蛇的感知输入信号在更大程度上可能被抑制(甚至视而不见)。由此,Yon 等指出预测模型事实上存在较强的顽固性(stubbornness),其先验信念并不会轻易通过预测误差获得修正。总之,两种信号相互依赖而存在。

3 主动预测加工

3.1 主动推断:行动的最优解

在表征-计算观中,行动仅作为认知的手段或中介在理论建构中被忽略了,认知活动的目标是产生内部表征。预测加工较之其理论前身层级预测编码,其实质性的纲领突破在于 Friston 等人将后者改造为行动引导的理论部件。这样,原本大致隶属于表征-计算观的层级预测编码具有与第二代认知科学纲领相兼容的可能性。该理论部件被称作“主动预测加工”或“主动推断”(active inference)(Friston, 2010)。主动推断是与自由能原理互补的行动决策理论部件,其基本理念为:预测模型通过对不同行动策略(policy)赋值,从而使认知主体基于贝叶斯模型均值尽可能给出行动选择的最优解,最终实施行动。赋值根本取决于每一行动在未来时刻导致的预期自由能(expected free energy)的变化。从主动预测加工的观点看,前运动的行动决策选择完全决定了实际行动与否,于是认知主体的行动策略选择被认为对实际行动存在必然因果作用,最终通过行动在某种意义上改变了外部世界。基于这一理念,外部环境(s)被重新规定为基于行动策略 π 在时刻t相应的改变状态($s_t | \pi$)。基于上文的数学表述,新的参数变量被再次代入变分自由能公式,以计算预期自由能。这即是主动推断的基本理念(Friston et al., 2017)。

回顾 Marr 与 Gibson 的理论关联, Gibson 对外部环境光阵不变量和生态学动允性框架的强调被认为欠缺内部视觉神经机制的过程分析;而 Marr 则轻视视觉神经机制设计并忽略了视觉行动的直接性,将问题引向表征-计算进路(Marr, 1982; Warren, 2012)。相较于前两者,主动预测加工既保留了第一代认知科学的层级计算加工理论部件,并将其改造为与第二代认知科学的知觉-行动观相兼容。

此外, Pezzulo 和 Cisek (2016) 指出,主动预测加工对多种可能行动策略的权衡从根本上符合多重动允性竞争机制。例如,当认知主体看见一只茶杯,他/她便会无意识地通过视觉双通道等大脑通路产生对该茶杯如何付诸行动(即动允性)的多种

可能性。不同动允性之间存在竞争关系,认知主体最终采取何种行动(例如,拿起茶杯喝水或者投掷茶杯)取决于大脑贝叶斯模型均值。

3.2 知觉-行动循环

一些学者提出预测加工的知觉-行动观更接近实用主义的认知科学(Engel et al., 2015)。不过预测加工在这一立场上可能与生成认知等派别不同。不同生成认知流派大致认同知觉即行动(O'Regan & Noë, 2001; Varela et al., 1991)。从行动的方面看,预测加工的认知过程可被视为从准备行动开始,到实现行动结束。但是从认知的方面看,预测加工类似于表征-计算观,同样将行动视为手段或媒介,通过获得关于外部环境状态的内部表征,以建构和更新贝叶斯模型。于是在预测加工中,知觉固然服务于行动,而行动同样服务于知觉。预测加工较之表征-计算观赋予行动更大的理论权重,但是行动同样服务于以表征为基础的贝叶斯模型的建构和更新。

在层级预测加工中,自上而下的预测与自下而上的预测误差之间体现出权衡。上行预测由基于表征的模型决定,而下行预测误差由行动引导的感知决定。在主动预测加工理论部件中,于是相应体现了知觉与行动之间的权衡,也被称作“认识-实用权衡”(epistemic-pragmatic tradeoff)(Friston et al., 2017)。总之,在预测加工中,知觉(认识价值维度)和行动(实用价值维度)根本服务于适应(适应价值维度)。

因为预测加工预设了前运动的行动决策完全决定了实际行动本身,所以预测加工的重点在于多重行动策略(或动允性竞争)权衡。如图2,人脑的各部位围绕惊异最小化/动允性竞争最优化这一总体目标构成了一个完整的循环回路。在目前的设计(Friston et al., 2017)中,该知觉-行动循环主要涉及脑的7个位置,根据图中顺时针顺序,其功能分别为:枕叶,感知输入;海马,基于合理策略的状态估值;基底神经节中的纹状体,策略选择;基底神经节中的腹侧背盖核和黑质,精确化腹侧前额叶,策略评估;背侧前额叶,下次输出的贝叶斯模型均值(预测值);运动皮层,行动输出。

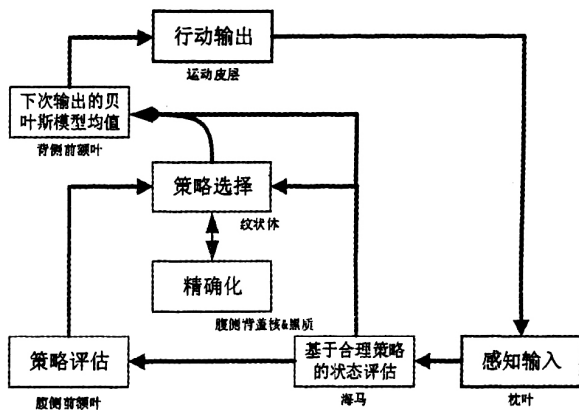


图2 脑中闭合的知觉-行动循环

历史上,预测加工的知觉-行动循环理念与 John Dewey 的反射弧理论相呼应 (Clark, 2016)。Dewey (1896) 基于看见与抓握的关系描述了一种全身性反射弧。在 Dewey 处,知觉与行动从来不是相互独立的,也并非一方比另一方更根本,而是同一个“协调”(coordination)的两个不同方面。一次看见固然意味着知觉对象的获得,一次抓握也反映了行动的实施,但是它们从根本上体现了一次知觉-行动协调的实现。预测加工基于可量化的贝叶斯权衡机制进一步发展了协调观念。Dewey 的知觉-行动循环是全身性的,可被视为一阶的;而预测加工的知觉-行动循环是基于-一阶具身性反射弧的二阶循环。该循环在脑中内化,构成新的闭合回路 (Clark, 2016; Friston et al., 2017)。

4 总结和展望:认知科学的新范式?

在当代,预测加工思潮的兴起不断点燃了认知科学哲学家渴望重塑认知科学新范式的热情。预测加工并非只是像表面反映的那样,是基于层级加工和主动知觉两大进路对两代认知科学分裂纲领的权宜式合拢。在认知的基本观念上,较之以往两代认知科学,预测加工可能已经取得了根本性超越。Marr 之所以称在诸多前人理论中, Gibson 理论与他本人的理念最接近,根本原因在于他们均认同知觉问题的本质在于知觉对象 (object) 的不变量 (性) 分析。诚然,在 Gibson 与 Marr 各自的理论中业已蕴涵了两种认知观的巨大分歧,他们走向了相反的道路: Gibson 走向了外在的生态学进路,将这一不变量归于外部环境光阵的固有性质;而 Marr 走向了内在的信息加工进路,将不变量归于抽象算

法。在预测加工中,认知作为适应系统的系统性功能服从最基本的适应性原则。从而,以往两代认知科学对知觉对象的不变量 (性) 分析让位于预测加工中的心智-世界适应性“关系”分析。

由此,在预测加工中,在上行预测与下行预测误差之间、在知觉与行动之间、在认识价值与实用价值之间处处体现出权衡。在哲学观上,预测加工被认为与一种过程观相适合 (Friston et al., 2017), 并在知觉对象上持有动态的可塑性观点。表征服务于贝叶斯模型的建构和更新,被认为是可变的和可塑的 (张静, 李恒威, 2016)。此外,伴随着认知主体对世界的主动的具身性耦合,世界也相应地被主体改变 (Clark, 2016)。总之,在预测加工中,认知是可变的,世界是可变的,唯一不变的是认知主体在世界之中的适应性关系。

通过比较,预测加工可能给出了第三种新范式 (表 1)。一般而言,表征-计算观的目标是产生内部表征,加工方式是抽象算法匹配,最终获得的知觉对象被认为是关于世界“如其所是”的精细的、不变的表征;具身行动观的目标是实施行动,加工方式是身体与环境之间的动力学耦合,该范式的激进立场反对头脑中存在关于外部世界的内部表征。预测加工范式大致体现出对前两种范式的融合态势,它的目标是实现知觉-行动循环,加工方式是基于层级神经网络的上行预测-下行预测误差权衡。

表 1 三种认知观

认知观	目标	加工方式	价值维度
表征-计算观	表征	算法	认识价值
具身行动观	行动	身体-环境动力耦合	实用价值
预测加工观	知觉-行动循环	上行预测-下行预测误差权衡	适应价值

作为一种发展中的理论思潮,预测加工在当代因认知科学领域的诸多学人出于各自立场而产生不同借用。其中,预测加工与以往两大认知科学范式的关系被视为最富争鸣的讨论热点之一 (Clark, 2013; Metzinger & Wiese, 2017)。最终,本文基于预测加工的理论现状和发展态势,尝试提出一种融合立场,认为预测加工在本时代主要承

担了融合表征 - 计算观与具身行动观这两大分裂派别的迫切时代任务。

参考文献:

[1] 李恒威, 黄华新. (2006). 表征与认知发展. 中国社会科学, 2, 34 - 44.

[2] 吕雪靖, 侯欣. (2019). 听觉预测编码: 对声音重复和变化的神经反应. 心理科学进展, 27(12), 1996 - 2006.

[3] 张静, 李恒威. (2016). 自我表征的可塑性: 基于橡胶手错觉的研究. 心理科学, 39(2), 299 - 304.

[4] Buckley, C. L., Kim, C. S., McGregor, S., & Seth, A. K. (2017). The free energy principle for action and perception: A mathematical review. Journal of Mathematical Psychology, 81, 55 - 79.

[5] Chemero, A. P. (2009). Radical embodied cognitive science. The MIT Press.

[6] Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behavioral and Brain Sciences, 36(3), 181 - 204.

[7] Clark, A. (2016). Surfing uncertainty: Prediction, action, and the embodied mind. Oxford University Press.

[8] Conant, R. C., & Ashby, W. R. (1970). Every good regulator of a system must be a model of that system. International Journal of Systems Science 1(2), 89 - 97.

[9] Dewey, J. (1896). The reflex arc concept in psychology. Psychological Review, 3(4), 357 - 370.

[10] Engel, A. K., Friston, K. J., & Kragic, D. (2015). The pragmatic turn: Toward action - oriented views in cognitive science. The MIT Press.

[11] Friston, K. (2005). A theory of cortical responses. Philosophical Transactions of the Royal Society B: Biological Sciences, 360(1456), 815 - 836.

[12] Friston, K. (2010). The free - energy principle: A unified brain theory? Nature Reviews Neuroscience, 11(2), 127 - 138.

[13] Friston, K., & Buzsáki, G. (2016). The functional anatomy of time: What and when in the brain. Trends in Cognitive Sciences, 20(7), 500 - 511.

[14] Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active inference: A process theory. Neural Computation, 29(1), 1 - 49.

[15] Friston, K., Thornton, C., & Clark, A. (2012). Free -

energy minimization and the dark - room problem. Frontiers in Psychology, 3, 130.

[16] Friston, K. J. (2019). Waves of prediction. PLoS Biology, 17(10), e3000426.

[17] Gibson, J. J. (1966). The senses considered as perceptual systems. Houghton Mifflin.

[18] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. The MIT Press.

[19] Hohwy, J. (2013). The predictive mind. Oxford University Press.

[20] Kirchhoff, M. D. (2018). Autopoiesis, free energy, and the life - mind continuity thesis. Synthese, 195(6), 2519 - 2540.

[21] Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information. W. H. Freeman and Company.

[22] Maturana, H. R., & Varela, F. J. (1980). Autopoiesis and cognition: The realization of the living. D. Reidel Publishing Company.

[23] Metzinger, T., & Wiese, W. (2017). Philosophy and predictive processing. MIND Group.

[24] O'Regan, J., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. Behavioral and Brain Sciences, 24(5), 939 - 973.

[25] Pezzulo, G., & Cisek, P. (2016). Navigating the affordance landscape: Feedback control as a process model of behavior and cognition. Trends in Cognitive Sciences, 20(6), 414 - 424.

[26] Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra - classical receptive - field effects. Nature Neuroscience, 2(1), 79 - 87.

[27] Shapiro, L. (2019). Embodied cognition. Routledge.

[28] Teufel, C., & Fletcher, P. C. (2020). Forms of prediction in the nervous system. Nature Reviews Neuroscience, 21(4), 231 - 242.

[29] Thompson, E. (2007). Mind in life: Biology, phenomenology, and the sciences of mind. Harvard University Press.

[30] Varela, F. J., Thompson, E., & Rosch, E. (1991). The embodied mind: Cognitive science and human experience. The MIT Press.

[31] Warren, W. H. (2012). Does this computational theory solve the right problem? Marr Gibson, and the goal of vision. Perception, 41(9), 1053 - 1060.

[32] Yon, D., de Lange, F. P., & Press, C. (2019). The predictive brain as a stubborn scientist. Trends in Cognitive Sciences, 23(1), 6 - 8.

Predictive Processing: Breaking the Impasse of Two Generations of Cognitive Science

Dong Da Chen Wei

Abstract: Nowadays, quarrels and fights between two generations of cognitive science almost result in an inevitable impasse. For instance, a deep theoretical relation between J. J. Gibson and David Marr have been carved off, and of which two theories each other have been pushed toward the opposite side; which could be viewed as a portrayal of these overt states of affairs. The central tenet of the First – Generation Cognitive Science summarizes as: The Representational – Computational Approach; and the central tenet of the Second – Generation Cognitive Science: The Embodied Action Approach. What is the nature of "mind" at all? These two approaches conflict each other, and the either/or choice seems the only game in the town.

In recent years, researchers attempt to integrate both strains of sources, hoping to provide a unified framework, thus of which could be viewed as prospectively breaking the impasse of two generations of cognitive science in the near future. It is called "Predictive Processing". This approach apparently originates in hierarchical predictive coding of visual and auditory studies (since 1980s), and thereby has been modified substantively by the neuroscientist Karl Friston, et al., into a mode of action – oriented. Predictive Processing combines with two indispensable components: Hierarchical Predictive Processing and Active Predictive Processing. In rough, the Hierarchical Predictive Processing mainly inherits from hierarchical predictive coding theories of the First – Generation Cognitive Science (e. g., Marr's multi – level computational theory of vision, hierarchical neural networks, etc.); and the Active Predictive Processing somehow advocates various versions of theories of the Second – Generation Cognitive Science related with action (e. g., Gibson's affordances theory, enactive cognition, etc.). Both components could be viewed as dual aspect of a single unified theory.

In Predictive Processing, the Free Energy Principle views adaptivity as the fundamental principle of the theory construction of neural and cognitive systems, and until self – organizing adaptive systems. The overall aim of Hierarchical Predictive Processing is to make the cognitive agency adaptable and sustainable as much as possible. Next, the optimization problem of adaptation turns into problems of surprising minimization or "prediction errors minimization" by means of Bayesian inference and Bayesian statistics. The basic working mechanism of Hierarchical Predictive Processing is the ascending prediction – descending prediction errors tradeoffs.

The so – called "Active Inference" is the complementary component of Free Energy Principle, which is seen by Paul Cisek et al., is confirmed with multiple affordances competition mechanism. The Active Predictive Processing shows perception – action tradeoffs, or "epistemic – pragmatic tradeoffs". Both perception (epistemic value) and action (pragmatic value) ultimately serve for adaptivity (adaptive value). Perception – action cycle of Predictive Processing could be regarded as a second – order cycle based on first – order embodied reflex arc (in Deweyan sense); which could also be seen as an internalized closed loop within the corporal brain.

Thus on the nature of cognition, contrasting with past two generations of cognitive science, Predictive Processing has made revolutionary breakthrough. Cognition as the systematic function of general adaptive system complies with the fundamental adaptivity principle. The conventional object – invariants analysis beneath paradigms of two generations of cognitive science, thus ought to give way to mind – world adaptivity analysis underlying the Predictive Processing paradigm. Philosophically speaking, Predictive Processing seems to be consistent with a process theory. In the end, on the basis of the currents and trends of Predictive Processing, we attempt to propose an integrated view. In our view, the pressing contemporary task of Predictive Processing is trying to integrate two generations of cognitive science – Representational – Computational Approach and Embodied Action Approach.

Key words: predictive processing; hierarchical predictive coding; free energy principle; active inference