



关于机器学习基本问题及 统一理论方面的思考

胡包钢

学术研究工作中的科学乐趣包括对根本性问题的长久思考。以机器学习研究为例,当我们研究工作分别应用了贝叶斯决策或信息论方法之后,导致了我们的困惑与思考:如何在两者中进行选择?两者中是否有一个更为基础?2011年本人在《中国人工智能学会通讯》中“关于贝叶斯决策与机器学习目标选择的讨论”^[1]一文结尾处,曾经提出以下疑问:“熵理论与贝叶斯理论分别代表了处理信息不确定性的两种不同理论体系,它们将有怎样的关联?两者的融合能否为我们对机器学习机理带来更为基础和本质性的数学理论方面的认知?”。随着不断的学习与思考。我最近发表了对该疑问进行初步回答的观点文章“Information Theory and its Relation to Machine Learning”^[2]。本文将简单介绍相关观点进展,并对人工智能研究未来走向提些个人看法。目的是增加同行的质疑与讨论,推动我们对机器学习或人工智能研究中基本问题及其统一理论方面的思考。

文献[2]借老子的名言:“道生一,一生二,二生三,三生万物”引发我们对认知自然万物的思考。中国古代哲人寥寥数语表达诸多智慧要点。包括宇宙万物演变由简至繁,“道”乃唯一源泉。可以理解对于“道”或“源”也成为科学家追求科学研究的目标之一。爱因斯坦(Albert Einstein)曾终生致力于发展物理世界中的统一理论^[3]。他特别强调“自然乃为最简洁且可理解的数学思想之实现(nature is the realization of the simplest conceivable mathematical ideas)”。

视觉计算理论开创者马尔(Davie Marr,于1980年去世,时年35岁)曾经提出计算视觉理论应用三个不同水平研究方式。它们分别为“Computational theory”、“Representation and algorithm”、“Hardware implementation”^[4]。与马尔曾经合作的Tomaso Poggio 2007年提出了(机器学习研究中的三个层次划分方法(three levels on learning)。这包括“Learning theory and algorithms”、“Engineering applications”、“Neuroscience: models and experiments”)^[5]。文献[2]在借鉴上述方法论后,提出了新的问题或层次分析方法(见图1)。可以理解研究视角不同,基本问题会随之改变^[6]。提出新的基本问题可以对解决问题带来新的

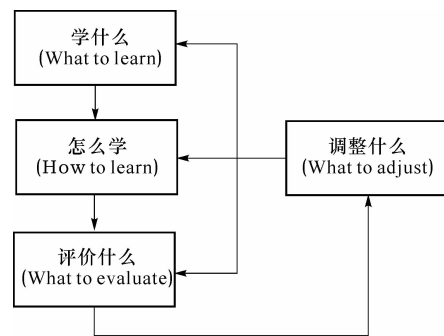


图1 机器学习研究中基本问题(或层次)流程示意图^[2]

驱动力。

图 1 中四个基本问题是基于机器学习理论研究中专题目标而划分。为突出基本问题,我们忽略其他可能的专题目标。马尔对问题划分方法论的见解是:尽管问题或层次之间存有耦合关联,这种划分有利于“逐步理解感知信息处理(eventual understanding of perceptual information processing)”原理。图 1 中“学什么”是机器学习中关于“学习目标选择”为专题的研究。所谓“学习目标”是对机器模型预期学习内容的“表达(Representation)”。该表达可以分解为两种,即“语义(Linguistic)”表达与“计算(Computational)”表达。它们分别反映为高层次中应用自然语言的定性叙述与低层次中应用计算语言的定量描述。在计算表达层面通常要涵盖应用数学语言描述的目标函数、约束条件,以及优化形式三个方面的内容。从模式识别研究领域看,以特征选择(Feature Selection)为专题的研究比比皆是。而在机器学习领域中似乎少有以学习目标选择为研究专题的提法。虽然许多研究工作也涉及了该方面内容的探讨(如“生成式(Generative)”与“判别式(Discriminative)”学习目标的比较研究),但是围绕该专题的系统性研究在机器学习理论研究中似乎尚未形成。“学什么”将直接涉及了认知科学的内容。该专题研究可以强化机器学习对学习机理的深刻认知,并强调包括语义表达与计算表达两个层面中显性知识的获取。

第一模块“学什么”与第三模块“评价什么”被归为机器学习研究领域“首要”的基本问题。“评价什么”模块是关于“评价指标选择”为专题的研究。当设定目标函数与评价指标应用完全相同的数学函数表达时,“评价指标选择”可以是“学习目标选择”研究中的子问题(无需选择优化形式)。该基本问题将消失。而实际工程中多数会采用单一学习目标设计的学习机器,但是应用多个指标对其进行综合评价的方式。多指标与单一学习目标通常情况下并非具有一致性,因此“评价什么”将成为基本问题。而第二模块“怎么学”是关于学习过程与实现的研究。该研究内容被列为“其次”的基本问题。也就是说,学习目标不正确,再好的“怎么学”方法也无法纠正这个错误,并很难实现正确学习目标下应有的结果。第四模块“调整什么”是关于前三个模块动态调整的研究。该模块专题目标是赋予机器具有“智能进化”功能。

图 1 中基本问题划分是期盼为机器学习理论研究带来新的视角,澄清目前研究中的不足并指明研究发展走向。如近几年一些研究工作应用 F1 准则为学习目标,给出的动机是源于不平衡数据学习中要保护小类^[7]。以“学什么”来考察,这个研究目标在语义表达转换为计算表达上存有缺陷。为实现计算考察,初始语义“保护小类”转化为可考察的具体语义为:“当小类趋于零时,其小类中的多半样本应该可以正确判断”。但是应用 F1 准则作为计算表达的学习目标,并不能满足具体语义表达的学习目标。文献[8]中我们应用反例表明了应用 F1 准则为学习目标无法保护小类。即“小类趋于零时,应用 F1 准则后小类中样本将趋于被全部误判为大类的情形”。除了上面提到的机器学习研究领域中对“学习目标选择”为专题的研究工作严重缺失外,图 1 中第四模块“调整什么”示意了未来机器学习需要关注的另一重要专题。目前机器学习研究多数只是对误差反馈的调整,而非包括知识反馈或知识更新的调整。

需要指出的是,对于任何复杂系统采取基本问题划分方式难免产生“还原论(Reductionism)”原理的内在弊端。为此,图1中应用多通道反馈方式来近似实现“整体论(Holism)”的哲学原理。这种处理方式是否符合人的智能行为与决策方式本身就是值得探讨的研究主题。“中庸之道”以及“兼容并包”这种东方哲学思想为处理复杂系统研究展示出智慧路径。文献[2]中最为主要的结论是提出了“学习目标选择猜想”,具体表述如下:

“机器学习中所有学习目标的计算表达均可以应用熵函数的优化形式来描述或解释”。

该猜想是基于前人的工作而大胆提出(参见文献[2]中具体叙述与引文)。该猜想也包含了 Zellner 于 1988 年导出的重要结论^[9]:信息熵函数优化乃为贝叶斯原理之源。如果将学习目标计算表达简单分为两类,一类基于优化熵原理的学习目标,一类是基于经验式的(非熵原理)学习目标。文献[2]中的另一个学术贡献是图示了两者之间非一一对应的关系以及发生的位置。上述猜想并非目标于否定应用基于经验式的计算表达。主旨是要回到“统一理论”为主题的思考。该主题能够促使我们尽快地切入到人工智能研究或类脑研究的本质:科学层面解释性的统一理论框架存在否,又会是什么?这里试举若干研究实例。

- 认知中的统一理论(Unified Theories of Cognition)。1975年图灵奖获得者 Allen Newell 历经 20 多年的努力提出并完善了该理论^[10]。这是针对人类所有认知行为可以应用单组机制(A single set of mechanisms for all cognitive behavior)描述的理论方法。Newell 认为人是一个符号处理系统,因此该理论是围绕符号处理为核心。

- 基于量子理论(Quantum Theory)的脑功能模型方法。该方法认为传统 Kolmogorov 概率体系无法描述脑功能,必须应用 von Neumann 量子概率系统来描述或解释。这之中主要包括量子脑智(Quantum Mind)或量子意识(Quantum Consciousness)学派^[11],以及量子认知(Quantum Cognition)学派^[12]。不同于前者假设脑功能存在微观物理层面上量子尺度的力学作用,后者是定义在宏观物理层面(非量子尺度)上脑认知功能的描述。故后者又被归为广义量子范式(Generalized Quantum Paradigm)。

- 基于自由能原理(Free Energy Principle)的脑统一理论。该理论学派认为大脑系统功能符合能量守恒定律,其中脑功能内部状态可以应用最小自由能原理来解释。该理论学派代表人物是英国神经科学家 Karl J. Friston 领导的团队^[13-14]。他们理论上证明了所谓贝叶斯大脑(Bayesian Brain)只是最小自由能原理导出诸多脑功能中的一种。应用该理论框架可以达到对记忆、注意、评价、再励、凸显(Saliency)等功能之间关联的理解。值得提及的是机器学习研究人员(包括深度学习开创者之一 Geoffrey Hinton)于 1995 年就开始应用该原理构造神经网络,并称之为“亥姆霍兹机(Helmholtz Machine)”^[15]。

由此可以质疑文献[2]中提出的猜想与上述三个统一理论框架的不同。我个人理解,熵仍然是计算层面中最为核心的概念。但是它自身数学层面定义及其导出的熵函数仍然需要不断地扩充与完善。例如钟义信曾提出了全信息理论,这是针对“熵”扩展到“信息-知识-智能”之间转换而提出的统一理论框架^[16]。计算表达与语义(知识)表达的关联与转换仍是开放问题。

“语义鸿沟”中“正向”与“反向”转换之间的差异^[2]表明各种统一理论还有很长的路要走。

在拟将开展的“中国脑科学计划”中,我们会不可避免地涉及到“统一理论”猜想的基础研究内容。这将会是百家争鸣与大浪淘沙的过程。有关研究工作极易掉入基本概念炒作或空洞理论框架的陷阱。然而面对“生物脑”与“机器脑”中重大科学问题,我们必须要有勇气并追求以数学语言对它们进行深刻而系统性的描述。期待中国学者能够在“脑统一理论”的继承、证伪、与发展中做出有国际影响力的创新贡献。

参考文献:

- [1] 胡包钢. 关于贝叶斯决策与机器学习目标选择的讨论[J]. 中国人工智能学会通讯, 2011,2(8): 1-4.
- [2] Hu Baogang. Information theory and its relation to machine learning[C]//ICAC2015. Fuzhou: [s. n.]. 2015: 1-11.
- [3] Parker B. Einstein's dream: the search for a unified theory of the universe[M]. [S. l.]:Basic Books, 2001.
- [4] Marr D. Vision;a computational investigation into the human representation and processing of visual information[M]. [S. l.]: The MIT Press,2000.
- [5] Poggio T. How the brain might work: the role of information and learning in understanding and replicating intelligence [C]//Information: Science and technology for the new century. [S. l.]: Lateran University Press, 2007:45-61.
- [6] 唐孝威,郭爱克,吴思,等主编. 神经信息学与计算神经科学[M]. 杭州:,浙江科技出版社, 2012.
- [7] Maratea A, Petrosino A, Manzo M. Adjusted F-measure and kernel scaling for imbalanced data learning[J]. Information Sciences, 2014, 257(1): 331-341.
- [8] Hu Baogang, Dong Weiming. A study on cost behaviors of binary classification measures in class-imbalanced problems [EB/OL]. (2014-03-26)[2015-08-12]. <http://arxiv.org/abs/1403.7100>.
- [9] Zellner A, Optimal information processing and Bayes's theorem[J]. The American Statistician. 1988, 42(4):278-284.
- [10] Newell A. Unified theories of cognition[M]. Cambridge, Massachusetts: Harvard University Press, 1990.
- [11] Koch C, Hepp K. Quantum mechanics in the brain[J]. Nature, 2006,440: 611-612.
- [12] Busemeyer J R, Bruza, P D. Quantum models of cognition and decision[M]. Cambridge: Cambridge University Press,2012.
- [13] Friston K, Kilner J, Harrison L. A free energy principle for the brain[J]. Journal of Physiology-Paris, 2006,100(1):70-87.
- [14] Friston K. The free-energy principle: a unified brain theory? [J]. Nat Rev Neurosci,2010,11 (2):127-38.
- [15] Dayan P, Hinton G E, Neal R M,et al. The Helmholtz machine[J]. Neural Computation, 1995,7(5): 889-904.
- [16] 钟义信. 信息转换原理: 信息, 知识, 智能的一体化理论[J]. 科学通报, 2013, 58 (14): 1300-1306.



作者简介: 胡包钢,中国科学院自动化研究所模式识别国家重点实验室研究员,博士生导师,主要研究方向为模式识别与植物生长建模。
Email: hubg@nlpr.ia.ac.cn