平行学习 — 机器学习的一个新型理论框架

李 力 [1] 林懿伦 [2,3] 曹东璞 [4,5] 郑南宁 [6] 王飞跃 [2,7]

摘 要 本文提出了一种新的机器学习理论框架. 该框架结合了现有多种机器学习理论框架的优点, 并针对如何使用软件定义的人工系统从大数据提取有效数据, 如何结合预测学习和集成学习, 以及如何利用默顿定律进行指示学习等目前机器学习领域面临的重要问题进行了特别设计.

关键词 机器学习, 人工智能, 平行学习, 平行智能, 平行系统及理论

引用格式 李力, 林懿伦, 曹东璞, 郑南宁, 王飞跃. 平行学习 — 机器学习的一个新型理论框架. 自动化学报, 2017, **43**(1): 1−8

DOI 10.16383/j.aas.2017.y000001

Parallel Learning — A New Framework for Machine Learning

LI Li[1] LIN Yi-Lun[2,3] CAO Dong-Pu[4,5] ZHENG Nan-Ning[6] WANG Fei-Yue[2,7]

Abstract In this paper, we propose a new framework of machine learning theory, parallel learning, which incorporates and inherits many elements from various existing machine learning theories. Special designs are also presented to deal with some important problems in the machine learning research field, e.g., useful data retrieval from big data using software defined artificial systems, combination of predictive learning and ensemble learning, application of Merton's law to prescriptive learning.

Key words Machine learning, artificial intelligence, parallel learning, parallel intelligence, parallel system and theory

Citation Li Li, Lin Yi-Lun, Cao Dong-Pu, Zheng Nan-Ning, Wang Fei-Yue. Parallel learning — a new framework for machine learning. *Acta Automatica Sinica*, 2017, **43**(1): 1−8

1 引言

随着计算能力的提高和计算理论的创新, 机器学习在过去 30 年中取得了长足的发展, 正受到越来越多人的关注, 并在生物、医药、能源、交通、环境等诸多领域中获得了成功的应用. 与此同时, 机器学习也面临越来越多的问题, 传统机器学习理论框架的不足被逐渐发现和确认, 新的机器学习理论框架不断被提出[1].

中科院自动化所王飞跃研究员于 2004 年提出了平行系统的思想, 试图用一种适合复杂系统的计算理论与方法解决社会经济系统中的重要问题. 其主要观点是利用大型计算模拟、预测并诱发引导复杂系统现象, 通过整合人工社会, 计算实验和平行系统等方法, 形成新的计算研究体系[2−5].

在过去的 10 多年中, 平行系统这一研究体系在实践中取得了大量的成果, 并不断丰富和完善起来[6−9]. 近年来, 我们尝试将平行系统的思想扩展并引入到机器学习领域建立一种新型理论框架以更好地解决数据取舍、行动选择等传统机器学习理论不能很好解决的问题.

以下我们将首先回顾常见的一些机器学习理论, 并比较它们在数据获取 — 行动选择这一核心问题上的处理方式. 接着, 我们将提出 "平行学习" 这一新型机器学习理论框架, 并着重分析其独特之处. 最后我们总结全文.

收稿日期 2016-11-25 录用日期 2017-01-01
Manuscript received November 25, 2016; accepted January 1, 2017
国家自然科学基金 (91520301) 资助
Supported by National Natural Science Foundation of China (91520301)
本文责任编委 刘德荣
Recommended by Associate Editor LIU De-Rong

1. 清华信息科学与技术国家实验室 (筹), 清华大学自动化系 北京 100084 中国 2. 中国科学院自动化研究所复杂系统管理与控制国家重点实验室 北京 100190 中国 3. 中国科学院大学 北京 100049 中国 4. 英国克兰菲尔德大学驾驶员认知与自动驾驶实验室 克兰菲尔德 MK43 0AL 英国 5. 青岛智能产业技术研究院 青岛 266000 中国 6. 西安交通大学人工智能与机器人研究所 西安 710049 中国 7. 国防科技大学军事计算实验与平行系统技术中心 长沙 410073 中国
1. National Laboratory for Information Science and Technology (TNList), Department of Automation, Tsinghua University, Beijing 100084, China 2. the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China 3. University of Chinese Academy of Sciences, Beijing 100049, China 4. Driver Cognition and Automated Driving Laboratory, Cranfield University, Cranfield MK43 0AL, UK 5. Qingdao Academy of Intelligent Industries, Qingdao 266000, China 6. Institute of Artificial Intelligence and Robotics (IAIR), Xi'an Jiaotong University, Xi'an 710049, China 7. Research Center of Military Computational Experiments and Parallel Systems, National University of Defense Technology, Changsha 410073, China

2 一些现有机器学习的理论框架

着眼于数据获取和行动选择之间的关系, 我们可以建立如下数学模型来描述常见的一些机器学习理论框架:

假设我们获取到一系列数据并构成集合 $X = \{x_i\}$, $i = 1, \cdots, I$. 若研究对象为复杂系统, 则这些数据通常为观测到的系统状态或输出. 针对这些数据, 我们可以采取一系列行动 a_k 并构成集合 $A = \{a_k(X')\}$, $k = 1, \cdots, J$, $X' \subseteq X$ 表示数据集 X 的一个子集. 每一个行动可以导致一个回报 $R(a_j)$, 且数据的获取和采取行动可以在时间上分离. 我们的目标是, 通过机器学习, 最大化长期回报

$$\max_{a_j, 1\le j\le J} \sum_{k=1}^{J} R(a_k) \tag{1}$$

如果我们关心的是每一个行动导致的损失 $L(a_k)$, 则目标函数可变为最小化长期损失

$$\min_{a_k, 1\le k\le J} \sum_{k=1}^{J} L(a_k) \tag{2}$$

对于常见的有监督学习 (Supervised learning)[10−11] 而言, 上述这一模型可以进一步简化为: 当所有数据已知且已经被正确分类后, 我们采取一个行动: 建立一个函数映射 (通常是分类函数) 来最小化分类误差. 一般而言, 我们预设数据服从独立同分布假设 (Independent and identically distributed, i.i.d.), 则目标函数可进一步写为

$$\min_{a} L(a(X)) = \min_{a} \sum_{i=1}^{I} L(x_i, a) \tag{3}$$

相对于有监督学习, 在线机器学习 (Online machine learning)[12−14] 强调了数据是逐渐获取的, 且每新获得一个数据, 系统可基于所有已经获取的数据采取一个行动. 对于在线机器学习的特例序贯学习 (Sequential learning), 我们每次仅仅获取一个数据 x_i, 依据映射函数 $f(\cdot)$ 产生一个预测行动 $f(x_i)$, 接着我们再获取 x_i 真正的标记数据 $y(x_i)$, 并计算由此产生的损失 $L[f(x_i), y(x_i)]$. 最终我们的目标函数为选取合适的映射 $f(\cdot)$, 最小化长期后悔值 (Regret value)

$$\min_{f} L(a(X)) = \min_{f} \sum_{i=1}^{I} L[y(x_i), f(x_i)] \tag{4}$$

这里我们有多个行动, 并随着所获数据的增多, 不断优化采取的行动.

与在线学习类似, 强化学习 (Reinforcement learning)[15−16] 依然假设数据逐渐获取, 但机器学习系统不再被动地接受数据, 而转为主动寻求. 系统在 t 时刻每获得一个数据 $x_i^{(t)}$ (在强化学习中 x_i 为系统状态), 可采取一系列行动 $a_i^{(t)}$. 为简化符号, 下文中系统状态和行动记为 x_t 和 a_t. 强化学习允许我们在 t 时刻的行动 a_t 影响到我们在 $(t+1)$ 时刻的获取的数据 x_{t+1}, 也即存在如下的 $T(\cdot,\cdot) : X \times A \to X$ 表示特定的状态转移函数

$$T(x_t, a(x_t), x_{t+1}) = \Pr(x_{t+1}|x_t, a(x_t)) \tag{5}$$

我们的目标函数为

$$\max_{a} \sum_{t=1}^{\infty} T(x_t, a(x_t), x_{t+1})[R_a(x_t, a(x_t)) + \gamma \cdot V(x_{t+1})] \tag{6}$$

其中 $R_a(x_t, a(x_t))$ 为 t 时刻系统处于状态 x_t、施加行动 $a(x_t)$ 所获得的即时回报, $V(x_{t+1})$ 为系统处于状态 x_{t+1} 的长期平均回报, γ 为折扣因子.

因此, 强化学习属于主动学习 (Active learning)[17] 的一种, 我们可以选取特定的行动来兼顾优化目标函数和探索输入数据集合 X. 这相对于在线机器学习是非常重要的改进.

然而, 经典的强化学习将数据获取和对应行动局限在马尔科夫决策过程 (Markov decision processes) 的框架中, 限制了其能力的发挥. 目前研究者提出了不少强化学习的变体, 如深度强化学习 (Deep reinforcement learning), 但基本沿用了马尔科夫决策过程这一框架. 这一做法虽然保证了一定范围内学习的有效性, 却不能很好地应用到非马尔科夫决策过程.

强化学习不需要传统意义上的有标签数据, 实际上其学习的过程就是不断更新数据标签的过程. 但是它的学习效率并不高, 需要跟环境进行大量交互从而获得反馈用以更新模型. 当面临复杂系统大数据处理时, 过高的系统状态维数常常使得可行解的探索变得十分困难[18−20].

3 平行学习的理论框架

为了进一步拓展学习能力, 特别是为了解决强化学习所面临的难题, 我们提出如下图 1 所示的平行学习的基本框架. 其大致可以分为数据处理和行动学习两个互相耦合关联的阶段.

在数据处理阶段, 平行学习首先从原始数据中选取特定的 “小数据”, 输入到软件定义的人工系统中, 并由人工系统产生大量新的数据. 然后这些人工数据和特定的原始小数据一起构成解决问题所需要学习的 “大数据” 集合, 用于更新机器学习模型[18].

在行动学习阶段, 平行学习沿用强化学习的思路, 使用状态迁移来刻画系统的动态变化, 从人工合

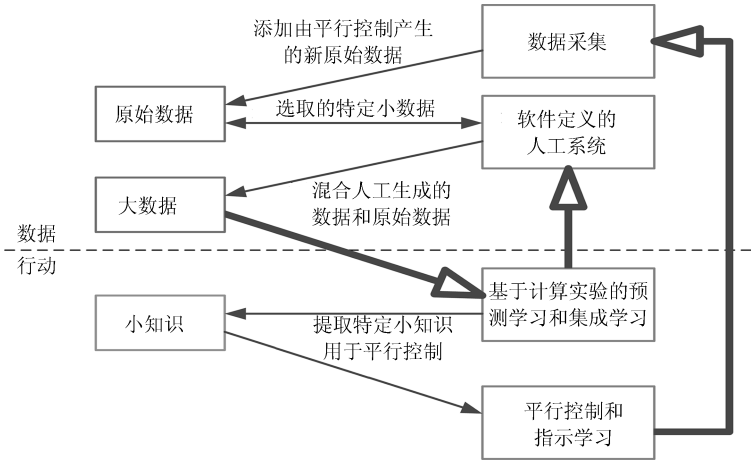

图 1　平行学习的理论框架图 (虚线上方为通过软件定义的人工系统进行大数据预处理, 虚线下方表示基于计算实验的预测学习和集成学习, 以及平行控制和指示学习. 细线箭头代表数据生成或数据学习, 粗线箭头代表行动和数据之间的交互.)

Fig. 1　The theoretical framework of parallel learning (The part above the dash line focuses on big data preprocessing using software defined artificial systems; the part beneath the dash line focuses on predictive learning and ensemble learning based computational experiments, as well as parallel control and prescriptive learning. The thin arrows represent either data generation or data learning; the thick arrows present interactions between data and actions.)

成大数据中学习, 并将学习到的知识存储在系统状态转移函数中. 但特别之处在于, 平行学习利用计算实验方法进行预测学习 (Predictive learning). 通过学习提取, 我们可以得到应用于某些具体场景或任务的 "小知识", 并用于平行控制和平行决策. 这里的 "小" 是针对所需解决具体问题的特定智能化的知识, 而不是指知识体量上的小.

而平行控制和平行决策将引导系统进行特定的数据采集, 获得新的原始数据, 并再次进行新的平行学习, 使系统在数据和行动之间构成一个闭环. 不仅如此, 我们还引入指示学习 (Prescriptive learning) 的思想, 从另一个角度来重新结合数据和行动.

4　平行学习的三大特色方法

在上述新型理论框架的基础上, 我们展开说明平行学习采用的特色方法.

4.1　通过软件定义的人工系统进行大数据预处理

杂乱无序的数据难以学习. 基于平行理论, 我们可以构建人工场景来模拟和表示复杂系统的特定场景, 并将选取的特定 "小数据" 在平行系统中演化和迭代, 以受控的形式产生更多因果关系明确、数据格式规整、便于探索利用的大数据, 再把大数据浓缩成小知识、小智慧和小定律[2−9, 18]. 这一点也符合美国物理学家费曼 (Richard Feynman) 所说的名言 "不是我创造的, 我就不能理解" (What I cannot create, I do not understand)[21].

以业界当前研究的新热点 "平行视觉" (Parallel vision) 为例[22−25]. 我们根据实际采集的少量图像数据, 提取特定要素, 然后在模拟环境条件中加入新的变化 (如改变摄像机朝向、光照和天气条件等) 以得到更加多样化的虚拟数据. 对虚拟实验产生的结果进行计算评估, 我们可以校正视觉模型, 并重新设计新计算实验, 产生新的虚拟数据. 这一迭代将反复执行直到收敛. 测试表明, 结合虚拟数据和真实数据可以有效提高模型性能.

需要指出的是, 平行学习中将用于产生虚拟数据的人工系统和分析数据的机器学习系统进行了一定程度的切分, 允许直接在数据处理阶段进行采样. 这为数据产生和数据分析添加了更多的灵活性, 和传统机器学习主要由行动来驱动数据采集是不同的.

由于我们采用的是软件定义的人工系统, 我们可以借助虚实互动的平行执行来在线优化人工系统模型, 更好地实现对复杂系统的智能理解和数据采样. 所有数据学习的结果都能用来对人工系统模型进行校正和升级. 因此, 我们并非割裂数据产生和数据分析这两部分, 而是将其看成一个整体的两面.

4.2　包含预测学习和集成学习的数据学习

在平行学习中, 我们强调使用预测学习和集成学习来拓展经典学习方法.

1) 我们允许多个智能体 (Agent) O^n, $n = 1, \cdots, N$ 共同学习, 每个智能体可以独立地获取到一系列观测数据并构成集合 $X^n = \{x_i^n\}$, $i = 1, \cdots, I^n$. 每个智能体还可以独立地采取一系列行动并构成集合 $A^n = \{a_k^n(X'^n)\}$, $k = 1, \cdots, J^n$, $X'^n \subseteq X^n$ 表示数据集 X^n 的一个子集.

2) 每个智能体获取的数据和采取行动的次数和

时间均独立. 首先, 我们允许一个行动可以产生多个新的数据, 而强化学习一次只能产生一个新的数据. 其次, 强化学习要求获取数据和完成行动必须依次间隔执行, 而平行学习允许获取数据和完成行动有着完全不同的频次和发生顺序.

3) 我们以平行世界的角度来看待系统状态的演化过程. 将新获得的数据映射到平行空间中, 我们可以通过大量长期的仿真迭代来预测和分析预期行动的结果, 并最终将最优动作返回现实空间[2−8, 18].

基于上述三点扩展, 我们可以放松数据和行动之间的耦合, 极大地扩展现有的强化学习方法[26−27]. 其实, 上述三点扩展已经在 AlphaGo 这一划时代的人工智能产品中得以体现[28]. 参见图 2 所示, AlphaGo 对于当前局面, 使用蒙特卡洛树 (Monte Carlo tree) 方法, 进行数盘 20 ∼ 30 步的模拟测试下来探求局部最优的下法, 这可以看作是使用智能体进行中长期仿真迭代来预测和分析预期行动的结果. 同时, 其数据的产生和行动的产生相对独立, 不需时间对齐. AlphaGo 利用输入的数万盘高手对局数据进行自我对战. 在和李世乭比赛之前, 自我对战了 3 000 多万盘. 这就是典型的实际小数据到虚拟大数据的过程.

LeCun 在最近几年的演讲中反复以 “蛋糕” 比喻整个机器学习领域. 而强化学习是蛋糕上的一粒樱桃, 监督学习是蛋糕外面的一层糖霜, 无监督学习 (Unsupervised learning)[29−30] 则是蛋糕胚. 然而目前, 糖霜和蛋糕胚之间还存在巨大的空白区域. 为此, LeCun 在 2016 年认为预测学习可以来填补这一空白[31].

预测学习起源于认知心理学对于儿童学习方式的解释[32]. 其后被用来解释智能体如何从与环境的交互中学习特定知识[33−34]. 深入比较不难发现, 预测学习和我们在过去 10 多年中所一直倡导的基于计算实验的平行系统方法[2−9] 本质是一样的. 简而言之, 平行系统和预测学习的核心就是用机器给真实环境建模, 仿真预测可能的未来, 并通过观察和演示来理解世界如何运行的能力. 其中仿真是无监督或半监督 (Semi-supervised learning)[35−36] 的, 而初始状态和最终结果是有监督的. 我们称这种学习方式为平行预测 + 指示学习 (Parallel predictive+ Prescriptive learning), 可以糅合无监督、半监督和有监督三种学习方式, 填补各自之间的空白.

4) 类比于多智能体系统[37−38], 平行学习还允许分散学习和集成学习两种机制.

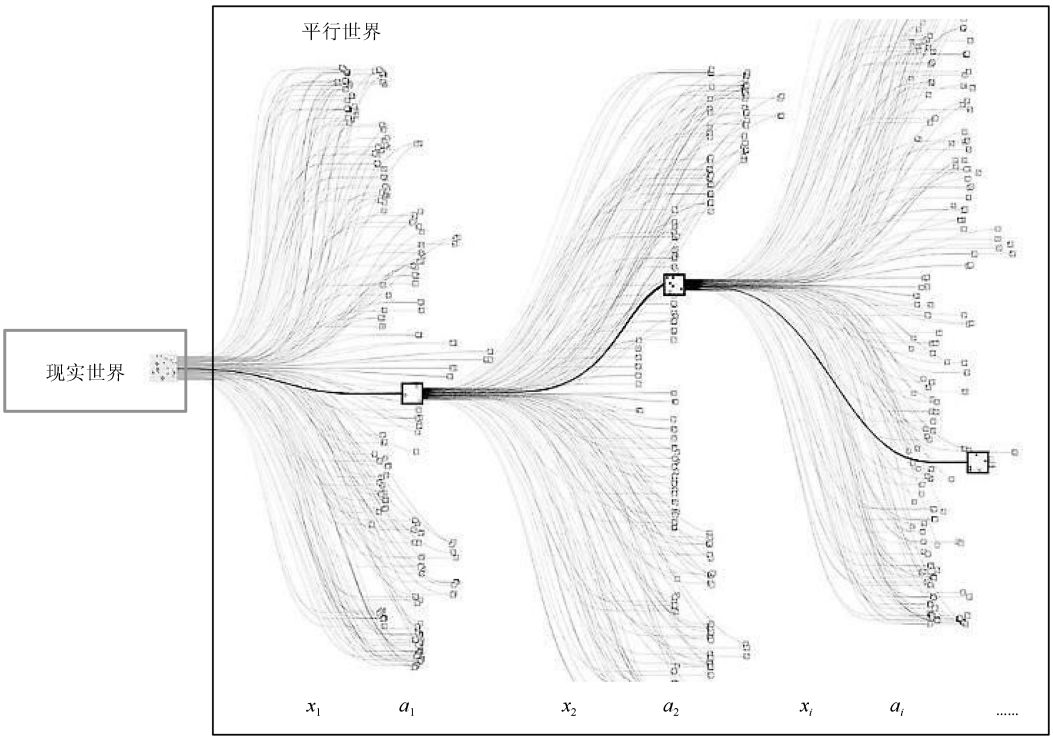

图 2 AlphaGo 将现实世界的数据映射到平行世界, 进行多线迭代来求取预期行动

Fig. 2 AlphaGo maps data in realistic world into parallel world and uses multithread iterations to determine the expected actions

分散学习机制要求每个智能体可以独立根据自身获取的数据来进行行动, 每一个行动可以导致一个回报 $R(a_k^n)$. 我们的目标是最大化所有智能体的总体长期回报

$$\max_{a_k^m,1\le k\le J,1\le m\le N}\sum_{m=1}^{N}\sum_{k=1}^{J}R^m(a_k^m) \tag{7}$$

该学习机制适合于分散控制等问题, 其中每个智能体获得的数据和执行的行动都是时空局部、甚至可以时间异步的. 例如, 我们在分布式平行交通控制系统中设定不同路口的控制器进行迭代和学习, 协同式产生和发现最优的交通信号灯控制策略[7−8].

集成学习机制则要求设置一个所谓代理智能体 (Surrogate agent). 每个智能体可以独立地根据自身获取的数据来进行行动, 每一个行动可以导致一个回报 $R^m(a_k^m)$. 我们的目标是选取所有这些智能体可能获得的回报中最大的那个动作, 并让代理智能体执行该动作.

$$\arg\max_{a_k^m,1\le k\le J,1\le m\le N}\sum_{m=1}^{N}R^m(a_k^m) \tag{8}$$

该学习机制类似于现有的集成学习 (Ensemble learning)[39], 可适合于静态数据的集中分类等问题. 而我们的研究表明, 更为广义的集成学习技术也可以应用于动态系统的建模和控制之中[40]. 其中每个智能体都致力于在学习中对复杂系统进行建模并执行特定的行动来达到某一优化目标. 各智能体执行的行动是时间同步, 但所获取的数据可以是时空局部的. 特别地, 我们可以设定策略使得每个 Agent 根据自身偏好来探索输入数据集的不同部分, 并建立适合特定输入数据子集的行动策略[40]. 这一想法亦被我们用在所谓平行动态规划 (Parallel dynamic programming) 之中, 以克服现有近似动态规划算法对于解空间探索不足的困难[41].

4.3 基于默顿定律实现数据 — 行动引导的指示学习

一般而言, 复杂系统可以大致分为牛顿系统和默顿系统两种. 无论我们怎么对牛顿系统进行分析, 都不会影响系统运行的结果. 例如天气预报明天下雨, 那么天气是下雨下雪还是刮风, 跟我们的分析调控无关. 而在默顿系统中, 存在双向影响通路, 我们对系统的调控会影响系统运行的结果. 例如, 著名分析师对于股市的评论会影响股市的波动. 因此, 我们可以设定预期的系统状态, 通过对于系统的描述、预测、引导来使得系统达到控制者所期望的状态. 美国社会学家默顿将这一长期行动称为预言的自我实现定律 (Self-fulfilling prophecy)[42]. 该定律也被后人称为默顿定律.

默顿定律希望通过改变行动的模式, 即实际系统与人工系统的平行互动, 促使实际系统运行到既定目标. 然而与牛顿系统不同, 默顿系统中的 “行动建模” 与 “目标建模” 相互独立, 且 “行动建模” 受到有限先验知识, 系统高度随机性等多方面因素的影响, 难以实现[43]. 如果用上面的学习框架来描述, 就是难以确知我们每一次探索所获得的即时回报 $R_a(x,a(x))$, 而且即时回报 $R_a(x,a(x))$ 和长期回报 $V(x)$ 之间的关系亦不明确.

幸运的是, 近 10 年来研究者先后提出了对抗学习 (Adversarial learning)[44−46]、对偶学习 (Dual learning)[47] 等全新的学习原则, 为解决上述问题提出了新的思路.

对抗学习通过构造相互竞争的生成器和辨别器来提高学习的效率[44−46]. 在图像学习中, 前者试图产生假的图像, 后者试图鉴别出真正的图像. 本质上, 这依然是费曼所称 “不是我创造的, 我就不能理解” 概念的体现, 即通过构建验证过的概念来理解事物. 但对抗学习最大的优点是系统的回报/损失函数不必显式给出, 而是通过生成器和辨别器的对抗来自动学习和挖掘来产生. 这是默顿定律所期待的, 在系统和行动的互动之中, 达成知识的 “泛化” (Generalization).

对偶学习[47] 的思路则更加偏重迭代演进. 假设学习过程中有两个智能体分别从事原任务 (从集合 X 到集合 Y 的学习任务) 和对偶任务 (从集合 Y 到集合 X 的学习任务). 假如我们首先把集合 X 用第一个智能体的模型 F 映射成集合 Y 的子集 Y', 再利用第二个智能体的模型 G 把集合 Y' 映射成集合 X 的子集 X'. 比较集合 X 和 X', 我们常常可以获得非常有用的反馈信号来改进映射模型 F 和 G. 以机器翻译为例, 上面的这个过程可以无限循环下去. 可以证明, 只要机器翻译模型 F 和 G 的解码部分都使用的是随机算法, 这一学习过程是收敛的, 最终会学到两个稳定有效的模型 F 和 G.

综合对抗学习和对偶学习, 我们可以提出如下更一般的指示学习. 指示学习关注如何设置引导, 使得我们获得预期的学习目的或者学习效果[48]. 具体到大数据环境中的数据获取 – 数据学习这一矛盾关系中, 指示学习可以被建模为反复循环的两阶段学习 – 探索过程.

假设我们不断地获取数据并构成集合 $X=\{x_i\}$, $i=1,\cdots,I$. 针对这些数据, 我们可以采取一系列行动并构成集合 $A=\{a_k(X)\}$, $k=1,\cdots,J$.

在第一阶段, 我们学习特定行动原则, 最大化阶段性回报 $V(X,a_k(X))$

$$\max_{a_k, 1\le k\le J} V(X, a_k(X)) \tag{9}$$

在第二阶段, 我们在此基础上通过数据生成算法产生新的数据集合 $X' = \{x'_i\}$, $i = 1, \cdots, I$, 并衡量阶段性损失 $L[X', a_k(X'))]$. 损失函数 $L[\cdot]$ 为已学会的行动原则在新数据集合 X' 上的损失, 该函数指示了我们下一步学习的方向. 对于对抗学习而言, $L[\cdot]$ 可以直接设为误分类比率. 对于对偶学习而言, $L[\cdot]$ 为迭代产生的映射误差. 对于平行系统控制而言, $L[\cdot]$ 通常可以设为当前系统状态和理想系统状态之间的差值[7−9].

当上述两个阶段过程完成后, 我们进入新的循环, 直到我们覆盖所有的数据集或系统状态, 又或者我们已经到达理想的系统状态.

目前, 强化学习已经为我们在第一阶段的数据学习提供了相当强的理论工具和方法. 但我们对于第二阶段的数据探索尚有很多未知的领域值得深入研究. 原始的对抗学习中, 数据产生的方式是无模型无监督学习, 但其效果有待改进[44]. 而对偶学习中, 数据产生的方式是有模型无监督学习, 理论证明清晰但限制较多[47]. 最近的研究表明, 类似于我们在上文中提及的平行仿真和预测学习这种半监督、具有限定条件和指示性目标的演化式学习 (例如利用信息熵的 InfoGAN[49]) 可能更加适合特定问题的解决. 这即是我们在上文中提到的平行预测 + 指示学习 (Parallel predictive + Prescriptive learning) 的思想的体现: 既有方向性指示, 亦不过多限制中间探索过程. 我们期待这一思想能带来更多新的机器学习方法.

5 总结

在本文中, 我们提出平行学习的基本思想和理论框架, 并阐述了平行学习的三大特色方法: 1) 通过软件定义的人工系统进行大数据预处理, 2) 包含预测学习、集成学习的数据学习, 和 3) 基于默顿定律实现数据 – 行动引导的指示学习.

概括而言, 平行学习理论框架强调: 使用预测学习解决如何随时间发展对数据进行探索; 使用集成学习解决如何在空间分布上对数据进行探索; 使用指示学习解决如何探索数据生成的方向. 目前我们的研究显示, 扩展现有的强化学习模型能够较好地和这三者结合. 但我们不排除今后有更好的模型来和我们提出的平行学习理论框架结合发展.

需要指出的是, 我们主要强调从理论框架层面改进和扩展已有机器学习方法, 并初步开展了相应模型和算法的研究. 例如我们将平行学习方法应用到了虚拟场景生成和无人驾驶车辆智能测试[50−51], 以及社会计算和情报处理[52−54]. 但很多细节之处尚需完备的理论证明. 期待本文抛砖引玉, 引起业内专家学者兴趣, 共同对机器学习理论做出更加深入的革新.

References

1 Zheng Nan-Ning. On challenges in artificial intelligence. *Acta Automatica Sinica*, 2016, **42**(5): 641−642
(郑南宁. 人工智能面临的挑战. 自动化学报, 2016, **42**(5): 641−642)

2 Wang Fei-Yue. Computational theory and method on complex system. *China Basic Science*, 2004, **6**(5): 3−10
(王飞跃. 关于复杂系统研究的计算理论与方法. 中国基础科学, 2004, **6**(5): 3−10)

3 Wang Fei-Yue. Artificial societies, computational experiments, and parallel systems: a discussion on computational theory of complex social-economic systems. *Complex Systems and Complexity Science*, 2004, **1**(4): 25−35
(王飞跃. 人工社会、计算实验、平行系统 — 关于复杂社会经济系统计算研究的讨论. 复杂系统与复杂性科学, 2004, **1**(4): 25−35)

4 Wang Fei-Yue. Parallel system methods for management and control of complex systems. *Control and Decision*, 2004, **19**(5): 485−489
(王飞跃. 平行系统方法与复杂系统的管理和控制. 控制与决策, 2004, **19**(5): 485−489)

5 Wang F Y, Lansing J S. From artificial life to artificial societies — new methods for studies of complex social systems. *Complex Systems and Complexity Science*, 2004, **1**(1): 33−41
(王飞跃, 史帝夫 · 兰森. 从人工生命到人工社会 — 复杂社会系统研究的现状与展望. 复杂系统与复杂科学, 2004, **1**(1): 33−41)

6 Zhang N, Wang F Y, Zhu F H, Zhao D B, Tang S M. DynaCAS: computational experiments and decision support for ITS. *IEEE Intelligent Systems*, 2008, **23**(6): 19−23

7 Wang F Y. Parallel control and management for intelligent transportation systems: concepts, architectures, and applications. *IEEE Transactions on Intelligent Transportation Systems*, 2010, **11**(3): 630−638

8 Li L, Wen D. Parallel systems for traffic control: a rethinking. *IEEE Transactions on Intelligent Transportation Systems*, 2016, **17**(4): 1179−1182

9 Wang F Y, Wang X, Li L X, Li L. Steps toward parallel intelligence. *IEEE/CAA Journal of Automatica Sinica*, 2016, **3**(4): 345−348

10 Jain A K, Duin R P W, Mao J C. Statistical pattern recognition: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(1): 4−37

11 Bishop C M. *Pattern Recognition and Machine Learning*. New York: Springer, 2006.

12 Shalev-Shwartz S. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 2011, **4**(2): 107−194

13 Rakhlin A, Sridharan K. Statistical learning and sequential prediction [Online], available: http://www-stat.wharton.upenn.edu/ rakhlin/courses/stat928/stat928_notes.pdf, January 1, 2017

14 Pérez-Sánchez B, Fontenla-Romero O, Guijarro-Berdiñas B. A review of adaptive online learning for artificial neural networks. *Artificial Intelligence Review*, 2016, DOI: 10.1007/s10462-016-9526-2

15 Wiering M, van Otterlo M. *Reinforcement Learning: State-of-the-Art*. Berlin Heidelberg: Springer, 2012.

16 Sutton R S, Barto A G. *Reinforcement Learning: An Introduction* (Second edition). Cambridge, MA: The MIT Press, 2017.

17 Settles B. *Active Learning*. San Rafael, CA: Morgan & Claypool Publishers, 2012.

18 Wang F Y. A big-data perspective on AI: Newton, Merton, and analytics intelligence. *IEEE Intelligent Systems*, 2012, **27**(5): 2−4

19 Zeng D, Lusch R. Big data analytics: perspective shifting from transactions to ecosystems. *IEEE Intelligent Systems*, 2013, **28**(2): 2−5

20 Zhou Z H, Chawla N V, Jin Y C, Williams G J. Big data opportunities and challenges: discussions from data analytics perspectives. *IEEE Computational Intelligence Magazine*, 2014, **9**(4): 62−74

21 Richard Feynman [Online], available: https://en.wikiquote.org/wiki/Richard_Feynman, January 1, 2017

22 Bainbridge W S. The scientific research potential of virtual worlds. *Science*, 2007, **317**(5837): 472−476

23 Wang Kun-Feng, Gou Chao, Wang Fei-Yue. Parallel vision: an ACP-based approach to intelligent vision computing. *Acta Automatica Sinica*, 2016, **42**(10): 1490−1500
(王坤峰, 苟超, 王飞跃. 平行视觉: 基于 ACP 的智能视觉计算方法. 自动化学报, 2016, **42**(10): 1490−1500)

24 Gaidon A, Wang Q, Cabon Y, Vig E. Virtual worlds as proxy for multi-object tracking analysis. In: Proceedings of the 2016 Computer Vision and Pattern Recognition. Las Vegas, NV: CVPR, 2016. 4340−4349

25 Wang K, Gou C, Wang F Y, Rehg J M. Parallel vision for perception and understanding of complex scenes: methods, framework, and perspectives. *Pattern Recognition*, to be published

26 Tian Yuan-Dong. A simple analysis of AlphaGo. *Acta Automatica Sinica*, 2016, **42**(5): 671−675
(田渊栋. 阿法狗围棋系统的简要分析. 自动化学报, 2016, **42**(5): 671−675)

27 Wang F Y, Zhang J J, Zheng X H, Wang X, Yuan Y, Dai X X, Zhang J, Yang L Q. Where does AlphaGo go: from church-turing thesis to AlphaGo thesis and beyond. *IEEE/CAA Journal of Automatica Sinica*, 2016, **3**(2): 113−120

28 Silver D, Huang A, Maddison C J, Guez A, Sifre L, van den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M, Dieleman S, Grewe D, Nham J, Kalchbrenner N, Sutskever I, Lillicrap T, Leach M, Kavukcuoglu K, Graepel T, Hassabis D. Mastering the game of Go with deep neural networks and tree search. *Nature*, 2016, **529**(7587): 484−489

29 Hinton G, Sejnowski T J. *Unsupervised Learning: Foundations of Neural Computation*. Cambridge: MIT Press, 1999.

30 Duda R O, Hart P E, Stork D G. Unsupervised learning and clustering. *Pattern Classification* (Second edition). New York: Wiley, 2001. 517−600

31 LeCun Y. *Predictive Learning*. Speech at NIPS, 2016. https://drive.google.com/file/d/0BxKBnD5y2M8NREZod0tVdW5FLTQ/view

32 Piaget J. *The Origins of Intelligence in Children* (Second edition). New York: International Universities Press Inc., 1974.

33 Drescher G L. *Made-Up Minds: A Constructivist Approach to Artificial Intelligence*. Cambridge: MIT Press, 1991.

34 Hawkins J, Blakeslee S. *On Intelligence: How a New Understanding of the Brain will Lead to the Creation of Truly Intelligent Machines*. New York: Times Books, 2004.

35 Zhu X J, Goldberg A B. *Introduction to Semi-Supervised Learning*. San Francisco, CA: Morgan & Claypool Publishers, 2009.

36 Schwenker F, Trentin E. Pattern classification and clustering: a review of partially supervised learning approaches. *Pattern Recognition Letters*, 2014, **37**: 4−14

37 Minsky M. *The Society of Mind*. New York: Simon & Schuster, 1988.

38 Shoham Y, Leyton-Brown K. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge: Cambridge University Press, 2008.

39 Zhou Z H. *Ensemble Methods: Foundations and Algorithms*. Boca Raton, FL, USA: CRC Press, 2012.

40 Li L, Chen X Q, Zhang L. Multimodel ensemble for freeway traffic state estimations. *IEEE Transactions on Intelligent Transportation Systems*, 2014, **15**(3): 1323−1336

41 Wang F Y, Zhang J, Wei Q, Li L. PDP: parallel dynamic programming. *IEEE/CAA Journal of Automatica Sinica*, 2017, **4**(1): 1−5

42 Merton R K. The unanticipated consequences of purposive social action. *American Sociological Review*, 1936, **1**(6): 894−904

43 Wang Fei-Yue. Software-defined systems and knowledge automation: a parallel paradigm shift from Newton to Merton. *Acta Automatica Sinica*, 2015, **41**(1): 1−8
(王飞跃. 软件定义的系统与知识自动化: 从牛顿到默顿的平行升华. 自动化学报, 2015, **41**(1): 1−8)

44 Lowd D, Meek C. Adversarial learning. In: Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery in Data Mining. Chicago, Illinois: ACM, 2005.

45 Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. In: Proceedings of Advances in Neural Information Processing Systems 27: 28th Annual Conference on Neural Information Processing Systems 2014. Montreal, Canada: NIPS, 2014. 2672−2680

46 Goodfellow I J, Shlens J, Szegedy C. Explaining and harnessing adversarial examples [Online], available: https://arxiv.org/abs/1412.6572, January 1, 2017

47 Xia Y C, He D, Qin T, Wang L W, Yu N H, Liu T Y, Ma W Y. Dual learning for machine translation [Online], available: https://arxiv.org/abs/1611.00179, January 1, 2017

48 Ullrich C. Descriptive and prescriptive learning theories. *Pedagogically Founded Courseware Generation for Web-Based Learning: Lecture Notes in Computer Science*. Berlin Heidelberg: Springer, 2008, **5260**: 37−42

49 Chen X, Duan Y, Houthooft R, Schulman J, Sutskever I, Abbeel P. InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets [Online], available: https://arxiv.org/abs/1606.03657, January 1, 2017

50 Li Y C, Liu Y H, Zhang C, Zhao D C, Zheng N N. The "floor-wall" traffic scenes construction for unmanned vehicle simulation evaluation. In: Proceedings of the 17th IEEE Conference on Intelligent Transportation Systems. Qingdao, China: IEEE, 2014. 1726−1731

51 Li L, Huang W L, Liu Y H, Zheng N N, Wang F Y. Intelligence testing for autonomous vehicles: a new approach. *IEEE Transactions on Intelligent Vehicles*, 2016, **1**(2): 158−166

52 Wang Fei-Yue. A framework for social signal processing and analysis: from social sensing networks to computational dialectical analytics. *Scientia Sinica Informationis*, 2013, **43**(12): 1598−1611
(王飞跃. 社会信号处理与分析的基本框架: 从社会传感网络到计算辩证解析方法. 中国科学: 信息科学, 2013, **43**(12): 1598−1611)

53 Wang Fei-Yue. Parallel control: a method for data-driven and computational control. *Acta Automatica Sinica*, 2013, **39**(4): 293−302
(王飞跃. 平行控制: 数据驱动的计算控制方法. 自动化学报, 2013, **39**(4): 293−302)

54 Wang Fei-Yue. The destiny: towards knowledge automation — preface of the special issue for the 50th anniversary of Acta Automatica Sinica. *Acta Automatica Sinica*, 2013, **39**(11): 1741−1743
(王飞跃. 天命唯新: 迈向知识自动化 —《自动化学报》创刊 50 周年专刊序. 自动化学报, 2013, **39**(11): 1741−1743)

李 力 清华大学自动化系副教授. 主要研究方向为人工智能和机器学习, 智能交通系统和智能汽车.
E-mail: li-li@tsinghua.edu.cn
(**LI Li** Associate professor at Department of Automation, Tsinghua University. His research interest covers artificial intelligence and machine learning, intelligent transportation systems and intelligent vehicles.)

林懿伦 中国科学院自动化研究所复杂系统管理与控制国家重点实验室博士研究生. 主要研究方向为社会计算, 智能交通系统, 深度学习和强化学习.
E-mail: linyilun2014@ia.ac.cn
(**LIN Yi-Lun** Ph. D at the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interest covers social computing, intelligent transportation systems, deep learning and reinforcement learning.)

曹东璞 英国克兰菲尔德大学驾驶员认知与自动驾驶实验室主任. 中科院自动化所客座研究员. 主要研究方向为自动驾驶, 人车协同, 与平行驾驶.
E-mail: d.cao@cranfield.ac.uk
(**CAO Dong-Pu** Director of Driver Cognition and Automated Driving Laboratory, Cranfield University. Visiting professor at Institute of Automation, Chinese Academy of Sciences. His research interest covers automated driving, driver-automation collaboration and parallel driving.)

郑南宁 西安交通大学人工智能与机器人研究所教授. 中国工程院院士. 主要研究方向为模式识别与智能系统, 机器视觉与图象处理.
E-mail: nnzheng@mail.xjtu.edu.cn
(**ZHENG Nan-Ning** Professor at Institute of Artificial Intelligence and Robotics (IAIR), Xi′an Jiaotong University. Member of Chinese Academy of Engineering. His research interest covers pattern recognition and intelligent systems, computer vision and image processing.)

王飞跃 中国科学院自动化研究所复杂系统管理与控制国家重点实验室研究员. 国防科学技术大学军事计算实验与平行系统技术研究中心主任. 主要研究方向为智能系统和复杂系统的建模、分析与控制. 本文通信作者.
E-mail: feiyue.wang@ia.ac.cn
(**WANG Fei-Yue** Professor at the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. Director of the Research Center for Computational Experiments and Parallel Systems Technology, National University of Defense Technology. His research interest covers modeling, analysis, and control of intelligent systems and complex systems. Corresponding author of this paper.)