基于模拟的主动策略学习算法(高斯过程代理模型,EI 采集函数),应用于机器人导航和不确定性地点探索[12].Schneider 讨论了嵌入式学习系统的挑战和贝叶斯优化应用到嵌入式学习系统的发展前景[13].Akrour 等人利用局部环境的贝叶斯优化,在高维度空间(70 维)中控制机器人臂运动[75].Torun 等人提出一种两阶段贝叶斯优化方法(第 1 阶段注重不确定性区域探索,第 2 阶段根据当前探索区域寻找最优)优化集成系统设计[76].

4) 环境监控与传感器网络

传感器设备用于测量速度、温度、湿度、空气质量、污染物含量等环境指标.由于不能在所有区域布置传感器,再加上噪声的干扰,传感器测量的数据常常存在不确定性.此外,激活传感器设备进行环境感知都会消耗能量,如电量和传输流量.Srinivas 等人使用高斯过程代理的贝叶斯优化,通过仅激活少量的传感器,便可找到室内温度极值位置或高速公路上最堵位置[23].Garnett 等人使用贝叶斯优化选择最优传感器子集,使其根据这些子集得到最优的预测效果[24].Marchant 等人把贝叶斯优化扩展到环境监控中,利用可移动机器人在环境中进行主动采样,得到对周围环境的精确感知[14].Morere 等人结合贝叶斯优化和部分观测的马尔可夫决策过程,以优化无人机采样策略监测周围环境[77].Colopy 等人利用贝叶斯优化调整基于个体的个性化监测模型,以个性化地监控病人生命体征[78].Candelieri 等人利用贝叶斯优化来优化控制给水管网系统中的泵,以达到在少量能量消耗的情况下得到理想的泵调度方案的目的[79].

5) 偏好学习与交互界面

在处理计算机图形与动画领域中的问题时,通常需要专业人员手动调整大量棘手的参数.例如,构造烟雾场景的粒子系统,需要调整速度、半径、涡环大小、长度尺度、旋度噪音等参数.通常情况下,这些参数十分复杂,非专业人员难以理解.Brochu 等人提出一种使用贝叶斯优化的迭代选择方法.该方法在处理图片时不需要专业人员手动调参,只需在每次迭代时从生成的两张对比图片(两张对比图片具有不同的参数配置)中选取与目标更像的图片作为反馈(此时,用户知道最终想要的图片效果),不需要用户理解复杂参数的具体含义.该方法通过返回的对比偏好信息更新代理模型,并根据完全随机、EI 等策略生成下一次迭代的两张对比图片,直到找到满足需求的目标图片[9,10].

6) 自动算法配置

构造一种优秀的算法通常需要经过大量的参数调节实验.若算法的参数调节都需要人工干预,将花费大量的时间和人力,甚至做无用功.因此,自动算法配置十分必要.这样不仅能减少人工干预,使得人们能够更专注于新模型构建等高层次问题,还能缩短大量的训练时间.相比人工经验或穷举,优化算法会自动选择合适的参数配置进行训练验证.贝叶斯优化能够胜任这类问题,并已取得了令人瞩目的成果.Bergstra 等人应用贝叶斯优化自动地调整神经网络和深度信念网络中的超参数[17].Snoek 等人应用贝叶斯优化自动调整卷积神经网络中的超参数[18,19].Mahendran 等人提出一种基于贝叶斯优化的自适应马尔可夫链蒙特卡洛算法[20].Thornton 等人应用贝叶斯优化提出一种针对分类算法的自动模型选择和超参数调节的方法:Auto-WEKA[21].Zhang 等人使用贝叶斯优化对卷积神经网络中的参数进行调整,解决目标识别问题[15].Wang 等人通过贝叶斯优化调整混合整数规划求解器的参数来提升求解器的效率[16].Klein 等人提出一种快速贝叶斯优化方法,能够调节大规模数据集上的机器学习算法的超参数[80].Xia 等人应用贝叶斯优化调节决策树中的超参数,提高信用评价精度[81].

7) 自然语言与文本处理

Wang 等人使用贝叶斯优化对文本进行术语提取(term extraction)[61].Yogatama 等人利用贝叶斯优化为不同类问题选择合适的文本表示,其实验结果表明,该方法能够使优化后的线性模型与未优化的复杂模型在主题分类问题上具有可比的效率[82].

8) 生物、化学及晶体学

贝叶斯优化同样可以胜任在生物、化学及晶体学等领域中的高代价优化任务.Carr 等人应用贝叶斯优化技术在晶体表面上寻找分子最稳定的吸附位置[83].Kriválk 等人用贝叶斯优化提升配体成键位置的预测质量[84].Tanaka 等人利用贝叶斯优化进行全基因组选择,能够在少量的模拟代价下找到较理想的基因型[85].在脑年龄分类预测任务中,Lancaster 等人利用贝叶斯优化调节对神经影像预处理时所采用重采样技术的参数,进而达到提

高分类精度的目的[86].

9) 迁移学习

Ruder 等人在迁移学习过程中,利用贝叶斯优化技术从多源或多领域数据中自动地选择有效数据作为训练集,以达到增强模型能力的目的,且与具体学习模型无关[87].

5 问题与挑战

前面详细地介绍了贝叶斯优化的研究现状.然而,随着大数据应用的发展,待优化目标的规模和复杂程度将会有所增加.作为处理评估代价大的复杂黑箱问题的有效解决方法,贝叶斯优化在未来发展中将面临下列问题与挑战.

一. 实时性和自适应性

贝叶斯优化每次迭代需要对概率代理模型进行更新,当问题维度高或存在大量历史数据时,更新概率模型需要高昂的计算量,尤其不能满足对实时性要求高的实际任务.针对该问题,研究者已经提出了一些解决策略.

1) 降维映射,见第 4.1.1 节.当贝叶斯优化处理高维度问题时,需要从高维度空间映射到低维度空间进行优化,虽然该方法加快了求解效率,但是需要假设问题存在低有效维度的性质;
2) 近似方法,见第 3.1 节.当模型的先验不为共轭先验时,需要使用变分贝叶斯近似推断或蒙特卡洛采样方法得到模型近似后验分布.当使用高斯过程代理目标函数时,精确推断需要 $O(t^3)$的时间复杂度,可使用 Cholesky 分解、SPGP、SSGP 等方法对高斯过程进行近似推断.虽然这些近似方法能够加快求解效率,但却具有求解精度不足的缺点;
3) 并行化,见第 4.1.2 节.通过对贝叶斯优化进行并行化扩展,能够同时评估多次目标函数,加快求解效率.该策略选择评估点时,根据部分未完成评估的采样点返回的虚拟观测值,而不是真实观测值,会在一定程度上影响求解精度;
4) 时间敏感性,见第 4.1.2 节.时间敏感性主动选择策略能够选择单位时间期望提升最大的点进行评估.但该方法在相同迭代预算下,与传统方法相比,存在精度差异.在提高贝叶斯优化求解效率时,难点在于如何解决精度和计算开销之间的平衡关系.

此外,贝叶斯优化在处理优化目标动态变化的问题时,应该具有自适应的调整能力.在已有规划解的基础上,针对问题变化,动态调整现有策略,而不需要推倒重来,从头计算.例如:在交通领域中,当车辆前方发生不可预测的事件(如车祸)造成拥堵时,需要优化程序能够自适应地、增量地调整规划路线.

二. 分布式

随着数据量的增加,复杂应用很难在一台终端上高效执行.因此,贝叶斯优化还需要具有分布式处理数据的能力.贝叶斯优化的分布式扩展应具有以下特点.

1) 负载均衡.能够有效地利用计算资源,避免资源过于集中和浪费;
2) 具有高效的计算效率.目前,贝叶斯优化并行技术是为了加快其求解效率,同时进行多次函数评估,本质上是对采集函数的并行化扩展(见第 4.1.2 节).该方法仍存在集中环节,即集中回收评估点返回的观测值集合,然后整合更新概率模型决策候选点集合;
3) 高容错性和强健壮性.分布式计算中一个任务往往存在多个备份,一个备份所在终端失效后,其余备份仍可继续执行,从而实现任务的健壮性.与之不同的是,贝叶斯优化过程所要求的高容错性和强健壮性应能有效处理没有备份的任务,根据需要动态地进行优化策略调整.例如:在无人机对抗情景中,将每个无人机看作节点,这些无人机基于自组织、不可靠的通信网进行协同作战.当一架无人机被击落时,该小组应能动态调整队形,继续执行作战任务.这种去中心化的优化策略可以避免出现击毁中心机使整个小组瘫痪的情况;
4) 多策略分布式协同求解.贝叶斯优化的分布式扩展可同时存在多个不同的策略(不同的概率模型和采集函数),并像深度学习中的对抗网络一样,各个策略相互促进、相互影响,从而达到理想的学习效果.

然而,对贝叶斯优化分布式扩展的难点在于分布式概率代理模型和采集函数的构建,并且需要处理各个分散节点之间的信息交互问题.

三. 多目标

贝叶斯优化的多任务扩展能够处理多个相关任务,根据相关性,将一个任务的信息应用到其他相关任务上,从而达到迁移学习的目的.例如:第 4.1.1 节中,Swersky 等人使用高斯过程同时处理多个相关的超参数优化任务,为每一个任务得到最优的超参数配置,从而使系统性能最大化.该方法的优化目标是最优化所有任务的平均性能.但在实际应用中,许多问题需要同时优化多个目标,这些目标可能会存在"冲突"关系.例如:在智能交通应用中,既要规划出最短路径,又要尽量多地收集未知区域的道路情况,但这两个目标很难同时满足.第 4.1.2 节中介绍的约束扩展方法将两个目标中的一个作为优化目标,另一个作为约束处理.当目标间存在冲突时,不存在绝对最优解,只存在有效解集合.当把多目标转换成带约束的单目标优化时,求得的优化解仅是单目标的最优解,忽略了转化为约束的目标的重要程度.Tesch 等人提出一种面向多目标的贝叶斯优化方法,尽管该方法能够得到帕累托集,但忽略了目标之间的依赖关系[88].多目标贝叶斯优化的难点在于处理多个目标之间的关系.为了保证所有目标的重要性并利用目标之间的依赖关系,贝叶斯优化在求解多目标问题时可考虑同时拥有多个概率代理模型和采集函数,在优化过程中,这些概率模型和采集函数相互促进、相互影响,达到优化学习的目的.

四. 模型选择问题

模型选择一直是贝叶斯方法面临的棘手问题.贝叶斯优化涉及的模型选择有观测模型选择、(非)参数模型先验选择以及超参数先验选择.观测模型需根据领域知识指导选择.合理的观测模型需对错误假设具有鲁棒性,即:当真实数据与模型假设不相符时,其仍具有良好的表现.不同问题具有不同的性质,因而具有不同的先验形式.例如:在监测城市道路状况时,由于人们有早出晚归的习惯,通常道路状况会出现早晚高峰周期性的表现,因此,可以选择存在周期性质的协方差函数构造先验模型.然而,极端环境监测问题不具备这样的周期性质,因此需要选择其他合适的先验模型.当使用贝叶斯方法估计超参数时,需要选择合理的超参数先验,增加超参数估计精度,提升模型预测准确率.

在贝叶斯优化中,选择合适的概率代理模型甚至比采集函数的选择还要关键.在一些领域中,如制药和传染病控制,需要更加谨慎地选择合适的模型,提高概率模型预测的准确度,降低评估过程的代价.尽管目前存在一些模型选择的方法[29],但这些方法都不具有通用性,仍需针对具体问题具体分析,运用领域专家的经验知识指导模型的选择.在贝叶斯优化研究和应用领域,如何针对具体问题选择合适的概率代理模型,仍是具有很大挑战性的问题.

6 总 结

作为求解非凸、多峰、评估代价高昂、黑箱的复杂优化问题的有效解决方案,贝叶斯优化近年来在多领域获得了广泛关注.本文综述了贝叶斯优化的研究现状.

- 首先,从其优化框架和优化原理入手,详细分析其优势与劣势,以帮助相关领域研究者深入理解贝叶斯优化;然后,从模型选择的角度介绍了贝叶斯优化两个核心部分:概率代理模型和采集函数,旨在为建模求解复杂优化问题进行模型选择时提供参考;
- 其次,介绍了贝叶斯优化涉及的近似与优化技术,并深入到技术细节;
- 最后,总结了贝叶斯优化的方法扩展和当前主要应用领域.

同时,本文也关注随着待优化目标的规模和复杂程度的增加,贝叶斯优化将面临实时性和自适应性、分布式、多目标以及模型选择等问题与挑战.此外,相比于其他优化技术,贝叶斯优化还存在一些局限性.本文通过对贝叶斯优化的详细分析和讨论,希望为相关领域的研究者予以帮助.

References:

[1] Shahriari B, Swersky K, Wang Z, Adams RP, Freitas ND. Taking the human out of the loop: A review of Bayesian optimization. Proc. of the IEEE, 2016,104(1):148−175.

[2] Kohavi R, Longbotham R, Dan S, Henne RM. Controlled experiments on the Web: Survey and practical guide. Data Mining and Knowledge Discovery, 2009,18(1):140−181.

[3] Scott SL. A modern Bayesian look at the multi-armed bandit. Applied Stochastic Models in Business and Industry, 2010,26(6): 639−658.

[4] Chapelle O, Li L. An empirical evaluation of Thompson sampling. Advances in Neural Information Processing Systems, 2011, 2249−2257.

[5] Khajah MM, Roads BD, Lindsey RV, Liu YE, Mozer MC. Designing engaging games using Bayesian optimization. In: Proc. of the ACM Conf. on Human Factors in Computing Systems. 2016. 5571−5582.

[6] Frazier PI, Wang J. Bayesian optimization for materials design. In: Proc. of the Mathematics. 2015.

[7] Li L, Chu W, Langford J, Schapire RE. A contextual-bandit approach to personalized news article recommendation. In: Proc. of the Int'l Conf. on World Wide Web. 2010. 661−670.

[8] Vanchinathan HP, Nikolic I, Bona FD, Krause A. Explore-Exploit in top-*n* recommender systems via Gaussian processes. In: Proc. of the ACM Conf. on Recommender Systems. 2014. 31.

[9] Brochu E, Brochu T, Freitas ND. A Bayesian interactive optimization approach to procedural animation design. In: Proc. of the ACM SIGGRAPH/Eurographics Symp. on Computer Animation. 2010. 103−112.

[10] Brochu E, Cora VM, Freitas ND. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. In: Proc. of the Computer Science. 2010.

[11] Lizotte D, Wang T, Bowling M, Schuurmans D. Automatic gait optimization with Gaussian process regression. In: Proc. of the Int'l Joint Conf. on Artifical Intelligence. 2007. 944−949.

[12] Martinez-Cantin R, Freitas ND, Doucet A, Castellanos JA. Active policy learning for robot planning and exploration under uncertainty. In: Proc. of the Robotics: Science and Systems III. 2007. 321−328.

[13] Schneider J. Bayesian optimization and embedded learning systems. In: Proc. of the ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. 2016. 413−413.

[14] Marchant R, Ramos F. Bayesian optimisation for intelligent environmental monitoring. In: Proc. of the IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems. 2012. 2242−2249.

[15] Zhang Y, Sohn K, Villegas R, Pan G, Lee H. Improving object detection with deep convolutional networks via Bayesian optimization and structured prediction. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2015. 132−132.

[16] Wang Z, Zoghi M, Hutter F, Matheson D, Freitas ND. Bayesian optimization in a high dimensions via random embeddings. In: Proc. of the Int'l Joint Conf. on Artificial Intelligence. 2013.

[17] Bergstra J, Bardenet R, Bengio Y, Kégl B. Algorithms for hyper-parameter optimization. Advances in Neural Information Processing Systems, 2011,24(24):2546−2554.

[18] Snoek J, Larochelle H, Adams RP. Practical Bayesian optimization of machine learning algorithms. Advances in Neural Information Processing Systems, 2012,4:2951−2959.

[19] Swersky K, Snoek J, Adams RP. Multi-Task Bayesian optimization. Advances in Neural Information Processing Systems, 2013, 2004−2012.

[20] Mahendran N, Wang Z, Hamze F, Freitas ND. Adaptive MCMC with Bayesian optimization. In: Proc. of the Int'l Conf. on Artificial Intelligence and Statistics. 2010.

[21] Thornton C, Hutter F, Hoos HH, Leyton-Brown K. Auto-WEKA: Combined selection and hyperparameter optimization of classification algorithms. In: Proc. of the Computer Science. 2013. 847−855.

[22] Hoffman MW, Shahriari B, Freitas ND. On correlation and budget constraints in model-based bandit optimization with application to automatic machine learning. In: Proc. of the Int'l Conf. on Artificial Intelligence and Statistics. 2014. 365−374.

[23] Srinivas N, Krause A, Kakade SM, Seeger M. Gaussian process optimization in the bandit setting: No regret and experimental design. In: Proc. of the Int'l Conf. on Machine Learning. 2010.

[24] Garnett R, Osborne MA, Roberts SJ. Bayesian optimization for sensor set selection. In: Proc. of the Int'l Conf. on Information Processing in Sensor Networks. 2010. 209−219.

[25] Ghahramani Z. Probabilistic machine learning and artificial intelligence. Nature, 2015,521:452−459.

[26] Jones DR, Schonlau M, Welch WJ. Efficient global optimization of expensive black-box functions. Journal of Global Optimization, 1998,13(4):455−492.

[27] Nelder JA, Baker RJ. Generalized linear models. Journal of the Royal Statistical Society, 1972,135(3):370−384.

[28] Sutskever I, Vinyals O, Le QV. Sequence to sequence learning with neural networks. Advances in Neural Information Processing Systems, 2014,4:3104−3112.

[29] Rasmussen CE, Williams CKI. Gaussian Processes for Machine Learning. The MIT Press, 2006.

[30] Lu C, Tang X. Surpassing human-level face verification performance on LFW with GaussianFace. In: Proc. of the Computer Science. 2014.

[31] Neal RM. Bayesian learning for neural networks [Ph.D. Thesis]. Toronto: University of Toronto, 1996.

[32] Paciorek CJ, Schervish MJ. Nonstationary covariance functions for Gaussian process regression. Advances in Neural Information Processing Systems, 2003,16:273−280.

[33] Hutter F, Hoos HH, Leyton-Brown K. Sequential model-based optimization for general algorithm configuration. In: Proc. of the Conf. on Learning and Intelligent Optimization. 2011. 507−523.

[34] Watson, GN. A Treatise on the Theory of Bessel Functions. 2nd ed., London: Cambridge University Press, 1966.

[35] Breiman L. Random forests. Machine Learning, 2001,45(1):5−32.

[36] Zhang Y, Chan W, Jaitly N. Very deep convolutional networks for end-to-end speech recognition. In: Proc. of the Int'l Conf. on Acoustics, Speech and Signal Processing. 2017.

[37] Karpathy A, Li FF. Deep visual-semantic alignments for generating image descriptions. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2014,39(4):664−676.

[38] Snoek J, Rippel O, Swersky K, Kiros R, Satish N, Sundaram N, Patwary MMA, Prabhat, Adams RP. Scalable Bayesian optimization using deep neural networks. In: Proc. of the Statistics. 2015. 1861−1869.

[39] Springenberg JT, Klein A, Falkner S, Hutter F. Bayesian optimization with robust Bayesian neural networks. Advances in Neural Information Processing Systems, 2016.

[40] Kushner HJ. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. Journal of Fluids Engineering, 1963,86(1).

[41] Jones DR. A taxonomy of global optimization methods based on response surfaces. Journal of Global Optimization, 2001,21(4): 345−383.

[42] Mockus J, Tiesis V, Zilinskas A. The application of Bayesian methods for seeking the extremum. In: Proc. of the Towards Global Optimisation 2. 1978. 117−129.

[43] Lizotte DJ. Practical Bayesian optimization [Ph.D. Thesis]. Alberta: University of Alberta, 2008.

[44] Lai TL, Robbins H. Asymptotically efficient adaptive allocation rules. Advances in Applied Mathematics, 1985,6(1):4−22.

[45] Thompson WR. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. Biometrika, 1933,25(3-4):285−294.

[46] Shahriari B, Wang Z, Hoffman MW, Bouchard-Côté A. An entropy search portfolio for Bayesian optimization. In: Proc. of the Conf. on Neural Information Processing Systems: Workshop on Bayesian Optimization in Academia and Industry. 2014.

[47] Lázaro-Gredilla M, QuiñOnero-Candela J, Rasmussen CE, Figueiras-Vidal AR. Sparse spectrum Gaussian process regression. Journal of Machine Learning Research, 2010,11(9):1865−1881.

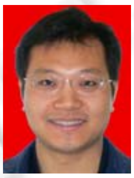
[48] Villemonteix J, Vazquez E, Walter E. An informational approach to the global optimization of expensive-to-evaluate functions. Journal of Global Optimization, 2009,44(4):509−534.

[49] Hernándezlobato JM, Hoffman MW, Ghahramani Z. Predictive entropy search for efficient global optimization of black-box functions. In: Proc. of the Conf. on Neural Information Processing Systems: Workshop on Bayesian Optimization in Academia and Industry. 2014.

[50] Brochu E, Hoffman M, Freitas ND. Portfolio allocation for Bayesian optimization. In: Proc. of the Conf. on Uncertainty in Artificial Intelligence. 2011.
[51] Tzikas DG, Likas CL, Galatsanos NP. The variational approximation for Bayesian inference. IEEE Signal Processing Magazine, 2008,25(6):131−146.
[52] Seeger M, Williams CKI, Lawrence ND. Fast forward selection to speed up sparse Gaussian process regression. In: Proc. of the Conf. on Artificial Intelligence and Statistics. 2003.
[53] Snelson E, Ghahramani Z. Sparse Gaussian process using pseudo-inputs. Advances in Neural Information Processing Systems, 2006,18(1):1257−1264.
[54] Martinez-Cantin R. BayesOpt: A Bayesian optimization library for nonlinear optimization, experimental design and bandits. Journal of Machine Learning Research, 2014,15:3735–3739.
[55] Osborne MA, Garnett R, Roberts SJ. Gaussian processes for global optimization. In: Proc. of the Int’l Conf. on Learning and Intelligent Optimization. 2009.
[56] Rasmussen CE, Ghahramani Z. Bayesian Monte Carlo. Advances in Neural Information Processing Systems, 2002.
[57] Osborne MA, Roberts SJ, Rogers A, Ramchurn SD, Jennings NR. Towards real-time information processing of sensor network data using computationally efficient multi-output Gaussian processes. In: Proc. of the Int’l Conf. on Information Processing in Sensor Networks. 2008. 109−120.
[58] Bardenet R, Kégl B. Surrogating the surrogate: Accelerating Gaussian-process-based global optimization with a mixture cross-entropy algorithm. In: Proc. of the Int’l Conf. on Machine Learning. 2010.
[59] Jones DR, Perttunen CD, Stuckman BE. Lipschitzian optimization without the Lipschitz constant. Journal of Optimization Theory and Applications, 1993,79(1):157−181.
[60] Hansen N, Ostermeier A. Completely derandomized self-adaptation in evolution strategies. IEEE Trans. on Evolutionary Computation, 2001,9(2):159−195.
[61] Wang Z, Shakibi B, Jin L, Freitas ND. Bayesian multi-scale optimistic optimization. In: Proc. of the Int’l Conf. on Artificial Intelligence and Statistics. 2014. 1005−1014.
[62] Qian H, Hu YQ, Yu Y. Derivative-Free optimization of high-dimensional non-convex functions by sequential random embeddings. In: Proc. of the Int’l Joint Conf. on Artificial Intelligence. 2016.
[63] Li CL, Kandasamy K, Poczos B, Schneider J. High dimensional Bayesian optimization via restricted projection pursuit models. In: Proc. of the Int’l Conf. on Artificial Intelligence and Statistics. 2016.
[64] Wang Z, Li C, Jegelka S, Kohli P. Batched high-dimensional Bayesian optimization via structural kernel learning. In: Proc. of the Int’l Conf. on Machine Learning. 2017.
[65] Gardner JR, Guo C, Weinberger KQ, Garnett R, Grosse R. Discovering and exploiting additive structure for Bayesian optimization. In: Proc. of the Int’l Conf. on Artificial Intelligence and Statistics. 2017.
[66] Li C, Gupta S, Rana S, Nguyen V, Venkatesh S, Shilton A. High dimensional Bayesian optimization using dropout. In: Proc. of the Int’l Joint Conf. on Artificial Intelligence. 2017.
[67] Bonilla EV, Agakov FV, Williams CKI. Kernel multi-task learning using task-specific features. In: Proc. of the Int’l Conf. on Artificial Intelligence and Statistics. 2007.
[68] Bonilla EV, Chai KMA, Williams CKI. Multi-Task Gaussian process prediction. Advances in Neural Information Processing Systems, 2007.
[69] Swersky K, Snoek J, Adams RP. Freeze-Thaw Bayesian optimization. Eprint Arxiv, 2014.
[70] Gelbart MA, Snoek J, Adams RP. Bayesian optimization with unknown constraints. In: Proc. of the Computer Science. 2014.
[71] Kandasamy K, Dasarathy G, Oliva J, Schneider J, Poczos B. Gaussian process bandit optimisation with multi-fidelity evaluations. Advances in Neural Information Processing Systems, 2016.
[72] Marco A, Berkenkamp F, Hennig P, Schoellig AP, Krause A, Schaal S, Trimpe S. Virtual vs. real: Trading off simulations and physical experiments in reinforcement learning with Bayesian optimization. In: Proc. of the Int’l Conf. on Robotics and Automation. 2017.

[73] Ginsbourger D, Riche RL, Carraro L. Kriging is well-suited to parallelize optimization. In: Proc. of the Computational Intelligence in Expensive Optimization Problems. 2010. 131−162.
[74] Hutter F, Hoos HH, Leyton-Brown K. Parallel algorithm configuration. In: Proc. of the Int'l Conf. on Learning and Intelligent Optimization. 2012. 55−70.
[75] Akrour R, Sorokin D, Peters J, Neumann G. Local Bayesian optimization of motor skills. In: Proc. of the Int'l Conf. on Machine Learning. 2017.
[76] Torun HM, Swaminathan M, Davis AK, Bellaredj MLF. A global Bayesian optimization algorithm and its application to integrated system design. IEEE Trans. on Very Large Scale Integration Systems, 2018, 1−11.
[77] Morere P, Marchant R, Ramos F. Sequential Bayesian optimization as a POMDP for environment monitoring with UAVs. In: Proc. of the Int'l Conf. on Robotics and Automation. 2017. 6381−6388.
[78] Colopy GW, Roberts SJ, Clifton DA. Bayesian optimization of personalized models for patient vital-sign monitoring. IEEE Journal of Biomedical and Health Informatics, 2018,22(2):301.
[79] Candelieri A, Perego R, Archetti F. Bayesian optimization of pump operations in water distribution systems. Journal of Global Optimization, 2018.
[80] Klein A, Falkner S, Bartels S, Henning P, Hutter F. Fast Bayesian optimization of machine learning hyperparameters on large datasets. In: Proc. of the Int'l Conf. on Artificial Intelligence and Statistics. 2017.
[81] Xia Y, Liu C, Li YY, Liu N. A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring. Expert Systems with Applications, 2017,78:225−241.
[82] Yogatama D, Kong L, Smith NA. Bayesian optimization of text representations. In: Proc. of the Conf. on Empirical Methods in Natural Language Processing. 2015. 2100−2105.
[83] Carr S, Garnett R, Lo C. BASC: Applying Bayesian optimization to the search for global minima on potential energy surfaces. In: Proc. of the Int'l Conf. on Machine Learning. 2016.
[84] Kriván R, Hoksza D, Škoda P. Improving quality of ligand-binding site prediction with Bayesian optimization. In: Proc. of the Int'l Conf. on Bioinformatics and Biomedicine. 2017. 2278−2279.
[85] Tanaka R, Iwata H. Bayesian optimization for genomic selection: A method for discovering the best genotype among a large number of candidates. Theoretical and Applied Genetics, 2017,131(1):1−13.
[86] Lancaster J, Lorenz R, Leech R, Cole JH. Bayesian optimization for neuroimaging pre-processing in brain age classification and prediction. In: Proc. of the Frontiers in Aging Neuroscience. 2018.
[87] Ruder S, Plank B. Learning to select data for transfer learning with Bayesian optimization. In: Proc. of the Conf. on Empirical Methods in Natural Language Processing. 2017. 372−382.
[88] Tesch M, Schneider J, Choset H. Expensive multiobjective optimization and validation with a robotics application. In: Proc. of the Conf. on Neural Information Processing Systems: Workshop on Bayesian Optimization and Decision Making. 2012.

崔佳旭(1990－),男,黑龙江齐齐哈尔人,博士生,主要研究领域为贝叶斯优化,复杂网络优化.



杨博(1974－),男,博士,教授,博士生导师,CCF 杰出会员,主要研究领域为知识发现与知识工程,网络分析理论与应用,推荐系统,多智能体系统.