

基于学习-推理的约束求解方法研究进展*

邹 悅^{1,3}, 赖家洋^{1,3}, 张永刚^{2,3}



¹(吉林大学 软件学院, 吉林 长春 130012)

²(吉林大学 计算机科学与技术学院, 吉林 长春 130012)

³(符号计算与知识工程教育部重点实验室(吉林大学), 吉林 长春 130012)

通信作者: 张永刚, E-mail: zhangyg@jlu.edu.cn

摘要: 机器学习与自动推理的融合是当前人工智能研究的新趋势。约束满足问题是人工智能研究的经典问题, 现实世界中大量的调度、规划和配置等问题均可以建模为约束满足问题, 高效的求解算法一直是研究热点。近年来涌现出众多将机器学习应用于约束满足问题求解的新方法, 这些基于“学习-推理”的新方法为约束满足问题求解开辟了新方向并展示出巨大发展潜力, 方法的突出优点是适应性强、可在线优化并具有更强的可扩展性。将当前的“学习-推理”方法分为基于消息传递神经网络、基于序列到序列和基于最优化等3类进行综述, 详细分析各类方法的特点和在不同的问题集上求解效果, 尤其对每类方法所涵盖的相关工作进行多角度的对比分析。最后, 对基于“学习-推理”的约束求解方法进行总结和展望。

关键词: 约束满足问题; 消息传递神经网络; 序列到序列; 强化学习; 最优化

中图法分类号: TP18

中文引用格式: 邹悦, 赖家洋, 张永刚. 基于学习-推理的约束求解方法研究进展. 软件学报, 2024, 35(1): 220–235. <http://www.jos.org.cn/1000-9825/6844.htm>

英文引用格式: Zou Y, Lai JY, Zhang YG. Methods for Constraint Solving Problems Based on Learn to Reason Model: A Review. Ruan Jian Xue Bao/Journal of Software, 2024, 35(1): 220–235 (in Chinese). <http://www.jos.org.cn/1000-9825/6844.htm>

Methods for Constraint Solving Problems Based on Learn to Reason Model: A Review

ZOU Yue^{1,3}, LAI Jia-Yang^{1,3}, ZHANG Yong-Gang^{2,3}

¹(College of Software, Jilin University, Changchun 130012, China)

²(College of Computer Science and Technology, Jilin University, Changchun 130012, China)

³(Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education (Jilin University), Changchun 130012, China)

Abstract: The integration of machine learning and automatic reasoning is a new trend in artificial intelligence. Constraint satisfaction is a classic problem in artificial intelligence. A large number of scheduling, planning, and configuration problems in the real world can be modeled as constraint satisfaction problems, and efficient solving algorithms have always been a research hotspot. In recent years, many new methods of applying machine learning to solve constraint satisfaction problems have emerged. These methods based on “learn to reason” open up new directions for solving constraint satisfaction problems and show great development potential. They are featured by better adaptability, strong scalability, and online optimization. This study divides the current “learn to reason” methods into three categories including message-passing neural network-based, sequence-to-sequence-based, and optimization-based methods. Additionally, the characteristics of various methods and their solution effects on different problem sets are analyzed in detail. In particular, a comparative analysis is conducted on relevant work involved in each type of method from multiple perspectives. Finally, the constraint solving method based on “learn to reason” is summarized and prospected.

* 基金项目: 国家自然科学基金(62076108, 61872159); 吉林省自然科学基金(20210101172JC)

收稿时间: 2022-05-18; 修改时间: 2022-06-25, 2022-09-18, 2022-11-14; 采用时间: 2022-12-02; jos 在线出版时间: 2023-07-28

CNKI 网络首发时间: 2023-07-31

Key words: constraint satisfaction problem (CSP); message passing neural network; sequence to sequence (Seq2Seq); reinforcement learning; optimization

1 引言

约束满足问题 (constraint satisfaction problems, CSPs)^[1]是人工智能领域研究的热点问题, 现实世界中大量时间表问题^[2]、资源分配问题^[3]和车辆路径规划问题^[4,5]等都可建模为约束满足问题。一个约束满足问题包括一组给定论域的变量集合和变量间的约束集合, 约束求解的目标是找到一组满足所有约束的变量的赋值, 即问题的解。经典的约束满足问题求解算法是回溯搜索结合约束传播的算法^[6], 回溯搜索算法执行深度优先搜索, 尝试通过迭代地选择变量并融合约束传播和变量值选择启发式来找到一个解。近年来, 局部搜索算法^[7]因适合规模较大的问题且求解效率高, 被广泛应用于求解约束满足问题。

然而, 这些具有鲜明“推理”特色的求解方法在实际应用中往往会受到限制, 如求解时间受限、启发式策略需要手动设置, 另外启发式的参数设置需要大量领域知识, 无形中增加了时间开销^[8,9]。随着约束满足问题在航空、航天、交通等领域的重要性日益凸显, 大规模约束满足问题的高效求解算法变得至关重要。同时, 学习和推理一直是人工智能中的核心研究内容, 学术界普遍认为: 两者的结合将是未来人工智能发展的新趋势, 著名人工智能学者周志华教授在首届 2020 年国际学习与推理联合大会 (IJCLR) 上从机器学习的角度做了“利用无标签数据: 从‘纯学习’到‘学习 + 推理’”的主旨演讲, 也强调了学习和推理融合研究的重要性^[10-12]。近年来, 在约束满足问题求解算法研究领域涌现出许多基于“学习-推理”^[13,14]的方法, 主要包含机器学习的技术和方法, 尤其是借助深度学习、强化学习和最优化的新方法^[15,16], 为提升约束满足问题求解效率开拓了新思路。“学习-推理”模型是将“学习”作为推理过程的一个重要组成部分, 学习和推理从多个角度融合在一起, 以提升约束求解的效率。求解约束满足问题的挑战之一是难以抓住具体领域问题的特征, 因为这需要专家经验和领域知识, 这在现实中是很难做到的, 而基于“学习-推理”的约束求解方法的基本原理是学习推理过程中的知识并利用这些知识进行推理。

本文第 2 节介绍约束满足问题和经典的求解方法, 并对学习-推理模型进行概述。第 3 节对当前主流的基于消息传递神经网络的约束求解方法进行综述, 并分析和对比各种方法的优缺点。第 4 节阐述基于序列到序列的约束求解方法, 并对不同方法的求解性能进行比较和分析。第 5 节总结几种基于最优化思想的约束求解方法。第 6 节对本文进行总结和展望。

2 背景知识

本节主要介绍约束满足问题的基本概念和经典的求解方法, 并对“学习-推理”模型的主要思想进行总结分析。

2.1 约束满足问题

约束满足问题^[1]定义为一个三元组 $P = \langle X, D, C \rangle$:

$$\begin{cases} X = \{x_1, x_2, \dots, x_n\} \\ D = \{D_1, D_2, \dots, D_n\}, \\ C = \{c_1, c_2, \dots, c_m\} \end{cases}$$

其中, X 表示 n 个变量的集合, i 是变量的索引, D 表示每个变量 x_i 的值域集合, D_i 不能为空集, C 表示 m 个约束的集合, 每个约束 $c_j \in C$ 都是一个变量对 $c_j = \langle t_j, R_j \rangle$, t_j 是 X 的子集, 包含 k 个变量, R_j 是 k 个变量间应满足的约束关系。以下是 3 类典型的约束满足(优化)问题。

(1) 图着色问题 (graph coloring problem, GCP)^[17]给定 k 种颜色, 约束条件是图中任意相邻的 2 个顶点涂不同颜色, 求解满足约束下的顶点着色方案。

(2) 布尔可满足性问题 (Boolean satisfiability problem, SAT)^[18]基本形式是一个表示为合取范式形式的布尔公式, 判断是否存在变量的一组赋值使布尔公式为真。如果存在, 则称该布尔公式是可满足的, 并求解出可满足的一组或多组赋值, 否则称该公式是不可满足的。

(3) 旅行商问题 (traveling salesman problem, TSP)^[19,20]是一个典型的约束优化问题。给定若干城市，由某一个城市出发，将所有城市都访问一遍并且最终返回起点，求解满足上述条件且距离最短的一条路径。

2.2 求解方法

经典的约束满足问题求解方法包括系统化搜索方法，如遍历部分解空间的回溯方法；相容性检查技术，如节点和弧相容技术^[21]；局部搜索方法，如探测完全赋值空间的爬山法和遗传算法等^[1,22]。本节将主要介绍回溯搜索方法、相容性检查技术和局部搜索方法。

回溯算法是对动态构造的搜索树进行深度优先遍历，它的思想是通过不断为变量选择一个与当前部分解相容的值来增量地将部分解扩展到完全解，首先根据某种变量排序启发式选择未被实例化的变量将其实例化，当与某条约约相关的变量都被实例化时，立即检查该约束的相容性。如果当前部分解违反了任何一条约束，则回溯到上一个变量，为其选择其他赋值，重复该过程直至所有变量都被赋值，即找到问题的一个解。

相容性技术的核心思想是过滤掉不可能成为解的取值，从而减少变量的值域。相容性算法又称约束传播算法，主要包括节点相容、弧相容、边界相容和路径相容算法。相容性检查技术是以约束传播^[23]的方式实现的，约束传播控制着相容性检查程度和检查顺序。约束传播是执行某种形式的局部相容性的方法。一般情况下，对于通用约束求解算法，相容性技术是普遍适用的，因此对弧相容检查等约束传播技术的研究一直是领域的热点问题。

局部搜索方法是一种不完备的求解算法，即在问题有解的情况下，也不能保证求解出问题的解。它的思想是为问题初始化一个完全的变量赋值，通过在不断地迭代中，更改部分变量的取值，从而增加当前变量赋值所能满足的约束数量^[24]。

经过多年发展，经典方法在约束满足问题的求解中都表现优异。例如著名的 TSP 求解器 Concorde^[25]利用分支定界法求解 TSP 问题，得到了全局最优解。当前研究致力于让模型自动学习约束满足问题的特征和约束，并根据约束对问题进行推理求解。

2.3 “学习-推理”模型

学习和推理一直是人工智能领域中两个重要的概念，两者的结合是人工智能发展的现实选择和发展趋势，研究空间巨大，为约束求解方法提出了新的思路。“学习”体现在从解决方案中提取约束和偏好，“推理”体现在从学习到的约束和偏好中求出全局最优解，二者相辅相成，互为补充。当前，“学习-推理”方法的研究主要表现为以下 3 种类型（具体模型如图 1 所示）。

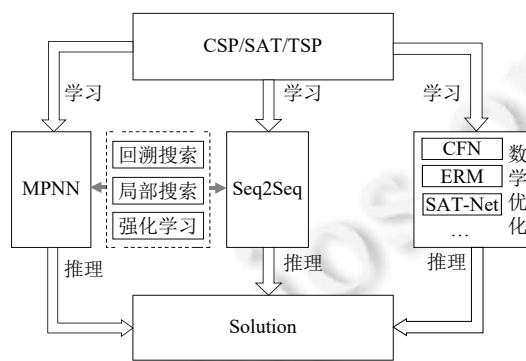


图 1 基于“学习-推理”的约束求解算法模型

(1) 采用二分图模型对约束满足问题进行建模并用消息传递神经网络学习图节点信息^[26]。这种模式既可以实现端到端的求解，也可以结合经典算法，采用强化学习方法学习搜索启发式。

(2) 将约束满足问题看作一个序列的输入到另一个序列的输出，序列的输入对应问题的学习与建模，序列的输出对应问题推理与预测。此外，基于策略和基于价值的 Actor-Critic 思想很好地解决了数据集的标签获取困难问题。

(3) 采用最优化模型对约束满足问题建模和求解，包括半正定规划方法、凸优化方法和经验风险最小化方法。

3 基于消息传递神经网络的约束求解方法

消息传递神经网络 (message passing neural network, MPNN) 是一个基于无向图的机器学习框架^[27], 其前向传递包括两个阶段: 消息传递阶段和读出阶段. 消息传递会进行多次迭代, 每次迭代过程定义如下.

$$\begin{cases} m_v^{t+1} = \sum_{w \in N(v)} M_t(h_v^t, h_w^t, e_{vw}) \\ h_v^{t+1} = U_t(h_v^t, m_v^{t+1}) \end{cases},$$

其中, M_t 是顶点接受消息函数, U_t 是顶点更新函数, 每一次迭代, 图中的每个顶点都根据其接受消息函数 M_t 来更新自己的隐藏状态 h_v . $N(v)$ 表示图 G 中顶点 v 的邻居顶点的集合, e_{vw} 表示顶点 v 和顶点 w 之间的边的特征.

读出阶段定义为:

$$\hat{y} = R(\{h_v^T \mid v \in G\}).$$

利用读出函数 R , 在读出阶段计算出从图中学习到的一个特征向量. 接受消息函数 M_t 、更新函数 U_t 和读出函数 R 都是被学习的可微函数.

目前约束求解领域的重要研究方向是将约束满足问题表示为图, 利用消息传递神经网络进行求解. 表 1 为现有基于消息传递神经网络的约束求解方法的总结.

表 1 基于 MPNN 算法模型、求解问题、测试数据及效果分析

序号	文献	基本思想	求解问题	测试数据集	效果分析
1	Selsam等人 ^[13] (2019)	MPNN	SAT	随机生成的平均40个变量、200个子句的SAT问题	实现85%的预测准确率. 与传统的高效求解器相比有一定差距, 但能求解规模更大、更难的问题
2	Liu等人 ^[28] (2020)	MPNN	PB	随机生成的0-1背包问题和、WIS问题、少于40个变量的SAT问题	预测准确率优于Selsam等人 ^[13] 提出的模型, 求解时间比CPLEX低10%
3	Lemos等人 ^[29] (2019)	MPNN	GCP	节点数在40和60之间的GCP问题	预测准确率优于Selsam等人 ^[13] 提出的模型
4	Cameron等人 ^[30] (2020)	MPNN	SAT	随机3-SAT	预测可满足性和赋值能实现75%–82%的准确率
5	Tönshoff等人 ^[31] (2021)	MPNN	MAXCSP	MAX-2-SAT、MAX-CUT、3-COLOR、MAX-IS	在MAX-2-SAT、MAX-CUT和3-COLOR问题上, 求解效果均优于部分传统的求解方法; 在MAX-IS问题上, 实现了和ReduMIS类似的效果, 明显优于贪婪启发式方法, 而在大规模问题上, 效果和ReduMIS有一定差距
6	Amizadeh等人 ^[32] (2019)	MPNN	SAT	随机4-SAT问题 (每个问题有100个变量)	实现无监督训练, 而Selsam等人 ^[13] 提出的模型在无监督训练集上的求解率为0; 与Glucose相比, 求解率相近, 但求解时间远慢于Glucose
7	Abboud等人 ^[33] (2020)	MPNN	#DNF	随机#DNF问题	与传统求解器KML相比, 求解速度更快, 求解率更高
8	Jaszczur等人 ^[34] (2019)	MPNN+CD CL(DPLL)+ Attention	SAT	随机SAT问题, 变量数在10–100之间	学习到的启发式求解效果优于DLIS启发式, 劣于JW-OS启发式
9	Yolcu等人 ^[35] (2019)	MPNN+SLS +RL	SAT	随机3-SAT、 k -clique、 k -cover、 k -coloring、 k -domset	训练数据与测试数据相同时, 求解率优于WalkSAT; 训练数据与测试数据不同时, 求解率劣于WalkSAT
10	Kurin等人 ^[36] (2020)	MPNN+CD CL+DQN	SAT	随机3-SAT	学习到的启发式比VSIDS启发式求解迭代次数更少; 泛化能力更强
11	Wen等人 ^[37] (2020)	MPNN+回溯 搜索+DDQN	CSP	随机CSP (变量数在15–40之间)	求解过程中的搜索节点数和搜索失败次数绝少于传统启发式MinDom、Dom/Ddeg和Impact, 但求解时间明显更长

3.1 基于消息传递神经网络的约束求解方法

NeuroSAT^[13]是典型的使用消息传递神经网络求解约束满足问题的方法, 其求解方法是端到端的, 即输入一个问题实例, 能通过预训练的模型直接输出结果, 而传统求解方法需要不断搜索得到问题的解。相比于传统方法, 端到端的求解方法泛化能力更强, 也为约束求解研究提供了一个崭新的方向。

如图 2 所示, NeuroSAT 的核心思想是用二分图表示布尔可满足性问题, 在表示文字的顶点和表示子句的顶点间进行消息传递, 从而得到整个图的特征向量, 基于该特征向量判断问题的可满足性。NeuroSAT 通过以下方式将 SAT 编码为一个无向图。

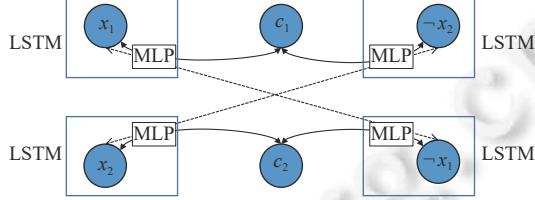


图 2 消息传递神经网络

每个文字表示一个顶点, 每个子句表示一个顶点, 文字顶点和子句顶点是两类不同的顶点。文字 x_i 和 $\neg x_i$ 之间存在一条边, 如果子句 C_j 中包含文字 x_i , 则 C_j 和 x_i 之间也存在一条边。

在 2017 年, Bünz 等人^[38]就提出了这种表示方法, 但并未与消息传递方法结合。为了使用图神经网络预测 SAT 问题的可满足性, 他们采用了简单的变量-变量图表表示法, 该方法预测准确率值达到 70% 左右。NeuroSAT 通过二分图边上的消息传递, 迭代地更新每个顶点的特征向量。每次迭代包括两个阶段, 首先, 每个子句顶点接受其邻接文字顶点的消息, 然后更新自己的特征向量。接下来, 每个文字顶点接受其邻接子句顶点和互补文字顶点的消息, 然后更新自己的特征向量。

每次迭代包括 2 次更新:

$$\left\{ \begin{array}{l} (C^{(t+1)}, C_h^{(t+1)}) \leftarrow C_u([C_h^{(t)}, M^\top L_{\text{msg}}(L^{(t)})]) \\ (L^{(t+1)}, L_h^{(t+1)}) \leftarrow L_u([L_h^{(t)}, \text{Flip}(L^{(t)}), MC_{\text{msg}}(C^{(t+1)})]) \end{array} \right.$$

模型的参数如下。

- (1) 初始特征向量: L_{init} 、 C_{init} 。
- (2) 多层感知机 (multilayer perceptron, MLP): L_{msg} 、 C_{msg} 、 L_{vote} 。
- (3) 长短期记忆神经网络 (long short term memory, LSTM): L_u 、 C_u 。

$L^{(t)}$ 是一个 $2n \times d$ 维的矩阵, 第 i 行是文字 l_i 的特征向量, $C^{(t)}$ 是一个 $m \times d$ 维的矩阵, 第 j 行是子句 C_j 的特征向量。 L_u 和 C_u 是长短期记忆神经网络^[39], $L_h^{(t)}$ 是 $2n \times d$ 维的向量, 表示 L_u 的隐藏状态, $C_h^{(t)}$ 是 $m \times d$ 维的向量, 表示 C_u 的隐藏状态。 M 是邻接矩阵, 如果 $l_i \in c_j$, 则 $M_{ij} = 1$ 。 Flip 将矩阵 M 中的每一行都替换成其互补文字的特征向量, 从而使互补文字之间进行消息传递。在 T 次迭代之后, 计算文字的平均投票结果作为最终判断结果。在预训练的过程中, 模型通过交叉熵损失函数不断学习参数。

该模型用图表示一个 SAT 问题, 通过消息传递神经网络计算顶点特征, 模型只需要一个二分邻接矩阵作为输入, 学习参数的过程与 SAT 问题规模的大小无关, 因此, 基于消息传递神经网络的方法理论上可用于求解各种规模的问题。

自将 MPNN 应用于 SAT 求解方法被提出之后, 多个基于 MPNN 的约束满足问题求解方法被相继提出, 已在伪布尔、图着色和旅行商等问题上取得了较好的效果。

在 NeuroSAT 模型被提出之后, Liu 等人^[28]在 2020 年采用二分图表示伪布尔问题, 利用 MPNN 学习顶点的特征表示。他们将伪布尔问题 (pseudo Boolean, PB) 正则化表示, 从而使约束只与子句、文字和文字系数相关, 在二分图表示法的基础上, 他们引入边的权重来表示变量的系数。同 NeuroSAT 一样, 该方法类似于二分类, 只能判断

问题的可满足性, 不能为可满足问题求解出变量的赋值。工作在 0-1 背包问题和加权独立数据集上进行实验, 结果表明对于变量个数小于 40 的 0-1 背包问题和加权独立集问题, 模型的测试准确率均能达到 85% 以上。

Lemos 等人^[29]提出了针对图着色问题的消息传递方法, 作者将每一种颜色作为一个顶点, 添加到图中, 每一个颜色顶点和所有城市顶点间都有一条边。每一次迭代都在颜色顶点和城市顶点间进行消息传递, 该模型在随机问题上实现了较高的预测准确率, 并在图分配问题上取得了比 NeuroSAT^[13]、Tabucol^[40]和贪婪算法更好的效果。

2020 年, Cameron 等人^[30]将 NeuroSAT 进行了扩展, 使得模型能够联合预测可满足性和解的赋值, 具体方法是在聚合操作过程中添加一个 MLP, 将文字的向量表示映射成每个变量的赋值, 并优化组合损失函数。实验采用随机 3-SAT 数据集, 实现了 75%–82% 的准确率。

2021 年 Tönshoff 等人^[31]将消息传递方法拓展到 MAXCSP 的求解中。其核心思想是实现变量的短期状态和长期状态间的消息传递, 从而预测每个变量取值的概率。不同于前面的端到端方法, 该方法不仅判断问题是否可满足, 还能求解出变量的赋值。此外, 作者提出了用于无监督训练的损失函数, 在大于 100 个变量的 MAX-2-SAT 问题上, 取得了比传统算法 Loandra^[41]和 WalkSAT^[42,43]更好的效果, 但在 MAX-CUT 上, 求解效果仍劣于极值优化方法^[44]。

Amizadeh 等人^[32]提出了求解 SAT 和 CSP 的通用框架, 并利用能量损失函数实现了无监督训练。该框架的求解步骤包含 3 个部分: 传播、抽取和预测。在传播和抽取阶段, 使用传播消息和抽取消息进行信息传递, 消息传递仍然使用循环神经网络。在预测阶段预测每个变量的赋值, 直到所有变量的赋值是可满足的。文献在 4-SAT 问题上进行了实验, 比较了解率和求解时间, 其求解性能优于传统的概率推理方法^[45]和 NeuroSAT 方法。但在求解时间和大规模问题的求解率上, 仍劣于传统的 SAT 求解器 Glucose^[46]。

2020 年, Abboud 等人^[33]将消息传递神经网络应用到#DNF 问题的求解中。他们将#DNF 问题表示为具有 3 类节点的图, 3 类节点分别为文字节点(包含正文字节点和负文字节点)、合取节点和析取节点。在每次消息传递中, 除了文字节点和合取节点间的消息传递, 还进行合取节点和析取节点间的消息传递。最终将析取节点输入多层次感知机, 得到预测高斯分布的平均值和标准差。论文将该方法与高效的传统方法 KML 进行对比, 证明该方法具有更高的求解率和更快的求解速度。

3.2 与传统算法相结合的约束求解方法

端到端的约束求解方法将约束满足问题看作一个二分类问题, 虽然能够直接得到问题的解, 但不具有可解释性。回溯搜索算法和局部搜索算法是求解约束满足问题的经典算法, 其本质都是使用人工启发式引导搜索方向, 启发式的设计是影响求解效果的关键因素^[47,48]。目前深度学习已经取得了一系列举世瞩目的成就, 尤其在以计算机视觉^[49]为代表的人工智能领域^[50]中发挥了巨大的作用。深度学习将人工智能研究从知识驱动转变为数据驱动, 当下人工智能的发展趋势是将知识驱动和数据驱动结合^[51]。近年来研究人员开始尝试利用机器学习方法来设计传统算法启发式。自 Selsam 等人^[13]提出将消息传递神经网络应用到约束满足问题求解后, 这种将消息传递神经网络融入经典约束求解算法研究成为一个具有巨大潜力的研究方向。

Jaszczerzak 等人^[34]提出将消息传递神经网络应用到回溯搜索算法(DPLL 和 CDCL)的选择变量阶段, 替代人工设计启发式。该模型采用和 NeuroSAT 相同的图结构表示, 每一次消息传递包含 3 个步骤: 顶点生成消息、顶点聚合消息和顶点更新特征, 在顶点聚合消息阶段, Jaszczerzak 等人引入了修改的 Attention 机制^[52], 能筛选需要接受的消息。论文在随机 SAT 实例上进行测试, 在小规模实例上, 这种神经启发式求解率仍低于人工启发式 JW-OS^[48], 但在 SR(90) 和 SR(100) 的大规模实例上, 神经启发式的求解率更高。Yolcu 等人^[35]在随机局部搜索算法(stochastic local search, SLS) 中引入了基于消息传递神经网络的神经启发式, 与传统的局部搜索算法相似, 先随机构造一个解, 通过局部搜索算法不断寻找一个可行解。作者仍然采用消息传递神经网络学习变量选择启发式, 此外, 将变量选择看作马尔可夫决策过程(Markov decision process, MDP), 利用强化学习训练神经网络的参数^[53](如图 3 所示)。与使用人工启发式的 WalkSAT^[42]相比, 其求解步数更少, 但求解时间却更长。

早在 2001 年, Lagoudakis 等人^[54]就将强化学习应用到回溯搜索算法中, 但该方法旨在从多个人工启发式中选择性能最好的, 不能实现完全学习启发式。此外, Liang 等人^[55]也提出过将变量分支启发式看作多臂老虎机问题。

受 Yolcu 等人的启发, Kurin 等人^[36]提出了融合消息传递神经网络和强化学习的 SAT 问题求解模型, 旨在学习传统的 CDCL 分支启发式算法^[56,57]. 该模型使用与 NeuroSAT 相同的 SAT 问题表示方法, 将神经网络作为函数近似器来提供对“状态-动作”空间的支持, 并使用基于 DQN (deep Q-networks)^[58]的强化学习方法进行训练. 他们证明该模型能使 SAT 问题的求解次数减少 1/3 至 1/2, 优于最常用的 CDCL 分支启发式 VSIDS^[59], 并可以泛化到比训练变量多 5 倍的问题上.

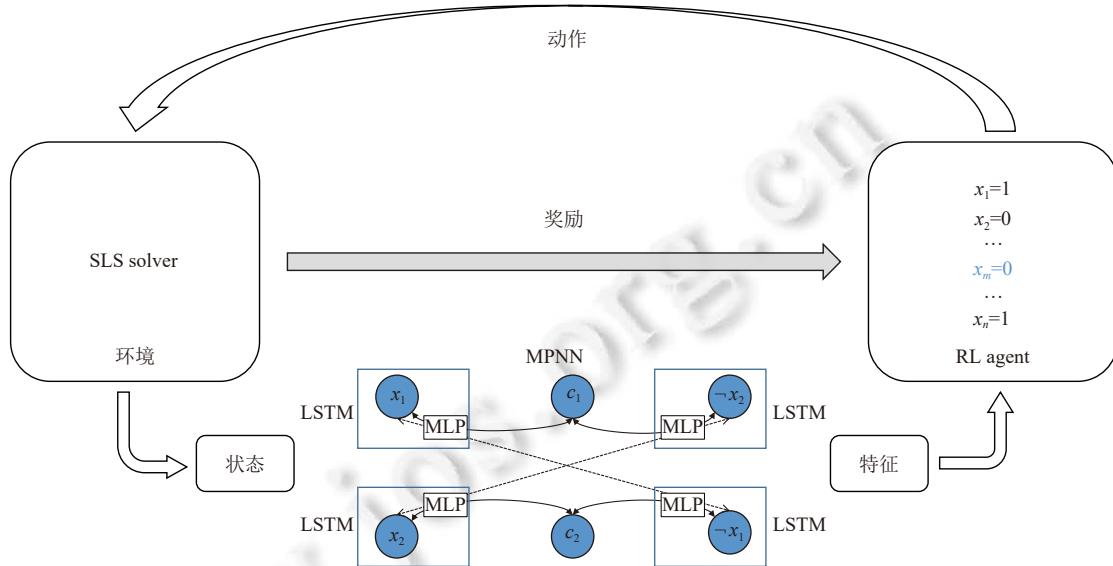


图 3 SLS 的强化学习训练过程

Wen 等人^[37]设计了一种基于消息传递神经网络的约束满足问题表示方法, 能够很好地表示变量和约束间的关系. 他们将回溯搜索算法中变量排序启发式建模为马尔可夫决策过程, 其优化策略是在每个决策点选择使搜索代价最小的变量, 并使用 DDQN^[60]强化学习方法训练消息传递神经网络. 该模型能够从已求解的 CSP 实例中学到解的分配, 从而实现无监督的训练. 在随机 CSP 实例上的实验结果表明, 与传统的手工启发式策略 MinDom^[61]、Dom/Ddeg^[62]和 Impact^[63]相比, 其学到的启发式策略能够更好地实现搜索树最小化, 并能有效地推广到比训练实例更大的实例中. 然而, 由于神经网络推理消耗过多时间, 该模型的求解时间远大于传统启发式算法.

4 基于序列到序列的约束求解方法

序列到序列 (sequence to sequence, Seq2Seq) 模型^[64]是一类特殊的递归神经网络体系结构, 通常用于求解复杂语言问题, 如机器翻译、问答、创建聊天机器人、文本摘要等. 近年来人们开始将 Seq2Seq 模型应用到约束满足问题中, 2015 年 Vinyals 等人^[65]提出的 Pointer Network 模型很好地解决了 Seq2Seq 问题中输入与输出不一致问题. Bello 等人^[14]在 Pointer Network 的基础上进行改进, 利用强化学习解决了约束满足问题标签获取困难的瓶颈问题. 此后又有多项对 Pointer Network 的改进研究, 如采用 Attention 机制^[66]和分层强化学习等. 本文通过后文表 2 对现阶段基于 Seq2Seq 算法模型求解约束满足问题进行总结与分析.

4.1 基于 Pointer Network 的约束求解方法

对于约束满足问题输入结构的处理, 我们通常使用一种将输入数据转化为高维空间向量 R^d 的结构——编码器 (encoder). 编码器的不同取决于输入数据的类型, 往年使用最多的编码器是由循环神经网络 (recurrent neural network, RNN) 构成的, 近年开始使用 Transformer 机制对 RNN 进行替换, 然后将输入序列编码成一个向量传入解

码器 (decoder) 中。RNN 由输入层、隐藏层、输出层组成, 隐藏层将序列的当前元素和前一个元素的输出作为输入, 并输出一个传递给序列下一个元素的向量。例如, 在旅行商问题 (traveling salesman problem, TSP) 中, 可将 RNN 应用到当前城市和上一个城市的输出来编码 TSP^[67]的城市序列。我们可将解码器看作编码的逆过程, 在这个阶段, 我们根据编码器传入的向量和生成的一个输出序列来预测下一个输出的序列。约束满足问题应用到 Encoder-Decoder 框架中, 也是将问题转化为序列的输入到序列的输出的问题, 即 Seq2Seq 模型。

表 2 基于 Seq2Seq 算法模型、求解问题、测试数据以及效果分析

序号	作者	基本思想	求解问题	测试数据集	效果分析
1	Vinyals等人 ^[65] (2015)	Pointer Network	TSP Convex Hull	节点为5–50的TSP问题	在节点数为5–20的数据集中训练, 可以泛化到更大规模的问题中; 当节点数小于等于25时, 求解效果接近最优解; 当节点数超出25时表现较差
2	Bello等人 ^[14] (2017)	Ptr-Net+RL	TSP 0-1背包	随机生成1 000个节点20、50、100的图和200物品的背包问题	在节点为50的TSP问题上超越了Vinyals等人 ^[65] 的监督式训练; 在节点为100的TSP问题和200物品的背包问题上接近最优解
3	Nazari等人 ^[68] (2018)	Ptr-Net改进+RL	VRP	节点为10–100的VRP问题	在VRP问题中的求解效果优于传统Clarke-Wright和Sweep启发式, 接近Google的Or-tools
4	Deudon等人 ^[69] (2018)	Transformer改进Ptr-Net+RL+2-opt	TSP	随机生成节点为20–100的TSP问题	在节点为20–100的TSP上, 求解质量优于Bello等人 ^[14] 提出的模型
5	Kool等人 ^[70] (2019)	Transformer改进解码层	TSP	随机生成节点为20–100的TSP问题	在TSP问题上均优于之前的模型, 且求解质量接近最优TSP求解器Concorde
6	Ma等人 ^[71] (2019)	PGN+HRL	TSP TSPTW	节点为20–1 000的TSP问题和节点为20的TSPTW问题	在节点为20–50的TSP问题上, 求解质量接近Kool; 在节点为20–100的TSP问题中训练, 可泛化到节点为250–1 000的TSP问题中
7	Wu等人 ^[72] (2021)	RL学习启发式	TSP CVRP	节点为20–1 000的TSP问题和CVRP问题	在节点为20的TSP问题中能求得最优解; 在节点为50–100的TSP问题中与最优TSP求解器只有0.2%和1.42%的差距。

2015 年 Vinyals 等人^[65]提出了改进的指针网络 (pointer network, Ptr-net), 该网络将 Seq2Seq 模型和 Attention 机制结合。该神经网络结构使用链式法则将旅行路线的概率分解, Seq2Seq 模型使用学习参数为 θ 的 RNN 来计算链式规则的每一项, 即条件概率:

$$p(C^p | \mathcal{P}; \theta) = \prod_{i=1}^{m(p)} p_\theta(C_i | C_1, \dots, C_{i-1}; \theta),$$

其中, $\mathcal{P} = \{P_1, \dots, P_n\}$ 是 n 个向量的序列, $C^p = \{C_1, \dots, C_m\}$ 是索引 $m(p)$ 的序列。如果我们输入的序列非常长, 那么网络的性能会大幅下降, 因为处理过长的序列时, 会容易忽略较前和较后的序列信息, Attention 机制很好地解决了该问题。结合 Attention 机制的指针网络的工作原理为: (1) 将编码器的隐藏状态 e_j 和解码器的隐藏状态 d_i 传入 \tanh 函数中, 以求得编码器和解码器的序列 embedding 的相似性; (2) 通过 $Softmax$ 函数归一化; (3) 将前 $i-1$ 个输出的序列向量和权重和作为条件概率, 来生成第 i 个输出的序列节点。不同于 Attention 机制将输入信息通过解码器整合成 context vector, 指针网络是将 Attention 转化为一个指针 (pointer), 来选择原来输入序列中对它影响最大的元素。

$$\begin{aligned} u_j^i &= v^T \tanh(w_1 e_j + w_2 d_i) \quad j \in (1, \dots, n), \\ p(C_i | C_1, \dots, C_{i-1}, p) &= Softmax(u^i). \end{aligned}$$

指针网络可以学习 3 个具有挑战性的约束满足问题的近似解——旅行商问题、求凸包问题和计算 Delaunay 三角形问题, 此外指针网络学习到的模型泛化到比训练数据更大规模的问题上。指针网络可以学习一个小规模(节点小于 50) 的 TSP 问题的近似求解器, 在节点为 5–20 的 TSP 的最优数据集中训练的模型可以泛化到节点为 50 的 TSP 问题中, 并能得到一个近似最优解。

4.2 结合强化学习的约束求解方法

对于指针网络来说, 模型想要扩展到更大规模的问题中, 自然就需要更多更大规模的优质标签的数据集, 然而一个纯数据驱动的方法学习近似的解决方案是很难的。针对监督学习获取训练数据标签困难、指针网络求解规模小和求解质量不佳的问题, 我们考虑将马尔可夫决策过程应用到求解过程中, 其中最优解可被视为一系列决策, 利用强化学习增加解码器输出序列的概率来产生接近最优的解。一种朴素的方法是通过分别考虑每个实例来找到特定问题的解决方案, 显然, 这种方法在解决方案质量不佳时并不实用, 因为从一个 MDP 中抽样许多轨迹才能产生一个接近最优的解决方案。Bello 等人^[14]提出利用基于策略的方法与基于值的方法相结合对指针网络进行训练——Actor-Critic 方法^[73], 它可以在给定分布中抽样的任何问题上都表现良好, 而不是针对每个问题实例训练一个单独的模型。该模型的基本原理是让 Actor 模型逼近价值函数, 让 Critic 模型逼近策略。通过这种方式, Critic 可以衡量 Actor 所采取的动作对解决方案的影响情况, 从而可以适当地调整下一个训练步骤的可学习参数。

网络的训练目标定义为旅行路线的总长度, 奖励函数设置为负的旅行路线的总长度, 采用策略梯度法和随机梯度下降法^[74]对参数进行优化。此外作者还将强化学习预训练分别与主动搜索和采样^[74]结合起来, 应用到 TSP 问题和 0-1 背包问题中。在节点为 50 的 TSP 问题中求得的最优解优于指针网络得到的结果, 并且可以泛化到节点为 100 的 TSP 问题中, 得到的最优解接近于通过 Google 的 Or-tools^[75]工具计算的最佳路径长度。

Nazari 等人^[68]提出了一种使用端到端的学习方法对车辆路径规划问题 (vehicle routing problem, VRP) 进行求解, 利用指针网络的简化版有效的处理具有动态元素的约束满足问题。动态元素使神经网络复杂化, 因为每当输入发生变化时, 编码器和编辑器都需要进行更新, 因此我们对输入模型的序列保持不变, 任意更改其中两个的输入顺序就不会对网络产生影响。例如在车辆路径问题中, 一旦访问完一个节点 (即将货物送达指定目的地), 其实际需求将会变成 0。由于对于车辆路径问题的输入元素的顺序会对结果产生影响, 作者认为 RNN 编码器会增加网络的复杂性, 但实际上 RNN 编码器是非必须的, 因此作者提出省略指针网络中的编码器, 直接将编码的结果作为解码器的输入以取代 RNN 隐藏状态。通过这种修改, 降低了模型的计算复杂性, 又不降低模型的效率。如图 4 所示, 模型由两个部分组成, 第 1 个是将输入元素映射到 D 维向量空间; 第 2 个部分是解码器, 它指向每个步骤的输出。作者使用 RNN 来模拟解码器网络, 只将静态元素作为解码器的输入 (动态元素也可以作为解码器的输入), 动态元素仅在注意力机制中作为输入。该方法对于大小规模为 10 和 20 的车辆路径问题可求出最优解, 但对于规模较大的 50 和 100 的问题中难以求得最佳方案。

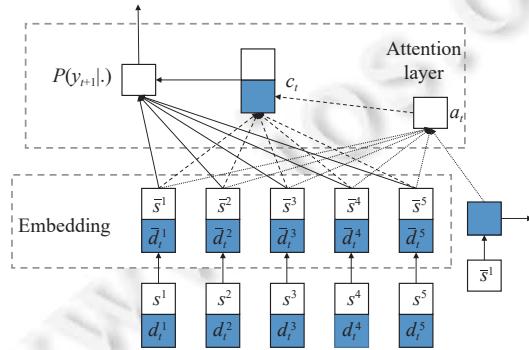


图 4 指针网络模型^[68]

之前提出的网络模型都是基于 RNN 的指针网络来实现的, Deudon 等人^[69]提出使用注意力机制替换 RNN, 将城市编码为一个集合而不是一个序列, 与长短期记忆神经网络将当前求解的节点与所有节点相关联不同的是, 该

方法只将当前节点与前 3 步的采样节点相关联, 即作者认为做出当前的最优决策时仅需考虑前 3 步的决策。此外作者将学到的模型与本地搜索算法 (2-opt^[76]) 相结合提高求解率, 而不增加推理过程中的运行时间。与 Deudon 等人^[69]不同的是 Kool 等人^[70]使用 Transformer^[77]机制提出了一个更强大的解码器层并且使用强化学习的 greedy rollout baseline 方法训练该模型^[70]在 TSP 问题的求解上取得了表现更加最优异的成果。Deudon 等人^[69]和 Kool 等人^[70]都将 structure2vec 模型^[78]替换为最先进的图注意网络 (graph attention network, GAN)^[52], 并使用一个经过强化学习训练的基于注意力的求解器自动构建 TSP 解决方案。Ma 等人^[71]研究大规模 TSP 问题和带时间约束 TSP 问题 (TSPTW), 作者在指针网络的基础上嵌入了图神经网络^[26,79,80], 得到的新体系结构称为图指针网络, 图指针网络的嵌入 (embedding) 由每个城市节点的嵌入和整个图 (所有城市) 的图嵌入组成, 使用分层强化学习^[81]对图指针网络 (graph pointer network, GPN) 进行训练。由 n 层组成, 每一层都设计了单独的策略和奖励函数, 从而可以稳定地进行训练。最底层的奖励函数的目的是确保解在约束优化问题的可行集内, 最高层的奖励函数以求解为最终的优化目标。在训练中, Actor 根据分布来采样, Critic 根据贪婪策略选择一个最大值, 两者共同训练 Actor 的分布。最终结果表明, 针对小规模节点为 50 和 100 的 TSP 问题训练的 GPN 可以很好地泛化到更大范围的 TSP 问题中, 例如节点大小为 500 和 1000 的 TSP 问题, 且具有较短的旅行长度和更快的计算时间。虽然利用深度学习 (deep learning, DL) 构造启发式算法求解路由问题的研究已经取得了很大的进展, 但大多数现有的基于 DL 的方法集中于学习构造启发式, 它通过在每个步骤中向部分解决方案添加节点来增量地创建完整的解决方案。虽然速度比较快, 但其结果与高度优化的传统求解器在客观值上仍有较大差距。为了缩小这一差距, Wu 等人^[72]提出了一个深度强化学习框架来学习路由问题的改进启发式, 它通过迭代执行基于特定局部算法的邻域搜索方法来改进初始解决方案, 朝着改进解决方案质量的方向前进, 利用强化学习自动发现更好的改进策略。首先提出改进启发式的 RL 公式, 其中 policy 指导下一个解决方案的选择。然后, 提出了一个基于自注意力的新架构来参数化策略, 通过该架构我们可以结合大量常用的成对局部操作符, 如 2-opt 和节点交换。最后, 将 RL 框架应用于两个具有代表性的路由问题, 即旅行商问题 (TSP) 和有能力车辆路径问题 (CVRP), 并设计了一个 Actor-Critic 算法来训练策略网络。

5 基于最优化的方法

采用最优化方法求解约束优化问题是当下的研究热点之一, 其基本思想是将约束优化问题转化成其他形式的最优化问题, 对具体的问题选用相应的优化方法, 并设计类神经网络的后向参数传递策略, 以有监督的学习方式求出最优解。本节主要介绍半正定规划、凸优化和基于 ERM 的 HASSLE 方法在约束优化问题上的具体应用。

5.1 基于半正定规划的方法

2019 年, Wang 等人^[15]提出了一个可微的 MAXSAT 求解器, 该求解器可以被集成到复杂的深度学习架构中并能学习 MAXSAT 问题中的逻辑关系。他们首先将 MAXSAT 问题转化成半正定规划 (semidefinite program, SDP) 公式, 然后采用基于快速坐标下降法的前向传递求解与 MAXSAT 问题相关的 SDP^[82]。该模型通过 SDP 的解来解析微分, 并有效地求解相关的后向传递。实验证明通过将该求解器集成到端到端学习系统中, 能以少量标签的样本学习具有挑战性的问题的逻辑结构。特别是深度网络难以求解的奇偶性校验问题^[83], SATNet 可以使用单位监督来学习。无需输入数独规则, 该模型能通过例子来学习如何求解 9×9 数独。另外, 该模型能够快速学习一个可行数独的约束, 在测试实例中正确地求解 98.3% 的谜题, 而不需要任何手工编码的问题结构知识。通过将可微求解器嵌入到更大的架构中, 该模型还求解了一个可视化数独问题, 其中输入是一个数独谜题的图像, 而不是二进制表示, 将 MAXSAT 求解器与传统的卷积架构^[84]相结合, 利用卷积架构将数独谜题的图像映射到 SATNet 能够求解相关的逻辑解决方案。

5.2 基于凸优化的方法

约束满足问题为建模和求解决策问题定义了一个强大的框架, 它通常被认为是计算机科学最接近编程“圣杯”的方法之一, 即用户提出问题, 计算机自动求解。然而, 问题的建模是困难的, 一方面约束求解语言过于复杂, 另一方面建模者可能无法完全理解实际问题的各个方面特征。在 2020 年 Brouard 等人^[16]提出了一种从数据中提取偏

好和约束的非完全可微架构,如图 5 所示,将模型学习到的代价函数作为约束网络的权重,也称为代价函数网络(cost function network, CFN).该方法首先利用一个基于交替方向乘子法算法的凸优化方法(黑盒优化算法)^[85]对约束满足问题建模和学习并有效地消除学习过程中固有的噪声,然后采用经验风险最小化方法对超参数进行调优,最后将学习到的代价函数网络传入 Toulbar2 求解器中求解,并对测试集进行预测以评估 CFN 的质量.此外,可以在学习到的 CFN 中添加额外的代价(约束)以应对问题的变化,这种约束扩展思路在求解著名的雷诺汽车配置问题中效果很好,对于包含 148 个变量和 173 个约束的中型汽车问题,仅需 0.1 s 即可生成 278 744 种可行配置.实验结果表明在未知约束的条件下,这种方法学习到的模型能有效地求解数独问题,且仅需 9 000 个训练样本进行训练,就可以完全求解最困难的 17-数独问题.该方法超越了最近提出的两种神经网络近似方法^[15]: SATNet 模型和循环关系网络模型^[86].

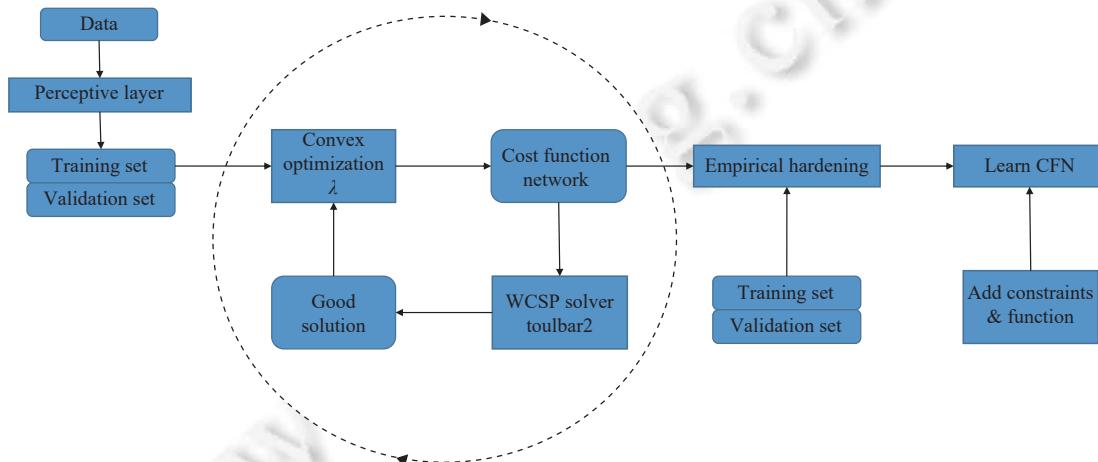


图 5 代价函数网络模型图

5.3 基于 ERM 的 HASSLE 方法

Kumar 等人^[87]从统计学习方法的角度出发,提出了一个基于经验风险最小化(empirical risk minimization, ERM)的 HASSLE(hard and soft constraint learning)算法^[88],该算法用于求解部分加权 MaxSAT 问题(partial weighted MaxSAT).主要思想是将问题中的子句划分为软、硬子句,为硬子句分配权重为 1,再根据重要程度为软子句分配不同的权重,目标是找到一组变量赋值使得所有硬子句是可满足的且可满足的软子句权重之和最大.此外,算法从指定上下文(context-specific)的样本中学习模型,上下文可以作为影响优化结果的一组约束,进而压缩可行解空间.根据 HASSLE 方法,将 MAX-SAT 问题转化为混合整数规划问题(MILP),采用 MILP 来学习 MAX-SAT 的软约束和硬约束的参数,最后采用 GUROBI 求解器进行求解.实验结果表明,通过增加样本数量 HASSLE 算法可以学习到更高精度的模型.

6 结语

2020 年以来,不断有各种新的“学习-推理”方法提出,代表性研究工作有:2020 年,Zhang 等人^[89]提出将 SLS 与神经网络解预测模型相结合,该模型使用神经网络预测初始化赋值,代替传统的随机赋值方法.作者将该初始化赋值方法与多种 SLS 求解器结合,求解率和运行时间均得到了显著提升.2020 年,Dudek 等人^[90]研究了多核和 GPU 对张量网络收缩加权模型计数的影响,采用多个核心实现并行组合的树分解求解器,以找到收缩张量的顺序,并采用 TensorFlow 来执行收缩.Jiang 等人^[91]在 2021 年提出利用组分布鲁棒性优化(group distributionally robust optimization, GDRO)来联合训练不同组实例的深度模型.为给图匹配(graph matching, GM)问题启发式创建更好的拓扑分布,Yu 等人^[92]在 2021 年提出在端到端学习方法中嵌入一个潜在拓扑模块.2022 年,Skryagin 等人^[93]为

深度概率规划语言提出了一种新的神经概率谓词形式, 允许特定任务的概率查询。对于机器学习在约束求解问题上的应用, 2022年, Popescu等人^[94]从求解问题的角度对机器学习在约束求解问题上的应用方法进行分类总结。本文从求解方法角度出发, 着重介绍如何根据约束满足问题的具体形式学习推理过程, 学习过程和推理过程相互融合、相辅相成, 并从消息传递神经网络、序列到序列和最优化3个方面对现有工作进行阐述和总结。

基于消息传递神经网络的约束满足问题求解方法可分为两类。一类是替代传统算法的端到端方法, 核心思想是将约束满足问题建模为二分类问题并采用消息传递神经网络对变量集合进行分类。这种方法在求解速度上具有一定优势, 但与传统方法相比, 目前的二分类方法在求解率上还有待改进。另一类方法是将消息传递神经网络与传统求解方法结合, 利用MPNN学习传统方法中的启发式, 取代人工设计启发式。一方面, 学习启发式能优化搜索策略, 减少回溯搜索的搜索空间和随机局部搜索算法的求解步数; 另一方面, 采用强化学习进行整体训练克服了难以获取大量标签数据的困难。由于结合传统算法需在每一次搜索时启动神经网络, 运行时间更长。目前亟待改进的问题是如何设计一个轻量级的求解模型, 从神经网络的角度, 可利用复杂的神经网络加速方法提高学习效率。此外, 采用更先进的训练机制提高训练性能也将成为未来重要的研究方向。

基于序列到序列的约束满足问题求解方法可分为两类, 一类是在循环神经网络上以监督学习的方式训练模型, 预测访问城市的序列。该方法为序列到序列求解约束满足问题开辟了新的方向。另一类是将强化学习应用到序列到序列模型中, 该方法应用了强化学习中的“Actor-Critic”思想, 通过Actor采样结果和Critic预测结果间的差值对优势函数进行无偏估计, 该方法训练的模型不再受限于数据集标签的质量, 并实现了搜索最优的解决方案而不是“复制”另一种算法的结果, 在小规模的约束满足问题上, 该方法训练的模型求解时间优于部分最优求解器, 甚至在大规模问题上也可媲美当前最优的求解器。但对于其他复杂的约束满足问题, 如CVRP和CVRPTW, 如何在小规模的数据集中训练并泛化到大规模的问题中, 以及如何让模型应对不同分布组合的数据集都是亟需解决的问题。

基于最优化方法的约束满足问题求解方法的核心思想是利用强大的数学思想将约束优化问题转化为其他形式的最优化问题, 以解决特定优化问题为目的的方式学习模型。该方法同样受限于高质量的样本标签, 但只需简单的修改便可从原始问题扩展到其他问题的求解中, 并可以将神经网络和传统求解器相结合, 相比于消息传递该方法具备更强大的推理能力, 尽管最优化方法在大量优秀数据集下训练的模型对大规模问题预测的结果接近甚至达到最优, 但如何将该方法应用到只有少量数据集又对精度要求极高的工程化问题上是未来需要推进的。

目前的“学习-推理”约束求解方法仍处于发展阶段, 对此我们提出以下几点展望。

- (1) 如何权衡可扩展性、表达能力和泛化性。目前的MPNN和Seq2Seq架构都存在可能遗漏数据中关键结构的问题, 而表达力更强的方法难以应用到更大规模的数据中, 即降低模型的泛化能力。此外, 训练数据通常分布单一, 这也导致交叉分布泛化能力的有待加强。因此, 理解可扩展性、表达性和泛化性间的权衡仍然是“学习-推理”模型的一个挑战。
- (2) 如何设计通用的实现框架。虽然目前已经实现了“学习-推理”框架, 但将学习方法集成到最先进的求解器中对实践者来说仍然很复杂。因此, 开发一个能整合机器学习的建模语言仍然是一个开放性的挑战。
- (3) 如何解决工程化问题。对约束满足问题的研究最终目的是要求解现实中的问题, 因此, 我们需要研究如何使用“学习-推理”方法高效求解现实中的工程化问题。

References:

- [1] Gent IP. Principles and Practice of Constraint Programming—CP 2009. Berlin: Springer, 2009. [doi: [10.1007/978-3-642-04244-7](https://doi.org/10.1007/978-3-642-04244-7)]
- [2] Nguyen S, Zhang MJ, Johnston M, Tan KC. Automatic programming via iterated local search for dynamic job shop scheduling. IEEE Trans. on Cybernetics, 2015, 45(1): 1–14. [doi: [10.1109/TCYB.2014.2317488](https://doi.org/10.1109/TCYB.2014.2317488)]
- [3] Song W, Kang DH, Zhang J, Xi H. Proactive project scheduling with time-dependent workability uncertainty. In: Proc. of the 16th Conf. on Autonomous Agents and Multiagent Systems. São Paulo: Int'l Foundation for Autonomous Agents and Multiagent Systems, 2017. 221–229.
- [4] Berbeglia G, Cordeau JF, Laporte G. A hybrid tabu search and constraint programming algorithm for the dynamic dial-a-ride problem. INFORMS Journal on Computing, 2012, 24(3): 343–355. [doi: [10.1287/ijoc.1110.0454](https://doi.org/10.1287/ijoc.1110.0454)]

- [5] Tang K, Wang J, Li XD, Yao X. A scalable approach to capacitated arc routing problems based on hierarchical decomposition. *IEEE Trans. on Cybernetics*, 2017, 47(11): 3928–3940. [doi: [10.1109/TCYB.2016.2590558](https://doi.org/10.1109/TCYB.2016.2590558)]
- [6] Ginsberg ML. Dynamic backtracking. *Journal of Artificial Intelligence Research*, 1993, 1(1): 25–46.
- [7] Dib M. Tabu-NG: Hybridization of constraint programming and local search for solving CSP [Ph.D. Thesis]. Franch: Université de Technologie de Belfort-Montbeliard, 2010.
- [8] Liberatore P. On the complexity of choosing the branching literal in DPLL. *Artificial Intelligence*, 2000, 116(1–2): 315–326. [doi: [10.1016/S0004-3702\(99\)00097-1](https://doi.org/10.1016/S0004-3702(99)00097-1)]
- [9] van Beek P. Backtracking search algorithms. *Foundations of Artificial Intelligence*, 2006, 2: 85–134. [doi: [10.1016/S1574-6526\(06\)80008-8](https://doi.org/10.1016/S1574-6526(06)80008-8)]
- [10] Dai WZ, Xu QL, Yu Y, Zhou ZH. Bridging machine learning and logical reasoning by abductive learning. In: Proc. of the 33rd Int'l Conf. on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2019. 253.
- [11] Zhou ZH, Huang YX. Abductive learning. In: Hitzler P, Sarker K, eds. *Neuro-symbolic Artificial Intelligence: The State of the Art*. IOP Press, 2021. 353–369. [doi: [10.3233/FAIA210363](https://doi.org/10.3233/FAIA210363)]
- [12] Huang YX, Dai WZ, Yang J, Cai LW, Cheng SF, Huang RZ, Li YF, Zhou ZH. Semi-supervised abductive learning and its application to theft judicial sentencing. In: Proc. of the 2020 IEEE Int'l Conf. on Data Mining. Sorrento: IEEE, 2020. 1070–1075. [doi: [10.1109/ICDM50108.2020.00127](https://doi.org/10.1109/ICDM50108.2020.00127)]
- [13] Selsam D, Lamm M, Bünz B, Liang P, de Moura L, Dill DL. Learning a SAT solver from single-bit supervision. In: Proc. of the 7th Int'l Conf. on Learning Representations. New Orleans: ICLR, 2018.
- [14] Bello I, Pham H, Le QV, Norouzi M, Bengio S. Neural combinatorial optimization with reinforcement learning. In: Proc. of the 5th Int'l Conf. on Learning Representations. Toulon: ICLR, 2016.
- [15] Wang PW, Donti PL, Wilder B, Kolter Z. SATNet: Bridging deep learning and logical reasoning using a differentiable satisfiability solver. In: Proc. of the 36th Int'l Conf. on Machine Learning. Long Beach: ICML, 2019. 6545–6554.
- [16] Brouard C, de Givry S, Schiex T. Pushing data into CP models using graphical model learning and solving. In: Proc. of the 26th Int'l Conf. on Principles and Practice of Constraint Programming. Louvain-la-Neuve: Springer, 2020. 811–827. [doi: [10.1007/978-3-030-58475-7_47](https://doi.org/10.1007/978-3-030-58475-7_47)]
- [17] Jensen TR, Toft B. *Graph Coloring Problems*. New York: John Wiley & Sons, 2011. 1–320. [doi: [10.1002/9781118032497](https://doi.org/10.1002/9781118032497)]
- [18] Cook SA. The complexity of theorem-proving procedures. In: Proc. of the 3rd Annual ACM Symp. on Theory of Computing. Shaker Heights: ACM, 1971. 151–158. [doi: [10.1145/800157.805047](https://doi.org/10.1145/800157.805047)]
- [19] Bellman R. Dynamic programming treatment of the travelling salesman problem. *Journal of the ACM*, 1962, 9(1): 61–63. [doi: [10.1145/321105.321111](https://doi.org/10.1145/321105.321111)]
- [20] Dorigo M, Gambardella LM. Ant colony system: A cooperative learning approach to the traveling salesman problem. *IEEE Trans. on Evolutionary Computation*, 1997, 1(1): 53–66. [doi: [10.1109/4235.585892](https://doi.org/10.1109/4235.585892)]
- [21] Sun JG, Zhu XJ, Zhang YG, Li Y. An approach of solving constraint satisfaction problem based on preprocessing. *Chinese Journal of Computers*, 2008, 31(6): 919–926 (in Chinese with English abstract). [doi: [10.3321/j.issn:0254-4164.2008.06.004](https://doi.org/10.3321/j.issn:0254-4164.2008.06.004)]
- [22] Dechter R. *Constraint Processing*. Amsterdam: Elsevier, 2003. 1–481. [doi: [10.1016/B978-1-55860-890-0.X5000-2](https://doi.org/10.1016/B978-1-55860-890-0.X5000-2)]
- [23] Bessiere C. Constraint propagation. In: Rossi F, van Beek P, Walsh T, eds. *Handbook of Constraint Programming*. Amsterdam: Elsevier, 2006. 29–83.
- [24] Zhu XJ, Sun JG, Zhang YG, Li Y. A new preprocessing technique based on entirety singleton consistency. *Acta Automatica Sinica*, 2009, 35(1): 71–76 (in Chinese with English abstract). [doi: [10.3724/SP.J.1004.2009.00071](https://doi.org/10.3724/SP.J.1004.2009.00071)]
- [25] Applegate D, Bixby R, Chvatal V, Cook W. Concorde TSP solver. 2006. <http://www.math.uwaterloo.ca/tsp/concorde/m>
- [26] Gori M, Monfardini G, Scarselli F. A new model for learning in graph domains. In: Proc. of the 2005 IEEE Int'l Joint Conf. on Neural Networks. Montreal: IEEE, 2005. 729–734. [doi: [10.1109/IJCNN.2005.1555942](https://doi.org/10.1109/IJCNN.2005.1555942)]
- [27] Gilmer J, Schoenholz SS, Riley PF, Vinyals O, Dahl GE. Neural message passing for quantum chemistry. In: Proc. of the 34th Int'l Conf. on Machine Learning. Sydney: ICML, 2017. 1263–1272.
- [28] Liu MH, Zhang F, Huang P, Niu SZ, Ma FF, Zhang J. Learning the satisfiability of pseudo-Boolean problem with graph neural networks. In: Proc. of the 26th Int'l Conf. on Principles and Practice of Constraint Programming. Louvain-la-Neuve: Springer, 2020. 885–898. [doi: [10.1007/978-3-030-58475-7_51](https://doi.org/10.1007/978-3-030-58475-7_51)]
- [29] Lemos H, Prates M, Avelar P, Lamb L. Graph colouring meets deep learning: Effective graph neural network models for combinatorial problems. In: Proc. of the 31st IEEE Int'l Conf. on Tools with Artificial Intelligence. Portland: IEEE, 2019. 879–885. [doi: [10.1109/ICTAI19.90125](https://doi.org/10.1109/ICTAI19.90125)]
- [30] Cameron C, Chen R, Hartford J, Leyton-Brown K. Predicting propositional satisfiability via end-to-end learning. In: Proc. of the 34th

- AAAI Conf. on Artificial Intelligence. New York: IEEE, 2020. 3324–3331. [doi: [10.1609/aaai.v34i04.5733](https://doi.org/10.1609/aaai.v34i04.5733)]
- [31] Tönshoff J, Ritzert M, Wolf H, Grohe M. Graph neural networks for maximum constraint satisfaction. *Frontiers in Artificial Intelligence*, 2021, 3: 580607. [doi: [10.3389/frai.2020.580607](https://doi.org/10.3389/frai.2020.580607)]
- [32] Amizadeh S, Matusevych S, Weimer M. PDP: A general neural framework for learning constraint satisfaction solvers. arXiv:1903.01969, 2019.
- [33] Abboud R, Ceylan I, Lukasiewicz T. Learning to reason: Leveraging neural networks for approximate DNF counting. In: Proc. of the 34th AAAI Conf. on Artificial Intelligence. New York: IEEE, 2019. 3097–3104. [doi: [10.1609/aaai.v34i04.5705](https://doi.org/10.1609/aaai.v34i04.5705)]
- [34] Jaszczyk S, Łuszczak M, Michalewski H. Neural heuristics for SAT solving. arXiv:2005.13406, 2020.
- [35] Yolcu E, Póczos B. Learning local search heuristics for boolean satisfiability. In: Proc. of the 33rd Int'l Conf. on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2019. 718.
- [36] Kurin V, Godil S, Whiteson S, Catanzaro B. Can Q-learning with graph networks learn a generalizable branching heuristic for a SAT solver? In: Proc. of the 34th Int'l Conf. on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2020. 9608–9621.
- [37] Wen S, Cao ZG, Zhang J, Xu C, Lim A. Learning variable ordering heuristics for solving constraint satisfaction problems. *Engineering Applications of Artificial Intelligence*, 2022, 109: 104603. [doi: [10.1016/j.engappai.2021.104603](https://doi.org/10.1016/j.engappai.2021.104603)]
- [38] Bünz B, Lamm M. Graph neural networks and Boolean satisfiability. arXiv:1702.03592, 2017.
- [39] Hochreite S, Schmidhuber J. Long short-term memory. *Neural Computation*, 1997, 9(8): 1735–1780. [doi: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735)]
- [40] Hertz A, de Werra D. Using tabu search techniques for graph coloring. *Computing*, 1987, 39(4): 345–351. [doi: [10.1007/BF02239976](https://doi.org/10.1007/BF02239976)]
- [41] Berg J, Demirović E, Stuckey PJ. Core-boosted linear search for incomplete MaxSAT. In: Proc. of the 16th Int'l Conf. on Integration of Constraint Programming, Artificial Intelligence, and Operations Research. Thessaloniki: Springer, 2019. 39–56. [doi: [10.1007/978-3-030-19212-9_3](https://doi.org/10.1007/978-3-030-19212-9_3)]
- [42] Selman B, Kautz H, Cohen B. Local search strategies for satisfiability testing. *Cliques, Coloring, and Satisfiability*, 1996, 26: 521–531.
- [43] Hoos HH. An adaptive noise mechanism for WalkSAT. In: Proc. of the 18th Conf. on Artificial Intelligence (AAAI 2002). Alberta: AAAI, 2002. 655–660.
- [44] Boettcher S, Percus AG. Extremal optimization for graph partitioning. *Physical Review E*, 2001, 64(2): 026114. [doi: [10.1103/PhysRevE.64.026114](https://doi.org/10.1103/PhysRevE.64.026114)]
- [45] Montanari A, Ricci-Tersenghi F, Semerjian G. Solving constraint satisfaction problems through belief propagation-guided decimation. arXiv:0709.1667, 2007.
- [46] Audemard G, Simon L. On the glucose SAT solver. *Int'l Journal on Artificial Intelligence Tools*, 2018, 27(1): 1840001. [doi: [10.1142/S0218213018400018](https://doi.org/10.1142/S0218213018400018)]
- [47] Gent IP, MacIntyre E, Presser P, Smith BM, Walsh T. An empirical study of dynamic variable ordering heuristics for the constraint satisfaction problem. In: Proc. of the 2nd Int'l Conf. on Principles and Practice of Constraint Programming. Cambridge: Springer, 1996. 179–193. [doi: [10.1007/3-540-61551-2_74](https://doi.org/10.1007/3-540-61551-2_74)]
- [48] Marques-Silva J. The impact of branching heuristics in propositional satisfiability algorithms. In: Proc. of the 9th Portuguese Conf. on Artificial Intelligence. Evora: Springer, 1999. 62–74. [doi: [10.1007/3-540-48159-1_5](https://doi.org/10.1007/3-540-48159-1_5)]
- [49] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 2017, 60(6): 84–90. [doi: [10.1145/3065386](https://doi.org/10.1145/3065386)]
- [50] Silver D, Hubert T, Schrittwieser J, Antonoglou I, Lai M, Guez A, Lanctot M, Sifre L, Kumaran D, Graepel T, Lillicrap T, Simonyan K, Hassabis D. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*, 2018, 362(6419): 1140–1144. [doi: [10.1126/science.aar6404](https://doi.org/10.1126/science.aar6404)]
- [51] Zhang B, Zhu J, Su H. Toward the third generation of artificial intelligence. *Scientia Sinica Informationis*, 2020, 50(9): 1281–1302 (in Chinese with English abstract). [doi: [10.1360/SSI-2020-0204](https://doi.org/10.1360/SSI-2020-0204)]
- [52] Veličković P, Cucurull G, Casanova A, Romero A, Liò P, Bengio Y. Graph attention networks. arXiv:1710.10903, 2017.
- [53] Kaelbling LP, Littman ML, Moore AW. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 1996(4): 237–285. [doi: [10.1613/jair.301](https://doi.org/10.1613/jair.301)]
- [54] Lagoudakis M G, Littman M L. Learning to select branching rules in the DPLL procedure for satisfiability. *Electronic Notes in Discrete Mathematics*, 2001, 9: 344–359. [doi: [10.1016/S1571-0653\(04\)00332-4](https://doi.org/10.1016/S1571-0653(04)00332-4)]
- [55] Liang JH, Ganesh V, Poupart P, Czarnecki K. Learning rate based branching heuristic for SAT solvers. In: Proc. of the 19th Int'l Conf. on Theory and Applications of Satisfiability Testing. Bordeaux: Springer, 2016. 123–140. [doi: [10.1007/978-3-319-40970-2_9](https://doi.org/10.1007/978-3-319-40970-2_9)]
- [56] Bayardo RJ Jr, Schrag R. Using CSP look-back techniques to solve real-world SAT instances. In: Proc. of the 14th National Conf. on Artificial Intelligence and the 9th Conf. on Innovative Applications of Artificial Intelligence. Providence: AAAI Press, 1997. 203–208.

- [57] Marques-Silva JP, Sakallah KA. GRASP: A search algorithm for propositional satisfiability. *IEEE Trans. on Computers*, 1999, 48(5): 506–521. [doi: [10.1109/12.769433](https://doi.org/10.1109/12.769433)]
- [58] Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, Graves A, Riedmiller M, Fidjeland AK, Ostrovski G, Petersen S, Beattie C, Sadik A, Antonoglou I, King H, Kumaran D, Wierstra D, Legg S, Hassabis D. Human-level control through deep reinforcement learning. *Nature*, 2015, 518(7540): 529–533. [doi: [10.1038/nature14236](https://doi.org/10.1038/nature14236)]
- [59] Moskewicz MW, Madigan CF, Zhao Y, Zhang LT, Malik S. Chaff: Engineering an efficient SAT solver. In: Proc. of the 38th Design Automation Conf. Las Vegas: IEEE, 2001. 530–535. [doi: [10.1145/378239.379017](https://doi.org/10.1145/378239.379017)]
- [60] van Hasselt H, Guez A, Silver D. Deep reinforcement learning with double Q-learning. In: Proc. of the 30th AAAI Conf. on Artificial Intelligence. Phoenix: AAAI, 2016. 2094–2100.
- [61] Haralick RM, Elliott GL. Increasing tree search efficiency for constraint satisfaction problems. *Artificial Intelligence*, 1980, 14(3): 263–313. [doi: [10.1016/0004-3702\(80\)90051-X](https://doi.org/10.1016/0004-3702(80)90051-X)]
- [62] Bessière C, Régin JC. MAC and combined heuristics: Two reasons to forsake FC (and CBJ?) on hard problems. In: Proc. of the 2nd Int'l Conf. on Principles and Practice of Constraint Programming. Cambridge: Springer, 1996. 61–75. [doi: [10.1007/3-540-61551-2_66](https://doi.org/10.1007/3-540-61551-2_66)]
- [63] Refalo P. Impact-based search strategies for constraint programming. In: Proc. of the 10th Int'l Conf. on Principles and Practice of Constraint Programming. Toronto: Springer, 2004. 557–571. [doi: [10.1007/978-3-540-30201-8_41](https://doi.org/10.1007/978-3-540-30201-8_41)]
- [64] Sutskever I, Vinyals O, Le OV. Sequence to sequence learning with neural networks. In: Proc. of the 27th Int'l Conf. on Neural Information Processing Systems. Montreal: MIT Press, 2014. 3104–3112.
- [65] Vinyals O, Fortunato M, Jaitly N. Pointer networks. In: Proc. of the 28th Int'l Conf. on Neural Information Processing Systems. Montreal: MIT Press, 2015. 2692–2700.
- [66] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate. In: Proc. of the 3rd Int'l Conf. on Learning Representations. San Diego: ICLR, 2015. 1–14. [doi: [10.48550/arXiv.1409.0473](https://doi.org/10.48550/arXiv.1409.0473)]
- [67] Papadimitriou CH. The euclidean travelling salesman problem is NP-complete. *Theoretical Computer Science*, 1977, 4(3): 237–244. [doi: [10.1016/0304-3975\(77\)90012-3](https://doi.org/10.1016/0304-3975(77)90012-3)]
- [68] Nazari M, Oroojlooy A, Takáč M, Snyder LV. Reinforcement learning for solving the vehicle routing problem. In: Proc. of the 32nd Int'l Conf. on Neural Information Processing Systems. Montréal: Curran Associates Inc., 2018. 9861–9871.
- [69] Deudon M, Courtnut P, Lacoste A, Adulyasak Y, Rousseau LM. Learning heuristics for the TSP by policy gradient. In: Proc. of the 15th Int'l Conf. on Integration of Constraint Programming, Artificial Intelligence, and Operations Research. Delft: Springer, 2018. 170–181. [doi: [10.1007/978-3-319-93031-2_12](https://doi.org/10.1007/978-3-319-93031-2_12)]
- [70] Kool W, van Hoof H, Welling M. Attention, learn to solve routing problems! In: Proc. of the 7th Int'l Conf. on Learning Representations. New Orleans: ICLR, 2019. 1–25.
- [71] Ma Q, Ge SW, He DY, Thaker D, Drori I. Combinatorial optimization by graph pointer networks and hierarchical reinforcement learning. arXiv:1911.04936, 2019.
- [72] Wu YX, Song W, Cao ZG, Zhang J, Lim A. Learning improvement heuristics for solving routing problems. *IEEE Trans. on Neural Networks and Learning Systems*, 2022, 33(9): 5057–5069. [doi: [10.1109/TNNLS.2021.3068828](https://doi.org/10.1109/TNNLS.2021.3068828)]
- [73] Mnih V, Badia AP, Mirza M, Graves A, Lillicrap T, Harley T, Silver D, Kavukcuoglu K. Asynchronous methods for deep reinforcement learning. In: Proc. of the 33rd Int'l Conf. on Machine Learning. New York City: PMLR, 2016. 1928–1937.
- [74] Williams RJ. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, 1992, 8(3): 229–256. [doi: [10.1007/BF00992696](https://doi.org/10.1007/BF00992696)]
- [75] Perron L, Furnon V. Or-tools. 2022. <https://developers.google.com/optimization/>
- [76] Helsgaun K. General k -opt submoves for the lin-kernighan TSP heuristic. *Mathematical Programming Computation*, 2009, 1(2–3): 119–163. [doi: [10.1007/s12532-009-0004-6](https://doi.org/10.1007/s12532-009-0004-6)]
- [77] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. In: Proc. of the 31st Int'l Conf. on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 6000–6010.
- [78] Ribeiro LFR, Saverese PHP, Figueiredo DR. struc2vec: Learning node representations from structural identity. In: Proc. of the 23rd ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. Halifax: ACM, 2017. 385–394. [doi: [10.1145/3097983.3098061](https://doi.org/10.1145/3097983.3098061)]
- [79] Scarselli F, Gori M, Tsoi AC, Hagenbuchner M, Monfardini G. The graph neural network model. *IEEE Trans. on Neural Networks*, 2009, 20(1): 61–80. [doi: [10.1109/TNN.2008.2005605](https://doi.org/10.1109/TNN.2008.2005605)]
- [80] Battaglia PW, Hamrick JB, Bapst V, et al. Relational inductive biases, deep learning, and graph networks. arXiv:1806.01261, 2018.
- [81] Haarnoja T, Zhou A, Abbeel P, Levine S. Soft Actor-Critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In: Proc. of the 35th Int'l Conf. on Machine Learning. Stockholm: ICML, 2018. 1861–1870.

- [82] Wang PW, Chang WC, Kolter JZ. The mixing method: Low-rank coordinate descent for semidefinite programming with diagonal constraints. arXiv:1706.00476, 2017.
- [83] Shalev-Shwartz S, Shamir O, Shammah S. Failures of gradient-based deep learning. In: Proc. of the 34th Int'l Conf. on Machine Learning. Sydney: ICML, 2017. 3067–3075.
- [84] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc. of the IEEE, 1998, 86(11): 2278–2324. [doi: [10.1109/5.726791](https://doi.org/10.1109/5.726791)]
- [85] Park Y, Hallac D, Boyd S, Leskovec J. Learning the network structure of heterogeneous data via pairwise exponential Markov random fields. In: Proc. of the 20th Int'l Conf. on Artificial Intelligence and Statistics. Fort Lauderdale: AISTATS, 2017. 1302–1310.
- [86] Palm RB, Paquet U, Winther O. Recurrent relational networks. In: Proc. of the 32nd Int'l Conf. on Neural Information Processing Systems. Montréal: Curran Associates Inc., 2018. 3372–3382.
- [87] Kumar M, Kolb S, Teso S, De Raedt L. Learning MAX-SAT from contextual examples for combinatorial optimisation. Artificial Intelligence, 2023, 314: 103794. [doi: [10.1016/j.artint.2022.103794](https://doi.org/10.1016/j.artint.2022.103794)]
- [88] Bessiere C, Daoudi A, Hebrard E, Katsirelos G, Lazaar N, Mechqrane Y, Narodytska N, Quimper CG, Walsh T. New approaches to constraint acquisition. In: Bessiere C, De Raedt L, Kotthoff L, Nijssen S, O'Sullivan B, Pedreschi D, eds. Data Mining and Constraint Programming. Cham: Springer, 2016. 51–76. [doi: [10.1007/978-3-319-50137-6_3](https://doi.org/10.1007/978-3-319-50137-6_3)]
- [89] Zhang WJ, Sun ZY, Zhu QH, Li G, Cai SW, Xiong YF, Zhang L. NLocalSAT: Boosting local search with solution prediction. In: Proc. of the 29th Int'l Joint Conf. on Artificial Intelligence. Yokohama: IJCAI, 2020. 1177–1183. [doi: [10.24963/ijcai.2020/164](https://doi.org/10.24963/ijcai.2020/164)]
- [90] Dudek JM, Vardi MY. Parallel weighted model counting with tensor networks. arXiv:2006.15512, 2020.
- [91] Jiang Y, Wu YX, Cao ZG, Zhang J. Learning to solve routing problems via distributionally robust optimization. In: Proc. of the 36th AAAI Conf. on Artificial Intelligence. AAAI, 2022. 9786–9794. [doi: [10.1609/aaai.v36i9.21214](https://doi.org/10.1609/aaai.v36i9.21214)]
- [92] Yu TS, Wang RZ, Yan JC, Li BX. Deep latent graph matching. In: Proc. of the 38th Int'l Conf. on Machine Learning. ICML, 2021. 12187–12197.
- [93] Skryagin A, Stammer W, Ochs D, Dhami DS, Kersting K. Neural-probabilistic answer set programming. In: Proc. of the 19th Int'l Conf. on Principles of Knowledge Representation and Reasoning. Haifa: KR, 2022. 463–473. [doi: [10.24963/kr.2022/48](https://doi.org/10.24963/kr.2022/48)]
- [94] Popescu A, Polat-Erdeniz S, Felfernig A, Uta M, Atas M, Le VM, Pilsl K, Enzelsberger M, Tran TNT. An overview of machine learning techniques in constraint solving. Journal of Intelligent Information Systems, 2022, 58(1): 91–118. [doi: [10.1007/s10844-021-00666-5](https://doi.org/10.1007/s10844-021-00666-5)]

附中文参考文献:

- [21] 孙吉贵, 朱兴军, 张永刚, 李莹. 一种基于预处理技术的约束满足问题求解算法. 计算机学报, 2008, 31(6): 919–926. [doi: [10.3321/j.issn:0254-4164.2008.06.004](https://doi.org/10.3321/j.issn:0254-4164.2008.06.004)]
- [24] 朱兴军, 孙吉贵, 张永刚, 李莹. 一种新的基于完全独立相容性的预处理技术. 自动化学报, 2009, 35(1): 71–76. [doi: [10.3724/SP.J.1004.2009.00071](https://doi.org/10.3724/SP.J.1004.2009.00071)]
- [51] 张锐, 朱军, 苏航. 迈向第三代人工智能. 中国科学: 信息科学, 2020, 50(9): 1281–1302. [doi: [10.1360/SSI-2020-0204](https://doi.org/10.1360/SSI-2020-0204)]



邹悦(1998—), 女, 硕士生, 主要研究领域为约束求解, 约束优化。



张永刚(1975—), 男, 博士, 教授, 博士生导师, CCF 高级会员, 主要研究领域为约束求解, 约束优化。



赖家洋(1997—), 男, 硕士生, 主要研究领域为约束求解, 约束优化。