

# 航空制造知识图谱构建研究综述\*

邱凌<sup>1a,1b</sup>, 张安思<sup>1b,1c†</sup>, 李少波<sup>1b,1c</sup>, 张仪宗<sup>1c</sup>, 沈明明<sup>1c,2</sup>, 周鹏<sup>1d</sup>

(1. 贵州大学 a. 计算机科学与技术学院; b. 公共大数据国家重点实验室; c. 机械工程学院; d. 现代制造技术教育部重点实验室, 贵阳 550025; 2. 贵州师范大学 机械与电气工程学院, 贵阳 550025)

**摘要:** 目前航空装备制造企业在设计、制造相关流程中积累了大量数据, 基于知识图谱技术可以对这些数据进行有效融合与管理, 对不断更新的制造知识进行挖掘, 为航空制造企业智慧化升级提供有力的知识支撑。为探明知识图谱在航空制造领域的理论支撑体系与实际应用情况, 通过文献调研分析航空制造知识图谱架构、定义及特点。阐明知识图谱领域构建过程中的核心技术并进行研究综述, 对比航空制造知识图谱与通用知识图谱构建技术上的异同, 并提出了三个切合实际的航空制造知识图谱应用方向及其解决方案。最后对未来航空制造知识图谱的挑战进行了分析及展望, 为后续该领域的研究提供一些思路。

**关键词:** 航空制造; 知识图谱; 知识抽取; 知识推理; 知识融合

**中图分类号:** TP3

**文献标志码:** A

**文章编号:** 1001-3695(2022)04-002-0968-10

doi:10.19734/j.issn.1001-3695.2021.09.0367

## Survey on building knowledge graphs for aerospace manufacturing

Qiu Ling<sup>1a,1b</sup>, Zhang Ansi<sup>1b,1c†</sup>, Li Shaobo<sup>1b,1c</sup>, Zhang Yizong<sup>1c</sup>, Shen Mingming<sup>1c,2</sup>, Zhou Peng<sup>1d</sup>

(1. a. School of Computer Science & Technology, b. State Key Laboratory of Public Big Data, c. School of Mechanical Engineering, d. Key Laboratory of Advanced Manufacturing Technology of the Ministry of Education, Guizhou University, Guiyang 550025, China; 2. School of Mechanical & Electrical Engineering, Guizhou Normal University, Guiyang 550025, China)

**Abstract:** At present, aviation equipment manufacturing enterprises have accumulated a large amount of data in the design and manufacturing-related processes. Based on the knowledge graph technology, these data can be effectively integrated and managed, and the manufacturing knowledge that is continuously updated can be mined, which will provide powerful knowledge support for the wisdom upgrade of aviation manufacturing enterprises. In order to explore the theoretical support system and practical application of knowledge graph in the field of aviation manufacturing, this paper firstly analyzed the architecture, definition and characteristics of aviation manufacturing knowledge graph through literature research. Secondly, it elucidated the core technologies in the process of constructing the knowledge graph domain and conducted a research review, then compared the similarities and differences in the construction technologies between the aerospace manufacturing knowledge graph and the general knowledge graph. In the light of the review, it proposed three practical directions for the application of the aerospace manufacturing knowledge graph and its solutions. Finally, this paper stated the challenges of the future aerospace manufacturing knowledge graph and foreseen, and hoped to provide some ideas for the subsequent research in this field.

**Key words:** aerospace manufacturing; knowledge graph; knowledge extraction; knowledge reasoning; knowledge fusion

## 0 引言

人类社会经历了数次生产力革命, 多次改变了社会结构。由万维网技术的发展引发的信息革命是最近一次的生产力革命, 其中万维网第一纪元的主要特征为文档互联, 而万维网第二纪元的主要特征为数据互联<sup>[1]</sup>, 现在人类即将走近一个崭新的以知识互联为特征的万维网第三纪元<sup>[2]</sup>。

为了实现知识互联的目标, 即为了创建一个人类和电脑都能知晓其中含义的互联网, 并使互联网更加智能, 科学家必须克服互联网来源众多的异质性和内部结构不紧凑这两种巨大挑战。在这一时期, 知识图谱以其高效的语义处理能力和开放式连接的特点进入人们的视野, 能够更立体生动地展现整个相互关联的人类的认知世界<sup>[3]</sup>, 为基于互联网的知识互联提供

了技术支持, 让 Web 3.0 提出的知识网络不再虚无缥缈。

谷歌最初于 2012 年正式提出了知识图谱的概念<sup>[4]</sup>, 其最初设想是利用知识图谱提高搜索引擎的搜索效率及能力, 以便用户的搜索体验得到更好的满足。与传统知识库相比, 知识图谱主要具有以下优点: a) 更丰富的语义关系, 知识图谱直接用连线表示实体之间的关系, 是帮助计算机理解物理世界的关键; b) 更高的知识质量, 知识图谱的知识来源是多样的, 不同的数据源可以交叉测试; c) 可视化<sup>[5]</sup>, 知识图谱支持可视化显示, 为人机交互提供了极大的方便。目前, 知识图谱已在医疗、财经、网络销售、教育等行业得到了广泛应用, 且获得了巨大的成功。研究航空制造领域的知识图谱具有重要的理论意义和实用价值。然而, 由于航空工业知识的复杂性, 其知识图谱的搭建以及合理使用变得极为困难。本文以构建航空制造知识图谱的三个关键手段(即知识表示与抽取、知识融合和知识推

**收稿日期:** 2021-09-23; **修回日期:** 2021-11-09 **基金项目:** 国家重点研发计划资助项目(2020YFB1713300); 贵州省高等学校集成攻关大平台资助项目(黔教合 KY 字[2020]005); 贵州省高等学校人才培养基地资助项目(黔教合 KY 字[2020]009); 贵州省科技计划资助项目(黔科合重大专项字[2019]3003, 黔科合人才[2015]4011, 黔科合平台人才[2017]5788)

**作者简介:** 邱凌(1998-), 男, 江西鹰潭人, 硕士研究生, 主要研究方向为知识图谱; 张安思(1991-), 男(通信作者), 贵州贵阳人, 讲师, 博士, 主要研究方向为工业大数据(zhangas@gzu.edu.cn); 李少波(1973-), 男, 贵州贵阳人, 院长, 教授, 主要研究方向为机械设备智能故障诊断; 张仪宗(1996-), 男, 硕士研究生, 主要研究方向为机械设备智能故障诊断; 沈明明(1988-), 女, 博士研究生, 主要研究方向为机器学习、知识图谱、工业大数据分析等; 周鹏(1988-), 男, 讲师, 博士研究生, 主要研究方向为复杂制造过程建模与监控、复杂装备数字化设计与制造。

理)为线索,在前人研究的基础上介绍相关研究和技术细节,浅析知识图谱在航空制造领域的应用,并结合相关研究进行说明。

## 1 航空制造知识图谱概述

### 1.1 航空制造知识图谱的概念

航空制造知识图谱属于特定领域知识图谱。认识特定领域的知识图谱,首先需要清楚通用知识图谱与领域知识图谱之间的联系与区别<sup>[6]</sup>。文献[7]这样阐明知识图谱:它是内部自成体系的信息库,利用象征语义的形式表示现实生活中物体的概念与彼此联系。其最初由谷歌用于其核心业务搜索引擎,实现了高效的智能搜索,基于知识图谱的搜索功能如图1所示。知识图谱的本质其实就是展现实体之间语义关系的图结构,揭示实体之间关系的语义网络<sup>[8]</sup>,而用来描述其内部概念相互关系的形式主要有资源的描述框架(resource description framework, RDF)和网络的本体语言(Web ontology language, OWL)。目前通用知识图谱比比皆是,如国外研究领域的 DBpedia<sup>[9]</sup>、YAGO<sup>[10]</sup>、Freebase<sup>[11]</sup>,国内目前的通用知识图谱项目如 OpenKG、中文知识图谱 CN-DBpedia<sup>[12]</sup>、zhishi.me<sup>[13]</sup>等;领域知识图谱也有许多项目落地,如欧盟重大联合攻关项目的医疗知识图谱 OPEN PHACTS、企业领域知识图谱的天眼查等。



图 1 知识图谱在搜索系统中的使用  
Fig. 1 Use of knowledge graphs in search systems

通识知识图谱尽管包含了现实世界中的大量常识知识,涵盖面极广,可以形象地描述为服务于通识领域的结构化的百科全书,但其精确度不够高,主要着眼于知识的广度。本体库对公理、规则和约束条件的支持能力可以用来规范其实体、属性和实体之间的密切联系<sup>[14]</sup>。领域知识图谱可以看做知识图谱的一个分系,不同于通用知识图谱,它专注于某一特定领域,通常用以帮助不同类型的复杂的分析应用以及智能决策支撑,因此其对该领域知识的深度和精度都有很高的要求<sup>[15]</sup>。

而航空制造知识图谱则是融合航空制造数据形成的知识图谱,其与之前两类知识图谱的关系如图 2 所示。

航空制造数据包括航空专业知识、航空文献、工艺性能、装备设计数据、图纸、生产线、使用说明书、应用软件等。航空制造数据有如下四个特点:a)数据来源广泛,包括内部数据库数据、互联网数据以及第三方来源数据;b)数据类型复杂,包括结构化数据(格式固定如数据库文件)、半结构化数据(具有相关标记如网页)以及非结构化数据(没有预先的组织方式如视频)<sup>[16]</sup>;c)数据模式未定,模式必须在数据出现后才可以确定下来,并且数据模式会随着数据的增加而不停变化;d)数据量巨大,随着大数据技术的广泛应用,特定领域的数据轻而易举

就能达到 TB、PB 甚至更高级别。图 3 展现了航空制造数据的模型的特征,表 1 总结了航空制造领域中部分实体属性的特征。

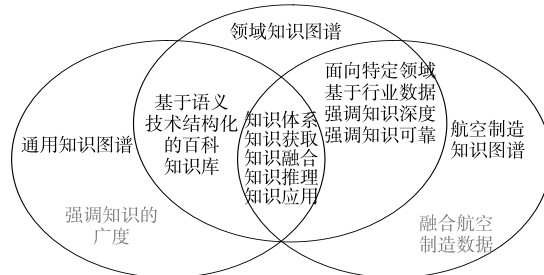


图2 通用、领域、航空制造知识图谱的关系  
Fig.2 Relationship between generic, domain, and aerospace manufacturing knowledge graphs

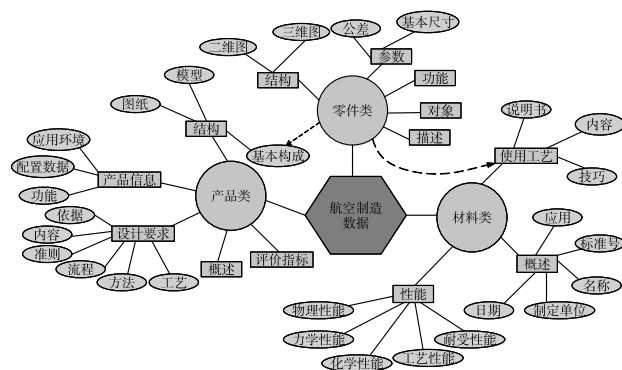


图3 航空制造数据特征模型

Fig.3 Aviation manufacturing data characterization model

表1 航空制造业领域知识的属性特征  
Tab.1 Attribute characteristics of domain knowledge in aviation manufacturing

| 属性类型 | 释义                  | 基本构成             | 特征说明                 | 规则约束               |
|------|---------------------|------------------|----------------------|--------------------|
| 材料属性 | 航空制造业设计制造过程中涉及的材料数据 | 物理、化学、生产日期、使用工艺等 | 航空设计制造活动中的材料以及所满足的属性 | 材料供货商按相关规格要求提供     |
| 产品属性 | 航空设计制造生产的产品数据       | 功能、工艺、配置环境、流程等   | 航空制造业产品从设计到制造的一系列规格  | 产品供销商提供使用说明手册等     |
| 零件属性 | 航空设计制造企业生产的零件数据     | 尺寸、模型、参数等        | 航空制造业零件从设计到制造的一系列规格  | 生产商提供详细参数及注意事项     |
| 工艺属性 | 航空设计制造生产等所用到的工艺要求   | 方法、描述、架构         | 航空设计制造采用的工艺具体解析      | 生产线工作人员与设计师大需掌握的技能 |
| 评价指标 | 航空设计制造企业所用到的评价指标    | 效用、使用期限、功能       | 零件具备的规格              | 设计制造应满足相应产品要求      |

航空制造知识图谱对于知识的精度与深度要求很高,换句话说就是要求具有更细的知识粒度,知识粒度代表着基本知识单元的大小。不同领域的知识图谱对于知识粒度的大小要求也不相同。在知识图谱技术融入搜索领域后,能够通过目标问题中本体的相邻关系给出最符合的答案,说明本体甚至本体的特征属性已经成为知识图谱的最小知识单元。而航空制造知识图谱更需要将知识划分成更小的粒度,更加注重命名实体识别的精确度,加重了航空制造知识图谱构建的难度。

从知识获取层面看,领域知识图谱对知识的全面性和精确程度提出了更高的相关要求,这是由于行业中所有层级人员都是系统的客户对象,不同的用户需要的知识和功能也不相同,这就对知识图谱的精度和完整性要求很高。严格的品质要求自然就需要更多的专家权威组织,这也是领域知识图谱科学

性和准确性的关键所在,但也不能完全由专家来负责构建,所以如何构建一个质量较高、规模较大的领域知识图谱一直是一个难题。杨玉基等人<sup>[17]</sup>提出的四步法是一种精准且快捷的领域知识图谱构建方法,它能够很好地平衡自动化方法和人工参与。由于航空制造知识图谱是为了航空制造专业领域而构建的,其应用复杂度自然更复杂一些。

## 1.2 航空制造知识图谱的构建框架

目前针对航空制造领域的知识图谱构建比较匮乏,在大数据和信息技术飞速发展的今天,大量的航空制造行业知识潜藏在非结构化文本数据、半结构化表格数据、互联网网页和少数航空企业内部管理系统的数据库中,如果成功构建出完整的航空制造业知识图谱,从业者可以通过其功能强大的应用界面,如何快速、通透地操作和维护航空制造知识图谱对于有关部门各类型数据的使用和内部数据的治理具有重大意义。构建航空制造知识图谱的主要目的是抽取海量的语义关系,转换为计算机语言的航空制造知识。构建知识图谱可以采用自底向上(bottom-up)的构建方法或者自顶向下(top-down)的构建方法<sup>[18]</sup>。自底向上的构建方法是先对数据源进行知识抽取,将所得实体、关系以及属性等经审查合格后加入到知识库中,而自顶向下方式首先创建顶层概念本体,然后从数据源中抽取实体以及关系匹配并更新到最初建立的顶层本体中。由于领域知识图谱要求严苛的知识专业性和准确性,同时又需要具备完

整的本体层模型,所以领域知识图谱一般采用两者相互结合的方法进行构建<sup>[19]</sup>。如图4所示,首先利用提取和总结的知识创建模式层,之后不断挖掘新的有价值的知识与信息反馈迭代模式层。

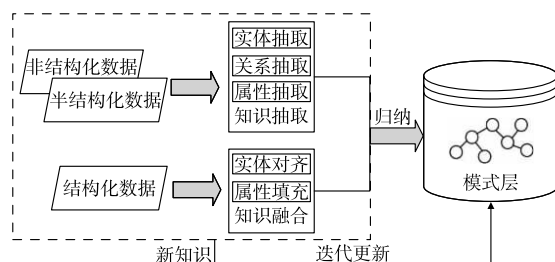


图4 自顶向下与自底向上结合的构建方法

Fig.4 Combined top-down and bottom-up construction method

领域知识图谱的架构则可考虑目标领域所具有的独特因素,如航空制造产业具有完备的生产流程,所以其图谱架构也应从此角度出发。如图5所示,本文基于航空制造全周期数据给出了航空制造知识图谱的构建模式结构。知识成分的关联关系以及关联形式共同组成了航空制造领域知识的整体结构。所以要建立航空制造知识模型,必须基于航空制造领域特有知识的分层体系、信息数据和知识元结构,其过程如图6所示。将航空制造专业知识形式化成航空制造知识本体,结合特定化的航空制造特征知识归纳推理出知识结构最终进行模型化操作。

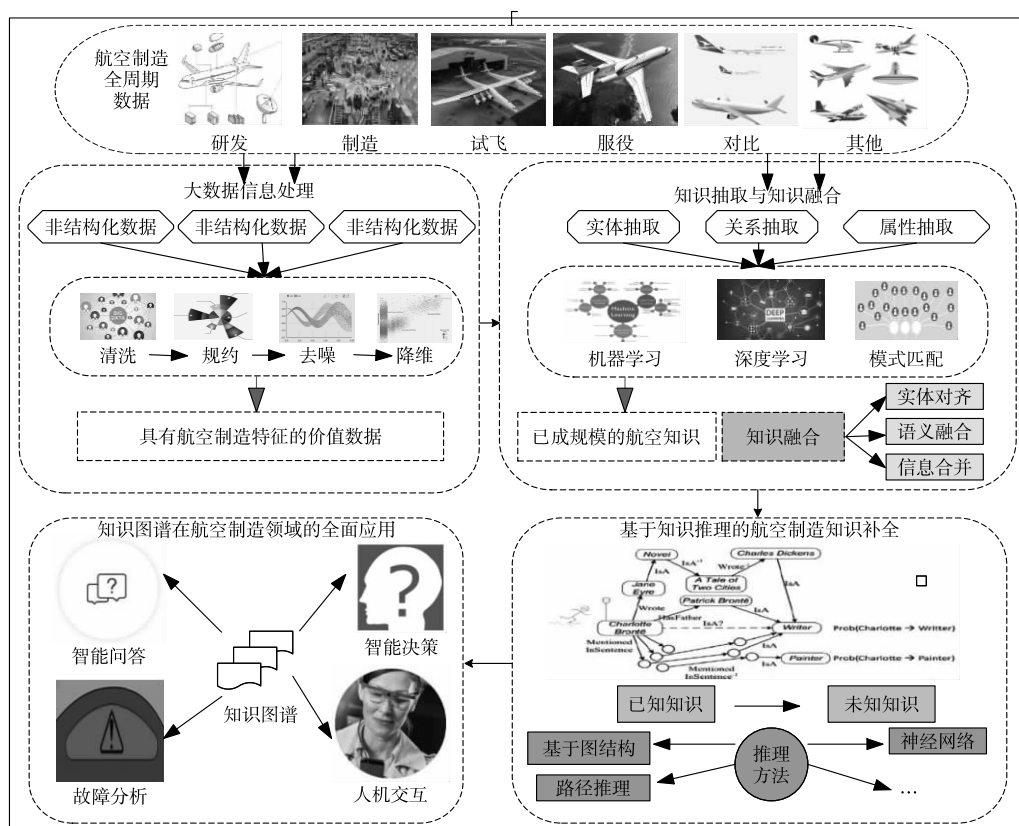


图5 基于航空制造全周期数据的知识图谱构建框架

Fig.5 Framework for building knowledge graphs based on full-cycle data of aerospace manufacturing

由于目前的通识知识图谱模型不足以表达航空制造业的知识,本文将对航空制造知识图谱构建过程中的三大关键技术,即知识抽取、知识融合和知识推理进行综述。航空制造知识抽取可从已有的结构化数据集、半结构化数据集以及非结构化数据集抽取实体信息、实体间关系以及属性等知识要素<sup>[20]</sup>;利用知识融合来解决航空制造知识库中实体信息、关系以及属性等一系列代词和实体对象之间模糊不清的问题,使知识图谱逻辑更为清晰、结构更为紧密;通过知识推理可以对航空制造

知识图谱不完整的知识进行补全,同时能从已有信息中挖掘更多有价值的信息<sup>[21]</sup>。

## 2 航空制造知识图谱构建关键技术

### 2.1 航空制造知识表示与抽取

航空制造实体中的语义信息可以以符号的形式存储,以便在知识抽取、融合、推理过程中发挥作用,这对于航空制造知识

库的构建流程有重大意义。这些年来人工智能、机器学习以及深度学习等智能学习技术取得了重大突破<sup>[22]</sup>,如今的知识表示方法主要分为利用符号逻辑进行知识表示以及基于语义网络的知识表示方法。目前对于知识图谱,一般而言采用第二种方法里的 RDF 描述知识,形式上将有效信息表示为(主—谓—宾)三元组的结构,由于 RDF 具有完善的数据描述体系,不必再进行消歧,有利于实现不同知识的互通性及标准化;而第一种方法面对规模庞大的领域知识库建设、面对具有挑战性的制造设计数据和装置数据时,仅作为辅助形式存在<sup>[23]</sup>。RDF 描述歼 20 及其关系的形式如图 7 所示,不仅是字符串构成的符号,还包含了语义信息,阐明了主体与客体之间的关系。

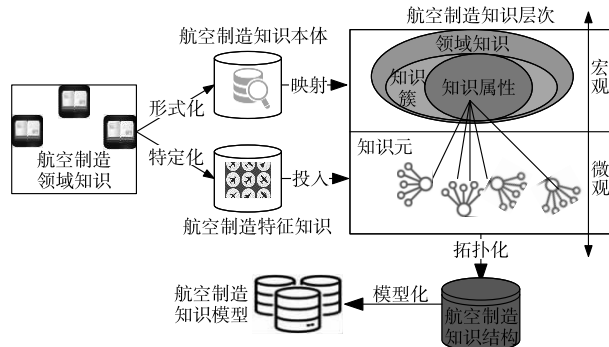


图6 航空制造知识层次结构  
Fig. 6 Aerospace manufacturing knowledge hierarchy

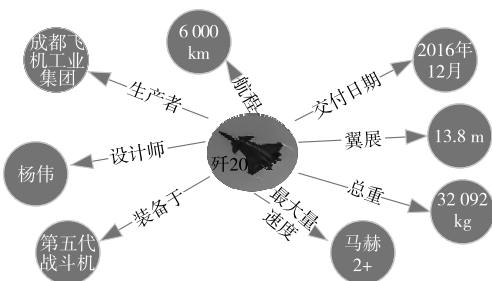


图7 RDF 对歼20的描述形式  
Fig. 7 Description of J-20 using RDF

对于最初航空制造企业内部浩如烟海的日志文件、设计说明、设备维修记录及百科数据,如要构建航空制造知识图谱首先要进行知识抽取,其目的主要是从源数据中提取特定类型的信息(如实体、关系和属性),并将这些信息以特定的形式表示和存储。如图8所示,根据知识抽取目标对象的不同,知识抽取又可分成三个子任务:实体抽取、关系抽取以及属性抽取<sup>[24]</sup>。

### 2.1.1 实体抽取

实体抽取的关键是从目标文本范围中对命名实体进行检测并分类,实体作为组成航空制造知识库的原子信息元素,知识图谱的优劣将由命名实体识别的完整度以及精确度直接决定。下面将对实体抽取的两种主要方法进行介绍:

a) 基于传统机器学习的实体抽取。传统机器学习方法是利用具有象征性的训练文档,为所有代表性的词打上语义标注,之后设计适用于该类型文本数据的特征提取策略来训练一个分类器以自动标注数据的标签。Singh 等人<sup>[25]</sup>利用基于约束的半监督学习将指定的轻量级监督注入到实体抽取的马尔可夫条件随机场(Markov conditional random field)模型中,实现了接近完全监督分类器的精度。

b) 基于深度学习的实体抽取。深度学习脱胎于机器学习领域,其建立、模拟人脑的运行模式进行分析学习的神经网络,很多学者尝试将深度学习引入实体关系抽取中。其中最具有代表性的就是 LSTM + CRF<sup>[26]</sup>,LSTM 能够使用以往的输入特征,CRF 能够利用句向量的标注信息,两者组成一种端对端的判

别式模型,能够充分地结合未来以及过去的标注来输出现在的标注。Abhyuday 等人<sup>[27]</sup>基于 CRF 的结构化学习模型和递归神经网络扩展了 LSTM-CRF 模型,并对成对势进行了显式建模,提出了一种具有 RNN 势的跳链 CRF 推理的近似形式,实验表明其对于实体抽取准确率有显著提升。

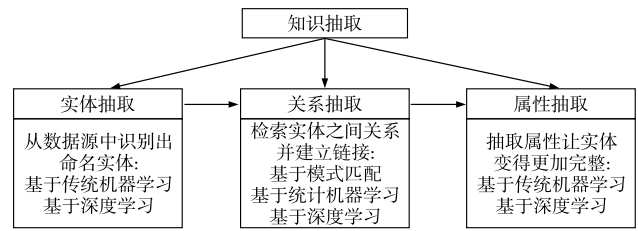


图8 知识抽取方法  
Fig. 8 Knowledge extraction methods

### 2.1.2 关系抽取

通过实体抽取得到的实体之间通常没有关联且被分割开的,关系抽取的目标就是在已完成命名实体识别的基础上挖掘实体之间的关联关系,搭建起实体间内部的语义桥梁。

#### 1) 基于模式匹配的实体关系抽取

关系抽取早期的研究是通过模式匹配(schema matching)算法来完成识别实体的语义关系,实现关系抽取任务,最初是用来设计数据库以及进行分布式数据库的研究<sup>[28]</sup>。语言学和自然语言处理学的知识在模式匹配的过程中起到了关键作用,在进行实体关系抽取任务前,通过人工构造实体的特征词典或规则并将它们存储下来。在实体关系抽取任务中,将规则与预处理后的非结构文本相匹配,提取出了三元关系组<sup>[29]</sup>。

张治等人<sup>[30]</sup>基于模式同态的框架将模式匹配问题转换为泛代数里的结构同态问题,在泛代数的基础上建立模式匹配问题的描述架构和算法模型;邵堃等人<sup>[31]</sup>结合轻型领域本体分析 CGER 标题,通过概念序列判别 CGER,载入相应的领域知识库,从信息项中提取出结构化的知识,但这种方法要求规则和词典制作者必须具备该领域较高的语言学基础和对自然语言处理的深刻理解和研究,且难度大,耗时费力。

#### 2) 基于统计机器学习的实体关系抽取

机器学习基于数据解决问题,该方法利用标注好的语料数据作为输入,模型输出为期望结果,不断提高模型的鲁棒性,然后再通过训练完成的模型对陌生语料进行实体识别。其抽取方法可分为以下三种:

a) 有监督的关系抽取。其核心思想是将大量人工标注的数据送入模型中训练,甘丽新等人<sup>[32]</sup>提出了一种实体间关系抽取方法,其利用依存句法关系组合和最近句法依赖动词特征,基于支持向量机(SVM)算法验证了该方法在关系探测和关系抽取上的有效性。

b) 半监督或无监督的关系抽取。这类方法利用少量的人工标注数据甚至没有标注数据,使用高斯混合模型(Gaussian mixture model)等算法的半监督、无监督关系抽取方法实现关系抽取。徐庆伶等人<sup>[33]</sup>将 co-training 与 tri-training 算法的思想相结合,采用两个不同参数的 SVM 分类器对非结构化样本进行标记,在已标记样本集中加入置信度高的样本进行迭代学习和调整;Chen 等人<sup>[34]</sup>提出了一种无监督关系提取算法,采用判别类别匹配(DCM)来寻找典型的和判别的词来表示不同的关系,实验结果表明了该算法的有效性。

c) 远程监督的关系抽取。远程监督能够减少对标注数据的需求,因此被大量应用于从非结构化制造文本中进行关系抽取,Sun 等人<sup>[35]</sup>提出了一种基于强化学习的标签去噪方法,利用提取网(ENET)和策略网络(PNET)两个模块组成的模型将任务建模为一个迭代过程,用于远距离监督关系提取,所提出

的方法可以胜过同期最新的关系抽取系统。

### 3) 基于深度学习的实体关系抽取

利用统计机器学习的抽取方法会在特征提取环节存在一些如边缘模糊、计算量大等问题,对于提升模型的性能有很大程度的约束。研究人员发现基于深度学习的方法在自动学习高阶语义特征的问题上有更高的准确度,所以近些年来使用深度学习的实体关系抽取方法逐渐在实体关系抽取研究领域获得越来越多的关注。

a) 利用 RNN 架构的实体关系抽取。递归神经网络算法(recursive neural network, RNN)在处理序列化输入时比前馈网络更有效。Socher 等人<sup>[36]</sup>最早提出利用 RNN 进行实体关系抽取,其核心思想是对测试语料使用句法依存分析使文本的原本顺序转变成解析树结构,之后对解析树的所有节点都分配一个向量及一个矩阵,向量包含单词的内在含义,矩阵确定怎样改变其相邻的单词以及短句的含义。

b) 利用 CNN 架构的实体关系抽取。卷积神经网络(convolutional neural network, CNN)在实体关系抽取上有很好的效果,但无法对长句子分析建模。Zhu 等人<sup>[37]</sup>提出了一种 TCA-CNN 模型,采取注意力机制,根据构建句子表示时集中的语义关系对单个单词标注不同的权重, TCA-CNN 在关系分类任务上与最先进的模型相比毫不逊色。

c) 利用 LSTM 架构的实体关系抽取。长短期记忆人工神经网络(long short-term memory network, LSTM)可以通过与聚类技术结合来大幅度降低训练和测试时长,还对解决实体之间的长期依赖问题具有帮助<sup>[38]</sup>。孙紫阳等人<sup>[39]</sup>提出了一种用以表示文本的基于最短依存路径思想的关系抽取方法,基于 LSTM 门控单元的記憶功能保存最短依存路径的表达方式,之后利用 CNN 对 LSTM 的输出进行训练,提高分类模型能力。实验结果表明,该方法能够提升抽取性能及准确性。

### 2.1.3 属性抽取

属性抽取的概念是对于航空制造实体而言的,如材料的属性包括性能、工艺、厂商等,通过属性可以让实体变得更加完整,例如遥控无人机可在室内或室外飞行。由于实体与属性的关系被认为是两者之间的一种指代性关系,所以能够通过关系抽取的思路来解决。

## 2.2 航空制造知识融合

通过知识表示与抽取,初步获得了数量可观的形式化知识,但航空制造知识库中的知识来源驳杂,知识的质量良莠不齐,来自不同数据源的知识容易重复并且知识之间的关联关系容易模糊<sup>[40]</sup>,想要获得一个优秀的航空制造知识图谱需经过知识融合,通过知识融合提升知识图谱的完整性以及质量。如图9所示,目前知识融合方法主要有实体对齐、语义融合与信息融合。

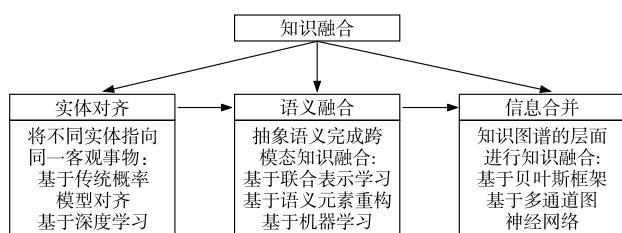


图9 知识融合方法

Fig.9 Knowledge fusion method

### 2.2.1 实体对齐

在现实世界里,一事物往往会有多种称呼,比方人们常吃的土豆也可称为马铃薯,它们都代表同一个客观事物。在航空制造知识图谱中也会出现这种现象,而通过实体对齐就能够消除歧义,将这些名称指向同一客观事物。

基于传统概率模型的对齐方法是一种基于属性相似程度

的成对相互比较的方法,没有将匹配实体对之间的关系纳入考虑,文献<sup>[41,42]</sup>将基于属性相似程度评分的实体匹配问题转换成匹配分类问题(分为匹配、能匹配以及不匹配的类型)建立起概率模型,该模型为实体对齐方法作出了巨大贡献。Chakrabarti 等人<sup>[43]</sup>基于一个同义挖掘框架将属性相似性作为输入,输出为满足朴素自然属性的近义词,且提出了两个新颖的属性相似程度计量法,并在必应系统上得到了实际应用,实验结果表明其对于识别同义实体非常有效。但基于传统概率模型的对齐方法在实体对齐过程中会产生两种错误:a)相同实体被错误归为不相等;b)不同实体被错误归为相等<sup>[44]</sup>,而且这两种情况错误的代价经常不相等。着眼于这一观察,Verykios 等人<sup>[45]</sup>提出了一种将决策成本降到最小的决策模型,它是基于传统模型再将不同的权重附加给不同的匹配状态,该模型能够在数据中存在不一致、错误或缺失值的情况下针对数据库记录对的匹配或连接问题的最优解决方案。

Mudgal 等人<sup>[46]</sup>对基于深度学习的实体对齐方法进行了综述,定义了深度学习在实体对齐领域的解决方法,并将这些解决方法分为注意力(attention)、平滑倒词频(smooth inverse frequency, SIF)、循环神经网络(RNN)和混合(hybrid)四种具体的表示。而在监督学习方面,Chochinwala 等人<sup>[47]</sup>使用经典的分类回归决策树算法(CART)以及矢量量化方法,将机器学习和统计技术结合起来进行实体的判别与分析,通过识别和消除冗余或无用的参数来降低匹配算法的复杂性;Chen 等人<sup>[48]</sup>研究了跨语言实体对齐问题,提出了一种迭代协同训练方法(Crea)来训练一对独立模型。这两种模型可以分别提取多语种知识图谱的属性和关系特征,进一步过滤了动态阈值,增强了对两种模型的监控能力。

### 2.2.2 语义融合

航空装备设计制造过程中产生的数据为多模态数据,是通过不同领域或视角获取到的数据,描述这些数据的每一个领域或视角叫做一个模态<sup>[49]</sup>。对于多模态数据来说,模态内部具有一定程度的关联性,所有模态都可以为其余模态提供或多或少的信息。基于语义的知识融合方法对于解决多模态数据融合问题有其独到之处,语义融合算法能够理解不同模态数据之间的特征关系以及每个模态数据的内在含义,利用人类大脑思考的方式在知识融合过程中抽象各类模态的语义内涵从而完成多模态数据的知识融合。

Zhao 等人<sup>[50]</sup>提出了一种框架,结合了线性判别分析(LDA)和K-均值聚类算法的简单性,在一个联合训练方案中利用在一个模态中智能学习获得的标签来挖掘另一个模态中的判别子式,该算法很容易地适应半监督学习和非线性问题。苏佳林等人<sup>[51]</sup>基于联合训练框架(co-training),从结构视角以及语义视角下分别训练基于两个不同视角的联合表示学习(joint embedding)的语义融合模型,并从中选出最可靠的实体对齐输出用以辅助训练另一个视角下的模型,达到提升模型能力的效果,并提出基于属性强约束联合训练过程中可能产生的语义漂移;Wang 等人<sup>[52]</sup>提出了一种基于语义元素重构的多源创新知识融合算法,为保证创新知识提炼的可信性和有序性,提出了相应的语义冲突解决规则和基于知识社会信任度的协同编辑策略;Xie 等人<sup>[53]</sup>对来自农业信息源的信息结果进行分析、合并,基于农业领域本体介绍了知识融合方法,并给出了一个基于知识融合的信息访问接口。

### 2.2.3 信息合并

信息合并不再是关注于知识图谱内的实体,更多是从整个知识图谱的层次上进行大范围的知识融合,利用已经成熟的优越的现有知识图谱来丰富目标知识图谱的内容,完善其中包含的知识体系。信息合并算法在知识融合中得到了继承和发展<sup>[54]</sup>,但是现如今的知识图谱或者知识库都是为了满足研究

人员或者构建者特定需求而去创建构造的,其中知识复杂程度高,知识冗余且包含许多错误知识,并不能直接融入目标知识图谱,信息合并算法可以有效解决这些问题。

Sun 等人<sup>[55]</sup>提出了一种新的数学方法,将 Dempster-Shafer 理论推广到组合型证据的融合,能够将来自领域专家和数据来源的异质知识整合到深度挖掘和决策支持中;Coussemont 等人<sup>[56]</sup>详细阐述了贝叶斯框架在决策支持系统(decision support system)中融合多个信息源的自然优势,提出了一种贝叶斯决策支持框架,该框架将主观的人类专家意见与更多的人的主观意见正式融合在一起;Cao 等人<sup>[57]</sup>提出了一种新的具有多通道的图神经网络模型(MuGNN),利用多个通道将两个不同的知识图谱进行鲁棒编码之后对知识图谱学习对齐,并提升知识图谱质量,其在五个公开可用的数据集进行了广泛的实验,证明了其优越性能。

### 2.3 航空制造知识推理

知识推理就是在现有航空制造领域知识库的成果上继续发掘隐含知识与知识之间未知的关系,达到扩大、完善和丰富知识库的目的,保证知识图谱的多样性与完整性。利用航空制造知识图谱,工作人员能够基于知识推理完成生产线数据搜集、故障诊断、制造方法、避免装备设计差错等。但是航空制造领域由于其特殊性,面对数据量大、关系复杂的特点,并不能使用传统的知识推理方法,大大增加了构建航空制造知识图谱模型的难度。近年来随着知识图谱成为研究热点,也有许多适用于知识图谱的知识推理方法逐渐涌现。

#### 2.3.1 基于神经网络的知识推理

利用神经网络进行知识推理的主要流程为:先将知识图谱内部的信息元组表示为向量(词向量、句向量),将其送入神经网络内部,利用神经网络的泛化能力和学习能力对知识图谱中的信息元组建模,并通过训练一直提升信息元组的得分,最终选择候选实体完成知识推理。更具体地说,按照推理依据的区别可以进一步细分为基于图谱结构的推理、基于语义信息的推理等<sup>[58]</sup>。

Khan 等人<sup>[59]</sup>利用人工神经网络(artificial neural network)将癌症分类为特定的诊断类别,该研究证明了这些方法在肿瘤诊断和确定治疗候选靶点方面的潜在应用。Das 等人<sup>[60]</sup>提出了三个重要的建模进展:对关系、实体以及实体类型进行协同推理;在单层 RNN 中共享参数对全体关系的逻辑结构进行表达;基于神经网络的注意力机制建模来合并多条路径。Graves 等人<sup>[61]</sup>分享了一种机器学习模型,名为可微分型神经计算机(differentiable neural computer, DNC),其就像一个神经网络可以从数据中学习如何去运作,能够成功地回答设计用于模拟自然语言推理和推理问题的综合问题;Neelakantan 等人<sup>[62]</sup>利用递归神经网络(RNN)将二元关系作为输入向量嵌入路径中,使其能够推广到训练时看不见的路径,而且还可以使用单个高容量 RNN 来预测在训练组合模型时没有看到的新关系类型。

#### 2.3.2 基于图结构的知识推理

受到传统知识推理的启发,一些研究者创造性地将知识推理延伸到机器视觉领域,基于知识图谱的图结构的特点将推理与机器视觉相结合,将潜藏在知识图谱中的规则挖掘出来加以利用完成推理,对于知识推理任务取得了不错的效果。

Wang 等人<sup>[63]</sup>将图像识别与知识推理相结合,通过端到端可训练图形推理模型(graph reasoning model, GRM),将场景中社会关系和语义上下文线索之间相关性的常识知识结合到深层神经网络中,以解决社会关系识别的问题;Kampffmeyer 等人<sup>[64]</sup>提出了一种远节点之间带有直接链路的密度图传播(dense graph propagation, DGP)模型, DGP 模型基于知识图谱的层次图结构,通过额外连接按照节点的距离为它们分配不同的权重,便于改善图中的信息传播,发挥了图结构的便利性,并

且解决了知识图谱内部知识过于稀疏的问题;Li 等人<sup>[65]</sup>基于门控递归单元和现代优化技术提出了一种门控图形序列神经网络(gated graph sequence neural network, GGS-NN),该模型在一系列问题上具有理想的归纳偏差;Velickovic 等人<sup>[66]</sup>提出了一种图形注意网络(graph attention network, GAT),它基于图数据结构进行工作,利用自我注意层解决现存的近似方法的缺点, GAT 模型已经在 Cora、Citeseer 和 Pubmed 引文网络数据集上取得了较先进的结果。

## 3 航空制造知识图谱对比及应用

### 3.1 航空制造知识图谱构建对比

经过上文的综述可以看出航空制造领域知识图谱的构建因其领域的特殊性,在构建过程中自然也有其特点,与通用知识图谱在构建技术上的不同主要有以下几点:a)由于航空制造领域的知识聚焦于航空制造领域的应用主体,如设备类型、制造工艺、材料属性等,不同于通用知识图谱面向所有领域,航空制造知识图谱在概念化的过程中,需要专家知识来帮助构建特定的领域本体和确定主题领域;b)航空制造知识图谱要求严苛的知识专业性和准确性,同时又需要具备完整的本体层模型,所以将自顶向下与自底向上方法相结合,不同于通用知识图谱常用的自底向上构建方法;c)通用知识图谱数据来源广泛且基本不作要求,而航空知识图谱则需要采用企业内部积累的工程数据及政府部门发布的工业标准等,几乎无法从公开数据取得;d)面向通用知识图谱的知识抽取技术已经较为成熟,而面向航空制造领域的抽取算法还在起步阶段,算法准确率和召回率相对较低;e)通用领域的知识推理技术大多关注于静态数据,忽略了时间信息,而在航空制造领域有些时候应该关注动态信息,如生产线装配过程对产品质量的影响。两者构建技术上的异同如表 2 所示。

表2 航空制造知识图谱与通用知识图谱构建技术异同点  
Tab.2 Similarities and differences of construction techniques between aerospace manufacturing knowledge graph and general knowledge graph

| 对比项  | 通用知识图谱   | 航空制造知识图谱    |
|------|----------|-------------|
| 数据   | 公开数据     | 专家知识        |
| 方法   | 自底向上     | 自底向上与自顶向下结合 |
| 本体   | 需要       | 需要          |
| 知识抽取 | 技术比较成熟   | 技术不成熟       |
| 知识融合 | 融合其他知识图谱 | 融合其他知识图谱    |
| 知识推理 | 静态信息     | 有时候关注动态信息   |
| 数据来源 | 公开       | 极少公开        |

### 3.2 航空制造知识图谱应用

知识图谱可以做到直接明了、清晰有效地表达出实体内部的联系。换言之,就是通过把众多不同种类的知识联系在一起形成一个知识网络,让人们能够做到基于知识间的联系来分析问题。利用知识图谱可以为航空制造信息系统中海量、多源、动态的航空制造大数据的表示、架构、管理及使用提供一种更加先进、更加高效的方法,让系统实现智能化,系统的使用更符合人类的思维模式,提升系统使用者的使用舒适感。目前制造业知识图谱技术主要用于航空智能制造问答系统、设备故障上报系统、制造领域智能搜索、制造决策支持系统等。

#### 3.2.1 航空制造智能问答

随着深度学习的迅速发展及其在自然语言领域的普遍应用,基于深度学习的知识图谱问答研究逐渐成为人们关注的焦点。在此之后,学者着眼于将自然语言查询转换为结构化查询句,以识别用户自然语言查询中的实体和属性,并使用指定的实体和属性检索知识图谱的答案。航空智能制造问答系统是制造信息检索系统的一种高级形式,将收到的大问题拆解成为多个小的子问题,对每个小问题在知识库中查找对应的答案后



进行答案合并,之后以简洁精确的人类语言作为输出进行直观的展示并给用户问题的解答。现如今已有一些产品与知识图谱结合得十分成熟,如阿里巴巴的淘宝客服机器人就是其中的代表。但在航空制造领域,因为现有航空制造知识图谱的推理能力的不足,导致目前市场上还未出现较为完善的航空制造问答系统。作者在此给出合理的系统构建框架,对用户问题进行自然语言处理,基于航空制造知识图谱搜索合适答案,具体过程如图 10 所示。

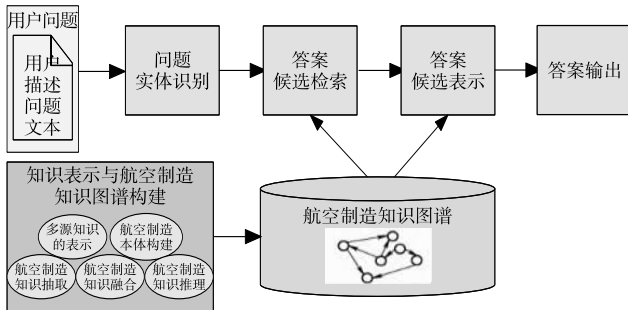


图 10 航空制造智能问答  
Fig. 10 Aerospace manufacturing intelligence Q&A

国内外对于将知识图谱与机械制造问答系统的结合运用展开了许多研究。Shi<sup>[67]</sup>提出了一种基于自然语言处理的机械智能问答系统,系统分为三个数据加载部分、SPARQL 语句分析部分和数据查询部分,其实现的知识图谱存储访问系统在保证查询性能的前提下做到了均衡负载的知识图谱的有效存储。

### 3.2.2 航空制造智能决策

随着信息的爆炸式增长,对信息处理的要求不断提高。如何快速准确地整合大量分散、孤立的信息,将信息的优势充分利用,从而转变为知识的优势,最后为智能决策制造优势、为产业链的安全高效保驾护航;同时构建专业的航空知识图谱,缓解信息超量负载的问题,可以为决策者提供更多的可利用的信息。航空制造行业可基于知识图谱进行智能分析以及决策,根据生产线温度、材料供给、机床工作效率等数据自动生成合理的生产方案,还能对工作人员的各类操作进行智能分析,及时发出提醒,避免损害设备的不合理操作。如图 11 所示,智能决策系统基于航空制造数据的特点构建决策模型,通过大数据分析技术实现实时决策功能,并根据方案表现逆向优化模型。

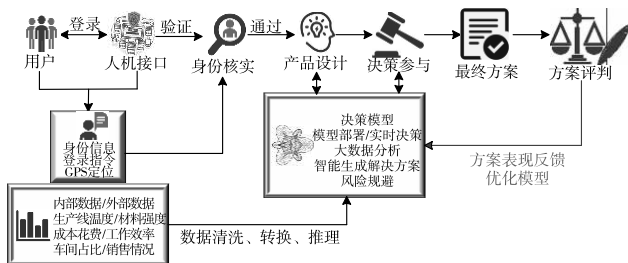


图 11 智能决策系统框架  
Fig. 11 Intelligent decision system framework

Wang 等人<sup>[68]</sup>构造了一个空战知识图谱实例,通过知识图谱进行资源共享和信息挖掘,发现任务与机动之间的隐藏关系,可用于任务优化和机动预测,辅助空战决策,有助于更好地实现领域专家与计算机系统之间的协调,实现最优任务规划;Shah 等人<sup>[69]</sup>提出了一种基于价值风险图谱的制造决策工具,用于制造场景的性能评估,该图谱汇集了与绩效、风险和决策有关的不同概念,以在决策背景下提供总体解决方案。

### 3.2.3 航空设备智能故障诊断

设备故障诊断是指诊断故障设备或产品的故障机理、影响、发展变化规律和发生概率等内容<sup>[70]</sup>。航空制造机电部件数量繁多,高强度的工作以及使用年限的增加导致故障频繁发

生,且故障原因多样化,甚至很多器件的故障因缺乏专业人员协助分析很难找到真正的故障原因。但知识图谱可以拓展故障事件以及故障部件的描述形式,将存在于各类故障诊断报表和故障案例库中的数据整合筛选,从中抽取有用的信息并整理作为知识构建的基础,利用该知识库更加精准地定位系统失效部位及其原因。因此航空制造设备故障分析作为当前航空技术研究的一个热点问题<sup>[71]</sup>,通过知识图谱技术建立航空制造故障快速分析手段的需求日益迫切。

杨浩伟<sup>[72]</sup>通过构建制造领域知识图谱,在其基础上开发带有实时补全以及修改功能的系统软件,帮助解决非专业人士在手工录入问题描述文本时,容易出现用词不准确以及语言不规范等问题导致的隐患排查出现误差,提高了故障排查的准确度以及效率。基于知识图谱的故障分析系统可在用户录入故障描述时,通过输入的故障特征条件检索关联信息进行实时提示,帮助用户正确录入故障信息;同时通过知识图谱提供的后端服务对录入的实体关系正确性进行判断,实现对用户输入内容的实时更正和补全<sup>[73]</sup>,其工作过程如图 12 所示。

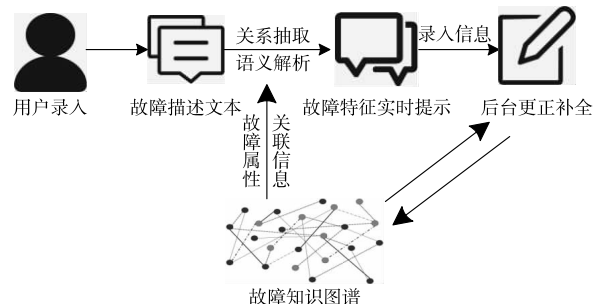


图 12 故障分析上报系统流程  
Fig. 12 Flow chart of fault analysis and reporting system

## 4 航空制造知识图谱的挑战与展望

### 4.1 航空制造知识图谱面临的挑战

知识图谱技术为基于互联网的知识互联提供了技术支持,从谷歌最初提出知识图谱的概念到现在,研究人员对其的关注度依然居高不下。知识图谱集合了人工智能、自然语言处理、信息抓取、可视化分析等多种研究热点以及应用技术,其处于多种学科的交叉领域,是吸取了各领域学科知识成果的实用技术。虽然目前知识图谱已成功应用于医疗、金融以及电商等领域,但仍有几个不小的障碍和挑战伫立在航空制造知识图谱构建的关键环节。

a) 航空制造数据来源有限。构建知识图谱首先需要海量的数据,而航空制造业领域数据来源不像如医疗领域,从网页上就能爬取大量可用知识。航空制造数据的来源主要是相关培训资料、教材或者企业内部数据,培训资料或教材包含了众多领域的概念和专业术语,同时具备极高的权威性和准确性,但数据量略显欠缺,而企业内部数据需要获得企业许可,一般很难获得。没有足够庞大的数据作为支撑,航空制造知识图谱则不能做到知识的准确性以及多样性,如何模仿人类思维进行小样本学习也是未来构建知识图谱研究的热点。

b) 知识图谱的高效存储问题。如今大部分知识图谱主要存储在图数据库中,虽然图数据库的查询结果十分直观,但图数据库也有其缺点,例如不能使用 SQL 查询、并发性要求高、查询效率较低。并且知识图谱对于数据量的需求过高,由于如今数据越来越庞大,将所有数据存放在一个地方不切实际,实现知识图谱的分布式存储意义重大,但如何在不影响知识图谱功能与组织结构的前提下进行分割存储,实现其负载均衡也是航空制造知识图谱应用需要解决的关键问题。

c) 知识图谱的应用。对于航空制造领域知识图谱的应用

场景和方式还比较有限,就算是在知识图谱运用比较成熟的行业也只是停留在智能搜索、智慧问答、辅助决策等初级阶段,仍具有广阔的探索空间。虽然航空制造管理系统运行前期产生的数据较少,但当数据积累到一定程度之后,可以对这些数据进行大数据挖掘、处理以及分析<sup>[74]</sup>,从中挖掘到对设计制造有用的信息。在此基础上促进人工智能在制造业领域的发展,智能分析并满足人们的需求,可为巨型知识图谱在航空制造领域的应用提供更加深远的道路。

## 4.2 展望

如何将制造知识成熟、准确地运用于智能制造生命周期以解决实际问题引入知识图谱技术的契机, Kim 等人<sup>[75]</sup>提出了一个知识图谱模型来获取并表示制造企业组织知识并成功提高企业绩效; Zhang 等人<sup>[76]</sup>提出了一种基于图的知识重用方法以辅助决策新产品的研发; Ruiz 等人<sup>[77]</sup>通过知识图谱实现专家知识共享从而支持工业维护管理的协作决策。无论是在制造业的生产、维护、决策及管理等方面,合理的知识图谱技术都能提供有效的解决方案,实现制造智能化。

而在应用范围更精细的航空制造领域,研究者可以从航空制造对于安全性能的高度需求出发,充分发挥知识推理技术的优势。Sarazin 等人<sup>[78]</sup>提出了一种异构信息网络构成的专家系统并应用于航空系统的异常检测;聂同攀等人<sup>[79]</sup>提出的无人机电源系统的知识图谱实现了无人机电源故障的高效诊断。

笔者认为未来要使知识图谱技术在航空制造领域的应用更进一步,必须吸取其他成熟行业的经验,如医疗行业的智能问诊、金融行业的风控分析等。结合航空制造领域特点,在不断提高通用技术(知识抽取等)精确度的同时找到航空制造领域独特的应用方案,这也将是未来研究的探索方向。

## 5 结束语

本文首先对航空制造知识图谱的定义和框架进行了详细的阐述,较为全面地对知识图谱构建的三大关键手段进行了综述,并结合航空制造行业自身特点与目前的相关研究对知识图谱在航空制造领域的应用进行了展望与分析,提出了三点切合实际的应用方向及其解决方案;概括了目前伫立在航空制造知识图谱构建关键环节的困难与障碍,并对其今后的发展道路提出了建议。希望可以为进一步展开对知识图谱在航空制造领域的研究者提供启发,同时为相关企业针对知识图谱的实际应用提供参考。

### 参考文献:

- [1] 徐增林, 盛泳潘, 贺丽荣, 等. 知识图谱技术综述[J]. 电子科技大学学报, 2016, 45(4): 589-606. (Xu Zenglin, Sheng Yongpan, He Lirong, et al. Review on knowledge graph techniques[J]. Journal of University of Electronic Science and Technology of China, 2016, 45(4): 589-606.)
- [2] Sheth A, Thirunarayan K. Semantics empowered Web 3. 0: managing enterprise, social, sensor, and cloud-based data and services for advanced applications[M]. San Rafael, CA: Morgan & Claypool Publishers, 2012: 35-36.
- [3] Mayer-Schönberger V, Cukier K. Big data: a revolution that will transform how we live, work and think [M]. London: Hodder & Stoughton, 2013.
- [4] Pujara J, Miao Hui, Getoor L, et al. Ontology-aware partitioning for knowledge graph identification[C]//Proc of Workshop on Automated Knowledge Base Construction. New York: ACM Press, 2013: 19-24.
- [5] Xu Qiuxue, Niu Na, Quan Yongmin, et al. Research on the development of text mining technology based on bibliometrics and knowledge map visualization[J]. Scientific Journal of Information Engineering, 2017, 7(1): 15-26.
- [6] 林明. 基于知识图谱的交互关系浏览与分析: 可视化模型与系统实现[D]. 杭州: 浙江大学, 2017. (Lin Ming. Interactive relation exploration and analysis based on knowledge graph: visualization model and system implementation[D]. Hangzhou: Zhejiang University, 2017.)
- [7] Pan J Z, Vetere G, Gomez-Perez J M, et al. Exploiting linked data and knowledge graphs in large organisations[M]. Cham: Springer, 2017.
- [8] 姚萍, 李坤伟, 张一帆, 等. 知识图谱构建技术综述[J]. 信息系统工程, 2020(5): 121-123. (Yao Ping, Li Kunwei, Zhang Yifan. Review of knowledge graph construction techniques [J]. China CIO News, 2020(5): 121-123.)
- [9] Bizer C, Lehmann J, Kobilarov G, et al. DBpedia: a crystallization point for the Web of data[J]. Journal of Web Semantics, 2009, 7(3): 154-165.
- [10] Suchanek F M, Kasneci G, Weikum G. YAGO: a large ontology from Wikipedia and WordNet[J]. Journal of Web Semantics, 2008, 6(3): 203-217.
- [11] Bollacker K, Evans C, Paritosh P, et al. Freebase: a collaboratively created graph database for structuring human knowledge[C]//Proc of ACM SIGMOD International Conference on Management of Data. New York: ACM Press, 2008: 1247-1250.
- [12] Xu Bo, Liang Jiaqing, Xie Chenhao, et al. CN-DBpedia2: an extraction and verification framework for enriching Chinese encyclopedia knowledge base[J]. Data Intelligence, 2019, 1(3): 271-288.
- [13] Niu Xing, Sun Xinruo, Wang Haofen. Zhishi.me: weaving Chinese linking open data[C]//Proc of the 10th International Semantic Web Conference. Berlin: Springer, 2011: 205-220.
- [14] 中国电子技术标准化研究院. 知识图谱标准化白皮书(2019版)[R/OL]. (2019-08). <http://www.199it.com/archives/940537.html>. (China Electronics Standardization Institute. Knowledge graph standardization white paper (2019) [R/OL]. (2019-08). <http://www.199it.com/archives/940537.html>.)
- [15] 刘烨宸, 李华昱. 领域知识图谱研究综述[J]. 计算机系统应用, 2020, 29(6): 1-12. (Liu Yechen, Li Huayu. Survey on domain knowledge graph research[J]. Computer Systems & Applications, 2020, 29(6): 1-12.)
- [16] 唐锚, 高凯丽, 张小娟. 面向大数据的北京水务数据融合技术研究[J]. 水利信息化, 2019(6): 9-17, 22. (Tang Mao, Gao Kaili, Zhang Xiaojuan. Research on big data oriented data fusion technology of Beijing water affair[J]. Water Resources Informatization, 2019(6): 9-17, 22.)
- [17] 杨玉基, 许斌, 胡家威, 等. 一种准确而高效的领域知识图谱构建方法[J]. 软件学报, 2018, 29(10): 2931-2947. (Yang Yuji, Xu Bin, Hu Jiawei, et al. Accurate and efficient method for constructing domain knowledge graph[J]. Journal of Software, 2018, 29(10): 2931-2947.)
- [18] 石刚. 一种基于知识图谱的用户搜索意图挖掘方法的研究[D]. 北京: 国际关系学院, 2016. (Shi Gang. Research on a method of mining user's search intent based on knowledge graph[D]. Beijing: University of International Relations, 2016.)
- [19] 侯梦薇, 卫荣, 陆亮, 等. 知识图谱研究综述及其在医疗领域的应用[J]. 计算机研究与发展, 2018, 55(12): 2587-2599. (Hou Mengwei, Wei Rong, Lu Liang, et al. Research review of knowledge graph and its application in medical domain[J]. Journal of Computer Research and Development, 2018, 55(12): 2587-2599.)
- [20] 梅发国, 戴大伟, 张冀. 基于知识图谱的战场目标关系融合技术[J]. 指挥信息系统与技术, 2017, 8(5): 81-86. (Mei Faguo, Dai Dawei, Zhang Ji. Battlefield target relation fusion technology based on knowledge mapping[J]. Command Information System and Technology, 2017, 8(5): 81-86.)
- [21] 赵彤宇. 医学知识图谱构建方法研究及在疾病诊断中的应用[D]. 济南: 齐鲁工业大学, 2020. (Zhao Tongyu. Research on medical knowledge graph construction method and its application in disease diagnosis[D]. Jinan: Qilu University of Technology, 2020.)
- [22] 杨晓慧, 万睿, 张海滨, 等. 基于符号语义映射的知识图谱表示学习算法[J]. 计算机研究与发展, 2018, 55(8): 1773-1784. (Yang Xiaohui, Wan Rui, Zhang Haibin, et al. Semantical symbol mapping embedding learning algorithm for knowledge graph[J]. Journal of Computer Research and Development, 2018, 55(8): 1773-1784.)
- [23] 袁凯琦, 邓扬, 陈道源, 等. 医学知识图谱构建技术与研究进展[J]. 计算机应用研究, 2018, 35(7): 1929-1936. (Yuan Kaiqi, Deng Yang, Chen Daoyuan, et al. Construction techniques and re-



- search development of medical knowledge graph[J]. *Application Research of Computers*, 2018, 35(7):1929-1936. )
- [24] 吴超. 面向突发事件领域的事理图谱平台的设计与实现[D]. 成都:电子科技大学, 2020. (Wu Chao. Design and implementation of event graph platform for the field of emergency[D]. Chengdu: University of Electronic Science & Technology of China, 2020:32-34. )
- [25] Singh S, Hillard D, Leggetter C. Minimally-supervised extraction of entities from text advertisements[C]//Proc of Annual Conference of the North American Chapter of the Association for Computational Linguistics. Stroudsburg, PA: Association for Computational Linguistics, 2010:73-81.
- [26] Wintaka D C, Bijaksana M A, Asror I. Named-entity recognition on Indonesian tweets using bidirectional LSTM-CRF [J]. *Procedia Computer Science*, 2019, 157:221-228.
- [27] Abhyuday N J, Hong Yu. Structured prediction models for RNN based sequence labeling in clinical text[C]//Proc of Conference on Empirical Methods in Natural Language Processing. Stroudsburg, PA: Association for Computational Linguistics, 2016:856-865.
- [28] Batini C, Lenzerini M, Navathe S B. A comparative analysis of methodologies for database schema integration[J]. *ACM Computing Surveys*, 1986, 18(4):323-364.
- [29] 刘辉, 江千军, 桂前进, 等. 实体关系抽取技术研究进展综述[J]. 计算机应用研究, 2020, 37(22):1-5. (Liu Hui, Jiang Qianjun, Gui Qianjin, et al. Review of research progress of entity relationship extraction[J]. *Application Research of Computers*, 2020, 37(22):1-5. )
- [30] 张治, 车皓阳, 施鹏飞, 等. 模式匹配问题的描述框架与算法模型[J]. 模式识别与人工智能, 2006, 19(6):715-721. (Zhang Zhi, Che Haoyang, Shi Pengfei. Framework and algorithm model of schema matching problem[J]. *Pattern Recognition and Artificial Intelligence*, 2006, 19(6):715-721. )
- [31] 邵望, 杨春磊, 钱立宾, 等. 基于模式匹配的结构化信息抽取[J]. 模式识别与人工智能, 2014, 27(8):758-768. (Shao Kun, Yang Chunlei, Qian Libing, et al. Structured information extraction based on pattern matching[J]. *Pattern Recognition and Artificial Intelligence*, 2014, 27(8):758-768. )
- [32] 甘丽新, 万常选, 刘德喜, 等. 基于句法语义特征的中文实体关系抽取[J]. 计算机研究与发展, 2016, 53(2):284-302. (Gan Lixin, Wan Cangxuan, Liu Dexi, et al. Chinese named entity relation extraction based on syntactic and semantic features[J]. *Journal of Computer Research and Development*, 2016, 53(2):284-302. )
- [33] 徐庆伶, 汪西莉. 一种基于支持向量机的半监督分类方法[J]. 计算机技术与发展, 2010, 20(10):115-117, 121. (Xu Qingling, Wang Xili. A novel semi-supervised classification method based on SVM[J]. *Computer Technology and Development*, 2010, 20(10):115-117, 121. )
- [34] Chen Jinxiu, Ji Donghong, Tan C L, et al. Unsupervised feature selection for relation extraction[C]//Proc of Conference Including Posters/Demos and Tutorial Abstracts. Stroudsburg, PA: Association for Computational Linguistics, 2002:262-267.
- [35] Sun Tingting, Zhang Chunhong, Ji Yang, et al. Reinforcement learning for distantly supervised relation extraction[J]. *IEEE Access*, 2019, 7:98023-98033.
- [36] Socher R, Huval B, Manning C D, et al. Semantic compositionality through recursive matrix-vector spaces[C]//Proc of Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. Stroudsburg, PA: Association for Computational Linguistics, 2012:1201-1211.
- [37] Zhu Jizhao, Qiao Jianzhong, Dai Xinxiao, et al. Relation classification via target-concentrated attention CNNs[C]//Proc of the 24th International Conference on Neural Information Processing. Cham: Springer, 2017:137-146.
- [38] Sundermeyer M, Schlüter R, Ney H. LSTM neural networks for language modeling[C]//Proc of the 13th Conference in the Annual Series of InterSpeech Events. 2012:194-197.
- [39] 孙紫阳, 顾君忠, 杨静. 基于深度学习的中文实体关系抽取方法[J]. 计算机工程, 2018, 44(9):164-170. (Sun Ziyang, Gu Junzhong, Yang Jing. Chinese entity relation extraction method based on deep learning[J]. *Computer Engineering*, 2018, 44(9):164-170. )
- [40] 孙郑煌, 鄂海红, 宋美娜, 等. 基于大数据技术的医学知识图谱构建方法[J]. 软件, 2020, 41(1):13-17. (Sun Zhengyu, E Haihong, Song Meina, et al. Method of medical knowledge graphs construction based on big data technology[J]. *Computer Engineering & Software*, 2020, 41(1):13-17. )
- [41] Newcombe H B, Kennedy J M, Axford S J, et al. Automatic linkage of vital records[J]. *Science*, 1959, 130(3381):954-959.
- [42] Fellegi I P, Sunter A B. A theory for record linkage[J]. *Journal of the American Statistical Association*, 2012, 64(328):1183-1210.
- [43] Chakrabarti K, Chaudhuri S, Cheng Tao, et al. A framework for robust discovery of entity synonyms[C]//Proc of ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2012:1384-1392.
- [44] 庄严, 李国良, 冯建华, 等. 知识库实体对齐技术综述[J]. 计算机研究与发展, 2016, 53(1):165-192. (Zhuang Yan, Li Guoliang, Feng Jianhua, et al. Survey on entity alignment of knowledge base[J]. *Journal of Computer Research and Development*, 2016, 53(1):165-192. )
- [45] Verykios V S, Moustakides G V, Elfeky M G. A Bayesian decision model for cost optimal record matching[J]. *The VLDB Journal*, 2002, 12(1):28-40.
- [46] Mudgal S, Li Han, Rekatsinas T, et al. Deep learning for entity matching: a design space exploration[C]//Proc of International Conference on Management of Data. New York: ACM Press, 2018:19-34.
- [47] Cochinwala M, Kurien V, Lalk G, et al. Efficient data reconciliation[J]. *Information Sciences*, 2001, 137(1-4):1-15.
- [48] Chen Baiyang, Chen Xiaoliang, Lu Peng, et al. CAREA: cotraining attribute and relation embeddings for cross-lingual entity alignment in knowledge graphs[J]. *Discrete Dynamics in Nature and Society*, 2020, 2020:article ID 6831603.
- [49] 李秋岑. 基于超图的多模态融合算法研究与应用[D]. 大连:大连理工大学, 2018. (Li Qiucen. Research and application of multimodal fusion algorithm based on hypergraph[D]. Dalian: Dalian University of Technology, 2018. )
- [50] Zhao Xuran, Evans N, Dugelay J L. A subspace co-training framework for multi-view clustering[J]. *Pattern Recognition Letters*, 2014, 41(5):73-82.
- [51] 苏佳林, 王元卓, 靳小龙, 等. 融合语义和结构信息的知识图谱实体对齐[J]. 山西大学学报:自然科学版, 2019, 42(1):23-30. (Su Jialin, Wang Yuanzuo, Jin Xiaolong, et al. Knowledge graph entity alignment with semantic and structural information[J]. *Journal of Shanxi University: Natural Science Edition*, 2019, 42(1):23-30. )
- [52] Wang Gangfeng, Hu Yongbiao, Tian Xitian, et al. An integrated open approach to capturing systematic knowledge for manufacturing process innovation based on collective intelligence[J]. *Applied Sciences*, 2018, 8(3):DOI:10.3390/app8030340.
- [53] Xie Nengfu, Wang Wensheng, Yang Xiaorong, et al. Rule-based agricultural knowledge fusion in Web information integration[J]. *Sensor Letters*, 2012, 10(1):DOI:10.1016/j.njas.2010.10.026.
- [54] 朱祥, 张云秋. 近年来知识融合研究进展与趋势[J]. 图书情报工作, 2019, 63(16):143-150. (Zhu Xiang, Zhang Yunqiu. Progress and trends of knowledge fusion research in recent years[J]. *Library and Information Service*, 2019, 63(16):143-150. )
- [55] Sun Lin, Wang Yanzhang. A multi-attribute fusion approach extending Dempster-Shafer theory for combinatorial-type evidences[J]. *Expert Systems with Applications*, 2018, 96(4):218-229.
- [56] Coussement K, Benoit D F, Antioco M. A Bayesian approach for incorporating expert opinions into decision support systems: a case study of online consumer-satisfaction detection[J]. *Decision Support Systems*, 2015, 79(12):24-32.
- [57] Cao Yixin, Liu Zhiyuan, Li Chengjiang, et al. Multi-channel graph neural network for entity alignment[C]//Proc of the 57th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: Association for Computational Linguistics, 2019:1452-1461.
- [58] 张仲伟, 曹雷, 陈希亮, 等. 基于神经网络的知识推理研究综述[J]. 计算机工程与应用, 2019, 55(12):8-19, 36. (Zhang Zhongwei, Cao Lei, Chen Xiliang, et al. Survey of knowledge reasoning based on neural network[J]. *Computer Engineering and Applications*, 2019, 55(12):8-19, 36. )
- [59] Khan J, Wei J S, Ringnér M, et al. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural

- networks[J]. *Nature Medicine*, 2001, 7(6): 673-679.
- [60] Das R, Neelakantan A, Belanger D, *et al.* Chains of reasoning over entities, relations, and text using recurrent neural networks[C]//Proc of the 15th Conference of the European Chapter of the Association for Computational Linguistics. Stroudsburg, PA: Association for Computational Linguistics, 2017: 42-44.
- [61] Graves A, Wayne G, Reynolds M, *et al.* Hybrid computing using a neural network with dynamic external memory[J]. *Nature*, 2016, 538(7626): 471-476.
- [62] Neelakantan A, Roth B, McCallum A. Compositional vector space models for knowledge base completion[C]//Proc of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. Stroudsburg, PA: Association for Computational Linguistics, 2015: 156-166.
- [63] Wang Zhouxia, Chen Tianshui, Ren J, *et al.* Deep reasoning with knowledge graph for social relationship understanding[C]//Proc of the 27th International Joint Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2018: 1021-1028.
- [64] Kampffmeyer M, Chen Yinbo, Liang Xiaodan, *et al.* Rethinking knowledge graph propagation for zero-shot learning[C]//Proc of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2019.
- [65] Li Yujia, Zemel R, Brockschmidt M, *et al.* Gated graph sequence neural networks[EB/OL]. (2017-09-22). <https://arxiv.org/pdf/1511.05493v4.pdf>.
- [66] Veličković P, Cucurull G, Casanova A, *et al.* Graph attention networks[EB/OL]. (2018-02-04). <https://arxiv.org/pdf/1710.10903.pdf>.
- [67] Shi Miaoyuan. Knowledge graph question and answer system for mechanical intelligent manufacturing based on deep learning[J]. *Mathematical Problems in Engineering*, 2021, 2021: article ID 6627114.
- [68] Wang Xue, Yang Xizhong, Fu Jiapan, *et al.* Research on the construction technology of knowledge graph in aviation[C]//Proc of the 5th International Conference on Mechanical and Aeronautical Engineering. 2020: 12040.
- [69] Shah L A, Etienne A, Siadat A, *et al.* Decision-making in the manufacturing environment using a value-risk graph[J]. *Journal of Intelligent Manufacturing*, 2016, 27(3): 617-630.
- [70] 张栋豪, 刘振宇, 郝维强, 等. 知识图谱在智能制造领域的研究现状及其应用前景综述[J]. *机械工程学报*, 2021, 57(5): 91-113. (Zhang Donghao, Liu Zhengyu, Jia Weiqiang, *et al.* Review on knowledge graph and its application prospects to intelligent manufacturing[J]. *Journal of Mechanical Engineering*, 2021, 57(5): 91-113.)
- [71] 李林. 航空制造设备故障预测与健康管理系统分析[J]. *中国设备工程*, 2020(9): 52-53. (Li Lin. Fault prediction and health management analysis of aviation manufacturing equipment[J]. *China Plant Engineering*, 2020(9): 52-53.)
- [72] 杨浩伟. 制造业领域知识图谱的构建与应用[D]. 杭州: 浙江工商大学, 2019. (Yang Haowei. Construction and application on knowledge graph of manufacturing industry[D]. Hangzhou: Zhejiang Gongshang University, 2019.)
- [73] Zhao Jun, Liu Feifan. Product named entity recognition in Chinese text[J]. *Language Resources and Evaluation*, 2008, 42(5): 197-217.
- [74] Khan F H, Qamar U, Bashir S. Senti-CS: building a lexical resource for sentiment analysis using subjective feature selection and normalized Chi-Square-based feature weight generation[J]. *Expert Systems*, 2016, 33(5): 489-500.
- [75] Kim S, Suh E, Hwang H. Building the knowledge map: an industrial case study[J]. *Journal of Knowledge Management*, 2003, 7(2): 34-45.
- [76] Zhang Chao, Zhou Guanghui, Lu Qi, *et al.* Graph-based knowledge reuse for supporting knowledge-driven decision-making in new product development[J]. *International Journal of Production Research*, 2017, 55(23): 7187-7203.
- [77] Ruiz P A P, Kamsu-Foguem B, Noyes D. Knowledge reuse integrating the collaboration from experts in industrial maintenance management[J]. *Knowledge-Based Systems*, 2013, 50(9): 171-186.
- [78] Sarazin A, Bascans J, Sciau J B, *et al.* Expert system dedicated to condition-based maintenance based on a knowledge graph approach: Application to an aeronautic system[J]. *Expert Systems with Applications*, 2021, 186(12): 115767.
- [79] 聂同攀, 曾继炎, 程玉杰, 等. 面向飞机电源系统故障诊断的知识图谱构建技术及应用[J]. *航空学报*, 2021, doi: 10.7527/S1000-6893.2021.25499. (Nie Tongpan, Zeng Jiyang, Cheng Yujie, *et al.* Knowledge graph construction technology and its application in aircraft power system fault diagnosis[J]. *Acta Aeronautica et Astronautica Sinica*, 2021, doi: 10.7527/S1000-6893.2021.25499.)
- (上接第967页)
- [37] Li Yaguang, Yu Rose, Shahabi C, *et al.* Diffusion convolutional recurrent neural network: data-driven traffic forecasting[EB/OL]. (2018-02-22). <https://arxiv.org/abs/1707.01926>.
- [38] Rusek K, Suárez-Varela J, Almasan P, *et al.* RouteNet: leveraging graph neural networks for network modeling and optimization in SDN[J]. *IEEE Journal on Selected Areas in Communications*, 2020, 38(10): 2260-2270.
- [39] Sutton R S, Barto A G. Reinforcement learning: an introduction[M]. Cambridge, MA: MIT Press, 2018.
- [40] Farahnakian F, Ebrahimi M, Daneshmand M, *et al.* Q-learning based congestion-aware routing algorithm for on-chip network[C]//Proc of the 2nd IEEE International Conference on Networked Embedded Systems for Enterprise Applications. Piscataway, NJ: IEEE Press, 2011: 1-7.
- [41] Mnih V, Kavukcuoglu K, Silver D, *et al.* Playing Atari with deep reinforcement learning[EB/OL]. (2013-12-19). <https://arxiv.org/abs/1312.5602>.
- [42] Hou Yuenan, Liu Lifeng, Wei Qing, *et al.* A novel DDPG method with prioritized experience replay[C]//Proc of IEEE International Conference on Systems, Man, and Cybernetics. Piscataway, NJ: IEEE Press, 2017: 316-321.
- [43] Henderson P, Islam R, Bachman P, *et al.* Deep reinforcement learning that matters[EB/OL]. (2019-01-30). <https://arxiv.org/abs/1709.06560>.
- [44] Sun Penghao, Hu Yuxiang, Lan Julong, *et al.* TIDE: time-relevant deep reinforcement learning for routing optimization[J]. *Future Generation Computer Systems*, 2019, 99: 401-409.
- [45] Xu Zhiyuan, Tang Jian, Meng Jingsong, *et al.* Experience-driven networking: a deep reinforcement learning based approach[C]//Proc of IEEE Conference on Computer Communications. Piscataway, NJ: IEEE Press, 2018: 1871-1879.
- [46] Wikipedia. National science foundation network[EB/OL]. (2018). [https://nl.wikipedia.org/wiki/National\\_Science\\_Foundation\\_Network](https://nl.wikipedia.org/wiki/National_Science_Foundation_Network).
- [47] Guo Yingya, Wang Weipeng, Zhang Han, *et al.* Traffic engineering in hybrid software defined network via reinforcement learning[J]. *Journal of Network and Computer Applications*, 2021, 189: 103116.
- [48] Almasan P, Suárez-Varela J, Badia-Sampera A, *et al.* Deep reinforcement learning meets graph neural networks: exploring a routing optimization use case[EB/OL]. (2020-02-14). <https://arxiv.org/abs/1910.07421>.
- [49] Zhang Junjie, Ye Minghao, Guo Zehua, *et al.* CFR-RL: traffic engineering with reinforcement learning in SDN[J]. *IEEE Journal on Selected Areas in Communications*, 2020, 38(10): 2249-2259.
- [50] Sun Penghao, Guo Zehua, Liu Sen, *et al.* SmartFCT: improving power-efficiency for data center networks with deep reinforcement learning[J]. *Computer Networks*, 2020, 179: 107255.
- [51] Sun Penghao, Guo Zehua, Wang Gang, *et al.* MARVEL: enabling controller load balancing in software-defined networks with multi-agent reinforcement learning[J]. *Computer Networks*, 2020, 177: 107230.
- [52] Hussein A, Gaber M M, Elyan E, *et al.* Imitation learning: a survey of learning methods[J]. *ACM Computing Surveys*, 2017, 50(2): 1-35.
- [53] Davoli L, Veltri L, Ventre P L, *et al.* Traffic engineering with segment routing: SDN-based architectural design and open source implementation[C]//Proc of the 4th European Workshop on Software Defined Networks. 2015: 111-112.
- [54] Tian Ying, Wang Zhiliang, Yin Xia, *et al.* Traffic engineering in partially deployed segment routing over IPv6 network with deep reinforcement learning[J]. *IEEE/ACM Trans on Networking*, 2020, 28(4): 1573-158.