

# 面向工业装配的知识图谱构建与应用研究

徐进

电子科技大学数学科学学院 成都 611731

**摘要** 在新时代智能制造的背景下,传统的工业装配设计方法已经无法满足现代用户追求智能、高效、高精的需求,推进工业设计的智能化成为目前工业领域研究的热点之一。文章通过在现有的工业装配设计方法上,开展面向装配设计图谱的构建,通过装配设计规范构建了装配设计本体模型,从三维图面档案中零件数据获取、零件实体的识别、零件间关系的抽取以及零件知识的融合等方向入手,将获取到的装配数据存入图数据库中构建以汽车发动机领域为例的工业装配知识图谱。实验结果验证了装配设计图谱的可行性。

**关键词**:知识图谱;实体抽取;关系抽取;知识融合;工业图谱

**中图分类号** TP182

## Construction and Application of Knowledge Graph for Industrial Assembly

XU Jin

School of Mathematical Science, University of Electronic Science and Technology of China, Chengdu 611731, China

**Abstract** In the context of smart manufacturing in the new era, traditional industrial assembly design methods have been unable to meet the needs of modern users in pursuit of intelligence, efficiency, and precision. Promoting the intelligent of industrial design as a research hotspot has become a top priority in the industrial field. This paper develops assembly design-oriented atlas through existing industrial assembly design methods, constructs assembly design ontology models through assembly design specifications, acquiring part data from three-dimensional drawing files, part entities identification, and relationships between parts. Starting from the extraction and the fusion of parts knowledge, the acquired assembly data is stored in a graph database to construct an industrial assembly knowledge map taking the automobile engine field as an example. The results of experimental verify the feasibility to use knowledge graph into assembly.

**Keywords** Knowledge graph, Entity extraction, Relationship extraction, Knowledge fusion, Industrial graph

知识作为人类在客观世界中通过实践总结凝练获取到的成果,成为人类不断进步的要素,这也正是智能化的要求。传统的人工智能在没有学习理解知识的前提下,仅是对数据简单叠加与推理,面向单连接的关系极大限制了人工智能的发展,通过将知识与人工智能的融合,可以使得人工智能解决复杂问题的能力得到提升。

现代工业装配大多采用了智能制造的方式,通过3D仿真,设计装配图形,利用数字化双胞胎技术,在工业设计软件中以数字化的方式拷贝一个物理对象,模拟对象在现实环境中的行为,对产品乃至整个工厂进行虚拟仿真,从而提高制造企业产品研发、制造的生产效率<sup>[1]</sup>。此外装配的效率与质量则取决于装配的设计方式,不同零件间的关系成为装配设计的关键,传统的单一连接关系无法全面地描绘三维图形,因此在基于深度学习的辅助决策方面效果欠佳。知识图谱作为用于表示、处理与运用知识的关键技术,则可以较好地处理该类问题,通过装配中零件实体的抽取、零件关系的抽取以及零件知识的融合来构建装配设计的知识图谱,通过面向多连接的深度推理,对装配的设计进行辅助决策,进而提升装配设计的效率与质量。因此,构建工业装配的知识图谱成为必要。

## 1 相关研究

装配是指将零件按规定的技术要求组装起来,并经过调试、检验使之成为合格产品的过程<sup>[2]</sup>,而装配实体是通过计算机辅助技术进行设计实现。

计算机辅助技术包括计算机辅助设计(CAD)、计算机辅助制造(CAM)、计算机辅助教学(CAI)、计算机辅助质量控制(CAQ)、计算机辅助测试(CAT)等<sup>[3]</sup>。诸多计算机辅助技术首先在大型制造业中广泛应用,其后逐步推广到微电子、轻工业等产品制造业以及城建等工程项目中。1990年以来,计算机辅助技术更加强调收集信息以及致力于成为资源平台,在此推动下,涌现出了许多成熟的集各项辅助技术于一身的商业软件,如I-DEAS,UG等。随着技术的不断发展、工业设计经验的丰富以及工业设计模式需求的升级,传统设计软件所持有的技术正在经历着前所未有的发展机遇和挑战,也正在积极地向着基于大数据基础的人工智能发展。

传统的装配设计通常分为以下4个步骤:功能设计、材料选取、工艺设计、结构设计,因此设计一个装配需要设计人员有丰富的机械知识与设计经验,与此同时,还需要掌握设计软件中大量复杂的操作,从而使得产品的设计效率和质量上强

依赖于设计人员。因此传统装配设计的弊端有两个:1)强依赖于涉及人员的设计经验,设计人员设计不同领域设计产品受限;2)设计软件操作步骤复杂,缺乏智能化。

通过构建装配设计的知识图谱能够有效解决以上两个弊端,进而提升设计的效率与质量。

## 2 知识图谱概述

### 2.1 概念

知识图谱最初由 Google 于 2012 年正式提出<sup>[4]</sup>,图谱的初衷是提升搜索的性能与智能化,优化用户的搜索体验。王昊奋等<sup>[5]</sup>提出,知识图谱本质上是一种语义网络(semantic-network),网络中的结点代表实例,边代表实例间的各种语义关系。另外,董登奎等<sup>[6]</sup>提出,知识图谱是指呈现知识发展流程以及属性关系的诸多图形,再辅以可视化技术将这些图形代表的知识实体之间、知识实体与属性之间等的关系呈现出来。

### 2.2 分类

知识图谱根据服务的领域可以分为两类:通用知识图谱和行业知识图谱。通用知识图谱是为横向需求提供服务,注重的是横向覆盖面,拥有各个领域的大量实体,通过领域间的关系形成多维度的知识图谱。行业知识图谱是为纵向需求提供服务,注重纵深覆盖面,由于涉及专业领域的业务场景,实体数量较少,实体间的关系需要在纵向维度得到充分抽取。通用知识图谱被应用于搜索、推荐和问答等业务场景,行业知识图谱被广泛应用于金融、图书情报、医疗等行业。

### 2.3 构建过程

不同类型的知识图谱构建方式有所不同。其根据构建的形式分为两种:自顶向下的构建方式和自底向上的构建方式,前者的构建模型与规则先于知识,后者的构建知识先于模型。图谱的构建主要分为 3 个步骤:知识抽取、知识融合和知识存储与更新。

#### 2.3.1 知识抽取

构建知识图谱的首要问题就是知识从哪里来,如何才能从大量的相关数据中提取出有效有意义的知识,这被称为知识抽取。知识抽取通常是指通过从不同领域收集该领域数据,将结构化和非结构化的数据进行预处理,形成知识、获取属性以及枚举出关系列表。文献<sup>[10]</sup>提出实体抽取的主要方法分为 3 种:1)使用规则进行抽取。预先定义好所需目标实体的规则,使用该规则在原始语料中进行匹配,最终抽取满足需求的实体。2)使用统计机器学习进行抽取。通过训练好的模型对数据进行筛选,从而抽取实体。Liu 等<sup>[11-12]</sup>将线性条件随机场模型和 K 紧邻算法进行结合来识别实体。3)通过开放域进行抽取。Jain 等<sup>[13]</sup>提出通过已知实体的语义特征从开放的网络数据中抽取实体。

关系抽取方式同样可分为 3 种:1)基于监督学习的关系抽取。首先通过对数据进行标注,进而进行模型的训练,最后在数据集中进行关系的抽取。2)基于无监督或半监督的关系

抽取。将一部分标注和未标注的数据作为训练源进行训练,从而在数据集中进行关系的抽取。

3)基于模板的关系抽取。预先定义好抽取关系的模板,根据该规则或模板进行关系的抽取,但该方式难以维护且仅适用于少量的数据。

#### 2.3.2 知识融合

王昊奋等<sup>[5]</sup>指出,知识融合是提供用户基于业务配置融合规则与自动算法相结合的知识融合等能力,自动进行冲突检测并能够根据策略进行解决。

知识融合通常涉及到实体消歧、实体对齐以及知识合并这 3 个重要的部分。

##### 1) 实体消歧

对应于某个实体而言,应该明确指向现实中某个实物,而不是一个实体对应现实世界中多个真实实物。可以使用聚类、实体链接等来进行实体消歧。Han 等<sup>[14]</sup>提出以维基百科作为知识背景,通过维基百科的语义来更加精确地找出实体间的相似性,从而实现更好的消歧效果。

##### 2) 实体对齐

某些实体在真实世界中指代的是相同的实体,这类数据需要进行实体对齐的处理。Mudgal 等<sup>[15]</sup>提出了使用深度学习来进行实体对齐。相对于嵌入表示的实体对齐,Guan 等<sup>[16]</sup>通过监督学习来实现实体对齐,该方法依赖于数据的标注,通常会耗费大量的时间,而无监督的学习过于依赖验证集相似度的计算方式。

##### 3) 知识合并

将不同源的数据进行合并处理,通常分为两个层次:数据层的融合和模式层的融合。对于数据层融合而言,关键性问题是避免实例冲突以及关系冲突,避免冗余的产生。而模式层的融合就是将新获得的本体融入已有的本体库中。

#### 2.3.3 知识存储与更新

大规模的知识图谱的存储以三元组存储为核心,同时还包括其他类型知识的存储。三元组知识的存储主要有资源描述框架(Resource Description Framework, RDF)存储和图数据库两种类型,前者以 RDF 图模型为基础,后者大多数采用属性图数据模型<sup>[5]</sup>。

知识更新通过新数据的生成,不断地生成新的三元组以及更新旧的三元组来实现图谱的动态更新。

## 3 工业装配知识图谱构建

工业装配知识图谱属于行业知识图谱,具有较高的专业性,本文选择工业中汽车发动机装配作为细分领域来构建装配的知识图谱。装配相关数据的来源分为 3 种:结构化的装配本体数据、半结构化的装配异常错误数据、非结构化的装配设计规则数据。因此,汽车发动机装配图谱的构建使用自顶向下和自底向上相融合的方式。其构建架构如图 1 所示。

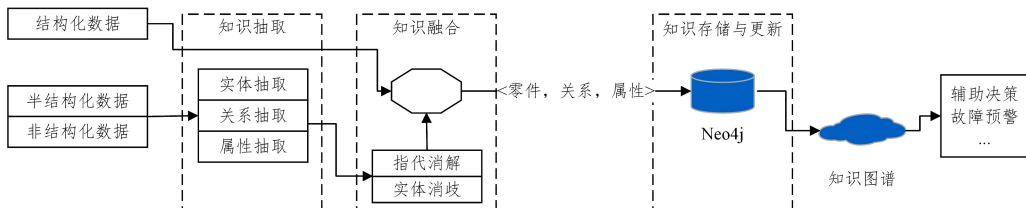


图 1 装配图谱构建架构图

Fig. 1 Assembly graph construction architecture diagram



## 4 工业装配知识图谱的应用

知识图谱在工业装配领域的应用可以在机械设计方面进行效率以及质量上的提升,目前知识图谱在金融、医疗、互联网以及电商领域都得到了广泛的应用,但是在制造业和工业领域还处在初探阶段。本文提出的装配知识图谱主要可应用于如下两个方面。

### 1) 装配设计辅助决策

现代工业装配的设计多数通过计算机辅助技术来进行设计,设计过程依赖于设计人员的专业知识与设计经验,装配知识图谱的构建,可以在装配设计过程中根据知识库中存储的装配知识以及关联关系,给当前装配设计予以辅助推荐和决策,减少设计人员的强依赖性,提高设计效率。

### 2) 装配设计异常警告

装配在设计过程中常常会出现设计细节的问题,通过装配知识图谱,可以根据当前设计零件存储的关联关系知识,预判出当前操作是否有可能发生异常,从而给设计人员以提示,在一定程度上避免了装配设计缺陷,提升了装配的设计质量。

**结束语** 通用领域中,知识图谱已经成为至关重要的角色。专业领域的知识图谱具有行业壁垒高、专业性强等特点。本文将知识图谱在装配设计中进行了初步应用,尝试了知识图谱在专业领域垂直深化的可能性,为工业装配的智能化提供了新的方向。当前知识图谱已经在搜索、智能问答、情报分析、金融等领域发挥了举足轻重的作用,但是对于工业领域的应用尚未成型,还有很大的发挥空间。对于装配本身而言,知识图谱的应用可以改变原有的设计方式,更加友好地服务于非专业人士,也同样可以提升专业人士的设计效率,不再是简单的根据用户行为习惯进行操作的推荐,而是赋能人工智能,增强机器的理解能力,从而更好地服务于人类社会。这种认知智能的产生,将会是机器生产力的一次解放,也必将导致生产关系的改变。未来我们还会针对图谱在工业中的深度应用做进一步的研究。

在未来的几年内,知识图谱在工业垂直领域的应用,必将成为工业智能化的一个重要研究方向与课题,也期待更多的研究者加入到推进工业智能化的队伍中来。

## 参 考 文 献

- [1] 孙焯. 基于GPS的肤面模型构建及公差分析技术研究[D]. 郑州: 郑州大学, 2019.
- [2] 谢佳志. 智能四驱汽车分动器数字化设计系统开发[D]. 合肥: 安徽农业大学, 2012.
- [3] 吴泉源. 计算机应用技术[J]. 计算机工程与科学, 2000(3): 3-7.

- [4] SINGHAL A. Introducing the knowledge graph: things, not strings[J]. Official Google Blog, 2012, 16: 1-10.
- [5] 王昊奋, 丁军, 胡芳槐, 等. 大规模企业级知识图谱实践综述[J/OL]. 计算机工程. [2020-06-07]. <https://doi.org/10.19678/j.issn.1000-3428.0057869>.
- [6] 董登奎, 王清. 基于图数据库的知识图谱管理系统构建分析[J]. 信息系统工程, 2020(4): 47-48.
- [7] 乔骥, 王新迎, 闵睿, 等. 面向电网调度故障处理的知识图谱框架与关键技术初探[J/OL]. 中国电机工程学报. [2020-06-07]. <https://doi.org/10.13334/j.0258-8013.pcsee.200033>.
- [8] 刘峤, 李杨, 段宏, 等. 知识图谱构建技术综述[J]. 计算机研究与发展, 2016, 53(3): 582-600.
- [9] 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.
- [10] ZHAO J, LIU K, ZHOU G Y, et al. Open information extraction[J]. J. Chin. Inform. Process, 2011, 25(6): 98.
- [11] 马忠贵, 倪润宇, 余开航. 知识图谱的最新进展、关键技术和挑战[J/OL]. 工程科学学报. [2020-09-23]. <http://kns.cnki.net/kcms/detail/10.1297.TF.20200918.0856.002.html>.
- [12] LIU X H, ZHANG S D, WEI F R, et al. Recognizing named entities in tweets[C]//Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. Association for Computational Linguistics, Portland, 2011: 359.
- [13] JAIN A, PENNACCHIOTTI M. Open entity extraction from web search query logs[C]//Proceedings of the 23rd International Conference on Computational Linguistics. Beijing, 2010: 510.
- [14] HAN X P, ZHAO J. Named entity disambiguation by leveraging wikipedia semantic knowledge[C]//Proceedings of the 18th ACM Conference on Information and Knowledge Management. Hong Kong, 2009: 215.
- [15] MUDGAL S, LI H, REKATSINAS T, et al. Deep learning for entity matching: a design space exploration[C]//Proceedings of the 2018 International Conference on Management of Data. Houston, 2018: 19.
- [16] GUAN S P, JIN X L, WANG Y Z, et al. Self-learning and embedding based entity alignment[J]. Knowl. Inform. Syst., 2019, 59(2): 361.



**XU Jin**, born in 1994, Ph.D. His main research interests include artificial intelligence and big data.