

基于大模型的武器装备故障知识图谱构建研究

李卓伦^{1,2}, 梁少军^{1,2}, 摆卫兵¹, 刘毅¹, 郑英²

(1. 陆军工程大学 军械士官学校, 武汉 430075;

2. 华中科技大学 人工智能与自动化学院, 武汉 430074)

摘要: 武器装备故障知识图谱不仅可以存储装备故障知识, 而且可结合图上推理技术挖掘新的故障模式并进行故障因果溯源, 有助于实现武器装备健康管理; 提出一种基于大模型的武器装备知识图谱构建方法: 围绕系统结构原理及故障案例数据设计本体框架, 结合离线部署大模型对命名实体与关系进行识别, 实现基于大模型的知识抽取, 在知识融合基础上以 Neo4j 图数据库实现故障知识图谱的存储与可视化; 结合专家思维方式以 Cypher 语言为框架进行了故障推理; 以某型旋翼无人机装备为对象进行了故障知识图谱构建与推理, 验证了所提方法的可行性与有效性。

关键词: 武器装备; 故障知识图谱; 结构原理; 大模型; 知识抽取

Large Model-based Fault Knowledge Graph Construction Method for Weapon Equipment

LI Zhuolun^{1,2}, LIANG Shaojun^{1,2}, BAI Weibing¹, LIU Yi¹, ZHENG Ying²

(1. School of Ordnance Sergeant, Army Engineering University, Wuhan 430075, China;

2. School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract: The knowledge graph of weapon equipment faults can not only store weapon equipment knowledge, but also combine graph-based reasoning techniques to explore new fault modes and analyze faults, which helps to achieve weapon equipment management. This paper proposes a weapon equipment knowledge graph construction method based on a large model. The basic framework is designed by the principles of system structure and fault case data, which is combined with off-line deployment of the large model to identify the named entities and relationships, achieving the knowledge extraction based on the large model, on the basics of knowledge fusion the storage and visualization of fault knowledge graph is achieved on the Neo4j graph database, and the fault inference is carried out using the Cypher language framework in combination with expert thinking modes. Fault knowledge graph is constructed and inferred for a certain type of rotary wing UAV, and the feasibility and effectiveness of the proposed method are verified.

Keywords: weapon equipment; fault knowledge graph; structural principle; large model; knowledge extraction

0 引言

武器装备的快速发展既为新军事变革提供了动力, 同时也在装备维修保障方面提出了新挑战。新型武器装备科技含量高, 系统复杂, 维修难度较大。装备保障人员在数量上和能力上都难以同步匹配。因此, 如何结合先进技术改进装备保障模式, 提高维修保障效率是值得研究的重要问题。

在武器装备故障诊断的研究与实践中, 主要可分为内嵌式和外延式两个方向。内嵌式故障诊断是指嵌入在武器装备中, 帮助操作人员进行维修作业的故障诊断系统。内嵌式故障诊断的研究通常以武器装备研发生产单位为需求主体, 从装备结构原理出发, 设计并集成关键节点传感器, 为维修保障工作提供自动故障诊断^[1-2]。外延式故障诊断是指以装备使用单位为需求主体, 基于现有装备研究维修保障。目前来说, 外延式故障诊断的

收稿日期:2024-06-29; 修回日期:2024-09-05。

基金项目:华中科技大学交叉研究支持计划(2024JCYJ031)。

作者简介:李卓伦(1988-),男,硕士,讲师。

引用格式:李卓伦,梁少军,摆卫兵,等. 基于大模型的武器装备故障知识图谱构建研究[J]. 计算机测量与控制, 2025, 33(9): 176-181, 190.

研究一部分集中在单一型号，扩展性不强^[3-4]；另一部分基于故障树^[5]或专家系统^[6]，需要大量的经验性知识积累，难以应对目前武器装备快速迭代的现状。

知识图谱技术利用自然语言处理技术直接从文本中挖掘知识并构图，具有天然的可解释性^[7]，为解决以上问题提供了新的思路。近年来大模型的快速发展简化了知识图谱的构建，特别是大大提高了文本中知识抽取的效率。而开源大模型陆续支持离线部署满足了依托私有装备数据构建领域故障知识图谱的迫切需求。

针对上述需求，本文提出了一种武器装备故障知识图谱的构建方案：基于专业知识搭建了以结构原理为核心，故障案例为抓手的本体模型；基于本机部署的大模型完成知识抽取并对齐消融，将输出的三元组数据储存在 neo4j 图数据库^[8]中，从而完成故障知识图谱的构建。最后以某型无人机武器装备为对象，采用上述方法构建了故障知识图谱，并结合专家维修思路设计了知识图谱推理方法，验证了故障知识图谱在武器装备故障诊断中的可行性。

本文的主要贡献如下：

- 1) 将大模型应用于故障知识图谱的构建过程，探索了高效构建装备故障知识图谱的新模式。
- 2) 提出了以结构原理为核心的故障知识图谱构建新思路，解决了新型武器装备故障案例不足情况下难以构建故障知识图谱的问题；
- 3) 以所提方法为基础，完成了某型旋翼无人机武器装备故障知识图谱构建与推理实践，为武器装备可解释性故障诊断的提供了实践样例。

1 故障知识图谱相关研究

1.1 故障知识图谱

2012 年，谷歌公司推出了知识图谱（KG，Knowledge Graph），标志着知识图谱正式进入公众视野，并迅速普及。此后，知识图谱成为了大数据和人工智能领域的重要组成部分，广泛应用于推荐系统、问答系统、语义搜索、知识管理和智能分析等多个场景^[9-11]。随着技术的不断进步，知识图谱的构建、管理和应用技术也在持续发展，包括自动化知识抽取、图神经网络在知识图谱上的应用等，进一步推动了知识图谱技术的深化和拓展。知识图谱通常以三元组的形式来表达知识，每个三元组包含实体、关系和值或另一个实体，可以用公式（1）表示：

$$(h,r,t) \tag{1}$$

式中， h 表示主实体， t 表示客实体， r 表示关系。实体可以是人、实物、事件等任何实际对象；关系则是连接两个实体的纽带，描述了主体和客体之间的某种联系^[8]。

故障知识图谱是一种特殊类型的知识图谱，专注于

表达和组织关于各种系统、设备、网络或产品故障的信息^[12]。它通过图的形式将故障现象、故障部位、故障原因、解决措施、相关组件、故障案例、维护策略等多维度的知识点连接起来，形成一个相互关联的知识网络。这种结构化的知识存储方式便于故障诊断、预防性维护、质量控制和智能决策支持。故障知识图谱的实体一般代表故障系统中的各个元素，如设备部件、故障类型、故障原因、解决方案等；关系用于定义实体间的联系，如障“A 由原因 B 引起”，“解决方案 C 适用于故障 D”；属性为实体和关系附加详细信息，如故障频率、解决方法的难易程度等。

在数据存储时，三元组常常存储为数据库中的行，或者在图数据库中表现为节点（实体）之间的带标签的边（关系）。众多三元组储存在图数据库中，以节点和边的形式相连，从而构成知识图谱。

目前来说，故障知识图谱的构建主要基于大量的故障案例，其对文本数据质与量具有较高要求，对推理能力的要求较低。针对知识图谱的构建，目前主流的方法步骤总结为：本体构建、知识抽取、对齐消融、图谱绘制等^[13-14]。

1.2 知识抽取技术

从文本数据中执行知识抽取是知识图谱构建过程中最具挑战性的任务。知识图谱中的知识抽取可分为命名实体识别（NER，Named Entity Recognition）和关系抽取（Relation Extraction，RE）两项任务^[15]。前者主要负责抽取具有特定实体类型的命名实体，而后者则负责抽取命名实体间的预定义关系并生成实体关系三元组。近年来，得益于自然语言处理技术的飞速发展，知识抽取模型的能力越来越强。

近年来，以 BERT^[17]、GPT 为代表的预训练自然语言大模型广泛应用于众多领域，其强大的语言理解和生成能力，在知识抽取任务中发挥了重要作用。自然语言大模型具有良好的上下文理解能力，可提取长时间序列文本中的隐藏规律，从而提高 NER 和 RE 任务的准确性。这些模型由于经过大量文本数据的预训练，其针对领域进行知识抽取前需要的训练文本标注工作可以大量节省，甚至不用标注。但是，由于之前大模型对硬件的要求高，且须使用互联网，存在较大的安全风险，所以不能依托自然语言大模型实现武器装备的知识抽取。

1.3 离线部署的大模型

与 GPT 类似，通义千问大模型（Qwen）是一种大规模预训练模型，旨在解决智能对话、知识图谱推理和其他多模态任务。相比于 GPT，Qwen 在中文文本中的表现更加强大。随着 Qwen 开源并支持本机部署，其对硬件的要求主要集中在 GPU 性能上，对其它硬件的要求并不苛刻。其各版本硬件要求如表 1 所示。

表 1 Qwen 本机部署的硬件要求

版本/B	最大上下文长度/K	预训练 token 数/T	实现微调的最小 GPU/GB
Qwen-1.8	32	2.2	5.8
Qwen-7	32	2.4	11.5
Qwen-14	8	3.0	18.7
Qwen-72	32	3.0	61.4

大模型的一个关键特性是能力泛化，能够适应各种新情景和任务，而不仅仅局限于训练时所遇到的特定任务。这一点使得 Qwen 在各种不同的应用场景中都可以发挥其优势，即强大的知识理解和获取能力。

由于现有大模型的开源及本机部署，使得知识抽取工作变得简单高效。相对于以往的命名实体识别和关系识别的技术，应用大模型从文本中执行知识抽取省略了分句、分词的步骤，大大简化了操作。另一方面，大模型可以很好的兼顾上下文关系，解决了传统知识抽取方法上下文关系难以全面顾及的问题。

2 武器装备故障知识图谱构建方法

故障知识图谱的构建流程采用自上向下与自下而上相结合的方式。首先采用本体建模技术定义知识图谱本体层，其次基于本体层采用不同手段获取多源异构数据，然后采用自然语言处理技术、人工筛选等方法对数据进行知识抽取与融合以构建知识图谱数据层，最终通过图数据库技术实现知识存储和知识图谱可视化。具体流程如图 1 所示。

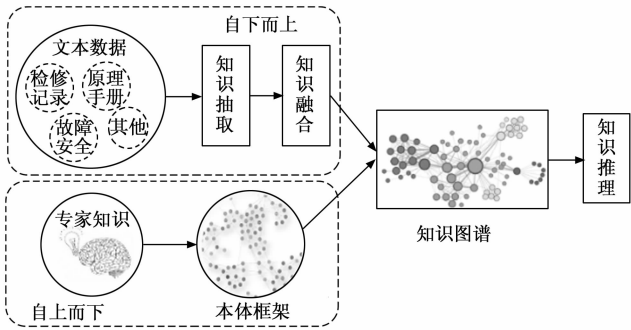


图 1 故障知识图谱构建流程

2.1 本体建模

由于武器装备故障知识专业性强、来源多样，因此在专家经验的指导下，采用改进的 IDEF5 法构建本体。IDEF5 法^[21]的本体构建方法流程分为 5 步包括：组织和范围、数据收集、数据分析、初始化本体和本体的确认。

武器装备的文本数据主要来源于原理手册、相关教学资料、故障案例和检修记录等。其中原理手册和相关教学资料以装备结构原理内容为主，故障案例和检修记录包含大量故障诊断的经验性知识。目前的武器装备由于更新换代周期短，故障案例和检修记录的知识量并不

充足。所以常见的以故障案例等经验性知识为核心的故障知识图谱构建方法在武器装备领域并不适用。从人工维修作业中进行总结，除了依靠经验外，更重要的是深入理解系统结构原理和故障成因，从而分析故障，最后达到故障排除的目的。所以，在本体层面以武器装备系统结构原理为核心，深挖各故障单元间的机械、电气、信号联系，并与武器装备自带的监测数据相结合，构建较全面的结构原理知识图谱。再辅以部分故障案例、维修日志等作为经验，补充到知识图谱中。采用 IDEF5 建立的故障知识图谱的本体结构如图 2 所示。

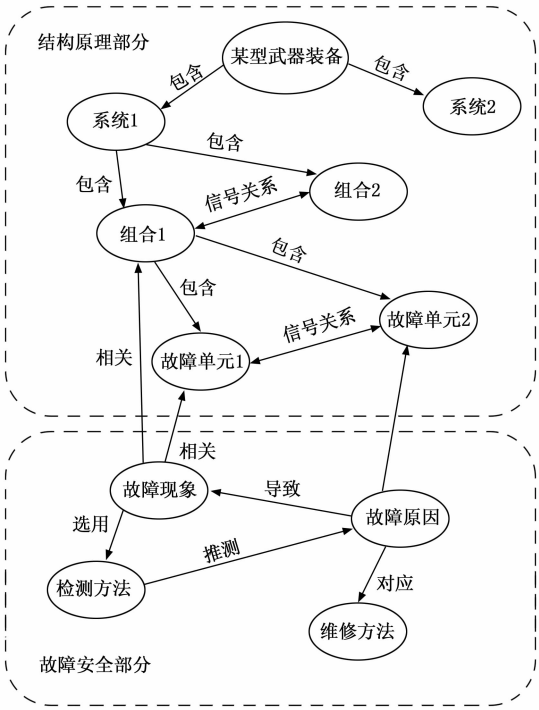


图 2 故障知识图谱本体框架

2.2 武器装备知识抽取

2.2.1 结构原理文本的知识抽取

结构原理文本的知识抽取主要分为如下步骤：1) 根据专家经验对武器装备基本结构建模，一方面可搭建结构原理知识图谱基本框架，另一方面用于文本知识提取的关键词正则匹配；2) 人工阅读文本资料，将与故障相关的文本段落提取；3) 用本机部署的大模型对筛选完毕的文本进行知识抽取，并以三元组形式输出；4) 人工检查三元组的内容，去除明显的错误，并将三元组储存在 CSV 文件中。

2.2.2 故障案例知识抽取

故障案例数据往往有相对固定的格式，如表格。这种故障案例的一般为半结构化数据，其知识图谱的转化比纯文本的非结构化数据要简便。表格中的多数信息是规范的、格式化的，但存在部分需要清洗的半结构化数据，此时直接通过标志字符或停用词检查对长子句进行

结构化抽取与语句合并。在本体的指导下, 对合并结果进行归纳整理, 然后将结果中的实体、关系以及属性与本体相对应。

表 2 所示是无人机装备某故障案例。以表 2 为例, 其故障现象、故障原因、检测方法和维修方法均可直接转存为实体。由于故障诊断的需求, 故障原因实体也要进行文本识别, 与相应的结构原理实体产生关系, 一并加入到知识图谱中, 以方便之后的图谱推理。

表 2 故障案例示例

实体类别	内容
故障现象	起飞前检查,地面站显示“数传故障”。
故障原因	机上数传接插件松动。
检测方法	执行功能自检中的“数传系统自检测”,发现机体数传设备报故障;再对机体数传设备及其电缆进行检查。
维修方法	将数传接插件紧固后再执行起飞前检查,如仍出现故障,须更换数传接插件。注意禁止带电插拔。

2.3 融合消歧

由于武器装备故障知识来自于不同数据源, 必须进行知识对齐消融。装备故障诊断知识融合是将不同数据源中不同标识实体的语义理解链接到同一实体上, 实现对同名、多名和缩写等多种实体语义的消歧和共指消解。从方法上, 一方面需要基于业务知识建立同名库, 另一方面可用自然语言处理技术将文本向量化后进行相似度比较。本文将以文本向量相似度对比为主, 建立同名库为辅的方式来进行融合消歧。

在武器装备领域, 大部分多名共指现象是由于专有名词在不同语境下的简化, 例如“电动机”与“电机”、“遥控发射机”与“发射机”等。此时, 采用如 (2) 所示的 cosine 相似度算法公式^[22]来进行文本相似度比较, 其中 A_i 和 B_i 分别是两个中文实体的向量化表示。可以发现, cosine 相似度算法与向量的模无关, 只比较向量之间方向的相似度, 这与专有名词简化的思维一致。设定一个阈值进行判定, 大于阈值, 则合并实体, 小于阈值则实体独立:

$$\cos\theta = \frac{\sum^n i = 1(A_i \times B_i)}{\sqrt{\sum^n i = 1A_i^2} \times \sqrt{\sum^n i = 1B_i^2}} \quad (2)$$

此外, 还有完全不同文本的多名现象。例如, 旋翼无人机领域名词中, “旋翼”与“桨叶”通常同指一个设备, “地面站”与“计算机”也经常是同指。这种特殊情况只能基于业务知识建立同名库。

2.4 故障知识图谱检索

Neo4j 图数据库内置的 Cypher 语言具备强大的检索功能, 支持在查询条件中设置多层次和串联的约束条件, 这为本文探讨的故障知识图谱内的信息检索提供了灵活性。在利用 Cypher 进行检索时, 检索策略应模拟维修专

家采用的诊断推理过程。本文从知识图谱的结构出发, 归纳了三大故障诊断指导原则。这些原则初步指导了 Cypher 查询语句的设计用于故障知识图谱的查询, 确保查询逻辑既贴近实际, 又充分利用知识图谱的结构优势。

1) 故障现象到潜在原因的映射原则:

模拟维修专家从直观故障现象逆推可能故障源的逻辑, 设计查询以快速关联故障现象与潜在部件故障或系统异常;

2) 因果关联分析原则:

借鉴实际维修中对因果链的追溯, 构建查询语句以探索故障间的因果关系, 识别故障的根本原因;

3) 结构导向排查原则:

利用知识图谱中部件间、系统层级的关联信息, 设计查询路径以遵循系统结构逻辑, 逐级定位故障点。

规则 1: 从症状节点出发, 识别所有相关节点及关系。

在 Cypher 中, 这一概念转化为寻找与已知节点相连的所有关系和节点。这可通过使用 MATCH 语句来实现, 具体如下所示:

```
-
MATCH (a:Symptom) --> (b)
RETURN a, b;
-
```

这样的查询策略直接体现了从故障现象出发, 逐步展开关联分析的核心思想, 与规则 1 相符。

规则 2: 将症状节点与相似组件节点及其关系进行对比分析。

在 Cypher 中, 执行这一策略意味着识别具有相似属性或标签的组件节点, 并进一步探索这些节点及其关联关系与当前症状节点之间的异同。具体实现可通过以下 Cypher 脚本片段达成:

```
-
MATCH (symptom:Symptom {name: '* * '}),
(component:Component) - [:BELONGS_TO*] -> (:Subsystem),
(similarComponent:Component)
WHERE component.label = similarComponent.label
AND component <> similarComponent
AND similarComponent - [:INVOLVES | CAUSES] -> (:Symptom)
WITH symptom, collect(similarComponent) AS similarComponents
UNWIND similarComponents AS sc
MATCH (sc) -[rel]-> (:Symptom)
RETURN symptom, sc, rel;
-
```

这段脚本首先定位到特定的 Symptom 节点, 随后查找与其相关的所有组件 Component 所属的子系统路径。通过标签匹配发现具有相同标签但不同的 similar-

3.3 故障知识图谱检索实验

为了验证基于结构原理的故障知识图谱的有效性和实用性, 选择了 3 个真实的故障案例, 并设计并实施了一组实验。通过对比分析, 我们在真实诊断推理的框架下, 展示了利用故障知识图谱的诊断过程, 从而验证其实用性和准确性。

1) 故障现象: 地面站无人机状态监控信息未显示。
依据故障表征, 此故障应源于无人机下行数据链的异常。该故障案例与规则 1 的基本理念相吻合, 体现了故障根源在于数据传输环节的中断。通过运用 Cypher 语言执行的关系查询操作, 其结果如图 4 (a) 所示。经对比, 查询结果与人工分析结论高度一致, 验证了 Cypher 查询在识别此类故障根源中的有效性和准确性。

2) 故障现象: 螺旋桨 1 不旋转。
在维修实践中, 除故障描述外的其他内容一般默认为正常状态。据此分析, 故障原因可初步锁定为电机 1 或螺旋桨 1 的问题。此故障实例与规则 2 的基本思路相契合, 即从直接相关的组件入手进行故障溯源。借助知识图谱的多跳查询功能, 查询结果如图 4 (b) 所示。通过在知识图谱中实施多层关联的查询策略, 我们不仅能够定位到与“螺旋桨 1 不旋转”这一症状直接相关的组件 (电机 1 和螺旋桨 1), 还能进一步探索这些组件之间的因果关系链及潜在的辅助诊断信息, 这一过程与维修专家依据经验进行的逐步排查逻辑相呼应。图 4 (b) 所展示的查询结果, 实质上是对规则 3 指导下的故障推理过程的一种可视化确认。

3) 故障现象: 遥控指令接收失败。
与遥控指令相关的设备组件众多, 故障的可能性亦十分复杂, 这一故障实例恰与规则 3 的核心思想相吻合, 强调了从受影响的子系统出发, 全面审视系统内部关联的重要性。通过在知识图谱中实施基于节点的查询策略, 涵盖了与遥控信号处理、通信模块直至指令解析等多环节相关的所有潜在故障节点。其结果显示于图 4 (c), 可见知识图谱中的节点关系结构大体上与人工分析的逻辑相匹配, 进一步证实了该方法的有效性。该结果不仅揭示了这些组件间的复杂交互网, 还清晰地映射出人工分析时所考虑的故障排查路径, 展现了知识图谱在处理多因素交织故障时的高度适应性和推理能力。

针对这 3 个各具特色的故障案例所开展的实验, 证明了所构建的故障知识图谱具有依据系统结构原理进行推理的内在能力。

4 结束语

武器装备故障知识图谱可将专家知识与客观知识整合, 用可解释方式解决故障检测、故障隔离和故障归因等问题, 为智能故障诊断打下了基础。本文针对新型武器装备故障案例较少的问题, 提出了以结构原理为核心

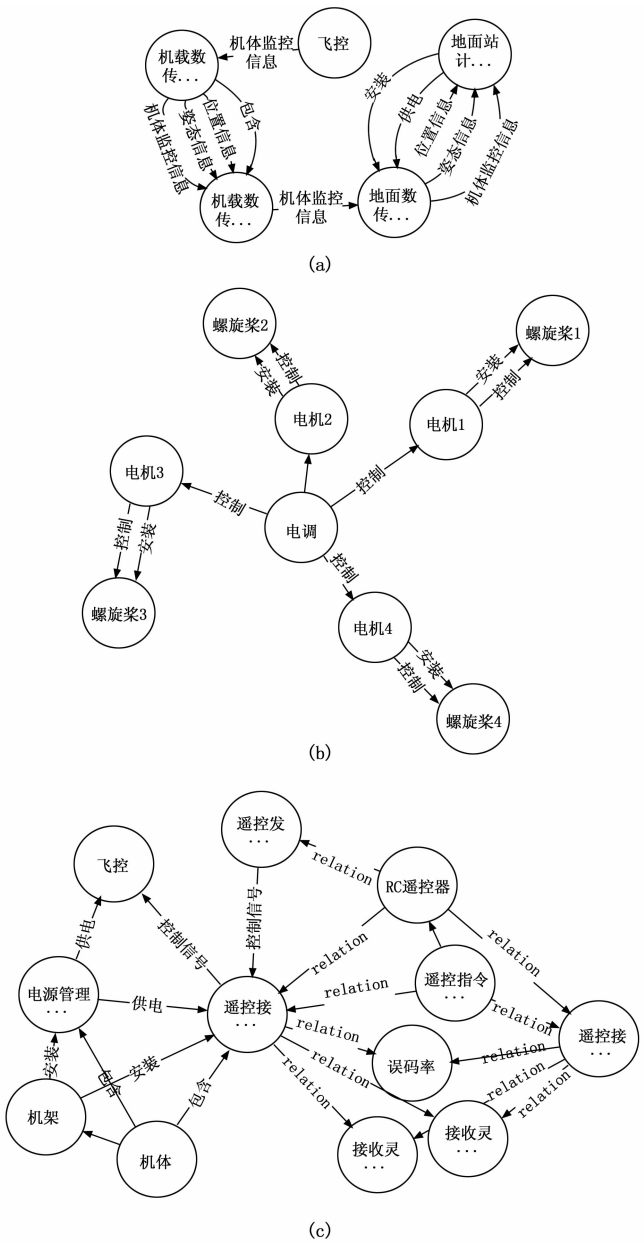


图 4 故障知识图谱检索结果

的故障知识图谱构建方法。该方法借鉴专家排查的思路, 以结构原理知识为主体, 以少量故障案例为辅助, 利用大模型构建了装备故障知识图谱, 并结合特定推理语言实现了抽象排查思路的具象化, 拓宽了图谱推理的方式方法。并在某型旋翼无人机装备上进行了应用验证。
值得注意的是, 本文的重点聚焦于故障知识图谱的构建, 对于故障推理的探讨并未深入, 仅将其作为一种验证知识图谱构建合理性的手段。下一步将更全面地解析维修人员的诊断推理过程, 并将其融入构建更加复杂精细的故障知识图谱检索规则中。开拓设计出更加精密且实用性强的故障推理架构, 进一步提升自动化故障诊断的智慧化水平和实际应用效能。

(下转第 190 页)