

北京航空航天大学学报
Journal of Beijing University of Aeronautics and Astronautics
ISSN 1001-5965,CN 11-2625/V

《北京航空航天大学学报》网络首发论文

题目：基于 GRPO 优化的大模型飞机故障文本结构化信息抽取方法
作者：谢小月，唐希浪，崔利杰
DOI：10.13700/j.bh.1001-5965.2026.0531
收稿日期：2026-07-24
网络首发日期：2026-11-13
引用格式：谢小月，唐希浪，崔利杰．基于 GRPO 优化的大模型飞机故障文本结构化信息抽取方法[J/OL]．北京航空航天大学学报．
https://doi.org/10.13700/j.bh.1001-5965.2026.0531



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于 GRPO 优化的大模型飞机故障文本结构化信息抽取方法

谢小月，唐希浪✉，崔利杰

（空军工程大学装备管理与无人机工程学院，西安 710051）
*通信作者　E-mail：tangxilang@sina.com

摘要　针对飞机故障文本中结构化信息抽取准确性和输出规范性不足的问题，提出一种基于组相对策略优化（GRPO）的大模型结构化信息抽取（SIE）方法 GRPO-SIE。结合飞机故障知识图谱模式设计结构化任务链提示，引导模型完成实体识别、属性抽取、关系判定和结果组织；利用领域标注数据对基础模型进行监督微调，使其学习飞机故障知识类型及结构化输出规则；构建由实体奖励、属性奖励、关系奖励和格式奖励组成的复合奖励函数，并采用 GRPO 对模型生成策略进行优化，使模型生成结果更符合抽取任务和结构化输出要求。在领域标注的飞机故障文本数据集上开展实验，结果表明：GRPO-SIE 在实体、属性和关系抽取任务上的平均 F_1 值分别达到 0.844、0.782 和 0.794，三项指标均优于各对比方法，并在不同知识密度的文本中保持了较好的抽取性能；消融实验进一步验证了方法各组成部分的有效性。所提方法可为故障知识图谱构建及故障知识检索、关联分析和后续图谱推理提供结构化数据支撑。

关键词　飞机故障文本；结构化信息抽取；提示学习；大语言模型；GRPO
中图分类号　V267；TP391.1；TP181　　文献标志码　A　　DOI: 10.13700/j.bh.1001-5965.2026.0531

A GRPO-optimized large language model method for structured information extraction from aircraft fault texts

XIE Xiaoyue, TANG Xilang✉, CUI Lijie

(Equipment Management and UAV Engineering College, Air Force Engineering University, Xi'an 710051,China)
*E-mail: tangxilang@sina.com

Abstract　To address the insufficient accuracy and output standardization in structured information extraction from aircraft fault texts, a large language model-based structured information extraction method based on group relative policy optimization (GRPO), termed GRPO-SIE, is proposed. First, a structured task-chain prompt is designed in conjunction with the aircraft fault knowledge graph schema to guide the model in entity recognition, attribute extraction, relation determination, and result organization. Second, the base model is supervised fine-tuned using domain-annotated data to learn aircraft fault knowledge types and structured output rules. Finally, a composite reward function consisting of entity, attribute, relation, and format rewards is constructed, and GRPO is employed to optimize the model generation policy, so that the generated results better conform to extraction-task and structured-output requirements. Experiments conducted on a domain-annotated aircraft fault text dataset show that GRPO-SIE achieves mean F_1 scores of 0.844, 0.782, and 0.794 for entity, attribute, and relation extraction, respectively. All three metrics outperform those of the baseline methods, and the method maintains good extraction performance on texts with different knowledge densities. Ablation experiments further verify the effectiveness of the method components. The proposed method provides structured data support for fault knowledge graph construction, fault knowledge retrieval, association analysis, and subsequent graph reasoning.
Key words　aircraft fault texts; structured information extraction; prompt learning; large language models; GRPO

收稿日期：2026-07-24
基金项目：国家自然科学基金（72201276）；中国博士后科学基金（2025M774456）
Fund：　National Natural Science Foundation of China (72201276); China Postdoctoral Science Foundation(2025M774456)
网络首发时间：2026-11-13 13:28:28　网络首发地址：https://link.cnki.net/urlid/11.2625.V.20261113.1146.002

飞机运行与维修过程中会产生大量的故障记录、检查报告、排故手册和案例资料等文本数据，其中包含组成单元、状态参数、故障现象、故障原因、检查方法及维修措施等重要信息。这些信息大多以非结构化的形式存在，难以直接用于故障知识关联、推理分析和维修决策。因此，为了充分利用飞机故障文本中蕴含的诊断知识，有必要从文本中识别和抽取关键知识要素，并将其组织为统一的结构化表示，从而提高故障知识的检索与利用效率[1]。

围绕飞机故障文本结构化抽取，现有研究主要采用规则匹配、机器学习和深度学习方法，对故障、组成单元、信号参数等实体及其语义关系进行识别[2-5]。例如，吴闯等[2]面向航空发动机润滑系统故障诊断，结合专家知识构建故障知识图谱本体，并采用双向长短期记忆网络-条件随机场（bidirectional long short-term memory–conditional random field，BiLSTM-CRF）模型实现故障知识抽取；Meng 等[3]面向飞机故障预测与健康管理（prognostics and health management，PHM）场景构建故障知识图谱平台，采用知识增强语义表示（enhanced representation through knowledge integration，ERNIE）-BiLSTM-CRF 模型进行实体识别，并结合 TreeBiLSTM 模型开展关系抽取；Xie 等[5]提出水平-垂直-级联关系联合抽取模型，通过融合横向与纵向上下文信息增强头实体表示，以提高故障文本中重叠三元组的抽取效果。这些方法在特定故障类型和预定义模式下取得了一定效果，但往往依赖领域标注数据，而且需要针对具体抽取任务设计专用的模型，在面对不同类型、不同表达形式的故障文本时泛化能力不足。

随着大模型技术的发展，基于提示学习、上下文学习和指令微调的生成式抽取方法逐渐受到关注。相关研究通过任务分解、示例检索、结构化提示和领域微调等方式，实现了多类结构化信息的统一抽取与生成，并按照 JavaScript 对象表示法（JavaScript object notation，JSON）等预定义格式输出结构化结果[6-13]。例如，Lu 等[6]提出统一结构生成框架，将实体、关系和事件等信息抽取任务统一为文本到结构的生成过程；Dagdelen 等[10]利用微调大模型从科学文本中联合抽取实体及其关系，并将结果组织为 JSON 结构；Yan 等[13]面向无人机系统训练知识图谱构建，通过设计领域提示指令，引导大语言模型抽取文本中的实体、属性及实体间关系，并按照预定义的 JSON 格式输出结构化结果。然而，飞机故障文本中的知识分布较为密集，单个文本片段中往往同时包含较多的实体、属性及实体间关系，直接采用提示或监督微调方法进行结构化信息抽取时，容易出现属性对应错误、关系方向混淆和输出格式不规范等问题。

为提高大模型结构化信息抽取结果的准确性和规范性，研究者们开始将强化学习引入信息抽取任务，利用任务相关奖励对模型生成结果进行优化[14-18]。其中，近端策略优化（proximal policy optimization，PPO）[19]和组相对策略优化（group relative policy optimization，GRPO）[20]为大模型强化学习提供了相应的策略优化框架。例如，Xu 等[14]在统一信息抽取框架中引入强化学习，以改善易混淆任务和有限样本条件下的抽取效果；Dai 等[17]将关系抽取转化为标注规范引导的推理过程，并采用可验证奖励强化学习提升模型的跨领域关系识别能力；Li 等[18]针对科学信息抽取设计相关性奖励和规则奖励，通过 GRPO 优化模型的关系抽取性能。这些基于任务奖励的策略优化能够为大模型信息抽取提供直接的反馈信号。然而，现有强化学习信息抽取方法大多以关系三元组正确性、整体生成质量或规则符合性作为主要奖励，对实体遗漏、属性与实体之间的归属以及结构化格式等多类错误的区分反馈不足，将其直接应用于飞机故障文本的结构化信息抽取时，难以满足其细粒度优化需求。

对此，本文面向飞机故障文本，提出一种基于 GRPO 的大模型结构化信息抽取方法，记为 GRPO-SIE。该方法首先结合飞机故障知识图谱模式设计结构化任务链提示，引导模型依次完成实体识别、属性抽取、关系判定和结果组织；然后利用领域标注数据对基础模型进行监督微调，使其学习飞机故障知识类型及 JSON 结构化输出规则；最后构建由实体奖励、属性奖励、关系奖励和格式奖励组成的复合奖励函数，并采用 GRPO 对模型进行策略优化，以提高飞机故障文本结构化信息抽取的准确性和输出规范性。

1 抽取方法

GRPO-SIE 的整体框架如图 1 所示，主要由结构化任务链提示、监督微调、复合奖励函数和 GRPO 策略优化四个部分组成。

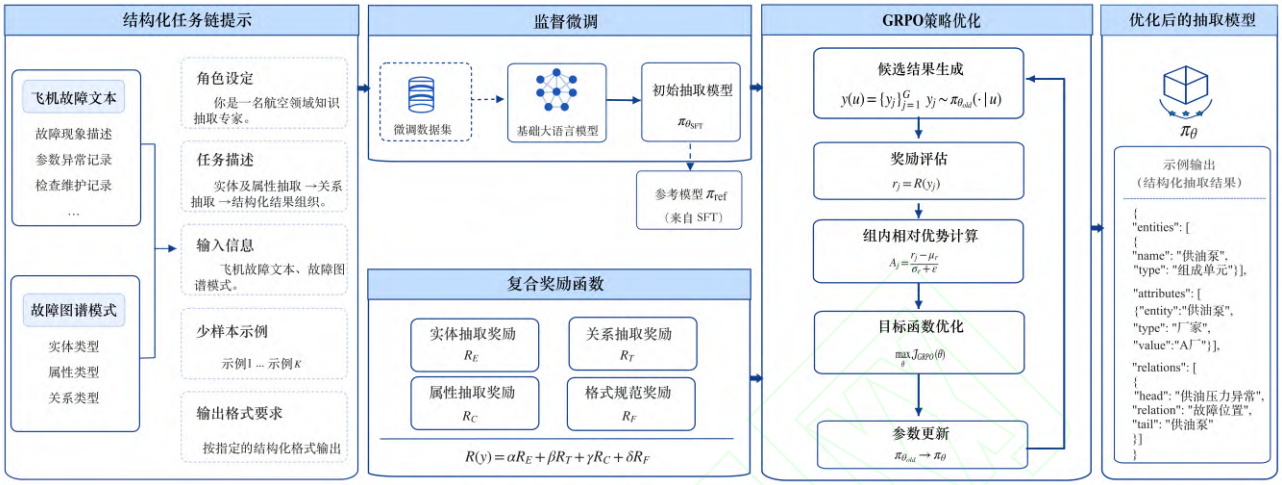

图 1 基于 GRPO 优化的飞机故障文本结构化信息抽取方法框架

Fig. 1 Framework of the GRPO-optimized structured information extraction method for aircraft fault texts

1.1 结构化任务链提示

结构化任务链（structured task chain，STC）提示用于在图谱模式约束下规范模型的抽取过程和输出形式。本文借鉴思维链提示中的任务分解思想[21]，将飞机故障文本结构化抽取任务按照“先识别知识要素、再判断语义关系、最后形成结构化输出”的思路进行设计，引导模型逐步完成实体、属性和关系信息的抽取。设计的结构化任务链提示如图 2 所示。



图 2 结构化任务链提示
Fig. 2 Structured task-chain prompt

从图 2 可以看出，结构化任务链提示包括角色设定、任务描述、输入信息、少样本示例和输出格式要求五个部分。其中，角色设定用于限定模型在飞机故障诊断和故障知识抽取语境下理解文本内容；任务描述用于明确实体抽取、属性抽取、关系抽取和结构化结果组织的子任务，使模型按照预设逻辑顺序完成知识要素识别及语义关系判断；输入信息包括待抽取故障文本以及图谱模式中给定的实体类型、属性类型和关系类型集合，用于限定抽取范围并减少模式外结果生成；少样本示例采用“示例输入—示例输出”的形式，为模型提供实体边界、属性归属、关系方向和结构化表达的参考；输出格式要求则约束模型按照统一 JSON 结构返回抽取结果，从而保证实体、属性和关系信息能够以规范的形式输出。

在监督微调阶段，结构化任务链提示、图谱模式与故障文本共同构成输入，人工标注的结构化结果作为输出目标，用于引导模型学习飞机故障文本到标准抽取结果的映射关系；在 GRPO 策略优化阶段，策略模型基于相同输入形式生成候选抽取结果，并由复合奖励函数对候选结果进行评价。通过保持输入形式一致，可减少不同训练阶段之间的任务偏差，提高结构化输出的一致性。

1.2 监督微调

虽然结构化任务链提示能够明确抽取步骤和输出格式，但仅依靠提示难以使模型充分掌握故障文本与标准结构化结果之间的映射关系。因此，在 GRPO 策略优化前，首先利用标注数据对基础大语言模型进行监督微调（supervised fine-tuning，SFT），以获得一个具备初步结构化信息抽取能力的初始模型。

具体地，该过程将结构化任务链提示、图谱模式与故障文本共同作为输入，将人工标注的实体、属性及关系结果作为目标输出。设第 i 条训练样本的输入为 $\boldsymbol{u}_i$，标准输出为 $\boldsymbol{y}_i^*$，则监督微调数据集表示为

$$\boldsymbol{D}_{SFT}=(\boldsymbol{u}_i,\boldsymbol{y}_i^*)_{i=1}^{N} \tag{1}$$

式中：N 为监督微调样本数量。

模型训练时，以输入 $\boldsymbol{u}_i$ 为条件自回归地生成标准输出序列 $\boldsymbol{y}_i^*$，并通过最大化条件生成概率来更新模型参数。其损失函数为

$$L_{SFT}=-\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{|\boldsymbol{y}_i^*|}\log\pi_\theta\left(y_{i,t}^*\mid \boldsymbol{u}_i,\boldsymbol{y}_{i,<t}^*\right) \tag{2}$$

式中：π_θ 表示监督微调阶段的模型；θ 为待优化参数；$y_{i,t}^*$ 表示标准输出序列 $\boldsymbol{y}_i^*$ 中的第 t 个 token；$\boldsymbol{y}_{i,<t}^*$ 表示在第 t 个 token 之前已生成的历史 token 序列。通过最小化损失函数 $\mathcal{L}_{SFT}$，促使模型学习在给定输入条件下生成标准的结构化抽取结果。

1.3 复合奖励函数

复合奖励函数用于评价策略模型生成的候选抽取结果，并为 GRPO 策略优化提供奖励信号。针对飞机故障文本结构化信息抽取任务，本文分别从实体抽取、属性抽取、关系抽取和输出格式四个方面设计奖励函数，以同时约束抽取结果的语义正确性和格式规范性。

具体地，实体抽取奖励用于评价候选结果中的实体抽取质量，定义为

$$R_E(\boldsymbol{y}_j)=\begin{cases}1, & \text{当}\hat{E}_j=\varnothing,\ E^*=\varnothing\text{时},\\ \dfrac{2\,|\hat{E}_j\cap E^*|}{|\hat{E}_j|+|E^*|}, & \text{其他}.\end{cases} \tag{3}$$

式中：$\boldsymbol{y}_j$ 表示针对同一输入模型生成的第 j 个候选结果；$\hat{E}_j$ 和 E^* 分别表示候选结果 $\boldsymbol{y}_j$ 中的预测实体集合和对应的标准实体集合；$|\hat{E}_j|$ 和 $|E^*|$ 分别表示预测实体集合和标准实体集合中的实体数量；$|\hat{E}_j\cap E^*|$ 表示两个集合中实体名称和实体类型均一致的实体数量。

关系抽取奖励用于评价候选结果中有向关系三元组的抽取质量，定义为

$$R_T(\boldsymbol{y}_j)=\begin{cases}1, & \text{当}\hat{T}_j=\varnothing,\ T^*=\varnothing\text{时},\\ \dfrac{2\,|\hat{T}_j\cap T^*|}{|\hat{T}_j|+|T^*|}, & \text{其他}.\end{cases} \tag{4}$$

式中：$\hat{T}_j$ 和 T^* 分别表示候选结果 $\boldsymbol{y}_j$ 中的预测关系集合和对应的标准关系集合；$|\hat{T}_j|$ 和 $|T^*|$ 分别表示预测关系集合和标准关系集合中的关系数量；$|\hat{T}_j\cap T^*|$ 表示两个集合中头实体、关系类型、尾实体和关系方向均一致的关系数量。

属性抽取奖励用于评价候选结果中的属性抽取质量，定义为

$$R_C(\boldsymbol{y}_j)=\begin{cases}1, & \text{当}\hat{C}_j=\varnothing,C^*=\varnothing\text{时},\\ \dfrac{2\,|\hat{C}_j\cap C^*|}{|\hat{C}_j|+|C^*|}, & \text{其他}.\end{cases} \tag{5}$$

式中：$\hat{C}_j$ 和 C^* 分别表示候选结果 $\boldsymbol{y}_j$ 中的预测属性集合和对应的标准属性集合；$|\hat{C}_j|$ 和 $|C^*|$ 分别表示预测属性集合和标准属性集合中的属性数量；$|\hat{C}_j\cap C^*|$ 表示两个集合中所属实体、属性类型和属性值均一致的属性数量。

格式规范奖励主要从输出 JSON 格式的可解析性、字段完整性和结构一致性方面评价候选结果的输出格式规范化，其奖励函数定义为

$$R_F(\boldsymbol{y}_j)=\frac{V_{json}(\boldsymbol{y}_j)+V_{field}(\boldsymbol{y}_j)+V_{struct}(\boldsymbol{y}_j)}{3} \tag{6}$$

式中：$V_{\text{json}}(\boldsymbol{y}_j)$ 表示候选结果能否被解析为有效 JSON；$V_{\text{field}}(\boldsymbol{y}_j)$ 表示在 JSON 能够被正确解析的情况下实体、属性和关系字段是否完整；$V_{\text{struct}}(\boldsymbol{y}_j)$ 表示在字段完整的情况下各字段的数据类型、层级结

构及内部子字段是否符合预设模式。三个变量均为二值变量，满足相应条件时取 1，否则取 0。

需要注意的是，这里格式校验结果同时也是实体、关系和属性奖励的计算前提。当候选结果无法解析为有效 JSON 时，实体、关系和属性奖励均置为 0；当某类字段缺失或结构不符合预设模式时，将该类奖励置为 0。若相应字段存在并以空列表表示该类抽取结果为空，则不视为字段缺失，并按照前述分段函数计算相应奖励。

综合上述四类奖励，候选结果 y_j 的复合奖励定义为

$$R(\boldsymbol{y}_j)=\alpha R_E(\boldsymbol{y}_j)+\beta R_C(\boldsymbol{y}_j)+\gamma R_T(\boldsymbol{y}_j)+\delta R_F(\boldsymbol{y}_j) \tag{7}$$

式中：α、β、γ、δ 分别表示实体抽取奖励、属性抽取奖励、关系抽取奖励和格式规范奖励的权重。这里，各权重的取值主要考虑不同抽取任务的重要性和难度，一般遵循($\gamma>\alpha>\beta>\delta$)的设置原则，且满足：

$$\alpha+\beta+\gamma+\delta=1,\quad \alpha,\beta,\gamma,\delta\geq 0. \tag{8}$$

1.4 GRPO 策略优化

在监督微调得到初始结构化信息抽取模型后，接下来采用 GRPO[20]进一步优化模型生成策略。与仅依赖标准答案拟合的监督微调不同，GRPO 通过对同一输入下的多个候选抽取结果进行组内比较，然后根据候选结果的相对优劣更新模型参数，从而增强模型对实体、属性、关系和格式约束的响应能力。

具体地，首先进行策略模型与参考模型初始化。令监督微调后得到的模型为 $\pi_{\boldsymbol{\theta}_{SFT}}$，在 GRPO 训练阶段，将其作为待优化策略模型 $\pi_{\boldsymbol{\theta}}$ 的初始状态，并将其冻结为参考模型 π_{ref}，即

$$\pi_{\boldsymbol{\theta}}\leftarrow\pi_{\boldsymbol{\theta}_{SFT}};\ \pi_{ref}\leftarrow\pi_{\boldsymbol{\theta}_{SFT}} \tag{9}$$

式中：π_{ref} 在训练过程中保持不变，用于约束策略模型的更新幅度，避免模型分布过度偏离监督微调后的初始分布。

其次，对同一输入进行候选结果组采样。对于由结构化任务链提示、图谱模式和故障文本组成的输入 $\boldsymbol{u}$，由当前迭代更新前的旧策略模型 $\pi_{\boldsymbol{\theta}_{\text{old}}}$ 生成 G 个候选抽取结果，构成候选结果组：

$$y(\boldsymbol{u})=\{\boldsymbol{y}_j\}_{j=1}^{G},\ \boldsymbol{y}_j\sim\pi_{\boldsymbol{\theta}_{\text{old}}}(\cdot\mid\boldsymbol{u}) \tag{10}$$

式中：$\boldsymbol{y}_j$ 表示第 j 个候选抽取结果。根据 1.3 节定义的复合奖励函数，计算每个候选结果的奖励值，即

$$r_j=R(\boldsymbol{y}_j) \tag{11}$$

在此基础上，基于同组候选结果的奖励分布计算相对优势。设该组候选结果的奖励均值和标准差分别为

$$\mu_r=\frac{1}{G}\sum_{j=1}^{G}r_j,\ \sigma_r=\sqrt{\frac{1}{G}\sum_{j=1}^{G}(r_j-\mu_r)^2} \tag{12}$$

则候选结果 y_j 的组内相对优势定义为

$$A_j=\frac{r_j-\mu_r}{\sigma_r+\varepsilon} \tag{13}$$

式中：ε 为防止分母为 0 的常数。该值反映候选结果在同组样本中的相对质量，用于指导策略模型增加高奖励候选结果的生成概率。

最后，依据组内相对优势更新策略模型。策略更新时，GRPO 通过当前策略模型与旧策略模型之间的概率比值控制更新幅度，并通过相对熵（Kullback-Leibler divergence，KL）散度约束当前策略模型与参考模型之间的分布差异。其优化目标为

$$\max_{\boldsymbol{\theta}} J_{GRPO}(\boldsymbol{\theta}) = \boldsymbol{E}_{\boldsymbol{u},y(\boldsymbol{u})}[\frac{1}{G}\sum_{j=1}^{G}(\min[\rho_j(\boldsymbol{\theta})A_j, \text{clip}(\rho_j(\boldsymbol{\theta}),1-\eta,1+\eta)A_j]) - \lambda D_{KL}(\pi_{\boldsymbol{\theta}} \| \pi_{ref})] \quad (14)$$

式中：

$$\rho_j(\theta) = \frac{\pi_{\boldsymbol{\theta}}(\boldsymbol{y}_j \mid \boldsymbol{u})}{\pi_{\boldsymbol{\theta}_{\text{old}}}(\boldsymbol{y}_j \mid \boldsymbol{u})}$$

为当前策略模型与旧策略模型在候选结果 $\boldsymbol{y}_j$ 上的概率比，η为策略裁剪系数，λ 为 KL 散度约束系数，$\text{clip}(\rho_j(\boldsymbol{\theta}),1-\eta,1+\eta)$ 为截断函数，表示将概率比 $\rho_j(\boldsymbol{\theta})$ 限制在 $[1-\eta,1+\eta]$ 区间内。通过上述目标函数，模型能够在提升高奖励候选结果生成概率的同时，抑制策略分布的过度偏移，从而提升飞机故障文本结构化信息抽取结果的准确性和输出稳定性。

1.5 模型训练流程

在上述方法设计基础上，GRPO-SIE 的训练流程如算法 1 所示。

算法 1 GRPO-SIE 训练流程。

输入：飞机故障文本训练集 $\boldsymbol{D}$，结构化任务链提示 $\boldsymbol{P}$，基础大模型 $\pi_{\boldsymbol{\theta}_0}$，候选结果数量 G，奖励权重(α、β、γ、δ)。

输出：优化后的结构化信息抽取模型 $\pi_{\boldsymbol{\theta}}$。

1.根据结构化任务链提示构造监督微调样本，将结构化任务链提示、图谱模式和故障文本作为输入，将标注的实体、属性和关系结果作为标准输出。

2.采用监督微调训练基础大模型，得到初始结构化信息抽取模型 $\pi_{\boldsymbol{\theta}_{sft}}$。

3.将监督微调后的模型初始化为 GRPO 阶段的策略模型，并冻结为参考模型: $\pi_{\boldsymbol{\theta}} \leftarrow \pi_{\boldsymbol{\theta}_{sft}}$；$\pi_{ref} \leftarrow \pi_{\boldsymbol{\theta}_{sft}}$

4.从训练集中采样输入 $\boldsymbol{u}$，由旧策略模型 $\pi_{\boldsymbol{\theta}_{old}}$ 针对同一输入生成 G 个候选抽取结果: $y(\boldsymbol{u}) = \{\boldsymbol{y}_j\}_{j=1}^{G}$，$\boldsymbol{y}_j \sim \pi_{\boldsymbol{\theta}_{old}}(\cdot \mid \boldsymbol{u})$

5.根据复合奖励函数计算每个候选抽取结果的奖励值：$r_j = R(\boldsymbol{y}_j), j = 1,2,...,G$

6.计算同组候选结果的奖励均值 和标准差 ，并得到组内相对优势：

$$A_j = \frac{r_j - \mu_r}{\sigma_r + \varepsilon}$$

7.根据 GRPO 优化目标更新策略模型参数。

8.重复步骤 4）—7），直至达到设定训练轮数或满足收敛条件。

9.输出优化后的结构化信息抽取模型。

2 实验结果与分析

为验证 GRPO-SIE 方法的有效性，本节以飞机典型故障案例为对象，开展结构化信息抽取实验，并从整体性能、消融实验、知识密度影响方面进行分析。

2.1 实验数据

实验数据来源于《XX 年度飞机典型故障案例汇编》，共选取涉及四型机的 186 个不同故障案例。为获得适用于结构化信息抽取任务的文本样本，首先对原始案例文本进行格式统一和无关字符清理，并按照自然段落进行切分，从中选取 1000 个语义完整的故障文本块，并保留各文本块对应的原

始故障案例编号；随后，以文献[5]构建的飞机故障知识图谱模式为基础，结合故障文本的知识特点，对实体、属性和关系类型进行调整，最终形成包括 5 类实体、7 类关系和 7 类属性的故障知识图谱模式，如图 3 所示。各实体类型及其属性、关系类型的含义如表 1 所示。

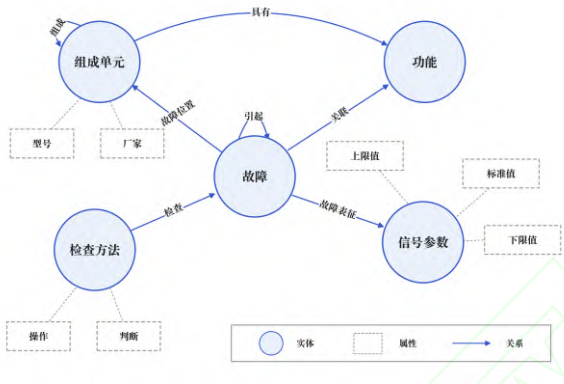

图 3 故障图谱模式
Fig. 3 Aircraft fault knowledge graph schema

在上述图谱模式约束下，借助大模型辅助标注工具，对 1000 个故障文本块中的实体名称、实体类型、属性信息及实体间关系进行初步标注，再由 2 名具有飞机故障诊断专业背景的人员依据图谱模式及各类标签的定义，分别对初始标注结果进行复核、补充和修正。随后，逐项比对两人的标注结果，检查其一致性；对于存在分歧的标注，两人结合原始文本、图谱模式及标签定义讨论确定最终结果，并统一保存为 JSON 格式。完成标注后，对文本长度、实体实例数量、属性实例数量、关系实例数量及类型覆盖情况进行统计，结果如表 2 所示。

表1 实体、属性及关系类型说明
Table 1 Description of entity, attribute, and relation types

实体类型	含义	属性说明	关系类型	连接对象	含义
组成单元	飞机系统、设备、部件等组成对象。	型号：组成单元的型号；厂家：组成单元的生产厂家。	组成	组成单元→组成单元	表示前一个组成单元是后一个组成单元的组成部分。
故障	运行或检查过程中出现的异常状态、失效现象或故障事件。	未设置属性	具有	组成单元→功能	表示组成单元承担某项功能。
功能	组成单元承担的作用或应实现的功能。	未设置属性	引起	故障→故障	表示前一故障导致后一故障。
信号参数	用于反映组成单元运行状态或故障特征的信号及参数。	上限值、下限值:参数允许取值的上下限；标准值:参数的标准或参考值。	故障位置	故障→组成单元	表示故障发生在相应组成单元。
检查方法	用于检查、定位或确认故障的方法。	操作：具体检查操作；判断：判定检查结果的条件或依据。	关联	故障→功能	表示故障与某项功能存在关联。
			检查	检查方法→故障	表示采用某种方法检查相应故障。
			故障表征	故障→信号参数	表示通过相应信号参数反映故障特征。

表2飞机故障文本标注数据统计
Table 2 Statistics of the annotated aircraft fault text dataset

指标	数值	单个文本块均值	单个文本块范围
文本块数量/个	1000	—	—
文本长度/字符	—	318.62	52～986
实体实例数量/个	11483	11.48	2～41
属性实例数量/个	7046	7.05	0～29
关系实例数量/个	13865	13.87	1～57
实体类型覆盖数/类	5	2.72	1～5
属性类型覆盖数/类	7	2.94	0～5
关系类型覆盖数/类	7	4.21	1～6

由表 2 可知，故障文本块的平均长度为 318.62 个字符，取值范围为 52～986 个字符；单个文本块平均包含 11.48 个实体、7.05 个属性和 13.87 条关系，其数量范围分别为 2～41 个、0～29 个和 1～57 个。同时，单个文本块平均覆盖 2.72 类实体、2.94 类属性和 4.21 类关系，表明不同文本块在知识规模和类型覆盖方面存在一定差异，为后续开展不同知识密度条件下的抽取性能分析提供了数据基础。

2.2 实验设置

实验选用 Qwen2.5-32B-Instruct[22]作为基础模型，并采用低秩适配（Low-Rank Adaptation，LoRA）[23]进行参数高效微调。根据文本块对应的原始故障案例编号，以故障案例为基本划分单位，按照案例数量 7:1:2 的比例随机划分训练集、验证集和测试集，三个数据子集分别包含 130、19 和 37 个故障案例，对应 703、110 和 187 个故障文本块。其中，训练集用于模型训练，验证集用于确定 GRPO 策略优化阶段的关键超参数，测试集用于模型最终性能评价。实验在配置 8 块 NVIDIA A100 80 GB GPU 的服务器上完成，模型训练采用 BF16 计算精度，最大序列长度设置为 3072。

GRPO-SIE 的训练过程包括监督微调和 GRPO 策略优化两个阶段。监督微调阶段利用训练集中的标注样本对基础模型进行参数更新，使模型学习图谱模式约束下的实体、属性和关系抽取规则，并按照规定的 JSON 格式生成结构化结果；GRPO 策略优化阶段以监督微调后的模型作为初始策略模型，在复合奖励函数的引导下进一步优化模型生成策略，以提高结构化信息抽取性能和输出规范性。主要实验参数设置如表 3 所示。其中，学习率、训练轮数、批量大小（Batch size）及 LoRA 相关参数依据模型规模、计算资源和实验经验进行设置；复合奖励权重、候选结果数量和 KL 散度系数则根据模型在验证集上的性能表现确定。

表3实验参数设置

Table 3 Experimental parameter settings

参数类别	参数	值	参数	值	参数	值
LoRA 参数	秩	16	缩放系数	32	-	-
SFT 参数	训练轮数	3	学习率	2×10^{-5}	Batch size	16
GRPO 参数	训练轮数	2	学习率	5×10^{-6}	Batch size	16
	候选结果数量	4	KL 散度系数	0.04	策略裁剪系数	0.2
	采样温度	0.8	实体/属性奖励权重	0.30/0.25	关系/格式奖励权重	0.37/0.08

2.3 评价指标

实验选取 F_1 值和格式有效率（format effectiveness，F_E）作为评估指标分别用来评价抽取结果的准确性与格式规范化。

1）F_1 值。采用 F_1 值评价实体、属性和关系抽取性能，计算公式为

$$F_1=\frac{2\times P\times R}{P+R} \tag{15}$$

式中：

$$P=\frac{T_P}{T_P+F_P},$$

$$R=\frac{T_P}{T_P+F_N}$$

其中：T_P、F_P 和 F_N 分别表示正确抽取、错误抽取和遗漏抽取的实例数量。这里，实体、属性和关系实例的匹配规则与 1.3 节相应奖励函数中预测集合与标准集合交集的确定规则一致。在此基础上，分别汇总测试集中三类抽取任务的 T_P、F_P 和 F_N，并计算相应的 F_1 值。

2）格式有效率（F_E）。格式有效率（F_E）用于评价模型输出结果的格式规范性，计算公式为

$$F_E = \frac{N_{\text{valid}}}{N} \times 100\% \tag{16}$$

式中：N_{valid} 为能够被 JSON 正确解析且字段结构符合预设要求的样本数量；N 为测试样本总数。

2.4 对比分析

为评估本文所提 GRPO-SIE 方法的整体性能，将其与 4 种具有代表性的结构化信息抽取方法进行对比，主要对比方法如下：

1）UIE[6]：一种统一结构生成方法，采用结构化抽取语言统一表示实体、关系和事件等不同信息抽取任务，并通过模式引导机制生成相应的结构化结果。

2）LLM-Prompt[13]：一种融合思维链和多轮判断机制的大语言模型提示抽取方法，引导大语言模型从领域文本中抽取实体、属性及实体间关系，并按照规定的结构化格式输出抽取结果。

3）LLM-SFT[10]：一种基于监督微调的大模型结构化信息抽取方法，利用文本—结构化结果样本对大语言模型进行微调，使其按照预定义的抽取模式生成结构化结果。

4）ChatUIE[14]：一种基于对话式大模型的统一信息抽取方法，通过监督微调学习抽取任务，利用奖励模型对生成结果进行评价，并采用 PPO 进行策略优化，同时结合生成约束完成结构化信息抽取。

为保证实验结果的可比性，各方法均采用相同的数据划分、飞机故障知识图谱模式和评价标准，并根据本文定义的实体、属性和关系类型进行适配。具体地，UIE 采用 mT5-base 作为基础模型，根据抽取类型构造模式提示，并将领域标注结果转换为结构化抽取语言进行全参数微调，学习率为 1×10^{-4}，最大训练轮数为 30，Batch size 为 64，预测结果统一转换为 JSON 结构进行评价。LLM-Prompt、LLM-SFT、ChatUIE 和 GRPO-SIE 均以 Qwen2.5-32B-Instruct 为基础模型。LLM-Prompt 通过思维链提示和多轮判断完成抽取，不更新模型参数，最大判断轮数为 3；LLM-SFT 采用常规抽取指令，利用文本—结构化结果样本进行监督微调，不引入本文的结构化任务链提示和 GRPO 策略优化，训练参数与表 3 中的监督微调阶段一致；ChatUIE 采用监督微调、奖励模型训练和 PPO 策略优化流程，其中监督微调参数与 LLM-SFT 一致，奖励模型训练的学习率为 2×10^{-5}，训练轮数为 2，每个训练批次包含 4 组正负样本对，PPO 阶段的策略学习率为 1×10^{-6}，训练轮数为 2，全局采样批量为 4。各方法分别独立进行 10 次实验，统计测试集上实体、属性和关系抽取 F_1 值及格式有效率 F_E 的均值和标准差，结果如表 4 所示。

表4 不同方法的整体性能对比

Table 4 Overall performance comparison of different methods

方法	实体 F_1	属性 F_1	关系 F_1	F_E/%
UIE	0.790 ± 0.008	0.704 ± 0.012	0.698 ± 0.013	90.56 ± 0.83
LLM-Prompt	0.733 ± 0.015	0.619 ± 0.014	0.628 ± 0.016	88.72 ± 1.35
LLM-SFT	0.817 ± 0.006	0.748 ± 0.009	0.743 ± 0.011	94.35 ± 0.62
ChatUIE	0.826 ± 0.007	0.753 ± 0.010	0.766 ± 0.008	97.83 ± 0.57
GRPO-SIE	0.844 ± 0.004	0.782 ± 0.006	0.794 ± 0.007	98.20 ± 0.39

由表 4 可知，五种方法在三项抽取任务上的表现存在一定差异。总体而言，各方法在实体抽取任务上的性能相对较高，而属性和关系抽取任务的性能相对较低。这是因为实体名称及类型通常具有较为清晰的语义边界；相比之下，属性抽取还需识别属性类型、属性值及其所属实体，关系抽取还需判断实体之间的语义关系及方向，抽取难度相对较大。

进一步比较可见，GRPO-SIE 在实体、属性和关系抽取任务上的平均 F_1 值均为最高，分别达到 0.844、0.782 和 0.794。与对比方法中性能最优的 ChatUIE 相比，GRPO-SIE 的实体、属性和关系 F_1 值相对提高 2.18%、3.85%和 3.66%，其中属性和关系抽取的提升更为明显。由此可以看出，GRPO-SIE 在飞机故障文本的属性和关系抽取方面具有一定优势。

在输出规范性方面，UIE、LLM-Prompt、LLM-SFT、ChatUIE 和 GRPO-SIE 的平均 F_E 值分别为 90.56%、88.72%、94.35%、97.83%和 98.20%，其中 GRPO-SIE 的平均 F_E 值最高，表明该方法在提高抽取性能的同时，能够保持较好的结构化输出规范性。同时，GRPO-SIE 在各项指标上的标准差相对较小，说明其实验结果具有较好的稳定性。

2.5 消融实验

为分析各组成部分对模型性能的影响，本节在完整 GRPO-SIE 模型基础上开展消融实验，具体设置如下：（1）GRPO-SIE w/o STC，移除结构化任务链提示，仅保留角色设定、输入文本、少样本示例和 JSON 输出要求指令；（2）GRPO-SIE w/o SFT，跳过监督微调阶段，直接以基础模型初始化策略模型和参考模型，并采用复合奖励函数进行 GRPO 训练；（3）GRPO-SIE w/o GRPO，保留结构化任务链提示和监督微调，但不进行复合奖励驱动的 GRPO 策略优化。其余实验设置均保持不变，各消融方案分别独立进行 10 次实验，实体、属性和关系抽取 F_1 值及格式有效率 F_E 的均值和标准差如表 5 所示。

表5 不同消融方案的实验结果

Table 5 Experimental results of different ablation settings

方法	实体 F_1	属性 F_1	关系 F_1	F_E/%
GRPO-SIE w/o STC	0.831 ± 0.007	0.761 ± 0.014	0.768 ± 0.017	95.34 ± 0.66
GRPO-SIE w/o SFT	0.788 ± 0.015	0.682 ± 0.023	0.671 ± 0.034	92.15 ± 1.62
GRPO-SIE w/o GRPO	0.818 ± 0.009	0.741 ± 0.015	0.747 ± 0.018	94.86 ± 0.88
GRPO-SIE	0.844 ± 0.004	0.782 ± 0.006	0.794 ± 0.007	98.20 ± 0.39

由表 5 可知，移除任一组成部分后，模型在三项抽取任务上的平均 F_1 值和 F_E 均有所下降。其中，移除 STC 后，实体、属性和关系 F_1 值分别下降 0.013、0.021 和 0.026，说明结构化任务链提示有助于规范抽取过程，尤其能够改善属性和关系抽取性能；移除 SFT 后，三项 F_1 值分别下降 0.056、0.100 和 0.123，F_E 下降 6.05%，且各项指标的标准差增大，表明监督微调能够使模型初步掌握领域知识类型和结构化输出规则，为后续策略优化提供较为稳定的初始化。移除 GRPO 策略优化后，实体、属性和关系 F_1 值分别下降 0.026、0.041 和 0.047，F_E 下降 3.34%，说明复合奖励驱动的 GRPO 策略优化能够在监督微调基础上进一步改善抽取性能和输出规范性。由此可以看出，STC、SFT 和 GRPO 策略优化分别在抽取过程引导、领域能力初始化和生成结果优化中发挥作用。

为进一步验证复合奖励中各奖励项的作用，在完整 GRPO-SIE 基础上，分别移除实体奖励、属性奖励、关系奖励和格式奖励，其余奖励项及其权重设置保持不变；同时，以通用 F_1 奖励替代复合奖励作为对照。具体地，通用 F_1 奖励按照与评价指标一致的匹配规则计算，分别统计每条候选结果 $\boldsymbol{y}_j$ 中实体、属性和关系抽取的正确数、错误数和遗漏数，然后合并计算整体微平均 F_1，其计算公式如下：

$$R_{F_1}(\boldsymbol{y}_j) = \frac{2\sum_{k\in\{E,C,T\}} T_{P_{k,j}}}{2\sum_{k\in\{E,C,T\}} T_{P_{k,j}} + \sum_{k\in\{E,C,T\}} F_{P_{k,j}} + \sum_{k\in\{E,C,T\}} F_{N_{k,j}}} \tag{17}$$

式中：E,C,T 分别表示实体、属性和关系抽取任务；$T_{P_{k,j}}$、$F_{P_{k,j}}$ 和 $F_{N_{k,j}}$ 分别表示候选结果 $\boldsymbol{y}_j$ 中第 k 类抽取结果的正确数、错误数和遗漏数。若候选结果无法解析为有效 JSON，或存在字段缺失、结构不符合要求等情况，则将其奖励置为 0。各方案均保留相同的结构化任务链提示、监督微调初始化和 GRPO 训练参数，并分别独立进行 10 次实验，结果如表 6 所示。

表6 不同奖励设置下的实验结果

Table 6 Experimental results under different reward settings

奖励设置	实体 F_1	属性 F_1	关系 F_1	F_E/%
移除实体奖励 R_E	0.825±0.011	0.748 ± 0.013	0.755 ± 0.017	96.95 ± 0.48
移除属性奖励 R_C	0.832±0.004	0.726 ± 0.012	0.787 ± 0.009	95.88 ± 0.53
移除关系奖励 R_T	0.841±0.008	0.753 ± 0.006	0.758 ± 0.013	96.82 ± 0.95
移除格式奖励 R_F	0.835±0.005	0.764 ± 0.007	0.771 ± 0.006	94.85 ± 0.87
使用通用 F_1 奖励	0.838±0.006	0.769 ± 0.005	0.783 ± 0.008	95.60 ± 0.76
GRPO-SIE	0.844±0.004	0.782 ± 0.006	0.794 ± 0.007	98.20 ± 0.39

由表 6 可知，移除任一奖励项后，模型的抽取性能和格式有效率均有所下降。移除实体奖励后，实体、属性和关系抽取的平均 F_1 值分别下降 0.019、0.034 和 0.039，表明实体奖励不仅有助于实体抽取，也对属性和关系抽取具有一定促进作用；移除属性奖励和关系奖励后，对应任务的平均 F_1 值分别下降 0.056 和 0.036，进一步体现了分项奖励对相应抽取任务的作用；移除格式奖励后，F_E 下降 3.35%，三项平均 F_1 值分别下降 0.009、0.018 和 0.023，说明格式奖励不仅能够改善输出规范性，也为有效抽取提供基础；当复合奖励替换为通用 F_1 奖励后，三项平均 F_1 值分别下降 0.006、0.013 和 0.011，F_E 均值下降 2.60%，由此可见，相较于单一整体 F_1 奖励，复合奖励能够更好地兼顾实体、属性、关系抽取及结构化输出规范性。

2.6 不同知识密度下的性能分析

为分析模型在不同知识密度文本中的抽取性能，根据文本块中实体、属性和关系实例数量及文本长度计算知识密度。具体地，第 i 个文本块的知识密度定义为

$$D_i = \frac{N_i^E + N_i^C + N_i^T}{L_i} \times 100 \tag{18}$$

式中：N_i^E、N_i^C 和 N_i^T 分别表示第 i 个文本块中的实体、属性和关系实例数量；L_i 表示文本长度，D_i 表示每 100 个字符中包含的知识实例数量。D_i 越大，表明单位文本长度内需要抽取和组织的知识实例越多，即知识实例分布越密集。根据测试集中各文本块知识密度值的分布，计算得到其 33.3%和 66.7%分位点分别为 8.3 和 11.8，并以此作为划分阈值，将文本块划分为低、中、高 3 个知识密度组。其中，低密度组满足 $D_i < 8.3$，中密度组满足 $8.3 \le D_i < 11.8$，高密度组满足 $D_i \ge 11.8$。各组文本块数量、文本长度及知识实例数量统计结果如表 7 所示。

表7 不同知识密度组的样本统计

Table 7 Sample statistics for different knowledge density groups

知识密度组	文本块数	文本长度/字符	实体实例数/个	属性实例数/个	关系实例数/个
低密度	58	290.6±82.5	7.8± 2.6	3.4± 2.4	8.4± 4.8
中密度	43	315.3±76.1	10.5±3.5	6.3± 4.0	12.5±6.9
高密度	86	340.5±84.2	15.1±5.4	10.4±6.1	19.2±10.3

1）不同知识密度组的整体性能。利用 2.4 节各对比方法 10 次独立实验得到的测试集预测结果，分别统计各方法在低、中、高知识密度组上的实体、属性和关系抽取 F_1 值，并计算 10 次结果的平均值，结果如图 4 所示。

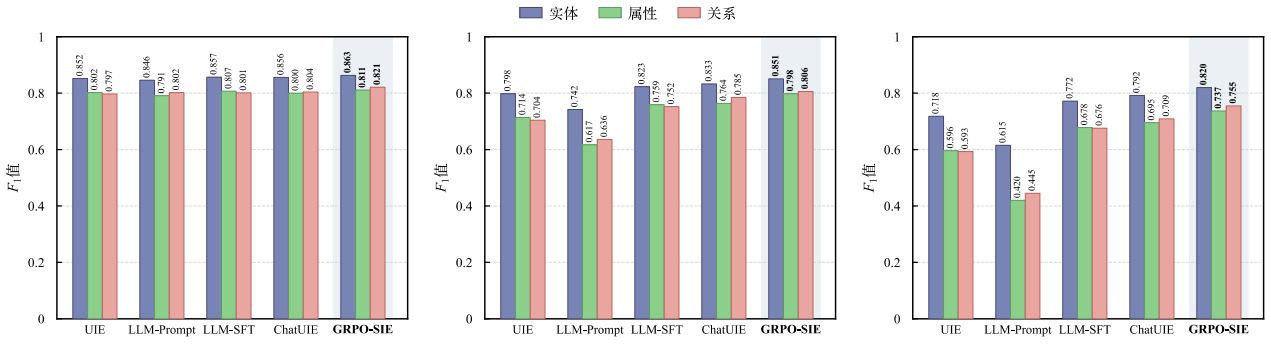

图 4 不同知识密度文本下各方法的抽取性能

Fig.4 Extraction performance of different methods on texts with different knowledge densities

由图 4 可知，随着知识密度的提高，各方法在实体、属性和关系抽取任务上的 F_1 值均呈下降趋势。在低知识密度组中，各方法的抽取性能相对较高，GRPO-SIE 与其他方法之间的差异较小；在中、高知识密度组中，GRPO-SIE 在三项抽取任务上的优势逐渐显现。特别是在高知识密度组中，与 ChatUIE 相比，GRPO-SIE 的实体、属性和关系 F_1 值相对提高 3.54%、6.04%和 6.49%。这表明，在当前知识密度分组条件下，GRPO-SIE 在知识实例分布较为密集的文本中能保持较好的结构化信息抽取性能。

2）文本长度匹配下的性能验证。考虑到不同知识密度组之间存在一定的文本长度差异，为减小文本长度差异对上述结果的影响，从低、高知识密度组中选取文本长度相近的文本块进行一一匹配，并基于各方法已有的测试集预测结果重新统计抽取性能。具体地，设低、高知识密度组中待匹配文本块的长度分别为 L_l 和 L_h，当其相对长度差满足 $|L_l - L_h| / ((L_l + L_h)/2) \le 0.05$ 时，将该文本块对视为长度相近；每个文本块仅参与一次匹配。最终得到 44 对满足条件的低、高知识密度文本块，统计结果如表 8 所示。

表8 长度匹配后的低、高知识密度文本块统计

Table 8 Statistics of low- and high-density text blocks after length matching

知识密度组	文本块数	文本长度/字符	实体实例数/个	属性实例数/个	关系实例数/个
低密度	44	326.2±30.1	8.2 ± 1.8	3.7±1.5	9.4 ± 2.3
高密度	44	332.9±30.6	13.8±2.6	9.3±1.9	18.0±4.5

基于各方法已有的测试集预测结果，重新计算匹配样本上的抽取性能，结果如图 5 所示。与图 4 所反映的整体趋势一致，GRPO-SIE 在低、高知识密度组的实体、属性和关系抽取 F_1 值均为最高；由低知识密度组到高知识密度组时，三项 F_1 值的下降幅度也最小，分别为 0.049、0.081 和 0.077。这表明，在文本长度相近的情况下，GRPO-SIE 仍能够较好地处理知识密度较高文本中的结构化信息抽取任务。

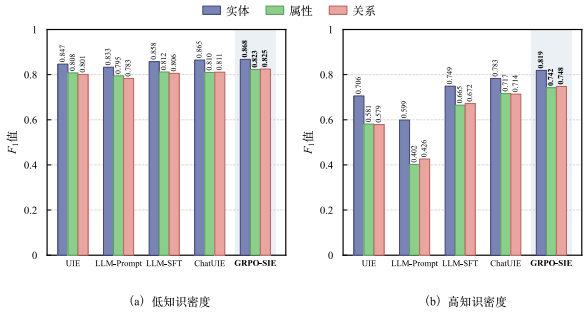

图 5 文本长度匹配条件下低、高知识密度文本的抽取性能

Fig.5 Extraction performance on low- and high-knowledge-density texts under length-matched conditions

3）高知识密度文本的典型案例分析。为进一步解释不同方法在高知识密度文本中的性能差异，从测试集高知识密度文本的逐样本预测结果中选取 3 个典型案例进行分析，结果如表 9 所示。

表9 高知识密度文本的典型预测结果分析

Table 9 Analysis of typical predictions on high-density texts

文本片段（节选）	标准标注结果（节选）	对比方法预测结果（节选）	GRPO-SIE 预测结果（节选）
...起飞前供电检查时，直流供电系统中的直流接触器（XX-01 型，厂家 A 生产）触点烧蚀，机载直流母线电压出现波动；采样支路（XX-02 型，厂家 B 生产）未见异常。采用电压测量方法进行检查，母线电压的标准值为 28 V，允许范围为 26～30 V...	属性：直流接触器—型号 XX-01/厂家 A；采样支路—型号 XX-02/厂家 B；母线电压—标准值 28/下限值 26/上限值 30。 关系：触点烧蚀—故障位置→直流接触器；触点烧蚀—引起→母线电压波动；母线电压波动—故障表征→母线电压；电压测量—检查→母线电压波动。	UIE：直流接触器—型号 XX-01/厂家 A/型号 XX-02/厂家 B；采样支路未关联型号 XX-02、厂家 B。 LLM-Prompt：“母线电压—上限值 26/下限值 30”，属性上下值混淆	采样支路—型号 XX-02/厂家 B；母线电压—下限值 26/上限值 30
...例行检查发现液压回油系统工作异常。对回油滤芯进行压差检查，结果显示滤芯前后压差超过规定范围，回油压差增大；随后拆检发现回油滤芯发生堵塞...	关系：回油滤芯堵塞—故障位置→回油滤芯；回油滤芯堵塞—引起→回油压差增大；回油压差增大—故障表征→滤芯前后压差；压差检查—检查→回油压差增大；拆检—检查→回油滤芯堵塞。	LLM-SFT：回油压差增大—引起→滤芯前后压差。 ChatUIE：压差检查—检查→回油滤芯堵塞，遗漏“拆检—检查→回油滤芯堵塞”。	回油压差增大—故障表征→滤芯前后压差；压差检查—检查→回油压差增大；拆检—检查→回油滤芯堵塞
...飞行过程中，燃油量指示发生跳变。...检查发现，燃油量系统中油量探头绝缘性能下降、连接支路接触电阻升高均可能导致该现象。随后采用绝缘测量和导通检查分别对前者和后者进行复核...	关系：油量探头绝缘下降—引起→油量指示跳变；连接支路接触电阻升高—引起→油量指示跳变；绝缘测量—检查→油量探头绝缘下降；导通检查—检查→连接支路接触电阻升高；油量探头—组成→燃油量系统；连接支路—组成→燃油量系统。	UIE：仅抽取“油量探头绝缘下降—引起→油量指示跳变”，遗漏连接支路异常对应的因果关系。 LLM-Prompt：绝缘测量—检查→连接支路接触电阻升高；导通检查—检查→油量探头绝缘下降。	油量探头绝缘下降—引起→油量指示跳变；连接支路接触电阻升高—引起→油量指示跳变；绝缘测量—检查→油量探头绝缘下降；导通检查—检查→连接支路接触电阻升高；油量探头—组成→燃油量系统。

由表 9 可知，案例 1 中，多个实体及其型号、厂家和参数属性在同一文本片段中连续出现，UIE 将采样支路的属性错误归属于直流接触器，LLM-Prompt 则混淆了母线电压的上下限属性。案例 2 中，文本包含“压差检查发现压差异常—拆检确认滤芯堵塞”的连续诊断过程，LLM-SFT 将故障表征关系误判为引起关系，ChatUIE 则混淆了压差检查与拆检所对应的检查对象。案例 3 中，“前者”、“后者”的跨句指代以及并列组成单元增加了关系对象匹配难度，导致对比方法出现因果关系遗漏或检查

对象错配。相比之下，GRPO-SIE 在案例 1 中能够正确关联实体及其属性，在案例 2 中能够较完整地抽取故障位置、引起、故障表征及检查关系；在案例 3 中能够正确抽取两条因果关系和两条检查关系，但仍遗漏“连接支路—组成→燃油量系统”这一关系。上述案例表明，GRPO-SIE 能够较好地处理高知识密度文本中属性信息密集、多类关系共现和跨句指代等情况，但在并列多对象关系同时出现时，仍可能存在局部关系遗漏。

3　结论

针对飞机故障文本结构化信息抽取中准确性和输出规范性不足的问题，本文提出了一种基于 GRPO 优化的大模型结构化信息抽取方法，记为 GRPO-SIE。主要结论如下：

1）提出了融合结构化任务链提示、监督微调和 GRPO 策略优化的结构化信息抽取方法。通过飞机故障知识图谱模式约束抽取范围，并利用由实体、属性、关系和格式奖励组成的复合奖励函数对模型生成结果进行细粒度优化，实现了飞机故障文本中实体、属性和关系信息的统一抽取与规范化输出。

2）实验结果表明，GRPO-SIE 在实体、属性和关系抽取任务上的平均 F_1 值分别达到 0.844、0.782 和 0.794，三项指标均优于各对比方法，并在不同知识密度的文本中保持了较好的抽取性能。消融实验进一步验证了结构化任务链提示、监督微调、GRPO 策略优化及复合奖励设计的作用。从工程应用角度看，所提方法能够将实际故障案例文本中分散描述的故障现象、组成单元、检查方法和信号参数等信息，按照统一的知识图谱模式组织为实体、属性和关系，实现故障知识的自动抽取与结构化表达。这为专业人员开展故障知识整理和校验提供了候选结果，也为故障知识图谱构建及后续知识检索、关联分析和图谱推理提供了数据基础。

本文方法仍依赖预定义的飞机故障知识图谱模式和领域标注数据，而且当前实验主要验证了结构化信息抽取性能，尚未对图谱推理和故障根因诊断开展系统级验证。后续将面向更多飞机故障场景扩展实验数据，探索逐字段、逐实体的细粒度奖励反馈机制，并进一步结合知识图谱构建、图谱推理和故障根因诊断模块，研究从原始故障文本到故障知识应用的完整技术流程。

参考文献（References）

[1] Tang X,Chi G,Cui L,et al.Exploring research on the construction and application of knowledge graphs for aircraft fault diagnosis[J].Sensors,2023,23(11):5295.

[2] 吴闯，张亮，唐希浪，等．航空发动机润滑系统故障知识图谱构建及应用[J]．北京航空航天大学学报，2024，50（4）：1336-1346.
Wu C,Zhang L,Tang X L,et al.Construction and application of fault knowledge graph for aero-engine lubrication system[J].Journal of Beijing University of Aeronautics and Astronautics,2024,50(4):1336-1346(in Chinese).

[3] Meng X,Jing B,Wang S,et al.Fault knowledge graph construction and platform development for aircraft PHM[J].Sensors,2024,24(1):231.

[4] Peng H,Chen Y,Li Y,et al.Knowledge graph construction method for commercial aircraft fault diagnosis based on logic diagram model[J].Aerospace,2024,11(9):773.

[5] Xie X Y,Tang X L,Gu S W,et al.An intelligent guided troubleshooting method for aircraft based on HybirdRAG[J].Scientific Reports,2025,15:17752.

[6] Lu Y,Liu Q,Dai D,et al.Unified structure generation for universal information extraction[C]//Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics.Stroudsburg:Association for Computational Linguistics,2022,1:5755-5772.

[7] Wang X,Zhou W,Zu C,et al.InstructUIE:multi-task instruction tuning for unified information extraction[EB/OL].(2023-04-17)[2026-07-06].https://arxiv.org/abs/2304.08085v1.

[8] Wei X,Cui X,Cheng N,et al.ChatIE:zero-shot information extraction via chatting with ChatGPT[EB/OL].(2023-02-20)[2026-07-06].https://arxiv.org/abs/2302.10205v1.

[9] Wan Z,Cheng F,Mao Z,et al.GPT-RE:in-context learning for relation extraction using large language models[C]//Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing.Stroudsburg:ACL,2023:3534-3547.

[10] Dagdelen J,Dunn A,Lee S,et al.Structured information extraction from scientific text with large language models[J].Nature Communications,2024,15:1418.

[11] Li Y,Ramprasad R,Zhang C.A simple but effective approach to improve structured language model output for information extraction[C]//Findings of the Association for Computational Linguistics:EMNLP2024.Stroudsburg: ACL,2024:5133-5148.

[12] 冯蕴雯，刘晚移，万春华，等．基于结构引导提示学习的试验实体关系抽取方法[J/OL]．北京航空航天大学学报，（2026-06-01）[2026-07-06]．https：//doi．org/10．13700/j．bh．1001-5965．2026．0093．
Feng Y W,Liu W Y,Wan C H,et al.An test entity relation extraction method based on structure-guided prompt learning[J/OL].Journal of Beijing University of Aeronautics and Astronautics,(2026-06-01)[2026-07-06].https://doi.org/10.13700/j.bh.1001-5965.2026.0093(in Chinese).

[13] Yan S,Xie X,Cui L,et al.An LLM-driven automatic construction method for knowledge graphs in UAV equipment training[C]//Proceedings of 5th 2025 International Conference on Autonomous Unmanned Systems.Singapore:Springer Nature Singapore,2026,1578:392-405.

[14] Xu J,Sun M,Zhang Z,et al.ChatUIE:exploring chat-based unified information extraction using large language models[C]//Proceedings of the 2024 Joint International Conference on Computational Linguistics,Language Resources and Evaluation.Paris:ELRA and ICCL,2024:3146-3152.
[15] Takanobu R,Zhang T,Liu J,et al.A hierarchical framework for relation extraction with reinforcement learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence.Palo Alto:AAAI Press,2019,33(1):7072-7079.
[16] Ding Z,Ke R,Huang W,et al.Adaptive reinforcement learning planning:harnessing large language models for complex information extraction[EB/OL].(2024-06-17)[2026-07-06].https://arxiv.org/abs/ 2406.11455v1.
[17] Dai R,Zheng T,Yang R,et al.R1-RE:cross-domain relation extraction with RLVR[EB/OL].(2025-07-07)[2026-07-06].https://arxiv.org/abs/2507.04642v1.
[18] Li R,Di S,Liu Y,et al.Beyond path selection:better LLMs for scientific information extraction with MimicSFT and relevance and rule-induced GRPO[EB/OL].(2025-05-28)[2026-07-06].https://arxiv.org/abs/2505.22068v1.
[19] Schulman J,Wolski F,Dhariwal P,et al.Proximal policy optimization algorithms[EB/OL].(2017-07-20)[2026-07-06].https://arxiv.org/abs/1707.06347v1.
[20] Shao Z,Wang P,Zhu Q,et al. DeepSeekMath: Pushing the limits of mathematical reasoning in open language models[EB/OL].(2024-02-05)[2026-07-06].https://arxiv.org/abs/2402.03300v1.
[21] Wei J,Wang X,Schuurmans D,et al.Chain-of-thought prompting elicits reasoning in large language models[C]//Advances in Neural Information Processing Systems.Red Hook:Curran Associates,2022,35:24824-24837.
[22] Yang A,Yang B,Zhang B,et al.Qwen2.5 technical report[EB/OL]. (2024-12-19)[2026-07-06].https://arxiv.org/abs/2412.15115v1.
[23] Hu E J, Shen Y, Wallis P, et al. LoRA: Low-rank adaptation of large language models[EB/OL].(2021-10-16)[2026-07-06]. https://arxiv.org/abs/2106.09685.