

基于知识图谱的智能决策辅助系统研究

钟昊¹, 郭永贞², 宗孝鹏¹

(1. 佳讯飞鸿智能科技研究院, 北京 100095; 2. 山东大莱龙铁路有限责任公司, 山东烟台 265400)

摘要: 信息化战争瞬息万变, 指挥人员的经验思维和人工分析方法已经无法满足速度、维度、精度、强度、广度的要求。通过分析作战领域态势认知、威胁分析、方案优选等智能决策的军事需求, 充分运用知识图谱的智能分析技术原理和方法。提出了决策智能引擎的框架、原理、特征和要求。研究了军事领域知识图谱构建方法, 多级混合知识推理框架构建, 以及智能决策辅助系统的实现流程。系统可广泛应用于作战指挥、情报分析、军事训练等领域。

关键词: 智能决策; 知识图谱; 知识推理; 机器学习

Study on Intelligent Decision Support System based on Knowledge Graph

ZHONG Hao¹, GUO Yong-zhen², ZONG Xiao-peng¹

(1. Jiaxun Feihong Intelligent Technology Institute, Beijing 100095;

2. Shandong Dalai Long Railway Co. Ltd., Yantai Shandong 265400)

Abstract: Information warfare is changing rapidly, and the empirical thinking and manual analysis methods of commanders can no longer meet the requirements of speed, dimensionality, precision, intensity, and breadth. By analyzing the military needs of intelligent decision-making such as situational awareness, threat analysis, and plan optimization in the operational domain, the principles and methods of intelligent analysis technology of the knowledge map are fully utilized. The framework, principles, characteristics and requirements of the decision-making intelligence engine are put forward. Study on the construction method of military domain knowledge graph, the construction of multi-level mixed knowledge reasoning framework, and the realization process of intelligent decision aid system. The system can be widely used in combat command, intelligence analysis, military training and other fields.

Keywords: Intelligent decision; knowledge graph; knowledge reasoning; machine learning

0 引言

随着信息化战争加速向智能化战争演化, 现代战争作战空间更加广阔、战场环境更加复杂、兵力构成更加多样、对抗节奏更加快捷、敌我博弈更加激烈, 对态势分析、方案优选、威胁分析等指挥决策活动提出了全新的要求和挑战, 仅靠指挥员和参谋人员的经验思维、简单计算和手工作业等传统方式, 已经很难适应智能化战争决策对速度、维度、精度、强度、广度的全新要求, 必须高度重视采用人工智能等先进的理念、技术、系统和工具, 研发

通用决策智能引擎, 不断提升军事领域态势实时研判和辅助决策的质量和水平。从技术层面来看, 智能决策辅助系统应当主要包括 7 个方面的能力需求 (图 1)。

1 研究现状

智能决策辅助系统是一项复杂的系统工程、体系工程, 涉及要素多、流程多、环节多, 系统集成难度大, 涉及问题识别、知识图谱、信息集成、知识推理、人机交互、智能问答、决策支持、自然语言理解等多个领域的技术现况。

作者简介: 钟昊 (1974—), 男 (汉), 四川隆昌人, 高级工程师, 博士研究生, 主要研究领域为人工智能应用技术, 邮箱: zhonghao@jxresearch.com。

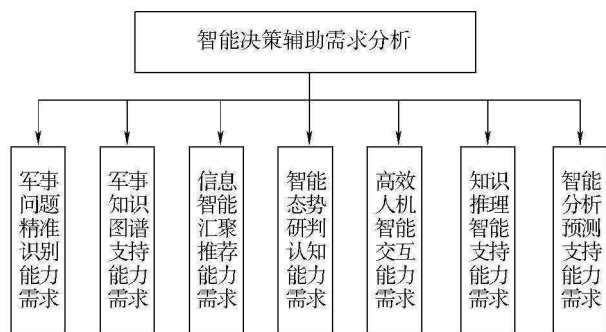


图1 能力需求

1.1 决策支持系统

决策支持系统 (Decision Support System, DSS) 的概念于 20 世纪 70 年代初由美国 Michael S. Scott Morton 在《管理决策系统》一文首次提出, 20 世纪 80 年代中期引入我国。20 多年来, DSS 已在理论研究、系统开发和实际应用诸方面取得了令人瞩目的进步, 并呈现出积极的多元化的发展态势。Holsapple^[1]等总结了系统的对决策过程的支持能力和学习能力, 把 IDSS 分为四类: 无自适应性被动支持; 无自适应性, 能提供主动支持; 自适应, 被动支持; 自适应, 主动支持。Mirchandani^[2]等从知识系统进化的角度讨论了 IDSS 模型, 从系统的知识、学习和进化能力来看, 系统的智能在逐步加深。

1.2 智能问答系统

智能问答系统起源于聊天机器人。聊天机器人, 是一种通过自然语言模拟人类, 进而与人进行对话的程序。它既可以在特定的软件平台 (如, PC 平台或者移动终端设备) 上运行, 也可以在类人的硬件机械体上运行。典型代表包括:

- 苹果公司的个人语音助理 Siri。
- IBM 公司的最强大脑 Watson。
- 谷歌公司的智能个人助理 Google Now。
- 微软公司的个人机器人助理 Cortana。

智能问答系统包括主动推荐系统、问答系统、面向任务的对话系统和闲聊系统, 都是综合集成了问题识别、智能人机交互、知识库、信息库、知识推理等技术, 实现了对人类决策的辅助与支持, 相关理念、框架和技术对于构建军事领域的通用决策智能引擎具有重要参考价值和借鉴意义。

1.3 知识图谱技术

知识图谱技术是人工智能技术的组成部分, 其强大的语义处理和互联组织能力, 为智能化信息应

用提供了基础。知识图谱 (Knowledge Graph) 以结构化的形式描述客观世界中概念、实体及其关系, 将信息表达成更接近人类认知世界的形式, 提供了一种更好地组织、管理和理解互联网海量信息的能力。具代表性大规模网络知识获取的工作包括 DBpedia、Freebase、KnowItAll、WikiTaxonomy 和 YAGO, 以及 BabelNet、ConceptNet、DeepDive、NELL、Probase、Wikidata、XLORE、Zhishi.me 等^[3]。这些知识图谱遵循 RDF 数据模型, 包含数以千万级或者亿级规模的实体, 以及数十亿或百亿事实 (即属性值和与其他实体的关系), 并且这些实体被组织在成千上万的由语义类体现的客观世界的概念结构中。知识图谱与大数据和深度学习一起, 成为推动互联网和人工智能发展的核心驱动力之一。

1.4 美军通用决策智能辅助发展情况

面对指挥信息系统的智能化问题, 美军特别是 DARPA 已经开展了很多基础性探索研究, 具体包括: 在基于大数据的辅助决策研究方面, 先后启动 Inisight 和 XDAT 项目, 主要针对实时性、分布式和不完整的战场数据信息开发自动化分析处理平台, 为战场快速决策和自动决策提供支持; 在基于深度学习的认知智能研究方面, 先后启动了深度学习 (DL)、文本深度发掘与过滤 (DEFT) 和高级机器学习概率编程 (PPAML) 等项目, 探索智能机器学习方法, 从不确定信息中理解数据、分析结果和推理关系, 实现辅助决策支持; 在面向战场管理的辅助决策研究方面, 于 2014 年提出分布式战场管理 (DBM) 项目^[4], 帮助飞行员监视战场态势和规划作战任务。同时 DARPA 认为, 未来智能项目重在人机结合, 即“半人马模式”, 应用重点在无人机群、智能雷达、网络作战及情报的数据洪流处理等领域。

2 系统框架

以军事领域通用决策中态势研判、方案优选、威胁分析典型决策需求为牵引, 以智能分析和推理通用技术为重点, 以构建问题库、信息库、情报库、知识库、地理空间库、自然语料库为基础, 重点研究问题分解、数据收集、信息汇聚、知识图谱、问题推理五个领域的关键技术, 设计和构建问题分解与求解规划系统, 实现问题的分解与求规划; 设计和构建面向语义理解的智能问答系统, 实现问题的交互输入; 设计和构建军事领域知识图谱, 支持知

识推理和求解；设计和构建统一的智能引擎支撑平台，实现各类资源统一调度与运行，实现智能决策辅助支持。基于对通用决策机理的分析，将智能分

析和智能引擎分为智能交互、智能汇聚、智能推理等要素，构建系统架构（图2）。

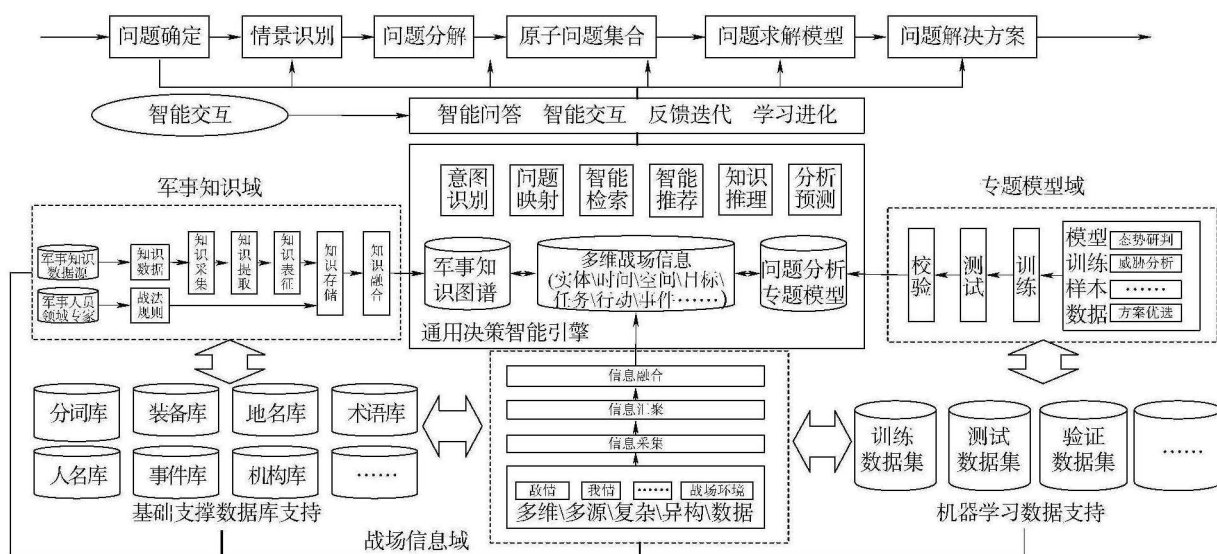


图2 系统框架

3 军事领域知识图谱构建

从本质上看，军事领域的知识图谱和通用领域的知识图谱，都是存储领域知识的结构化的语义知识库，都要是用于以符号形式描述物理世界中的概念及其相互关系。其基本组成单位是“实体—关系—实体”三元组，以及实体及其相关属性——值对，实体间通过关系相互联结，构成网状的知识结构。

构建知识图谱包括知识抽取、知识表示、知识融合、知识推理、知识验证等技术^[5,6]。如图3所示，知识图谱的构建过程，同时也是知识图谱更新的过程，知识图谱的构建过程是从原始数据出发，采用一系列自动或半自动的技术手段，从原始数据中提取出知识要素（即事实），并将其存入知识库的数据层和模式层的过程。这是一个迭代更新的过程，根据知识获取的逻辑，每一轮迭代包含三个阶段：信息抽取、知识融合以及知识加工。

3.1 知识数据获取

围绕军事领域知识分类体系框架，综合采集专业文献、互联网数据、内部数据库等数据源，构建知识数据源，并采用数据集、数据库的方式统一存储，构建支持知识数据源检索支持数据平台。

主要采取两种方式录入与采集知识：一是通过数据源采集知识。包括从军事领域的各种专题数据

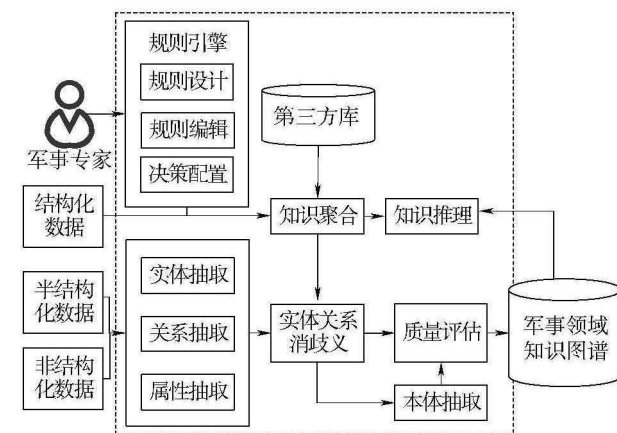


图3 军事领域知识图谱构建

库，各种数字化专业文献、著作、报告、教程等，以及图像数据、视频数据、音频数据等，以结构化、半结构化、非结构化的形式录入与采集数据。二是通过规则引擎采用人机交互的方式，支持直接从部队作战人员或领域专家中，采集相关规则、决策条件、决策经验、行动原则、处置策略等实践知识，并以形式化、规则的方式形成知识集合。

3.2 知识抽取

知识抽取，即从不同来源、不同结构的知识源数据中进行知识提取，形成知识（结构化数据）存入到知识图谱，知识抽取的任务主要包括：

一是命名实体识别。例如，华盛顿特区是美国首都。[华盛顿特区]：实体。

二是分类。例如，F35B 是一种短距起降战机。抽取结果 [F35B]: 战机。

三是术语抽取。从语料中发现多个单词组成的相关术语。

四是关系抽取。抽取各实体之间的关系。例如，[七舰队] <隶属关系> [美国海军太平洋舰队]。

五是事件抽取。例如从消息中抽取出事件发生是触发词、时间、地点等信息。

六是共指消解。弄清楚在一句话中的代词的指代对象。

3.3 知识表征

从海量数据知识源提取的知识，需要采用统一的表征技术。知识表征方法是研究系统中知识的组织形式，强调表示和控制之间的关系，表示与推理及其他研究领域的知识。知识表征与问题的性质和推理控制策略有密切的关系。任何一个给定的问题都有许多等价的表示方法，但它们可产生完全不同的效果。通过综合研究状态空间、与或图、谓词逻辑、产生式规则、语义网络、框架、情景分析，实现知识的表征。从实现来看，常用方法包括：状态空间表征方法、产生式规则表征方法、语义网络表征方法、框架表征方法、场景表示方法等，本文不再展开。

3.4 知识融合

由于知识图谱的实体、关系等知识单元往往源于多个数据源，可能存在有多个知识实体、关系，在不同数据源有对同一实体的不同表达，即使在一个数据源里也可能存在这种情况，需要通过一定手段将其合并、融合、对齐。知识融合中难点在于相似度、冲突检测与消解处理。

相似度计算：对实体的属性进行计算相似度，有了实体各个属性的相似度才容易进行下一步的实体相似度计算。实体对齐 (Object Alignment)，旨在发现具有不同 ID 但却代表真实世界中同一对象的那些实体，并将这些实体归并为一个具有全局唯一标识的实体对象添加到知识图谱中。常用计算方法包括：回归和聚类。

冲突检测与消解：当融合来自不同数据源的信息构成知识图谱时，有一些实体会同时属于两个互斥的类别或某个实体所对应的一个属性对应多个值。这样就会出现不一致性。研究中采取的基本方法是，充分考虑数据源的可靠性以及不同信息在各个数据源中出现的频度等因素来决定最终选用哪个类别或

哪个属性值。也就是说，优先采用那些可靠性高的数据源 (如信源置信度) 抽取得到的事实。另外，如果一个实体在多个数据源中都被识别为某个类别的实例，或实体某个功能属性在多个数据源中都对应相同的值，倾向于最终选择该类别和该值。在统计某个类别在数据源中出现的频率前需要完成类别对齐计算。类似地，对于数值型的属性值还需要额外统一使用的单位。

4 多级混合推理框架设计

知识推理是决策支撑系统的重要组成部分，也是决策智能引擎的核心。知识推理是在现有知识的基础上，使用一定的规则对数据和内容进行推导，得到数据中隐藏的有价值的信息或者数据之间的内在关联。在语义网中，本体作为人工智能中语义网研究的核心概念之一，在知识处理、共享和重用中扮演了重要的角色，已被广泛应用于特定领域知识的描述和组织，以本体为知识源的推理开辟了信息检索技术的新方向。而本体推理机是本体创建和辅助决策系统中必不可少的基础支撑工具之一。

推理可以分为通用推理机和专用推理机，通用推理应用面广，专用推理机效率高，但是支持本体语言有限。综合多推理机的特色和优点，将一个推理机构建在另一个推理机之上，也即把经由一个推理机处理后得到的推理模型作为另一个推理机的基本数据源，构建多级混合推理框架，如图 4 所示。该推理机综合了通用推理机和专用推理机的优点，

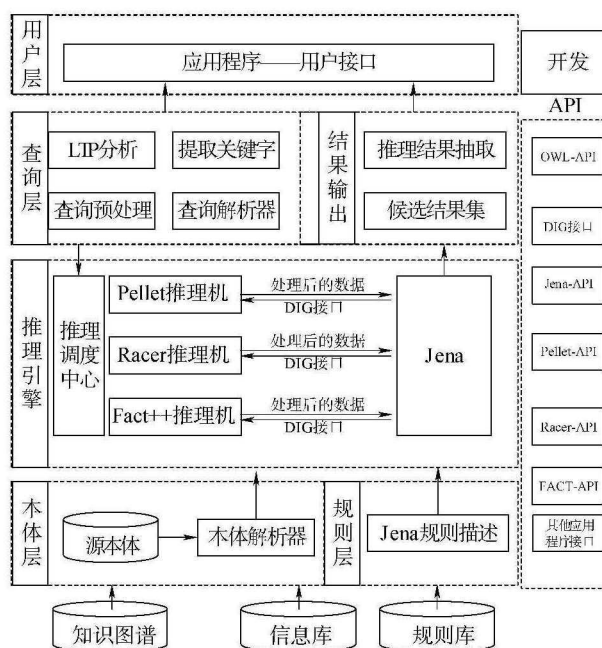


图 4 多级混合推理

并在实验中应用 Pellet 和 Jena 推理机^[7,8] 组建层叠式推理, 运用 Pellet 推理机和 Jena 自定义规则对语义进行推理。针对具体领域, 可以制定复杂的自定义规则, 通过层叠推理机, 可以构造功能完备的推理系统。该类型的推理机在本地一致性检测、推理效率、推理准确性等方面较原有单一的推理机都有所提高。通过层叠推理, 可以实现语义扩展, 关联搜索等智能推理的应用。

5 智能决策辅助系统

智能决策辅助系统的是基于决策智能引擎的知识图谱、战场信息库以及问题库、问题分解结果, 实现的问题交互式问答系统, 目的是提升辅助决策的质量和水平。系统的实现包括三个核心模块的集成: 自然语言理解、查询映射、答案生成, 系统技术框架如图 5 所示。

传统方式战场态势, 可表述为一定时空范围内敌我友各方作战实体名称、属性、状态、行为、相互关系以及战场环境特征的总体形态。战场态势智能感知, 通过多种传感器实现数据采集, 通过多种途径实现信息和情报采集。其中针对信息来源多、格式多样, 战场目标的识别与提取复杂困难, 可考虑采用基于深度学习框架, 分别针对图像、文本、语音、视频情报构建四类学习框架, 实现各类战场目标的自动识别与分类、融合、聚合, 完成对战场初级始态势的感知和描述, 实现对战场态势要素的提取, 构建战场态势状态空间, 框架见图 5。战场态势应包括时间、空间、实体(属性、状态、行为、关系)和战场环境数据, 可表示概念模型为: $B = \{BS|T, S, E(EA, EC, ES, ER), BE\}$, 可支持基于时间、空间、作战目标、作战任务的战场信息汇聚与展示。

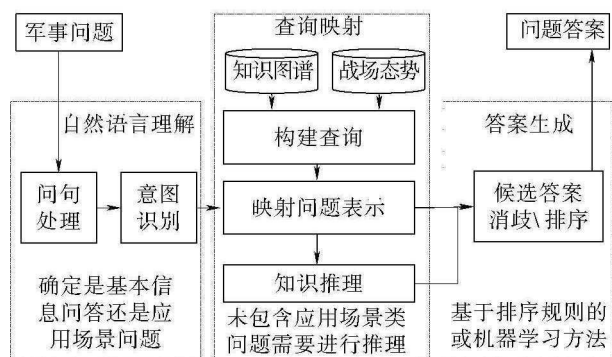


图 5 智能决策辅助系统

自然语言理解模块: 也称为问句分析模块, 采用模板匹配的方法提取问句中的实体等信息词。该步骤也可以采用自然语言处理领域的技术实现, 如中文分词、词性标注、命名实体识别、句法分析等。

查询映射模块: 根据自然语言理解模块提取的问句信息和用户意图将自然语言问题转换为相应的查询, 进行各领域主题数据接口调用及知识图谱调用。

答案生成模块: 候选答案消歧、排序等操作, 可以采用基于排序规则的或者机器学习的方法实现答案的生成。

6 结束语

针对智能决策辅助系统的构建, 着眼军事领域态势研判和辅助决策的需求, 充分运用知识图谱的智能分析技术原理和方法, 系统研究决策智能引擎的框架、原理、特征和要求, 突破相关智能分析与推理通用技术, 重点研究面向战场环境的复杂问题分解技术、面向语义理解的智能决策辅助技术、多源答案语义级融合技术; 研究跨媒体异构数据的知识提取与统一表征技术、多源军事知识融合对齐技术; 研究基于知识驱动的战场信息汇聚、知识推理和机器学习的问题推理技术。可广泛应用于作战指挥, 情报分析, 军事训练等领域^[9]。

参考文献

- [1] Holsapple CW, Jacob V S, Zaveri J S. Learning by problem processors: Adaptive decision support systems [J]. Decision Support Systems, 1993, 10 (2): 85-108.
- [2] Mirchandani D, Pakath R. Four models for a decision support system [J]. Information and Management, 1999, 35 (1): 31-42.
- [3] DBpedia A large-scale, multilingual knowledge base extracted from Wikipedia. Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo Mendes, Sebastian Hellmann, Mohamed Morsey [J]. Semantic Web, 2015 (2).
- [4] DARPA software lets manned/unmanned teams fight despite jamming. Graham Warwick [EB/OL]. <http://aviationweek.com/future-aerospace/darpa-software-lets-mannedunmanned-teams-fight-despite-jam-ming>. 2018.
- [5] 秦川, 祝恒书, 庄福振, 等. 基于知识图谱的推荐系统研究综述 [J]. 中国科学: 信息科学, 2020 (7).
- [6] 邢萌, 杨朝红, 毕建权. 军事领域知识图谱的构建及应用 [J]. 指挥控制与仿真, 2020 (4).

-
- [7] Deepika C, Jaiteg S. Hybrid DL and RL Based Reasoner for Optimum Entailments in Ontologies [J]. Indian Journal of Science and Technology, 2017 (10), 27.
- [8] Jinhyung K, Dongwon J. Ontology-based semantic recommendation system in home network environment [C]. IEEE Transactions on Consumer Electronics, 2009.
- [9] 简实松. 基于知识图谱的问答系统 [D]. 武汉: 华中科技大学, 2019.