

面向油藏地质领域的知识图谱构建研究

文必龙,薛广有

(东北石油大学 计算机与信息技术学院,黑龙江 大庆 163318)

摘要:经过多年的地质勘探,油藏地质领域积累了丰富的勘探成果以及地质知识。为了实现油藏地质领域知识的共享、传播及对知识进行有效的管理,油藏地质领域知识图谱成为了一种必然的选择。针对目前油藏地质领域知识图谱在实际构建中存在的已有本体不能直接构建知识图谱的模式层,抽取得到的多个实体指称项对应于同一个事实对象,难以从抽取出的多个属性值中得到最优属性值等问题,通过对构建领域知识图谱的方法和技术的研究,该文提出了一种以领域内的业务活动为核心的领域本体构建方法,并且依据此方法构建了油藏地质领域本体。改进了传统的孪生循环神经网络模型,解决了油藏地质领域实体对齐问题;针对不同类别的属性设计了不同的属性值融合方法,实现属性值的最优。

关键词:油藏地质领域;领域本体;知识图谱;循环神经网络;实体对齐;属性值融合

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2021)12-0204-07

doi:10.3969/j.issn.1673-629X.2021.12.034

Research on Knowledge Graph Construction in Reservoir Geology

WEN Bi-long, XUE Guang-you

(School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China)

Abstract: After years of geological exploration, rich exploration results and geological knowledge have been accumulated in the field of reservoir geology. In order to realize the sharing, dissemination and effective management of knowledge in the field of reservoir geology, the knowledge map of reservoir geology has become an inevitable choice. At present, the existing ontology in the actual construction of reservoir geological knowledge map cannot directly construct the mode layer of knowledge map, and the extracted multiple entity references correspond to the same fact object, so it is difficult to get the optimal attribute value from the extracted multiple attribute values. The method and technology of domain knowledge mapping are studied, and a domain ontology construction method with business activities as the core is proposed to realize the construction of reservoir geology domain ontology. The traditional twin neural network model is improved to solve the problem of entity alignment in the field of reservoir geology. Different attribute value fusion methods are designed for different types of attributes to achieve the optimal attribute value.

Key words: reservoir geology; domain ontology; knowledge graph; recurrent neural network; entity alignment; attribute value fusion

0 引言

油藏地质领域历经多年的勘探,积累了丰富的勘探研究成果及地质认识。长期以来,这些成果大多分散在档案馆存放的研究报告中、计算机内个人文件夹以及油藏工程师的经验中。使这些分散的、零碎的地质认识,转化为专业的、规范化的油藏知识,并对其进行有效管理,指导后续的勘探工作,让历史勘探研究成果发挥出更大的价值。知识图谱(knowledge graph)作为现阶段知识管理的有效手段,通过它可以方便、快捷地管理知识。因此,建立油藏地质领域知识图谱具有很大的实际意义。

为了实现油藏地质领域知识的共享、传播以及对知识进行有效的管理,油藏地质领域知识图谱成为了一种必然的选择。当前在油藏地质领域知识图谱构建的过程中存在一些不足,具体有以下几个方面。第一,现存的油藏地质领域本体数量不足,不能表示油藏地质领域的核心知识,即使是已经存在的油藏地质领域本体由于在构建时没有考虑到具体应用状况,导致最后得到的本体相差甚远,不能直接用来构建具有实用性的油藏地质领域知识图谱的模式层;第二,由于知识抽取所使用的数据源的多样性,会导致抽取出来的多个指称项对应于现实世界中的同一个实体的歧义,增

收稿日期:2020-12-31

修回日期:2021-05-05

基金项目:中石化科技攻关项目(33550000-20-ZC0613-0098);中石油科技攻关项目(JDYT-2020-JS-50311)

作者简介:文必龙(1967-),男,博士,教授,研究方向为大数据、软件工程等;薛广有(1993-),男,硕士研究生,研究方向为软件工程、软件集成技术。

加了知识规范化的难度;第三,知识在抽取的过程中所用数据源的不同及油藏地质领域实体属性自身的特点,会出现同一个属性抽取多个属性值的情况,并且难以从多个属性值中选取最优的属性值。

针对油藏地质领域知识图谱构建过程中存在的问题,该文对领域知识图谱构建的相关方法和技术展开了研究,在传统的领域本体构建方法的基础上,将实际的应用场景和油藏地质领域知识的特点等因素考虑进去,提出一种以领域业务活动为核心的领域本体构建方法;对传统的孪生循环神经网络模型进行了改进,使其适合油藏地质领域实体的对齐任务;针对不同类别的属性设计了不同的属性值融合方法,实现属性值的最优。

1 相关研究

Google 公司为了提高搜索引擎的检索能力,使其能够返回更加准确的检索结果,以此来达到改善用户使用体验的目的。在此种背景下,知识图谱的概念^[1]于 2012 年 5 月 17 日被 Google 公司率先提出。经过长期的发展,知识图谱的概念也发生了改变。目前,人们广泛接受的知识图谱的概念如下:“知识图谱 (Knowledge Graph) 是以图形的方式来描述真实世界中的各种概念以及它们之间关系的知识库^[2]”。知识图谱本质上是一种刻画实体之间关系的语义网络^[3],能够形式化地描述现实世界中的实体及其相互关系。现在的知识图谱已被用来泛指各种大规模的知识库。

目前国外的大型知识图谱主要有谷歌公司的 Knowledge Graph 和微软公司的 Satori。国内的中文知识图谱主要有搜狗的“知立方”、百度的“知心”、王元卓等人^[4]研发的可演化的 Open KN 以及上海复旦大学的中文知识图谱等^[5]。此外,人际知识图谱的应用较早,国外有 Facebook 公司以人际关系数据为基础建立的实体搜索服务,国内有杜亚军等人^[6]以微博数据为基础建立的微博社区知识图谱。

在行业知识图谱的研究上,周蓝璐^[7]采用各种方法从新浪网的海量数据中抽取了实体的关系,并实现了一个面向中文音乐领域的关系抽取平台。Patrick Ernst 等人^[8]将从医学网站中抽取出的数据进行融合,研发了一个面向医学领域的知识图谱系统。在通用知识图谱的研究上,蔡鸿博^[9]利用信息抽取技术从百科类网站中抽取了实体、关系等数据,建立了包含 75 万个实体和 900 万条关系的中文知识图谱。胡芳槐^[10]研究了传统的知识库和在线百科类网站,提出了一种能够从不同数据源中抽取出的数据进行高度融合的知识图谱构建方法。基于此方法构建了大规模的中文知识图谱,其中包含了 700 多万个实体以及 6 000 多万个句子。

在知识图谱的构建研究上,邓小亚^[11]提出了石油勘探开发领域本体的构建方法,并以采油工程为例说明此方法的可行性。杨宇飞^[12]实现了使用条件随机场模型从互动百科中抽取实体的属性关系。庄严等^[13]对知识库的实体对齐技术进行了总体概述。谢颖^[14]通过对信息融合的方法进行研究,提出可以利用数学中的相关工具来解决知识图谱的数据融合问题。

在知识图谱的应用领域,知识图谱已被广泛应用于语义搜索、智能问答、个性化推荐等领域。Zhao Shi 等人^[15]以用户的搜索日志为基础数据,采用无监督的机器学习算法获得数据的基本特征,并对其基本特征进行了综合分析以达到理解用户查询意图的目的。Bahareh Sarrafzadeh 等人^[16]提出了一个适用于知识图谱的信息检索框架,在该框架的帮助下用户能够在不同领域知识图谱中进行复杂的信息检索。

2 油藏地质领域本体构建

该文在传统的七步法基础上,结合油藏地质领域实际的业务需求将本体构建划分为:规划与设计、领域本体构建、评价与进化三个阶段。具体流程如图 1 所示。

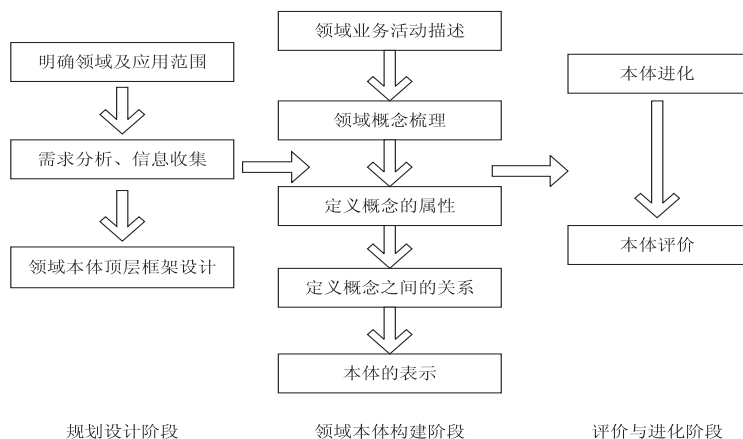


图 1 领域本体构建流程

2.1 规划与设计阶段

规划与设计阶段的主要任务是明确具体的领域以及应用范围、进行需求分析和信息收集、实现领域本体的顶层框架设计。

根据油藏地质领域的业务模型和数据模型所提供

的概念结合实际的业务需求,该文将油藏地质领域本体划分为活动(activity)、对象(object)、元信息(meta_class)三个一级本体,其中每个一级本体可以继续向下划分子本体。具体结构如图2所示,图2在一级本体的基础上,向下衍生了一层。

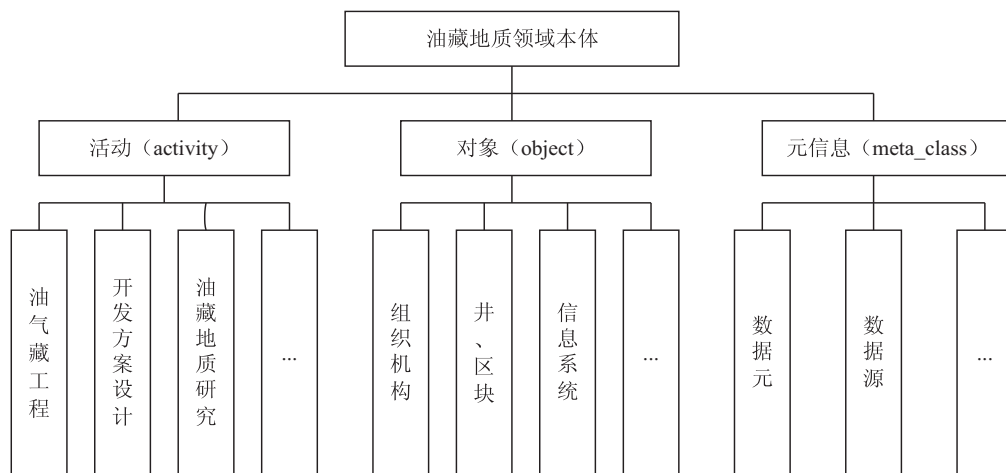


图2 油藏地质领域本体结构

活动(activity)表示一种能够被人们感知的可以让物理世界发生变化的行为。对于某个活动,为了知道活动发生的时间、地点、人物等要素,建立起活动发生时的具体场景,需要对活动进行标准化的描述。通过对领域中多个业务活动进行标准化描述,就可以构建起领域业务活动的流程。油藏开发评价、地质油藏、剩余油价值研究等都是活动的子类。

对象(object)表示客观世界中真实存在并可被人们感知到的任何事物。井、区块、油气藏、组织机构、油田各种资源等都是对象的子类。

元信息(meta_class)表示描述信息的信息。在某些情况下,借助元信息来描述的内容会比使用本体概念形式化地描述该内容更加简洁。数据元、数据库等都是元的子类。

2.2 领域本体构建阶段

领域本体构建阶段的主要任务是描述领域内的核心业务活动,梳理领域内的核心概念,定义概念的属性和相互关系,进行本体表示。

2.2.1 业务活动描述及领域概念梳理

首先,根据业务流程对业务进行逐级的划分,直至业务活动为止。业务活动即不可再分的最小业务单元。然后,从六个方面对业务活动进行标准化的描述,即活动是由何人(Who)何时(When)何处(Where)发起的、发起活动的原因(Why)、涉及到哪些(Which)对象、产生了什么结果(What)。最后,某个业务活动中所包含的相关信息可以通过上述几个方面的描述完全提取出来。

2.2.2 定义概念的属性

根据实际的业务需求和数据需求,将油藏地质领域本体划分为 activity、object、meta_class 三个一级类,分别定义每个一级类及子类的属性,其子类可直接继承父类的属性。

activity 类的属性:活动代码、活动原因、活动时间、活动地点、操作者、操作对象、产生结果、描述。

油藏地质领域的一级对象(object)类所包含的子类种类较多,而每个子类所含的属性不同,因此分别定义每个子类的属性。具体如下:井、区块、油气藏、盆地的属性:名称、代码、位置、类型、条件;组织机构的属性:名称、职能、岗位;信息系统的属性:名称、版本、供应商、管理部门、访问网址、部署时间;仪器设备的属性:名称、用途、编号、所属部门。

meta_class 类的属性:名称、代码、类型、长度、精度、计量单位。

2.2.3 定义概念的关系

根据油藏地质领域的业务模型和数据模型以及石油领域的相关书籍和字典,结合实际的需求,将油藏地质领域本体概念间的关系定义为通用关系和自定义关系两类。

(1) 通用关系。

一般而言,领域本体都包含 Kind-of、Part-of、Attribute-of、Instance-of 四种基本关系,在此基础上该文定义了五种基本的概念关系,见表1。

(2) 自定义关系。

自定义关系是油藏地质领域中为了满足实际的业务需求而建立的概念间的特殊关系,见表2。

表 1 油藏地质领域概念间的基本关系

| 关系名称 | 表示形式 | 解释说明 |
|---------|---------------|-----------------|
| 整体与部分关系 | Part-of | 表示概念之间是整体与部分的关系 |
| 实例关系 | Instance-of | 表示一个概念是另一个概念的实例 |
| 属性关系 | Attribute-of | 表示一个概念是另一个概念的属性 |
| 继承关系 | Kind-of | 表示概念之间的层次继承关系 |
| 等同关系 | Equivalent-of | 表示概念之间的同义关系 |

表 2 油藏地质领域概念间的自定义关系

| 关系名称 | 表示形式 | 解释说明 |
|--------|----------------|-------------------------------------|
| 所属关系 | Belong-of | 表示井、站、库等概念所属管辖单位或范围 |
| 关联关系 | Relation-of | 表示概念之间有一定的相关性,比如可采储量与油藏类型、储层物性等因素有关 |
| 顺序关系 | Sequence-of | 表示事件的发生有一定的先后顺序 |
| 步骤关系 | Sub process-of | 表示某个事件是另一个事件的步骤之一 |
| 依赖关系 | Affect-of | 表示某个概念依赖于另一个概念 |
| 组成部件关系 | Component-of | 表示业务活动过程中所用设备的组成部件 |

2.2.4 油藏地质领域本体构建

使用 protege 本体编辑工具构建油藏地质领域本体,将预先定义好的概念、关系、属性等利用软件进行

编辑,采用 owl 格式保存文件。部分本体的可视化效果如图 3 所示。

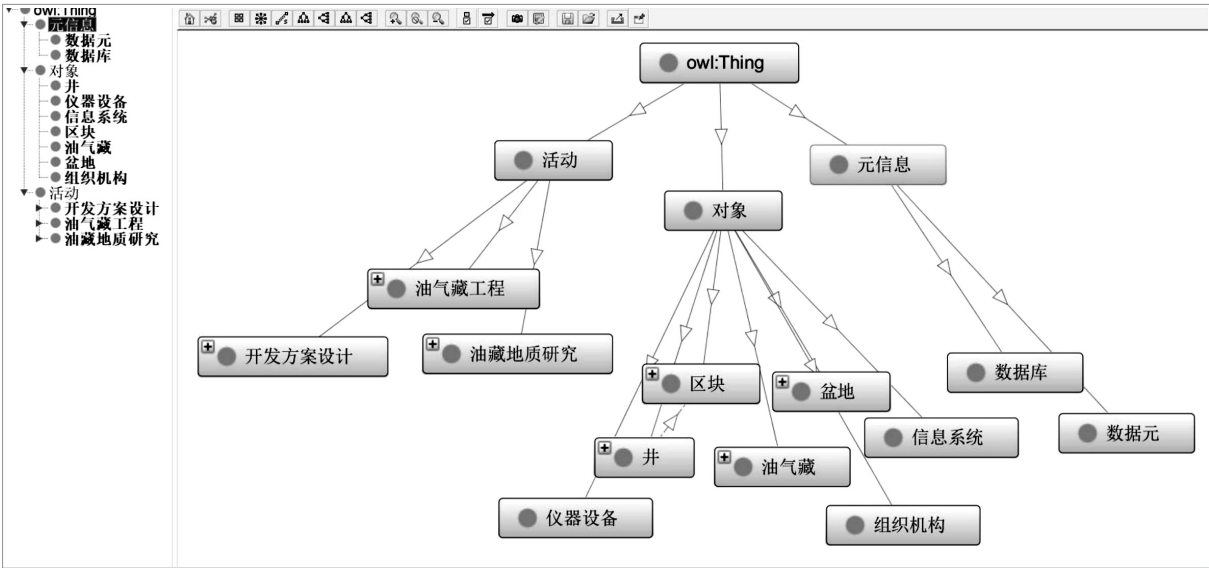


图 3 本体可视化效果图

2.3 评价与进化阶段

评价与进化阶段的主要任务就是针对所构建的本体进行质量评价以及对本体进行不断的完善与进化。本体构建的质量决定了知识图谱的质量,本体自身的缺陷会直接导致知识图谱在检索、推理等应用方面使用能力的下降,从而失去了知识图谱构建的意义。因此在油藏地质领域本体构建的过程中,应该及时地对所构建的本体进行评价,发现其不足。此外,本体的构建是一个长期的、反复的过程,需要不断地对本体进行维护和完善。在初步建立油藏地质领域的本体后,

需要将本体应用于实际并反复的进行评价、优化和扩展。这个过程应该始终存在于本体的生命周期之中。

3 油藏地质领域实体对齐方法

通过信息抽取技术从不同数据源中抽取出的数据存在着大量的冗余、重复、错误等问题。为了保证知识图谱数据层的质量,需要将抽取出的数据利用实体对齐技术进行处理。一些传统的实体对齐算法都存在着过度依赖属性域的问题,并且它们都忽略了文本的语义特征。为了更好地利用文本中上下文的语义特征,

该文以孪生循环神经网络为基础,对其进行改进,最终实现基于文本语义的相似度计算,解决油藏地质领域的实体对齐问题。

孪生神经网络是由两个 bi LSTM 组成并且这两个神经网络彼此共享权重。它的基本原理如下:首先将待对齐的实体当作字符串输入到 Embedding

层,Embedding 层会将其映射为向量;然后 bi LSTM 模型会将输入的向量编码为固定的长度,选择 bi LSTM 模型最后一次的输出结果作为最终的向量;最后计算两者的余弦相似度并与阈值进行比较,如果相似度超过阈值,就判定为匹配,反之,则不匹配。该模型的具体结构如图 4 所示。

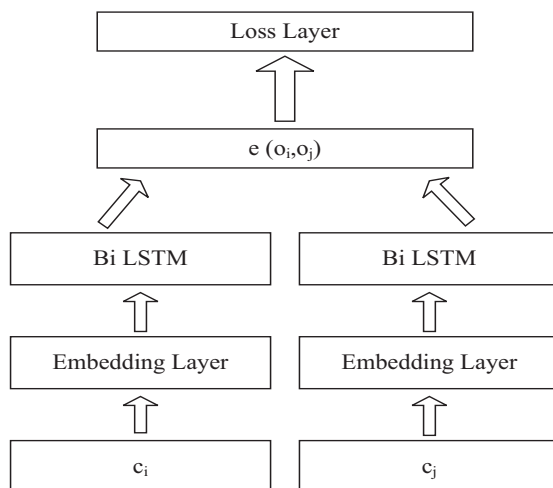


图 4 孪生神经网络结构

模型训练的详细步骤如下:

样本构造阶段。首先根据业务活动将抽取得到的数据进行分类,得到业务活动字典 $D = \{\text{activity1}:[c_1, c_2, \dots], \text{activity2}:[c_3, c_4], \dots\}$, 那么每个 activity 下的数据都指向相同的业务活动;然后将同一个 activity 下的数据两两相互配对得到正样本集合 P , 将不同 activity 下的数据两两相互配对得到负样本集合 N 。

经过处理之后,数据集中共包含 6 378 条样本,其中有 4 658 条正样本和 1 720 条负样本。训练样本由“记录对+标记”构成,例如:((record1, record2, 1), (record3, record4, 0), ...)。测试样本由“记录对”构成,例如:((record1, record2), (record3, record4), ...)。表 3 展示了构造出的部分训练样本。

表 3 训练样本示例

| 名称 | 操作对象 | 描述 | 名称 | 操作对象 | 描述 |
|---------|-------|----------------------------|---------|----------|----------------------------|
| 生储盖组合研究 | 储层、盖层 | 分析生、储、盖层分布规律,判断生、储、盖层的组合类型 | 生储盖特征分析 | 生层、储层、盖层 | 分析生、储、盖层分布规律,判断生、储、盖层的组合类型 |
| 储量计算 | 油藏 | 采用容积法计算油藏储量 | 油藏储量计算 | 油藏 | 计算油藏的储量 |
| 井号 | 井 | 由主管部门统一发布 | 井编号 | 井 | 由主管部门统一发布 |

数据预处理阶段。以字符为粒度对抽取得到的数据进行切分,去除重复无效的字符,得到新的字符序列。将不定长的字符序列通过设置 max_sequence_length 的值转化为定长的字符序列,再使用字符串“UNK”填充长度不能达到固定长度的字符序列。文中参数 max_sequence_length 的值设置为 60。

One-hot 词向量表示阶段。将数据预处理阶段得到的单词表示成 One-hot 向量的形式。由于在 LSTM 模型之前添加了嵌入层(Embedding Layer),所以只需要用每个词在词表中的索引号来替换这个单词即可,并不需要真的将每个单词都表示成 One-hot 向量的形

式。例如:{"断","层","闭","合","解","释"}数据集词表,则用 1 替换“断”,用 3 替换“闭”,用 0 替换填充字符串“UNK”即可。

训练样本批量构造阶段。模型训练的次数由参数 epochs 决定。epoch 即 batch 训练的次数。参数 batch_size 决定每批训练样本的数量。模型采用 mini-batch 的方式进行训练。例如:样本总量为 512, batch_size 的值为 64,则模型需要进行 8 次 batch 训练。这 8 次的 batch 训练称之为一次 epoch。文中 batch_size 的值设置为 128, epochs 的值设为 160。

模型训练阶段。首先将训练样本 C_i 和 C_j 输入到

共享权值的嵌入层得到新的词向量 X_i 和 X_j , 然后将 X_i 和 X_j 输入到对称的 LSTM 模型中进行训练, 取最后一次的计算结果作为输出, 得到 O_i 和 O_j 。将 O_i 和 O_j 向量看作为样本 C_i 和 C_j 新的特征表示。最后通过计算 O_i 和 O_j 的余弦相似度来得到样本 C_i 和 C_j 的相似度分值, 并与预先设定的阈值进行比较, 如果相似度分值大于阈值, 则判定匹配, 否则, 则不匹配。文中的阈值设置为 0.6, 并且使用交叉熵作为损失函数来进行模型参数的训练, 以及使用梯度下降进行参数优化。

4 油藏地质领域属性值融合方法

属性值融合的主要任务是从多个数据源中抽取得到的不同属性值进行融合, 得到最优的属性值结果。该文根据油藏地质领域自身的数据特点将基本属性划分成两大类, 具体如下:

单值属性: 在抽取出的多个候选属性值中只能有一个是正确的属性值的属性。例如: 在抽取“井”这个实体的“代码”属性时, 得到了“J00006”和“Y00008”这两个属性值。但由于“代码”是单值属性, 只能有一个属性值是正确的, 需要经过筛选得到最准确的结果。

表 4 三元组抽取来源详细信息

| 编号 | 来源 | 值 | 权重 |
|----|------------------|--------|----|
| 1 | 基于总部文件得到的三元组 | CN_PC | 10 |
| 2 | 基于勘探院文件得到的三元组 | JD_EXP | 8 |
| 3 | 基于开发院文件得到的三元组 | JD_DEV | 8 |
| 4 | 基于各个二级单位文件得到的三元组 | JD_SUB | 5 |

单值属性只能存在一个正确的属性值。因此, 需要在多个候选属性值中选出最准确的属性值作为单值属性的属性值。该文采用计算属性值信任度的方法来确定候选属性值的准确性, 候选属性值的信任度越大说明其准确性越高, 越有可能成为最终的属性值。信任度 con 的计算公式: $con = W_1 \times L + W_2 \times N + W_3 \times S$ 。参数的具体含义如下: L 表示属性值的长度, L 值越大说明候选属性值越详细, 它成为正确属性值的可能性就越大; N 表示属性值被抽取的次数, N 值越大说明候选属性值被抽取的次数越多, 它越有可能成为正确的属性值; S 表示属性值抽取来源的权重, S 值越大说明抽取属性值的文档越可信, 属性值的准确性就会越高; W_1 、 W_2 、 W_3 分别表示 L 、 N 、 S 在信任度计算公式中所占的权重, 是自定义的参数。文中参数设置为 $W_1 = 25\%$ 、 $W_2 = 25\%$ 、 $W_3 = 50\%$ 。对于某个实体的单值属性而言, 如果抽取多个候选属性值, 则需要通过信任度计算, 取信任度最大的候选属性值作为该属性的属性值。

多值属性: 在抽取出的多个候选属性值中存在多个正确的属性值的属性。例如: 在抽取“油气成藏综合研究”活动类实体的“产生结果”属性时, 得到了“成藏时间、成藏动力、成藏条件”三个正确的属性值, 则“油气成藏综合研究”实体的“产生结果”属性含有三个属性值。

该文将定义的基本属性依据油藏地质领域属性值的特点划分为单值属性和多值属性两类。单值属性包含: 活动时间、活动地点、活动原因、活动代码、名称、代码、位置、类型、条件、版本、供应商、管理部门、访问网址、部署时间、用途、编号、所属部门、长度、精度、计量单位; 多值属性包含: 操作者、操作对象、产生结果、描述、职能、岗位。

为实现信任度计算, 该文采用四元组 <实体, 属性, 属性值, 抽取来源>来存储信息。例如: 从勘探院提供的文件中抽取得到的三元组 <沉积构造特征描述, 操作对象, 沉积构造>可扩展为四元组 <沉积构造特征描述, 操作对象, 沉积构造, JD_EXP>。三元组抽取来源的详细信息如表 4 所示。

例如:

<井, 代码, J00006, CN_PC> con = 5.75
<井, 代码, J00006, JD_EXP> con = 4.75
<井, 代码, Y00008, JD_DEV> con = 4.50
<井, 代码, Z00006, JD_SUB> con = 4.50

如上例所示, “井”实体的“代码”属性经过抽取后得到了 4 个候选属性值。通过对每个属性值的信任度计算, 取“J00006”作为“井”实体的“代码”属性的最终属性值, 得到三元组 <井, 代码, J00006>。

多值属性可以存在多个正确的属性值。因此, 多值属性的属性值融合只需将多个正确的候选属性值去重后直接合并即可。针对多值属性的特点, 该文采用将抽取出的多个候选属性值按照抽取来源的权重进行排序、删除重复的属性值后直接合并的方法得到多值属性的属性值。由于展示内容有限, 不能将所有的属性值都保留。因此, 该文规定多值属性的属性值最多为 5 个。

例如:

<油气成藏综合研究,产生结果,成藏时间,CN_PC>

<油气成藏综合研究,产生结果,成藏动力,CN_EXP>

<油气成藏综合研究,产生结果,成藏条件,CN_DEV>

<油气成藏综合研究,产生结果,成藏时间,JD_SUB>

如上例所示,四元组经过排序、去重等步骤得到了“成藏时间、成藏动力、成藏条件”三个属性值。由于满足预先设定的阈值,所以这三个属性值应该全部被保留。最后,得到三元组<油气成藏综合研究,产生结果,成藏时间、成藏动力、成藏条件>。

5 结束语

该文以油藏地质领域知识的共享、传播以及对知识进行有效的管理为出发点,针对现阶段油藏地质领域知识图谱构建过程中存在的问题,对领域知识图谱构建的方法和技术进行研究;提出了以领域业务活动为核心的领域本体构建方法,实现了油藏地质领域本体的构建,完成了油藏地质领域知识图谱模式层的构建;改进了传统的孪生循环神经网络模型,使其能够满足油藏地质领域的实体对齐任务;针对不同类别的属性设计了不同的属性值融合方法,实现属性值的最优,完成了油藏地质领域知识图谱数据层的构建。最终,实现了高质量油藏地质领域知识图谱的构建。

参考文献:

- [1] AMIT S. Introducing the knowledge graph[R]. America: Official Blog of Google, 2012.
- [2] 黄恒琪,于娟,廖晓,等. 知识图谱研究综述[J]. 计算机系统应用, 2019, 28(6): 1-12.
- [3] BIZER C, HEATH T, BERNERS-LEE T. Linked data: the story so far[J]. International Journal on Semantic Web and Information Systems, 2009, 5(3): 1-22.
- [4] 王元卓,贾岩涛,赵泽亚,等. Open KN-网络大数据时代的知识计算引擎[J]. 中国计算机学会通讯, 2014, 10(11): 30-35.
- [5] 王元卓,贾岩涛,刘大伟,等. 基于开放网络知识的信息检索与数据挖掘[J]. 计算机研究与发展, 2015, 52(2): 456-474.
- [6] 杜亚军,吴越. 微博知识图谱构建方法研究[J]. 西华大学学报: 自然科学版, 2015, 34(1): 27-35.
- [7] 周蓝珺. 音乐领域中文实体关系抽取研究[D]. 哈尔滨: 哈尔滨工业大学, 2009.
- [8] ERNST P, SIU A, WEIKUM G. Know life: a versatile approach for constructing a large knowledge graph for biomedical sciences[J]. BMC Bioinformatics, 2015, 16: 157.
- [9] 蔡鸿博. 基于上位词的中文实体关系图谱构建[D]. 哈尔滨: 哈尔滨工业大学, 2014.
- [10] 胡芳槐. 基于多种数据源的中文知识图谱构建方法研究[D]. 上海: 华东理工大学, 2015.
- [11] 邓小亚. 计算机本体理论在石油勘探开发知识管理系统中的应用研究[J]. 计算机系统应用, 2013, 22(9): 91-96.
- [12] 杨宇飞. 面向中文网络百科的属性关系抽取研究[D]. 成都: 西南交通大学, 2014.
- [13] 庄严, 李国良, 冯建华. 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.
- [14] 谢颖. 信息融合算法研究及其应用[D]. 重庆: 重庆大学, 2008.
- [15] ZHAO Shi, ZHANG Yan. Tailor knowledge graph for query understanding: linking intent topics by propagation[J]. EMNLP, 2014, 6(2): 1070-1080.
- [16] SARRAFZADEH B, VECHTOMOVA O, JOKIC V. Exploring knowledge graphs for exploratory search[M]. New York: ACM, 2014: 135-144.
- [10] ZHANG Zhiyong, CHEN Mingxu. Advantages and disadvantages of third party payment method and traditional payment method[C]//Proceedings of the 2019 10th international conference on e-business, management and economics. New York: Association for Computing Machinery, 2019: 195-198.
- [11] 徐锋,叶镒娟. 移动互联背景下校园卡支付体系设计与实践[J]. 浙江大学学报: 理学版, 2018, 45(1): 60-64.
- [12] VASHISTHA A, ANDERSON R, MARE S. Examining the use and non-use of mobile payment systems for merchant payments in India[C]//Proceedings of the 2nd ACM SIG-CAS conference on computing and sustainable societies. Accra, Ghana: Association for Computing Machinery, 2019: 1-12.
- [13] 徐海勇,刘虹,张琳,等. 二维码扫码聚合支付模式研究[J]. 中国新通信, 2018, 20(3): 58-59.
- [14] SHAFIN M, HANSDA R, PALLAVI E. Partial face recognition: a survey[C]//Proceedings of the third international conference on advanced informatics for computing research. New York: Association for Computing Machinery, 2019: 1-6.
- [15] ALSKEINI N H, THANH K N, CHANDRAN V. Face recognition: sparse representation vs. deep learning[C]//Proceedings of the 2nd international conference on graphics and signal processing. New York: Association for Computing Machinery, 2018: 31-37.

(上接第 203 页)

融科技时代, 2017(10): 59-62.