

许多,鲁旺平,许瑞清,等.基于农业时空多模态知识图谱的水稻精准施肥决策方法[J].华中农业大学学报,2023,42(3):281-292.
DOI:10.13300/j.cnki.hnlkxb.2023.03.032

基于农业时空多模态知识图谱的 水稻精准施肥决策方法

许多^{1,2},鲁旺平^{1,2},许瑞清³,张红雨³,江洋^{1,4},游良志^{1,2,5},冯在文^{2,3,4}

1. 华中农业大学植物科学技术学院,武汉 430070; 2. 华中农业大学宏观农业研究院,武汉 430070;
3. 华中农业大学信息学院,武汉 430070; 4. 华中农业大学双水双绿研究院,武汉 430070;
5. 国际食物政策研究所,美国华盛顿 20005

摘要 为构建基于农业时空大数据的管理系统,实现田间养分精细化管理,提出了一种面向精准施肥的农业时空多模态知识图谱的构建及其控制与决策方法。通过基于深度学习的子图匹配方法,将地块待查询图和农业时空多模态知识图谱中的节点和关系嵌入表示;利用向量相似度计算获取候选子图,并从存储历史数据信息的子图中获取适合查询地块的施肥模型数据。结果显示,基于实例化后的待施肥地块查询图,在农业时空多模态知识图谱中可获取与给定地块查询图同构的子图,并从存储历史决策信息的子图中,获得适合当前地块的农业施肥模型。结果表明,基于农业时空多模态知识图谱的农业模型自动化选择结果可为精准施肥任务提供新思路 and 决策支撑。

关键词 时空;多模态知识图谱;智慧农业;精准施肥;水稻;农业大数据;知识共享

中图分类号 S511; F49; TP311.13 **文献标识码** A **文章编号** 1000-2421(2023)03-0281-12

农业产业正面临多重挑战,既要增加产量以满足不断增长的人口需要,又要解决气候不断变化带来的产能下降问题,在技术改革的同时还要减少对环境的影响。以大数据技术为支撑的精准农业被视为应对此类挑战的解决方案^[1]。在精准农业的发展中,存在的问题之一是整合、解释来自不同设备和数据源的异质数据^[2]。“天空地”农业信息采集系统在农业信息获取中的广泛应用,为施肥决策、农田灌溉、病虫害防控提供了多元化的数据。但获取到的农业数据存在标准不尽相同、数据冗杂、命名混乱等问题,使得农业数据呈现出异质、异构、多模态的特点^[3]。

精准施肥是精准农业的重要子任务,传统的精准施肥决策基于施肥模型实施^[4]。而进行大田实践时,存在施肥模型选择困难、模型应用效率低、农艺数据收集困难等问题;由于无法获得栽培领域专家的指导,难以真正使用农业施肥模型。因此,选择合

适的农业施肥模型、整合多源异构的农业大数据,建立统一的农业时空数据管理系统,是精准农业当前亟需解决的问题。

知识图谱为上述问题提供了解决方案。将基于多模态农业知识图谱的知识库管理技术引入精准农业,可实现农业种植管理的信息化、智能化和标准化。近年来,农业知识库的构建多有报道:联合国粮食及农业组织(Food and Agriculture Organization of the United Nations,FAO)主导创建的农业顶级本体AGROVOC^[5]涉及25个农业顶层概念,是涵盖作物、林业、渔业等相关子领域的多语种结构化知识库^[5-6]。法国的作物规划和生产过程本体(crop planning and production process ontology,C3PO)是面向国外农场条件的领域本体^[7]。农业知识图谱(AgriKG)可自动整合海量互联网农业数据,在农业命名实体识别、智能问答等任务中均有较好表现^[8]。尽管知识图谱在病虫害识别^[9-10]、生物育种^[11]、施肥

收稿日期:2023-01-13

基金项目:内蒙古自治区科技重大专项(2021SZZD0099);作物遗传改良全国重点实验室开放基金项目(ZK202203);湖北洪山实验室重大项目(2022HSZD031);武汉大学杂交水稻国家重点实验室开放基金项目(SKLHR202101)

许多,E-mail:Duo.Xu@webmail.hzau.edu.cn

通信作者:冯在文,E-mail:Zaiwen.Feng@mail.hzau.edu.cn

推荐^[12]等农业领域已经取得了一些成果,但目前还没有专门面向精准施肥的农业大数据时空多模态知识图谱的构建方法。

本研究从时空要素及时空关系入手,提出构建农业时空多模态知识图谱的方法,将离散性、多源异构的农业数据和事实信息转化为聚合性的知识。在此基础上,以水稻精准施肥为例,并结合现有的农业施肥模型,提出1种基于农业时空多模态知识图谱的施肥模型自动化选择方法,以整合历史数据信息,针对地块特征选择合适的农业施肥模型,并用以农业施肥决策,旨在为面向精准施肥的农业大数据标准化整合与辅助决策提供新思路与解决方案。

1 材料与方法

基于农业时空多模态知识图谱的精准施肥流程如图1所示,该流程由3个环节组成。

第一,数据采集。基于天空地一体化技术获取土壤属性、作物长势、水土营养等农业多模态数据。第二,基于农业时空多模态知识图谱的精准施肥的数据管理和决策分析。采用知识工程的方法,结合领域专家知识,构建细粒度精准施肥时空本体;进一步本体实例化,获得农业时空多模态知识图谱,通过子图匹配的方法,选择合适的农业施肥模型。第三,精准施肥的决策应用。输出农业时空多模态知识图谱的决策结果——农业施肥模型,用于指导施肥肥料配比。各环节依次进行,环环相扣,从而实现农业时空多模态知识图谱的构建与决策应用。

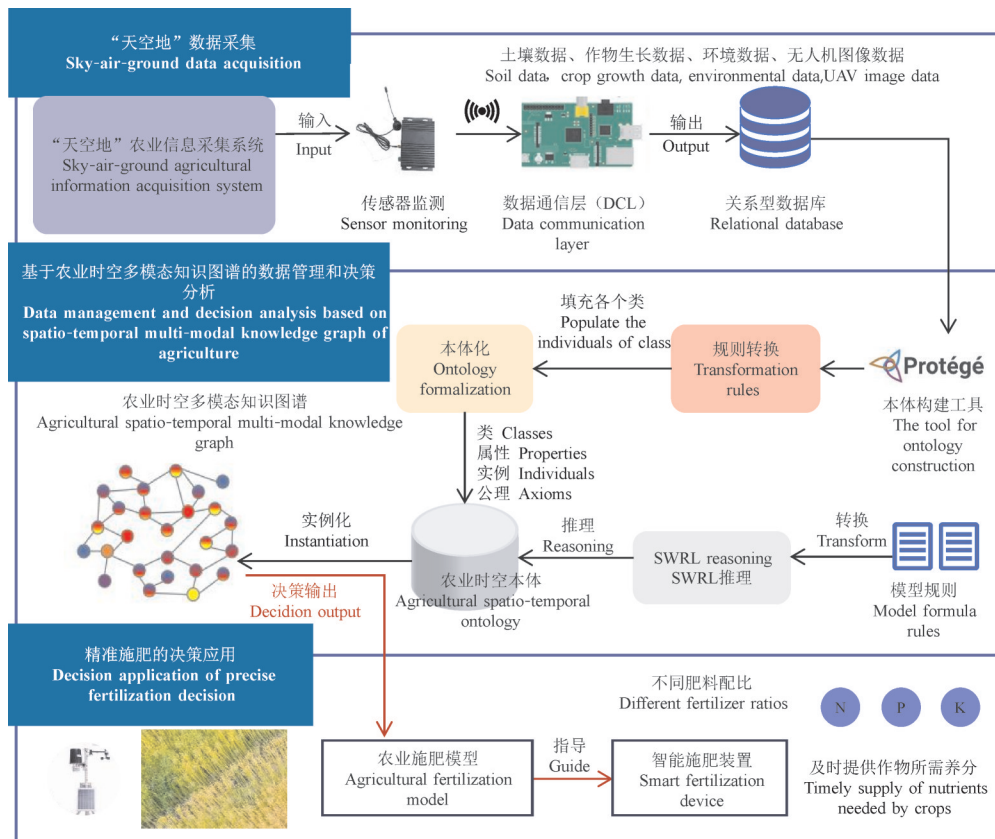


图1 基于农业时空多模态知识图谱的精准施肥决策流程

Fig. 1 Precise fertilization decision-making process based on agricultural spatio-temporal multi-modal knowledge graph

1.1 “天空地”农业信息采集

田间信息的采集是实现精准施肥的第一步。高密度、高速度、高精度的信息采集技术是实现精准施肥的基础^[13]。“天空地”数字化的信息采集是现代信息技术与农业深度融合形成的新型农业体系。建立

“天空地”农业信息采集系统,可实现高精度、立体化的农业时空信息获取^[14]。目前,“天空地”农业信息采集系统包括基于全球定位系统(global positioning system, GPS)定位的定期监测、农业遥感技术、地面物联网传感器自动监测等^[15]。本研究利用“天空地”

农业信息采集系统获取农田地块实时数据,并存储于关系型数据库中。

1.2 数据管理和决策分析

“天空地”信息管理和决策应用贯穿整个精准施肥应用场景。决策系统负责对采集的多模态农业数据进行合理分析和整合,并根据这些信息做出决策。通过时空多模态知识图谱的形式整合采集获得农田数据。

农业时空多模态知识图谱是一个包含大量农业领域事实知识的语义网络,使用描述农业实体及其属性值的节点和节点间的关系链接来表示农业知识信息。农业场景中的数据由不同设备采集获得,包含了表格、文本、图像等多模态信息,并呈现出多源异构的特性。

传统的知识图谱被定义为包含实体节点集合 E 、属性值节点集合 V 、实体关系链接集合 R 以及属性链接集合 A 的有向图 $G = \{E, R, A, V\}$ 。不同于以单一文本符号表示的传统知识图谱,本研究提出的农业时空多模态知识图谱中的节点信息由更为丰富的时空多模态知识表现。由此,可将农业时空多模态

知识图谱定义为由多模态实体节点集合 E_{Multi} 、多模态属性值节点集合 V_{Multi} 、多模态实体关系链接集合 R_{Multi} 以及多模态属性链接集合 A_{Multi} 表示的多模态有向图 $G_{Multi} = \{E_{Multi}, R_{Multi}, A_{Multi}, V_{Multi}, T_R, T_A\}$ 。其中, E_{Multi} 和 V_{Multi} 中分别包含了农业时间和空间多模态信息。 $T_R = E_{Multi} \times R_{Multi} \times E_{Multi}$ 为农业时空多模态知识图谱中的关系三元组集合,该集合由多模态头实体 $s \in E_{Multi}$ 、多模态尾实体 $o \in E_{Multi}$ 以及两实体间的多模态关系链接 $p \in R_{Multi}$ 表示的多模态三元组 $(s, p, o) \in T_R$ 组成。类似地, $T_A = E_{Multi} \times A_{Multi} \times V_{Multi}$ 为由多模态实体 $s \in E_{Multi}$ 、多模态属性值 $o \in V_{Multi}$ 以及实体与属性值间的多模态属性链接 $p \in A_{Multi}$ 表示的三元组 $(s, p, o) \in T_A$ 组成的多模态属性三元组集合。

由于传统的以单一文本符号表示的知识图谱无法满足对农业时空多模态知识信息的表示和应用。因此,本研究提出农业时空多模态知识图谱信息管理框架,通过标准化整合农业时空多模态信息,完成精准施肥大数据的决策应用。基于农业时空多模态知识图谱的数据管理和决策分析平台框架如图2所示。

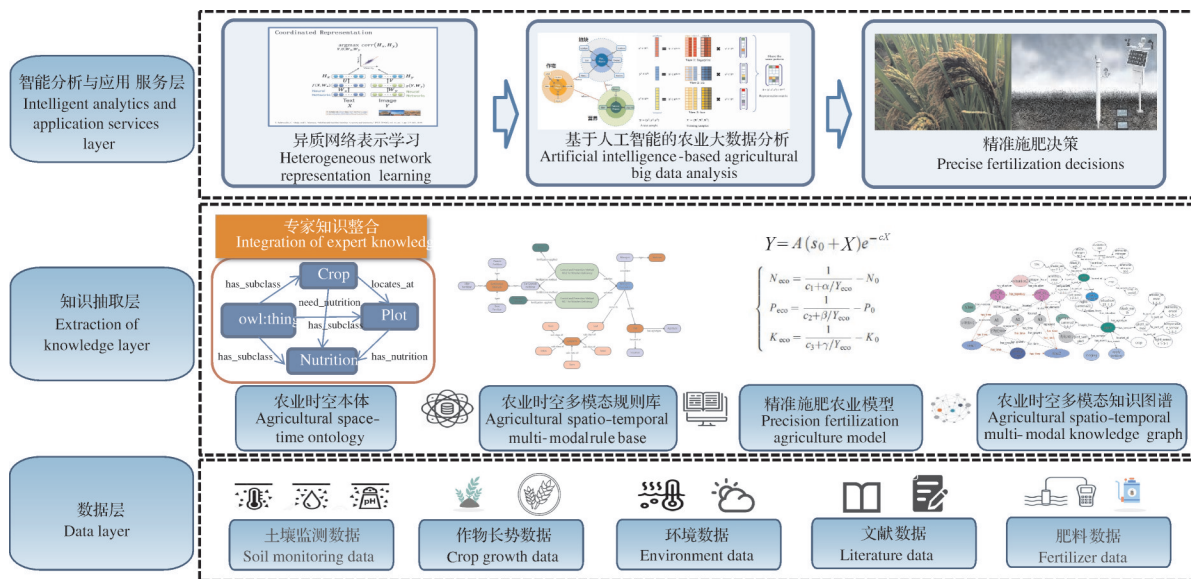


图2 基于农业时空多模态知识图谱的数据管理和决策分析框架

Fig.2 An framework for data management and decision analysis based on agricultural spatio-temporal multi-modal knowledge graph

该框架分为3层:数据层、知识抽取层和智能分析与应用服务层。数据层囊括了构建农业时空多模态知识图谱的所有数据来源,其中包括通过“天空地”信息采集系统获取的土壤、作物、环境等方面的农田大数据和在书籍、科技论文、研究报告中与作物生产相关的农业数据。对于数据层中的多源异构农

业时空多模态数据信息,在知识抽取层中通过知识建模与知识抽取方法,基于精准施肥本体对其进行标准化整合,并结合农业时空多模态规则库和精准农业施肥模型完成面向精准施肥的农业时空多模态知识图谱的构建。在智能分析与应用服务层的农业决策应用场景中,结合现有的农业模型及知识推理

方法最终实现农业精准施肥任务中的决策推理目标。

1.3 农业时空多模态知识图谱的构建与决策

农业时空多模态知识图谱的构建与决策环节包括4个步骤:农业时空知识建模、农业时空知识抽取、农业时空知识存储以及农业时空知识推理。农业精准施肥决策推理任务通过在农业时空多模态知识图谱上执行基于深度学习的图查询(图匹配)操作来实现。具体来讲,首先定义农业模型中施肥量、作物产量、施肥影响因素、地块理化性质间的关系,以抽象得到农业施肥模型查询模板;进而基于试验地块实时监测数据,将查询模板实例化以获得农业模型查询图Q;最后结合基于TransE的候选子图查找方法和基于深度学习的整图表示的图匹配相似度计算方法,以支持获取农业精准施肥的最终决策。

1)知识建模。知识建模是知识逻辑化、体系化的过程^[16],其本质是建立结构严谨、语义信息完善的本体(Ontology)模型,以全面、准确地表示领域知识,完成知识推理任务^[17]。将本体论用于大量多源、异构的农业数据和知识整合、转换,是将多种类型的农业信息转化成计算机可理解可计算资源的有效解决方案。农业知识建模是构建农业时空多模态知识图谱的基础。层次分明、结构严谨的农业知识模型不仅便于对农业知识的描述和整理,并且能够支持完成知识推理等工作。

本研究采用自顶向下的知识建模方法(Top-Down)^[16,18],首先由领域专家从最顶层的概念进行人工编制,逐步细化,以形成具有良好分类层次结构的领域本体^[16],进而基于此领域本体整合数据信息来构建获取知识图谱。领域本体建模采用斯坦福大学基于Java语言开发的本体编辑和知识获取软件Protégé 5.5.0进行开发,并部署于Windows 10操作系统下使用。

①精准施肥本体构建。精准施肥本体的建立需要大量的领域知识^[19-20]。采用自顶向下的构建方法。首先对水稻在施肥、营养失衡和作物农情方面的知识进行分析和整理;由农业专家进一步细化知识间的关系和层次;然后使用本体编辑器Protégé 5.5.0构建精准施肥时空本体模型。由于农业实体本身具有时空一体化的特征,因此,精准施肥时空本体定义的类具有时间属性和空间属性。农业实体的空间属性是指农田地块空间实体的几何特征(如位置、大小、形状等)及其与其他地块实体的空间关系(如

方位关系、拓扑关系、度量关系)^[21]。农业实体的时间属性是指作物或地块随着时间变化而呈现出不同形态的特征。针对精准施肥决策任务构建所得的本体旨在为时空信息支撑的专家决策系统提供规范的、概念性的知识表示。

该模型涵盖作物、地块、营养3个方面,涉及农业领域的27个类,具体如图3所示。其中,Crop类、Geo:Plot类、Nutrition类以及Temporal_entity类被定义为超类“Thing”的子类。

Crop类包含品种(Breed)、生长期(Growth)、症状(Symptom)、特征(Feature)、地块样本(Crop_sample)、器官(Organ)、水稻性状(Rice_trait)、时间(Event)这7个子类。Crop类及其子类描述了水稻生长期、品种等一般概念。其中,Event用以施肥、灌溉、作物倒伏等事件的抽象表示。水稻性状部分知识来自EMBL-EBI水稻本体Rice Ontology^[22]。Rice_trait类包含Agronomical、Biochemical、Physiological、Quality 4个子类,分别表示水稻农艺性状、生化性状、生理性状、数量性状。

Geo:Plot类是试验地地理位置、地块间拓扑关系、地质地势等空间要素和农田基本信息的抽象表示。Geo:Plot类包含地形地势(Terrain)、位置(Location)、土壤类型(Agrotype)、地块样本(Plot_sample)、土壤性质(Soil_properties)、农业模型(Agricultural_model)6个子类。

Nutrition类整合了水稻营养相关知识,如生长期施肥量和养分补充方法等,同时整合了地块土壤元素等墒情信息。Nutrition类包含元素(Element)、施肥类型(Fertilizer_category)2个子类。

时态实体Temporal_entity类包含时刻(Time_instant)、时间间隔(Time_interval)2个子类,在概念层描述施肥时间、作物生长期等时间信息。

②农业施肥模型本体的构建。所建模的农业施肥模型来自三元非结构肥效模型^[23]、水稻高产优化施肥模型^[24]等肥效模型。

$$Y = b_0 + b_1N + b_2P + b_3K + b_4N^2 + b_5P^2 + b_6K^2 + b_7NP + b_8NK + b_9PK \quad (1)$$

$$Y = 301.48 + 178.19X_1 + 27.64X_2 - 56.18X_3 - 22.08X_2^2 + 22.87X_1X_2 \quad (2)$$

$$Y = 301.48 + 178.19X_1 + 40.29X_3 - 56.18X_1^2 - 13.05X_3^2 + 6.14X_1X_3 \quad (3)$$

$$Y = 301.48 + 27.64X_2 + 40.29X_3 - 22.08X_2^2 - 13.05X_3^2 + 2.93X_2X_3 \quad (4)$$

式(1)为三元非结构肥效模型。Y表示模型拟

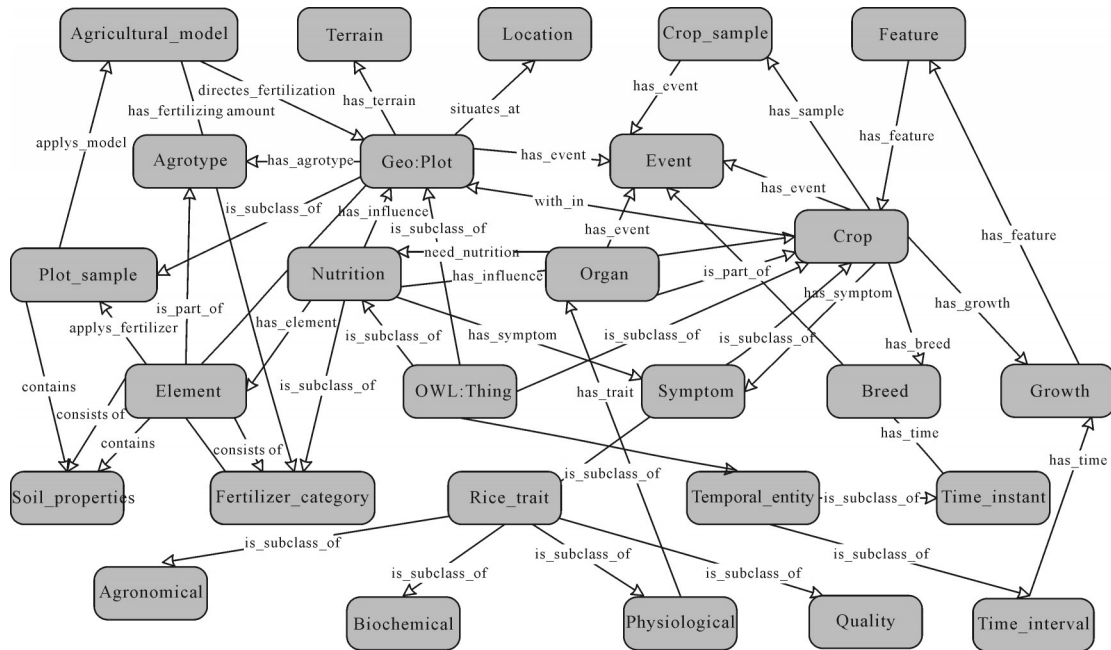


图 3 精准施肥时空本体模型

Fig. 3 A spatio-temporal ontology model for precise fertilization

合产量； N 、 P 、 K 分别表示 N 、 P_2O_5 、 K_2O 施肥量， kg/hm^2 ； $b_0 \sim b_9$ 表示模型的肥效系数^[23]。

式(2)、(3)、(4)分别表示水稻高效优化模型的氮、磷、钾交互作用^[24]。其中，式(2)表示将钾肥(X_3)固定在 0 水平，得到的氮肥(X_1)、磷肥(X_2)的交互作用^[24]；式(3)表示将磷肥(X_2)固定在 0 水平，得到的氮肥(X_1)、钾肥(X_3)交互作用^[24]；式(4)表示将氮肥(X_1)固定在 0 水平，得到的磷肥(X_2)、钾肥(X_3)交互作用^[24]。

参考本研究提出的精准施肥本体，用手工提取出农业施肥模型公式中的抽象概念：如 Plot(地块)、Soil_properties(土壤理化性质)等，然后利用本体编辑器 Protégé 5.5.0 构建如图 4 所示的农业施肥模型本体。其中，Plot(地块)表示可以使用农业施肥模型进行施肥决策的地块；Soil_properties(土壤性状)用以描述地块中的 Alkali-hydr.N(碱解氮)、Olsen-P(有效磷)、Avail K(速效钾)、pH 等土壤理化要素；Element(元素)代表组成土壤有机质含量及肥料的化学元素；Plot_sample(地块样本)表示施用农业施肥模型的具体地块实例的抽象类；Fertilizer_category(肥料种类)代表农业施肥模型涉及的肥料知识的抽象类；Agricultural_model(农业模型)代表用于精准施肥决策的农业模型。

2) 知识抽取。农业知识图谱决策系统是一个智能系统^[18]，通过使用领域中的实体概念和关系来整

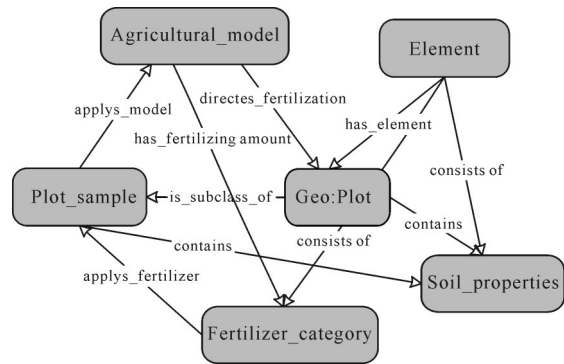


图 4 农业施肥模型本体

Fig. 4 Agricultural fertilization model ontology

合农业数据，以指导精准施肥等农业生产活动。

农业知识抽取任务是通过手动或自动的方式，从多源异构数据源中提取农业实体、关系和属性知识信息，并存储在多模态农业时空知识图谱中。人工进行农业领域知识抽取的方法通常采用农业专家定义的规则来整合施肥决策信息，并建立包含作物本体知识库和施肥决策规则知识库的农业本体专家知识库。农业领域知识自动抽取方法通常是指基于预先构建的精准施肥本体，结合机器学习和自然语言处理技术对非结构化的农业施肥知识进行知识提取^[25]。采用自动知识抽取方式，从现有农业领域的非结构化文献中抽取实体、关系、属性知识，并集成到面向精准施肥的农业时空多模态知识图谱中，用于提供施肥决策。

①实体抽取。实体抽取也称命名实体识别,是自然语言处理领域的一项基本任务,通过检测实体边界和类型来识别文本中有价值和有意义的实体信息^[26]。农业实体的抽取对于构建一个全面、精确的农业时空多模态知识图谱至关重要,其完整性、准确性和召回率都会直接影响知识图谱的质量^[25]。

对于农业文本信息,采用基于BERT+BiLSTM+CRF的命名实体识别模型来完成文本实体抽取任务^[26]。对于一个包含 n 个单词的农业文本知识 $T_i=(W_1, W_2, \dots, W_n)$,首先使用BERT^[26]。预训练模型学习每个单词的初始向量表示 $T_i^e=(W_1^e, W_2^e, \dots, W_n^e)$;然后输入到BiLSTM层以获取蕴含上下文信息的词向量表示;最后结合Softmax分类层的输出结果和CRF中的状态转移矩阵完成对农业文本信息中单词的BIO标记工作,以识别精准施肥领域非结构化文献数据中的实体信息。

②关系抽取。经过实体抽取,可从农业时空领域信息中识别得到一系列离散的农业命名实体。关系抽取任务用以识别实体间语义关系,并将识别出的关系编码到知识图谱中。采用基于Transformer^[27]的深度学习方法完成对农业命名实体间语义关系的识别抽取。对于输入的2个多模态农业实体 $Entity_1$ 和 $Entity_2$,首先使用预训练的Transformer特征提取器生成对应的向量表示 $Entity_1^e$ 和 $Entity_2^e$;然后拼接输入到全连接层以学习表示实体间的语义信息;最后,经过Biaffine分类器和Softmax层预测得到农业多模态实体 $Entity_1$ 和 $Entity_2$ 之间的关系。

农业时空领域关系的提取可分为两类:各实体间的层次语义关系提取和非层次语义关系的提取。其中,层级关系的提取主要涉及is-a和is类型的关系,这种关系相对简单。非层次语义关系包括以地块为单位的种植实例与作物-土壤养分情况间的连接关系。以及施肥量、基肥穗肥配比及施肥时间、产量与相应的农业模型间的连接关系。

③属性抽取。属性抽取的目标是从领域数据中抽取特定实体的属性信息,以补充描述实体的完整知识信息^[18]。提取得到的属性可看作是实体和属性值之间的关系,可将属性抽取任务转化为关系抽取任务^[10],并采用基于Transformer的深度学习方法实现对农业时空领域知识中的属性识别和目标抽取。具体而言,农业时空实体的属性信息包括地块的编号、土壤类型、作物的生长时期和营养状况等。

3)知识存储。农业时空知识图谱的存储可采用

基于图数据库的存储方式。图数据库不仅可以完成异构、多源的农业数据存储,还可以支持上层知识推理、知识快速查询、图计算等应用^[28]。为便于农业时空知识信息在后续任务中的有效应用,在知识存储阶段采用具有高性能、访问速度快、轻量级等特点的Neo4j图数据库完成对精准施肥时空多模态知识图谱中所蕴含的大量事实知识的存储和可视化^[29]。

4)知识推理。知识推理是根据知识图谱中蕴含的信息推断、获取新知识的工作,本质上是下游的分析决策应用提供全面可靠的知识支撑。从预先构建的农业时空多模态知识图谱中推理得到有效可靠的施肥决策知识,以用于支持精准施肥决策任务。

①基于专家规则的知识推理。基于专家规则的知识推理方法,根据领域专家经验知识抽象出一系列规则,并运用于知识图谱,对其中缺失的知识信息进行推理补充。

在精准施肥决策任务中,针对“底肥影响水稻养分水平”“实时养分水平决定施肥量”“作物养分需求的时态性变化”等自然语言描述的复杂逻辑规则;如图5所示,通过结合精准施肥时空本体中的TBox和ABox逻辑表示集合,采用语义网规则语言SWRL(semantic web rule language)进行表达^[30]。其中,TBox定义了本体中概念与概念间的关系,是描述概念的术语公理集;ABox用于指明个体的属性及个体间的关系,是有关实例断言的集合^[30]。SWRL是基于OWL语言、具有强表达能力的规则描述语言。通过使用SWRL Editor编辑器^[30],根据自然语言描述的专家施肥决策规则和施肥决策经验,从精准施肥本体概念中抽取涉及的推理元素,如Plot、Element等。并根据元素组建成机器可理解的SWRL推理规则^[30-31]。利用SWRL规则增强现有知识库在解决精准施肥决策问题中的知识推理能力。

水稻精准施肥决策规则围绕“通过性状、病症进而推断出施肥需求、施肥量”建立。从精准施肥时空本体的概念中抽取出施肥决策涉及的具体推理元素,进而组建施肥推理规则。为确保推理规则中使用的所有原子都源自所构建的农业时空本体,需对精准施肥时空本体中的类和属性进行筛选^[31]。

施肥规则的推理过程如图5所示。由于SWRL Editor编辑器不具备推理能力,采用Jess(Java expert system shell)作为推理引擎^[32]。通过集成SWRL Editor编辑器和Jess编辑器,将精准施肥本体和SWRL施肥规则编译成Jess事实库和Jess规则库,以

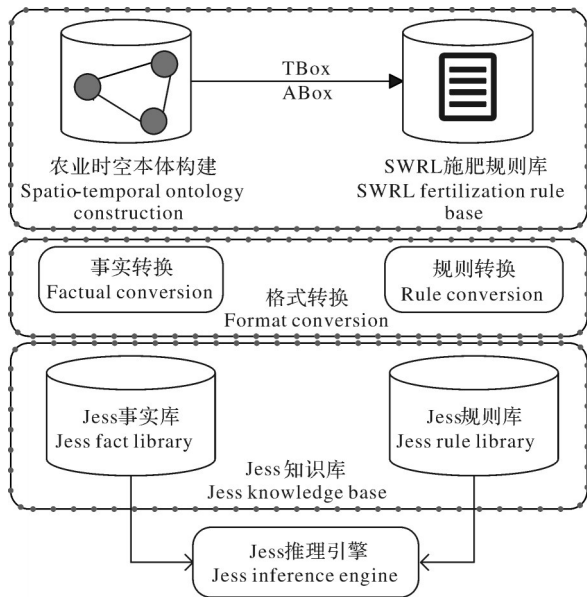


图 5 基于 SWRL 的推理过程

Fig. 5 Reasoning process based on SWRL

调用 Jess 推理引擎^[31-32]。对于推理得出的新的事实和规则,进一步经格式转换后补充进精准施肥时空本体,以达到农业时空多模态知识图谱的充实和更新。

②基于深度学习的知识推理。与传统的基于逻辑规则的推理方法相比,基于深度学习的关系推理方法能够更有效地挖掘出知识图谱中隐藏的语义信息^[25]。在结构上,知识推理可划分为全局任务和局部任务^[33],局部任务即节点分类、链路预测;全局任务包含子图匹配和子图分类等。利用农业时空多模态知识图谱进行精准施肥决策的方法通过基于深度学习的子图匹配完成。

对于给定的待查询图 Q,从知识图谱 KG 中获得正确的子图匹配结果,分别为候选子图查找和匹配相似度计算这两部分内容。在候选子图查找这一步骤中,首先使用 TransE 模型完成对查询图 Q 和知识图谱 KG 中节点和关系的嵌入表示^[34];然后基于向量的相似度计算结果,返回查询图 Q 中的节点和关系在知识图谱 KG 中置信度较高的 Top-K 个候选结果。对于由不同候选节点和关系组合得到的 N 个候选子图结果,在图匹配相似度计算这一步骤中使用基于表示学习的图匹配相似度计算方法完成对查询图 Q 和 N 个候选子图的相似度计算,以获取准确的匹配结果。

从嵌入对象角度上进行划分,基于表示学习的图匹配相似度计算方法可分为基于节点表示的图匹

配相似度计算方法和基于整图表示的图匹配相似度计算方法^[35]。前者通过图中节点的向量表示集合 $h_{Ve}=(h_{e1}, h_{e2}, \dots, h_{en})$ 来描述图的语义信息,其中 h_{ei} 表示图中第 i 个节点的向量表示。由此,图匹配相似度计算任务可转化为对两组节点嵌入集合的相似度衡量。与基于节点表示的图匹配相似度计算方法相比,基于整图表示的方法能够整合表示图中隐含的语义信息和结构信息。给定 2 个图 $G_1=(V_1, E_1)$ 和 $G_2=(V_2, E_2)$,首先生成对应的图嵌入表示 h_{G1} 和 h_{G2} ,然后通过相似度学习融合两图的嵌入表示 $h=(h_{G1}, h_{G2})$,以输出衡量 G_1 和 G_2 匹配度的相似性分数 $Sim=(G_1, G_2)$ 。本研究使用基于整图表示的图匹配相似度计算方法来完成对查询图和候选子图的相似度计算任务。

2 结果与分析

通过构建精准施肥时空本体,完成对精准施肥领域中概念、术语及其相互关系的规范化描述,同时有效整合、表示多模态农业数据中的知识信息;并进一步构建 SWRL 规则以更新完善精准施肥时空本体;并通过知识抽取等知识图谱构建技术获取描述水稻施肥时空知识信息的农业时空多模态知识图谱。最后结合基于农业施肥模型和实时采集数据构建所得的查询图,采用整图表示学习的子图匹配方法,完成从农业时空多模态知识图谱中推理获得精准施肥决策结果的任务。

2.1 基于 SWRL 的精准施肥规则设计

将构建完成的精准施肥本体和 SWRL 规则放入 Jess 推理机中进行推理,可以获得隐含的知识:

1) 已知作物性状,确定施肥类型。如:“监利市新沟镇双水双绿基地灰泥田试验点(地块编号 22-1)水稻叶片发黄,主要原因为根系发育不良,需要施加酸性肥料磷酸二氢钾促新根”的事实描述,可将上述自然语言形式化为 SWRL 并构建相应的施肥决策推理规则,具体的规则示例如下:

$$\begin{aligned} & \text{Plot}(? p) \wedge \text{Terrain}(? p, ? t) \wedge \text{Location}(? p, ? l) \wedge \\ & \text{Crop}(? p, ? c) \wedge \text{Organ}(? c, ? o) \wedge \text{Biochemical}(? o, ? b) \rightarrow \\ & \text{Fertilizer Category}(? p, ? f) \end{aligned} \quad (1)$$

其中, p 是地块的实例; $\text{Terrain}(? p, ? t)$ 表示 p 的土壤类型是 t ; $\text{Location}(? p, ? l)$ 表示地块 p 位于 Location l ; $\text{Crop}(? p, ? c)$ 表示位于地块 p 上的作物实例 c ;

Organ(? *c*, ? *o*)表示*c*具有器官*o*;Biochemical(? *o*, ? *b*)表示器官*o*上出现生化性状*b*; Fertilizer Category(? *p*, ? *f*)表示在地块*p*施加肥料*f*。

2)已知土壤初始养分含量,判断是否施肥。事实描述如“监测数据显示,试验点地块22-1土壤碱解氮含量低于54.39 mg/kg,不利于水稻生长,需要补充氮素”。规则实例如下:

$$\text{Plot}(\text{? } p) \wedge \text{Soil Properties}(\text{? } p, \text{? } s) \wedge \text{Crop}(\text{? } p, \text{? } c) \wedge \text{swrlb: less Than}(\text{? } s, 54.39) \rightarrow \text{Event}(\text{? } e) \wedge \text{Fertilizer Category}(\text{? } p, \text{? } f) \quad (2)$$

其中,Soil Properties(? *p*, ? *s*)表示地块*p*的初始养分*s*;swrlb: less Than(? *s*, 54.39)表示养分*s*的含量低于54.39 mg/kg;Event(? *e*)表示事件实例*e*。

3)已知作物生长时期,确定施肥类型。事实描述如“黄泥田试验地水稻处于分蘖期,为保证水稻生长期的养分供应,需要施加锌肥”。规则实例如下:

$$\text{Plot}(\text{? } p) \wedge \text{Plot Sample}(\text{? } p, \text{? } x) \wedge \text{Terrain}(\text{? } p, \text{? } t) \wedge \text{Crop}(\text{? } p, \text{? } c) \wedge \text{Growth}(\text{? } c, \text{? } g) \rightarrow \text{Fertilizer Category}(\text{? } p, \text{? } f) \quad (3)$$

其中,Plot Sample(? *p*, ? *x*)表示地块*p*的试验

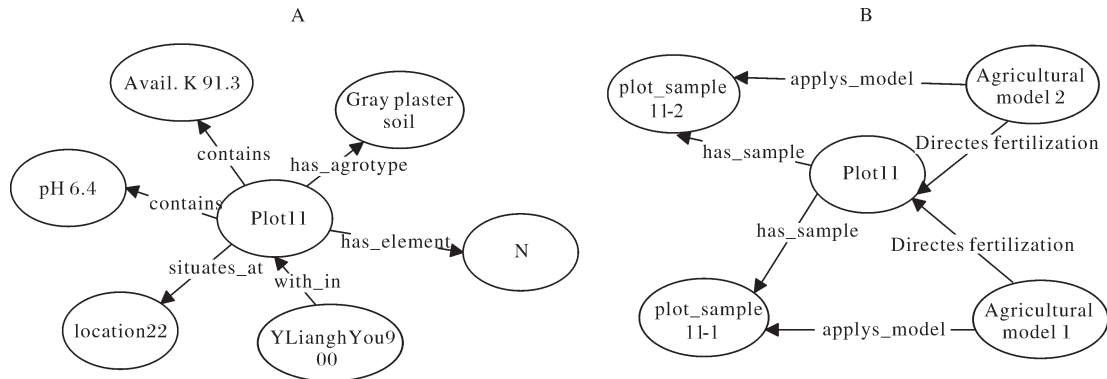
地*x*;Growth(? *c*, ? *g*)表示水稻*c*处于生长期*g*。

2.2 精准施肥农业模型的自动化选择

本研究所构建的农业时空多模态知识图谱,整合了大量农业施肥模型、农业施肥决策试验信息。基于本体层的地块-农业施肥模型抽象概念,在知识图谱的实例层,描述了具体地块实体对象、农业模型实体对象并建立了实体间的关系。

如图6A所示,在农业时空多模态知识图谱中,使用地块、位置等节点及节点间连边关系表示某一具体地块的事实和数据信息,即“地块(Plot11)有土壤类型(has_agrotype);灰泥田(gray plaster soil);地块(Plot11)含有土壤理化性质(contains)为pH 6.4”。

如图6B所示,农业模型节点(agricultural model)与地块节点(Plot11)为“多对多”的关系,即表示“一个地块可使用不同的农业施肥模型”“一个农业施肥模型在不同的地块均可使用”。而地块事例样本节点(plot_sample11-1)与农业模型节点(agricultural model)为“一对一”的关系,即表示“某一农业施肥模型使用在地块11上的具体地块施肥事例样本”。



A:以节点、连边关系表示的地块事实与数据信息;B:以节点、连边关系表示的地块、地块施肥事例样本、农业模型信息。A:Plot facts and data information represented by node and edge plot; B:Plot sample and agricultural model information represented by node and edge.

图6 农业时空多模态知识图谱中的部分实体节点及关系

Fig.6 Some entity nodes and relationships in agricultural spatio-temporal multi-modal knowledge graph

在精准施肥决策场景中,首先通过“天空地”农业信息采集系统采集地块Plot实时数据(包括地块位置信息、水土营养含量、地形地势等)并结合作物生长状况监测数据,发现作物存在潜在的养分缺乏问题。根据当前地块状况,利用农业多模态时空知识图谱,查询最适合当前地块的农业施肥模型。随后结合地块实时数据和农业施肥模型的查询图模板(图7A),实例化得到农业施肥模型查询图Q(图7B)。根据不同地块的实时数据,可实例化获得节点信息丰富度具有差异的查询图Q。图7B中,带有红

色边框的节点表示需要在知识图谱中匹配查询的节点;“? plot_sample”和“? y”节点信息分别表示待匹配获取的地块施肥事例样本和待查询的农业施肥模型,即精准施肥的决策信息。

以Plot1为例,对于实例化获得的农业施肥模型查询图Q(图7B中蓝色子图),利用TransE模型在农业时空多模态知识图谱(图7C)中匹配得到其多个候选子图,并采用基于整图表示的图匹配相似度计算方法从候选子图中计算得到查询图Q的置信度较高的匹配结果,即图7C中以蓝色显式标注的区域。最

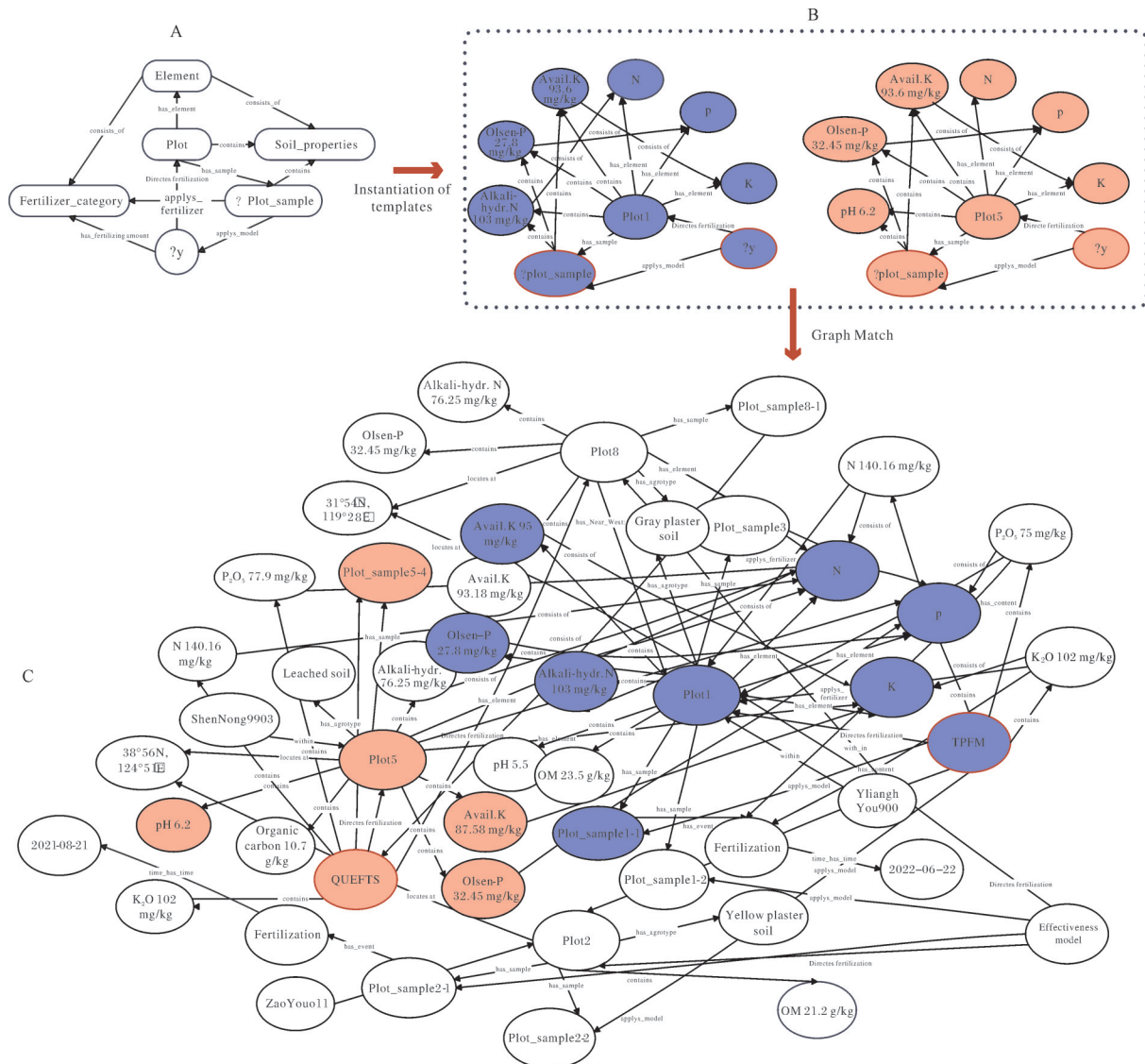


图 7 查询图模板(A)、实例化的查询图(B)和农业时空多模态知识图谱(C)

Fig. 7 Query graph template (A), instantiated query graph (B), and agricultural spatio-temporal multi-modal knowledge graph (C)

后,基于显示标注节点以及“Plot”类的实例节点 (“Plot1”节点)“Directs_fertilization”“has_sample”“applies_model”关系链接,在农业时空多模态知识图谱的匹配区域中获取到地块施肥事例样本节点 Plot_sample1-1 和最适合查询地块 Plot1 的农业施肥模型 TPFM。

3 讨论

知识图谱作为有效的知识表示模型,已在工业制造业^[38]、教育^[39]、医学^[40]、军工^[41]等多个领域广泛构建及应用,YAGO^[42]、Freebase^[43]和DBpedia^[44]等是目前国际上典型的领域知识图谱。农业领域与知识图谱技术的深度融合是重要的发展方向。

针对农业生产过程中数据孤立,知识融合不充分的问题。本文利用知识图谱网络状的结构以及图的形式,对多源异构数据进行整合。通过将农业时空场景下的农田地块、作物、营养与施肥措施等要素抽象表示并进行形式化的描述,在不同时空维度建立要素间的关系。同时,多模态知识图谱的引入,将农业领域的专家知识系统化、抽象化表达,并借助图结构实现施肥决策等功能。

本研究提出的农业时空多模态知识图谱,本质上是对领域内农业专家知识、模型知识的精准化描述。农业时空多模态知识图谱相较于传统的农业专家知识库,有着更高的知识覆盖率、更强的数据分析能力等优点,为精准施肥的农业大数据标准化整合

与辅助决策提供新思路与解决方案。

后续将在具体大田试验中,将获得的农业施肥模型用于农事管理,并根据肥效与产量结果,对农业时空多模态知识图谱进行迭代优化,以实现农业时空知识图谱的精准施肥决策。同时,进一步集成多种模态的农业信息,深度发掘农业大数据的潜在价值和用途。

参考文献References

- [1] ZHANG A R, HEATH R, MCROBERT K, et al. Who will benefit from big data? Farmers' perspective on willingness to share farm data[J]. *Journal of rural studies*, 2021, 88: 346-353.
- [2] KHANAL S, FULTON J, SHEARER S. An overview of current and potential applications of thermal remote sensing in precision agriculture [J]. *Computers and electronics in agriculture*, 2017, 139: 22-32.
- [3] 陈栋. 多源异构农业感知数据接入系统的设计与实现[D]. 泰安: 山东农业大学, 2014. CHEN D. Design and implementation of multi-source and heterogeneous agricultural sensor data access system[D]. Tai'an: Shandong Agricultural University, 2014 (in Chinese with English abstract).
- [4] 白由路. 精准施肥实施技术[M]. 北京: 中国农业科学技术出版社, 2019: 3. BAI Y L. Implementation technology of precision fertilization[M]. Beijing: China Agricultural Science and Technology Press, 2019: 3 (in Chinese).
- [5] SUBIRATS-COLL I, KOLSHUS K, TURBATI A, et al. AGROVOC: the linked data concept hub for food and agriculture [J/OL]. *Computers and electronics in agriculture*, 2022, 196: 105965 [2023-01-13]. <https://doi.org/10.1016/j.compag.2020.105965>.
- [6] RAJBHANDARI S, KEIZER J. The AGROVOC concept scheme: a walkthrough[J]. *Journal of integrative agriculture*, 2012, 11(5): 694-699.
- [7] DARNALA B, AMARDEILH F, ROUSSEY C, et al. Crop planning and production process ontology (C3PO), a new model to assist diversified crop production [C]//IFOW 2021-integrated food ontology workshop@12th ICBO, September 15-18, 2021, Bolzano, Italy. Bolzano: Free University of Bozen-Bolzano, 2021.
- [8] CHEN Y Z, KUANG J, CHENG D W, et al. AgriKG: an agricultural knowledge graph and its applications [M]//MENG W Y, FENG L, BRESSAN S, et al. Database systems for advanced applications. Cham: Springer International Publishing, 2019: 533-537.
- [9] RODRÍGUEZ-GARCÍA M Á, GARCÍA-SÁNCHEZ F, VALEN-CIA-GARCÍA R. Knowledge-based system for crop pests and diseases recognition [J/OL]. *Electronics*, 2021, 10(8): 905 [2023-01-13]. <https://doi.org/10.3390/electronics10080905>.
- [10] LIU X X, BAI X S, WANG L H, et al. Review and trend analysis of knowledge graphs for crop pest and diseases [J]. *IEEE access*, 2019, 7: 62251-62264.
- [11] HASSANI-PAK K, SINGH A, BRANDIZI M, et al. KnetMiner: a comprehensive approach for supporting evidence-based gene discovery and complex trait analysis across species [J]. *Plant biotechnology journal*, 2021, 19(8): 1670-1678.
- [12] 戈为溪, 周俊, 袁立存, 等. 基于知识图谱与案例推理的水稻精准施肥推荐模型[J]. *农业工程学报*, 2023, 39(2): 126-133. GE W X, ZHOU J, YUAN L C, et al. Recommendation model for rice precision fertilization using knowledge graph and case-based reasoning [J]. *Transactions of the CSAE*, 2023, 39(2): 126-133 (in Chinese with English abstract).
- [13] 赵春江, 薛绪掌, 王秀, 等. 精准农业技术体系的研究进展与展望[J]. *农业工程学报*, 2003, 19(4): 7-12. ZHAO C J, XUE X Z, WANG X, et al. Advance and prospects of precision agriculture technology system [J]. *Transactions of the CSAE*, 2003, 19(4): 7-12 (in Chinese with English abstract).
- [14] 吴文斌, 史云, 周清波, 等. 天空地数字农业管理系统框架设计与构建建议[J]. *智慧农业*, 2019, 1(2): 64-72. WU W B, SHI Y, ZHOU Q B, et al. Framework and recommendation for constructing the SAGI digital agriculture system [J]. *Smart agriculture*, 2019, 1(2): 64-72 (in Chinese with English abstract).
- [15] 何东健, 何勇, 李明赞, 等. 精准农业中信息相关科学问题研究进展[J]. *中国科学基金*, 2011, 25(1): 10-16. HE D J, HE Y, LI M Z, et al. Research progress of information science-related problems in precision agriculture [J]. *Bulletin of national natural science foundation of China*, 2011, 25(1): 10-16 (in Chinese with English abstract).
- [16] WANG Y, WANG Y. Citrus ontology development based on the eight-point charter of agriculture [J]. *Computers and electronics in agriculture*, 2018, 155: 359-370.
- [17] 杨洁. 基于本体的柑橘病虫害知识建模及推理研究[D]. 武汉: 华中师范大学, 2014. YANG J. Research on knowledge modeling and reasoning ontology-based of citrus disease and pests [D]. Wuhan: Central China Normal University, 2014 (in Chinese with English abstract).
- [18] ZHANG H, SI H J, MA X M, et al. Research and application of agriculture knowledge graph [C]//Proceedings of the 2021 5th international conference on electronic information technology and computer engineering. New York: ACM, 2021: 680-688.
- [19] RAFFAT S K, SIDDIQUI M S, SHAIKH Z A, et al. Ontology: a scientific classification technique [J]. *Sindh University research journal*, 2012, 44(2): 63-68.
- [20] WANG Y, WANG Y, WANG J, et al. An ontology-based approach to integration of hilly citrus production knowledge [J]. *Computers and electronics in agriculture*, 2015, 113: 24-43.
- [21] 王志华, 杨晓梅, 周成虎. 面向遥感大数据的地理知识图谱构想[J]. *地球信息科学学报*, 2021, 23(1): 16-28. WANG Z H, YANG X M, ZHOU C H. Geographic knowledge graph for re-

- mote sensing big data [J]. *Journal of Geo-information science*, 2021, 23(1): 16-28 (in Chinese with English abstract).
- [22] SHRESTHA R, MATTEIS L, SKOFIC M, et al. Bridging the phenotypic and genetic data useful for integrated breeding through a data annotation using the crop ontology developed by the crop communities of practice [J/OL]. *Frontiers in physiology*, 2012, 3: 326 [2023-01-13]. <https://doi.org/10.3389/fphys.2012.00326>. eCollection 2012.
- [23] 李娟, 章明清, 章赞德, 等. 三元非结构肥效模型提高水稻施肥推荐的可靠性 [J]. *植物营养与肥料学报*, 2019, 25(2): 311-320. LI J, ZHANG M Q, ZHANG Z D, et al. Increasing precision of fertilizer recommendation using ternary non-structural fertilizer response model [J]. *Journal of plant nutrition and fertilizers*, 2019, 25(2): 311-320 (in Chinese with English abstract).
- [24] 吴家琼, 丁亨虎, 刘克芝, 等. 潜江市水稻氮磷钾施肥模型的建立及其应用研究 [J]. *现代农业科技*, 2013(4): 9-11, 13. WU J Q, DING H H, LIU K Z, et al. Establishment and application of nitrogen, phosphorus and potassium fertilization model for rice in Qianjiang city [J]. *Modern agricultural science and technology*, 2013(4): 9-11, 13 (in Chinese).
- [25] 梁静茹, 鄂海红, 宋美娜. 基于属性图模型的领域知识图谱构建方法 [J]. *计算机科学*, 2022, 49(2): 174-181. LIANG J R, E H H, SONG M N. Domain knowledge graph construction method based on attribute graph model [J]. *Computer science*, 2022, 49(2): 174-181 (in Chinese with English abstract).
- [26] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding [DB/OL]. arXiv, 2018: 1810.04805 [2023-01-13]. <https://doi.org/10.48550/arXiv.1810.04805>.
- [27] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Proceedings of the 31st international conference on neural information processing systems. New York: ACM, 2017: 6000-6010.
- [28] 李岩, 胡文岭. 基于知识图谱的农业知识问答系统研究 [J]. *智慧农业导刊*, 2021, 1(11): 20-22. LI Y, HU W L. Research on agricultural knowledge question and answer system based on knowledge graph [J]. *Journal of smart agriculture*, 2021, 1(11): 20-22 (in Chinese with English abstract).
- [29] FERNANDES D, BERNARDINO J. Graph databases comparison: AllegroGraph, ArangoDB, InfiniteGraph, Neo4J, and OrientDB [C]//Proceedings of the 7th international conference on data science, technology and applications. Setubal: ScitePress, 2018: 373-380.
- [30] O'CONNOR M, KNUBLAUCH H, TU S, et al. Supporting rule system interoperability on the semantic web with SWRL [M]//HUTCHISON D, MITCHELL J C. *The Semantic Web-ISWC 2005*. Heidelberg: Springer, 2005: 974-986.
- [31] 王雪, 魏圆圆, 王儒敬, 等. 猪病本体的构建及 SWRL 推理研究 [J]. *江西农业学报*, 2012, 24(3): 103-107. WANG X, WEI Y Y, WANG R J, et al. Research on construction and SWRL reasoning of swine disease nomenclature [J]. *Acta agriculturae Jiangxi*, 2012, 24(3): 103-107 (in Chinese with English abstract).
- [32] 黄勇奇, 牛振国, 崔伟宏. 基于地理本体和 SWRL 的时空推理研究 [J]. *武汉理工大学学报(交通科学与工程版)*, 2009, 33(6): 1175-1178. HUANG Y Q, NIU Z G, CUI W H. Research on spatio-temporal reasoning based on geo-ontology and SWRL [J]. *Journal of Wuhan University of Technology (transportation science & engineering)*, 2009, 33(6): 1175-1178 (in Chinese with English abstract).
- [33] ROBERTS T L. Right product, right rate, right time and right place: the foundation of BMPs for fertilizer [J]. *Better crops with plant food*, 2007, 4: 14-15.
- [34] BORDES A, USUNIER N, GARCIA-DURÁN A, et al. Translating embeddings for modeling multi-relational data [C]//Proceedings of the 26th international conference on neural information processing systems. New York: ACM, 2013: 2787-2795.
- [35] CAI H Y, ZHENG V W, CHANG K C C. A comprehensive survey of graph embedding: problems, techniques, and applications [J]. *IEEE transactions on knowledge and data engineering*, 2018, 30(9): 1616-1637.
- [36] 许驹雄, 李敏波, 刘孟珂, 等. 发动机故障领域知识图谱构建与应用 [J]. *计算机系统应用*, 2022, 31(7): 66-76. XU J X, LI M B, LIU M K, et al. Construction and application of knowledge graph in diesel engine fault field [J]. *Computer systems and applications*, 2022, 31(7): 66-76 (in Chinese with English abstract).
- [37] 刘凯航. 基于知识图谱的个性化学习路径推荐研究 [D]. 南京: 南京邮电大学, 2022. LIU K H. Research on personalized learning path recommendation method based on knowledge graph [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2022 (in Chinese with English abstract).
- [38] 梁静, 文奕. 知识图谱在医学辅助诊断中的应用研究 [J]. *医学信息学杂志*, 2022, 43(11): 34-40. LIANG J, WEN Y. Study on the application of knowledge graph in medical auxiliary diagnosis [J]. *Journal of medical informatics*, 2022, 43(11): 34-40 (in Chinese with English abstract).
- [39] 赵永亮, 于倩, 邓博, 等. 基于知识图谱的装备故障诊断技术 [J]. *电子设计工程*, 2022, 30(9): 125-129. ZHAO Y L, YU Q, DENG B, et al. Equipment fault diagnosis technology based on knowledge graph [J]. *Electronic design engineering*, 2022, 30(9): 125-129 (in Chinese with English abstract).
- [40] PELLISSIER T T, WEIKUM G, SUCHANEK F. YAGO 4: a reasonable knowledge base [C]//European semantic web conference 31 May-4 June, 2020, Heraklion, Crete, Greece. Cham: Springer, 2020: 583-596.

- [41] BOLLACKER K, EVANS C, PARITOSH P, et al. FreeBase: a collaboratively created graph database for structuring human knowledge [C]//Proceedings of the 2008 ACM SIGMOD international conference on management of data. New York: ACM, 2008:1247-1250.
- [42] LEHMANN J, ISELE R, JAKOB M, et al. DBpedia: a large-scale, multilingual knowledge base extracted from Wikipedia [J]. Semantic web, 2015, 6(2):167-195.

A method of deciding precision fertilization of rice based on spatio-temporal multi-modal knowledge graph of agriculture

XU Duo^{1,2}, LU Wangping^{1,2}, XU Ruiqing³, ZHANG Hongyu³,
JIANG Yang^{1,4}, YOU Liangzhi^{1,2,5}, FENG Zaiwen^{2,3,4}

1. College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, China;
2. Macro Agricultural Research Institute, Huazhong Agricultural University, Wuhan 430070, China;
3. College of Informatics, Huazhong Agricultural University, Wuhan 430070, China;
4. Shuangshui Shuanglü Institute, Huazhong Agricultural University, Wuhan 430070, China;
5. International Food Policy Research Institute, Washington D.C. 20005, USA

Abstract Using information technology to realize the effective integration and application of multi-source heterogeneous spatio-temporal multi-modal big data of agriculture is a key issue that needs to be urgently solved in the precision agriculture. A method of constructing, controlling and decision-making for precision fertilization was proposed based on the spatio-temporal multi-modal knowledge graph of agriculture to construct a management system and realize the fine management of nutrients in field. The nodes and relationships in the plots to be queried and the spatio-temporal multi-modal knowledge graph of agriculture were embedded and represented through the subgraph matching method based on deep learning. Vector similarity calculation was used to obtain candidate subgraphs. The fertilization model data suitable for query plots were obtained from the information of subgraphs storing historical data. The results showed sub maps isomorphic to the given land query map were obtained in the spatio-temporal multi-modal knowledge graph of agriculture based on the instantiated query map of the land to be fertilized. An agricultural fertilization model suitable for the current plot was obtained from the information of subgraph storing historical decision. It is indicated that the automatic selection of model based on spatio-temporal multi-modal knowledge graph of agriculture is accurate and reliable. It will provide decision-making support for precision fertilization.

Keywords spatio-temporal; multi-modal knowledge graph; smart agriculture; precision fertilization; rice; agricultural big data; knowledge-sharing

(责任编辑:赵琳琳)