

基于 Jena 引擎的湿地食物链语义模拟研究

陈文君¹, 王子豪¹, 邹睿泓², 丁顾霏¹, 严 鸣¹, 徐尚瑜¹, 曲爱妍¹, 张 燕³

(1. 金陵科技学院软件工程学院, 江苏 南京 211169; 2. 金陵科技学院国际教育学院, 江苏 南京 211169;

3. 金陵科技学院, 江苏 南京 211169)

摘 要: 湿地生态系统生物种类繁多, 内部食物链关系复杂。为模拟食物链循环过程, 构建了基于种族关系的本体模型, 依靠 Jena 引擎与语义网规则语言 (semantic web rule language, SWRL) 技术实现了语义数据的高效存储和查询。同时, 以真实池塘样本数据为例, 借助 Java swing 和超文本标记语言 (hyper text markup language, HTML) 技术完成对模拟结果的可视化展示, 从稳态和非稳态两方面完善生态系统食物链的建模与仿真, 从而实现模拟过程的动态化与量化。

关键词: 语义模拟; 本体构建; 湿地食物链; 语义网规则语言; Jena 引擎

中图分类号: TP393; X56

文献标识码: A

文章编号: 1672-755X(2021)01-0008-06

Study on Semantic Simulation of Wetland Food Chain Based on Jena Engine

CHEN Wen-jun, WANG Zi-hao, ZOU Rui-hong, DING Gu-fei,

YAN Min, XU Shang-yu, QU Ai-yan, ZHANG Yan

(Jinling Institute of Technology, Nanjing 211169, China)

Abstract: The wetland ecosystem has a wide variety of organisms, and the internal food chain is complicated. In order to simulate the wetland food chain cycle process, this paper proposes an ontology model based on race relations, and achieves efficient storage and query of semantic data by relying on Jena engine and semantic web rule language technology. At the same time, taking real pond sample data as an example, with the help of Java swing and hyper text markup language to complete the visualization of simulation results, the research improves the modeling and simulation of the wetland food chain of the ecosystem from both the steady state and the unsteady state, and achieves dynamic and quantitative simulation process.

Key words: semantic simulation; ontology construction; wetland food chain; semantic web rule language; Jena engine

湿地是地球上极其重要的自然资源, 保护湿地需深入了解生态系统的动态过程, 以实现对湿地的评估、预测与管理^[1]。由于人力、物力等资源的限制, 在不同地区, 以不同尺度对整个湿地生态系统实行长久的现场观测仍存在一定的难度^[2], 因此需要构建语义模型来达到该目的^[3-4]。语义网是互联网技术飞速发展的产物, 已广泛应用于社会生产中。在语义网体系中, 本体占据主导地位, 是解决层次上的 Web 信息

收稿日期: 2020-07-14

基金项目: 江苏省自然科学基金 (BK20180115); 江苏省农业科技自主创新资金项目 (CX(17)2015); 金陵科技学院博士科研启动基金 (jit-b-201804); 金陵科技学院校级科研基金孵化项目 (jit-fhxm-201804); 金陵科技学院“科教融合”项目 (2020KJRH26); 江苏省高校中外合作办学高水平示范性建设工程项目

作者简介: 陈文君 (1986—), 男, 江苏南京人, 讲师, 博士, 主要从事数学生态农业、GIS 软件工程、中外合作办学研究。

共享和重用问题的基础。基于本体构建的语义模型^[5]需要编程语言的支持,在当今所有的编程语言中,由 W3C 组织定义的网络本体语言(ontology web language,OWL)是较为合适的描述语言。

本研究首先利用 Protege 来制作关系图谱,将湿地食物链通过关系图谱进行可视化展示;然后利用 SWRL 技术对湿地食物链进行规则推理,利用 Jena 引擎对关系图谱生成的 OWL 文件进行读取和解析,利用 SPARQL 语句对关系图谱进行查询,并将所有个体之间的关系以三元组的形式打印出来;最后运用实例对构建的湿地食物链模型进行验证,基于 Java swing 和 Jfreechart 工具包,对规则推理的结果进行可视化展示,对展示结果进行分析与比对,以查看运行是否达到预期效果。

1 湿地食物链的本体

1.1 湿地食物链的本体设计

关系图谱是结构化的语义知识库,它以符号的形式来描述物理世界中的概念及其相互关系,其基本组成单位是资源描述框架(resource description framework,RDF)三元组,为实体或资源的描述提供了统一的标准。大多数知识图谱的构建过程需要经历本体建模、信息抽取、信息融合、信息存储、信息推理五个阶段^[6],其中本体建模遵循精准的指导方法,选取最具权威的数据源,指定该领域内的对象模型,从不同层面明确分析对象与对象之间的关系^[7]。湿地食物链本体概念之间的通用关系如表 1 所示。

表 1 湿地食物链本体概念之间的通用关系

通用关系名称	语义关系的描述
Part-of	概念术语之间的继承机制,也可以理解成包含与被包含的关系
Kind-of	概念术语之间存在整体与部分的关系
Instance-of	概念术语与实例之间的关系,如“鲫鱼”是“鱼类”的一个实例
Attribute-of	概念术语与属性之间的关系,如“死亡”是“鲫鱼”的一个属性

经过对本体领域范围的分析,确定研究对象为湿地动物和湿地植物。根据实际情况,本试验构建了自上而下的本体层次体系,其中具体的本体核心概念为:1)湿地动物具体可分为兽类、鸟类、鱼类、两栖类、爬行类、甲壳类、脊椎类、无脊椎类等。2)湿地植物具体可分为沼生植物、湿生植物和水生植物。3)湿地动植物的子类本体概念需要根据具体种类的相应需求再进行划分。

在构建湿地食物链本体后,设计湿地属性,分别为:1)数据属性的设计。“生存状态”用来表示当前生物个体的生存状态,“即将死亡”用来表示当前生物个体即将死亡(该死亡为自然死亡)。2)对象属性的设计。“捕食”用来连接不同生物个体,从而决定个体在食物链中的顺序。

图 1 为 Protege 生物类间的关系图谱,实线箭头表示从属关系,如 Animal→Fish→Carp,表示动物包含鱼类,鲤鱼是鱼类的一个实例;虚线箭头表示捕食关系,如 Eagle→Rabbit,表示老鹰捕食兔子。每个实例标签内均保存生物的属性,如生存或死亡。

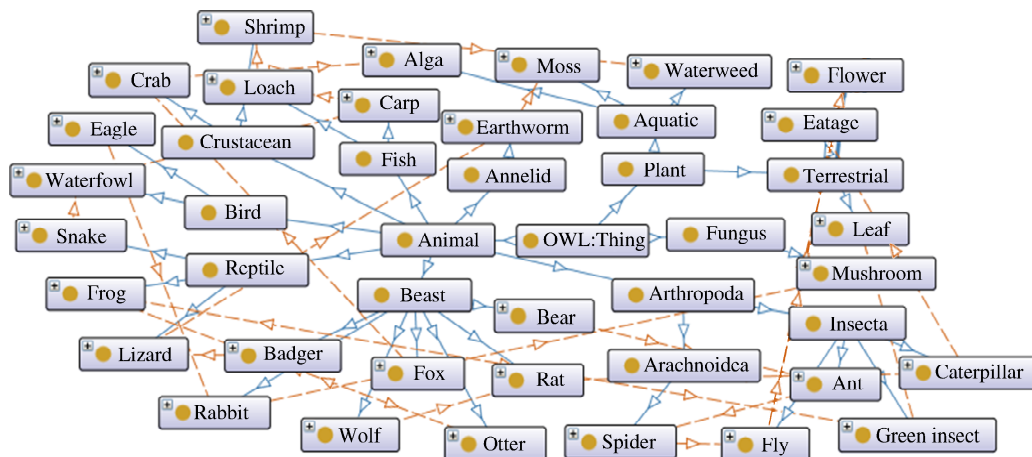


图 1 Protege 生物类间的关系图谱

1.2 基于 Jena 引擎的 OWL 文件解析

生态关系的推理机制是根据 Jena 引擎提供的推理机而建立的。Jena 引擎是惠普实验室开发的一个构建语义 Web 应用程序的框架,为 RDF、RDFS(RDF schema)、OWL 以及 SPARQL 提供了一个可编程的环境。此外,Jena 引擎还含有基于规则的推理引擎^[8],能够结合本体语言 OWL,实现文件解析。

OWL 文件主要表示对象类别以及对象相互关联的信息。本研究生成的 OWL 文件主要包含 3 种元素: Object property、Datatype property、Class。这里以鲫鱼的语法分析为例:

```
<OWL:Object property RDF:about = "Crucian, Carp;Eat"/>
```

```
<OWL:Datatype property RDF:about = "Crucian, Carp;Living state"/>
```

```
<OWL:Class RDF:about = "Crucian, Carp;Alga"/>
```

在 OWL 文件中对不同元素的定义都是以 RDF 三元组的形式(“<元素类别”“=元素名”“元素具体值>”)表示,其中“<元素类别”采用“RDF:about”连接。

本研究使用 Jena API 来处理以上 RDF 三元组。通常情况下,用户使用 Jena API 构建本体,可获取基于该语言设计的本体对象,但参数均默认为空,此时系统便会根据参数值来构建一个基于 OWL 文件的本体模型,此模型可操作 Jena API 所提供的任何信息,可避免对相应语言类型的设置。因此,Jena 框架不仅支持基于 OWL 文件的本体模型,并且实现了语言和推理的本体操作。

2 湿地生态推理

2.1 湿地简单推理查询

本研究为使推理查询具有更强的逻辑表达能力,采用 SWRL 技术以及 Jena 推理机来对推理查询进行完善^[9]。在湿地食物链中,本体采用 Jena 框架自带的 RDFS 推理机^[10],该推理机遵循自定义的湿地推理规则^[11]。

Jena RDFS 推理机结构如图 2 所示^[12]。其工作原理概括为:1)推理机是根据已经创建或已读入 RDF 三元组的信息资源和本体结构内所包含的信息而建立的,它通过 Jena 第三方库的 Reasoner 库进行操作;2)把推理机和需要进行查询推理的本体绑定在一起,从而得到检索的模型对象(InfGraph);3)借助本体结构和模型的 API 对已建立的模型对象进行操作和处理,并通过对概念的推理,完成基于语义的信息检索,从而得到想要的结果。

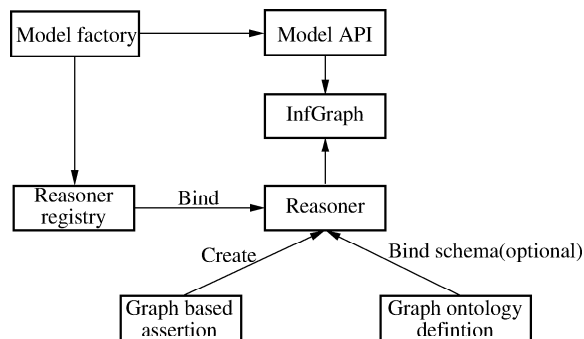


图 2 Jena RDFS 推理机结构

2.2 规则的确定与编写

通过多次模拟生物出生与死亡的过程来观察湿地食物链的发展状况,其算法设计的难点是物种出生与死亡的随机性,以及生物因捕食与被捕食导致的数量波动。借助 SWRL 设计规则,在原有本体架构上进行操作,具体的操作过程为:根据 OWL 文件得到本体 Model 对象,从而确定本体中生物的种类与具体数目,以贴合实际环境中生物数量的动态演变过程。部分规则展示如下:

自然出生规则为全部生物添加了 Data property 属性(Living_state),并将起始值都设置为 Live,再通过 For 循环以及 Rule.append()函数来对每种生物的个体数目进行随机计算。其规则推理格式为:

$[(? \text{ Subject RDF:type } ? \text{ Biological species}) \rightarrow (? \text{ Biological individual living_state live})]$

自然死亡规则为部分个体添加了新的 Data property 属性(Dead_state),并将它们设置为 Dying。其规则推理格式为:

$[(? \text{ Subject dead_state dying}) \rightarrow (? \text{ Subject living_state dead})]$

非自然死亡规则是提取本体中因被捕食而死亡的个体,检索食物链中含有 Eat 属性的生物。其规则推理格式为:

$[(? \text{ Subject eat } ? \text{ object}) \rightarrow (? \text{ Subject living_state dead})]$

以上规则通过检索生物种类,以获得生物属性,从而达到以更改属性来推理生物数量的目的。

2.3 稳态与非稳态

本研究旨在模拟湿地食物链的变化情况,主要包含稳态模拟和非稳态模拟两种类型。稳态模拟是指模拟食物链中因生物自然出生、死亡、捕食和被捕食而引起的数量变化。非稳态模拟建立在稳态基础之上,模拟食物链中一种生物消亡后剩余生物的数量变化。非稳态规则的设计是以单条食物链的变化来覆盖整个食物链的稳态局面^[13],并在推理结束后,借助 Jena 框架对推理结果进行解析并计算出每个生物的存活数目,从而来判断非稳态的模拟结果是否符合实际情况。上述功能实现的代码如下:

```
//定义 StringBuffer 类型的 Rule 对象
StringBuffer rule = new StringBuffer();
//编写 Rule 规则的具体内容,这里的规则具体是
//[ (? Subject RDF:type ? Biological species) → (? Subject living_state dead) ]
//
Rule.append("? Subject");
Rule.append("RDF:type");
//
Rule.append(montage(prefixString, slist1.get(i)));
Rule.append("? Subject");
//
Rule.append(montage(prefixString, "living_state"));
Rule.append(" 'dead'");
```

首先根据输入信息来获取希望消灭的物种,然后根据物种的生活习性和形态特征来检索该物种在食物链中所处的位置,并对高层级的捕食者信息进行归纳。通过递归算法来执行非自然死亡规则,将捕食者的 Living_state 值全部推导成 Dead 值,再将这些推导值覆盖至稳态模拟结果,最终导致单条食物链断链或消失,其余食物链依旧处于稳态,从而达到模拟的效果。

3 实验建模验证

为了更加直观地分析可视化界面,本研究使用基于 Java 的 GUI 工具包(Java swing 与 HTML)来设计界面(图 3),并组合 Jfreechart 对每次模拟后的生物数量进行折线图的绘制,以模拟前后的生物存活量作为比对标准,结合实际对稳态与非稳态的模拟结果进行比对分析。

本研究以南京市郊区某池塘为样本进行稳态和非稳态模拟实验。该池塘面积约 12 000 m²,平均水深 2.5 m,存在着较完整的生态系统食物网^[14]。

在该池塘设置 5 处采集水样点,分别位于池塘的四角与中心位置。对整块湿地及其周边的生物数量进行放大估算^[15],而对于部分大型生物则直接计算,具体如表 2 所示。



图 3 湿地生态系统的模拟界面

表 2 池塘生物统计情况

物种	所属类别	总量
硅藻	浮游植物类	2.7×10^{10} 株
轮虫	浮游动物类	1.5×10^8 只
黑藻	水生植物类	7.4×10^4 株
黑壳虾	甲壳动物类	6.0×10^4 只
草鱼	草食性鱼类	1 000 条
青蛙	两栖动物类	200 只
鲶鱼	肉食性鱼类	100 条
野鸭	水鸟类	6 只

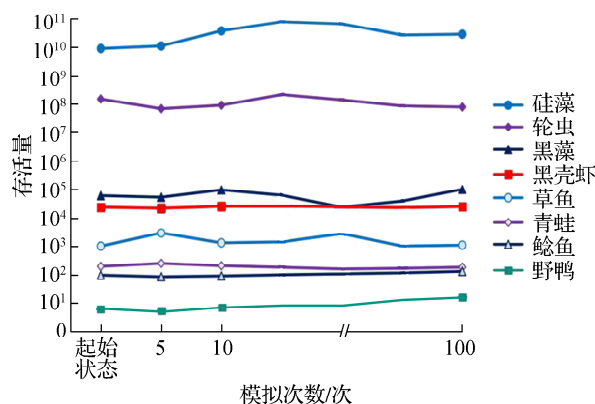


图 4 稳态模拟后的存活量变化情况

从图 4 可以看出,模拟后的湿地生物总量都在起始状态湿地生物总量的上下浮动。以黑壳虾为例,100 次模拟后的数量与起始状态数量相当,表明此模拟生物存活量的改变幅度不大,符合稳态模拟的预期效果。

当用户输入特定的消亡物种时,会依据递归算法来查询其所在食物链包含的所有物种,并通过折线图来表示。图 5 为非稳态模拟的结果,从图中可看出每进行一次演化,一个物种就会消亡,其他物种依旧处于稳态。以黑壳虾消亡为例,模拟后黑壳虾的生物数量归零,以其为主要食物的草鱼数量会受到影响,存活量大幅减少,5 次模拟后数量降至 700 条,比起始状态减少了 30%,100 次模拟后数量较起始状态下降了近 90%,仅剩余 100 条左右。这充分体现出底层物种的消亡对高层级物种的影响,此结果符合湿地非稳态的预期效果。

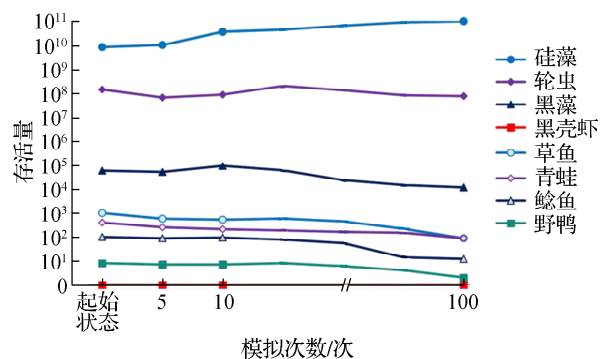


图 5 非稳态模拟后的存活量变化情况(黑壳虾消亡)

为验证本研究的模拟结果,2017 年初,将野鸭驱逐出该池塘,2 年后草鱼数量较 2017 年上涨了 90% 左右,数量增至近 1 900 条,与模拟结果相近,其契合程度在可接受范围之内,证实该模拟方法基本可行。此外,从应用角度来看,对湿地食物链的模拟为养殖池塘水质的管理提供了理论依据和参考。

4 结 语

本文基于 Jena 引擎从语义网和生态推理两方面完善了对湿地食物链语义的模拟,并借助可视化界面验证了模拟结果,创新性地实现了将语义网的应用技术拓展到湿地生态监测与管理领域。此外,运用实例对语义模型进行了验证,表明该模型能较好地支持稳态与非稳态的动态模拟,对当今湿地的环保监控具有现实的指导意义。但是湿地实体之间的关系错综复杂,后续实验中关于湿地本体和规则的研究还需进一步深化,以完善推理方式。

参考文献:

- [1] CHEN W J, HE B, ZHANG L, et al. Developing an integrated 2D and 3D WebGIS-based platform for effective landslide hazard management[J]. International Journal of Disaster Risk Reduction, 2016(2): 26 – 38
- [2] CUI B S, YANG Z F. Research advances in wetland ecosystem models[J]. Advance in Earth Sciences, 2001, 16(3): 352 – 358
- [3] WALKET B H. Conserving biological diversity through ecosystem resilience[J]. Conservation Biology, 1995(3): 9 – 12
- [4] CHEN W J, HE B, MA J T, et al. A WebGIS-based flood control management system for small reservoirs; a case study in the lower reaches of the Yangtze River[J]. Journal of Hydroinformatics, 2017, 19(2): 299 – 314
- [5] TIM B L. Design issues[EB/OL]. (2014 – 03 – 06)[2020 – 04 – 15]. <http://www.w3org/DesignIssues/>
- [6] RAZMERITA L, ANGEHM A, MAEDCHE A. Ontology-based user modeling for knowledge management systems[C]//JOHN A. Proceedings of the user modeling. New York: IEEE Access, 2003: 213 – 217
- [7] CHRISTIAN B, HEATH T, BERNERS L T. Linked data the story so far[J]. International Journal on Semantic Web and Information Systems, 2009(3): 1 – 22
- [8] 刘峤, 李杨, 段宏, 等. 知识图谱构建技术综述[J]. 计算机研究与发展, 2016, 53(3): 582 – 600
- [9] 袁凯琦, 邓扬, 陈道源, 等. 医学知识图谱构建技术与研究进展[J]. 计算机应用研究, 2018, 35(7): 1930 – 1936
- [10] 钱凌. 一个基于本体和规则推理的查询系统的设计与实现[D]. 南京: 东南大学, 2006
- [11] 张宁祥. 移动端关联式思维导图设计与实现[D]. 西安: 西安电子科技大学, 2018
- [12] 罗钧旻, 王英, 见伟平. 模糊 Jena 的研究[J]. 计算机与数字工程, 2011, 39(4): 71 – 73
- [13] 柴留祥, 何丰. 基于 Jena 及其本体推理的研究[J]. 计算机技术与发展, 2011, 21(11): 117 – 119
- [14] 于金海, 冷晓飞, 李春艳, 等. 黄海北部仿刺参养殖池塘关键环境因子的周年变化与管理研究[J]. 大连水产学院学报, 2009(6): 59 – 66
- [15] 王摆, 田甲申, 董颖, 等. 海蜇—对虾—缢蛏—牙鲆综合养殖池塘的食物网分析[J]. 水产科学, 2019, 38(3): 327 – 332

(责任编辑: 谭彩霞)