Advancing Underutilized Crops Knowledge using SWRL-enabled Ontologies - A survey and early experiment

Abba Lawan[1], Abdur Rakib[1], Natasha Alechina[2], and Asha Karunaratne[3]

[1] School of Computer Science
The University of Nottingham, Malaysia Campus
{khyx3alw, Abdur.Rakib}@nottingham.edu.my
[2] School of Computer Science
The University of Nottingham, UK
natasha.alechina@nottingham.ac.uk
[3] Crops For the Future Research Centre (CFFRC), Malaysia
asha.karunaratne@cffresearch.org

Abstract. Due to their powerful knowledge representation formalism and associated inference mechanisms, ontology-based approaches have been increasingly adopted to formally represent domain knowledge. We propose the use of ontologies to advance knowledge-sharing on underutilized crops and propose how to integrate those ontologies with rules for added expressiveness. We first present a survey on the use of ontologies in the field of life-sciences with emphasis on crop-related ontologies, and justify why we need a new formalism. We then present the UC-ONTO (an Underutilized Crops Ontology) as a case study showing the integration of OWL (Web Ontology Language) ontologies with Semantic Web Rule Language (SWRL) rules for added expressiveness.

Keywords: Crop Ontology, OWL Ontology, Semantic Web Rule Language, Underutilized Crops, Knowledge representation, Reasoning

1 Introduction

A shared concern among knowledge engineers and domain experts is the formalization of knowledge domains with minimum ambiguities. One possible solution is the use of ontologies, which serve as an explicit specification of terms that formally define and structure the concepts of a shared domain and the relationships that exist between them [6]. In essence, ontologies help to provide a common understanding of a domain while enabling knowledge-sharing among experts and software tools.

In the field of Life Sciences, ontologies have proven to be increasingly valuable by providing the semantic framework for defining domain concepts and their relationships coupled with automated reasoning and analysis tools that support knowledge organization and sharing [45, 51]. Thus, breathing new life into biological/agricultural classifications by providing common understanding of terms among researchers and bridging the gaps in semantic and organizational differences between tools and databases. In the Crops domain, various ontologies do exist, such as the Crop Ontology [7], the Plant Ontology [12], the Gene Ontology [5] and the popular AGROVOC [10, 11], among others. However, information on specific crops (categorized as Underutilized) hardly exist in these crops vocabularies and ontologies (see figure 1). *Underutilized Crops* [37], are those that are currently neglected though previously grown and consumed with considerable nutritional and/or market value [15].

Moreover, existing crop-related ontologies such as those listed above are usually available in the Open Biomedical Ontology (OBO) format, being developed using an open-source OBO-Edit environment [12]. Though, OWL versions of these ontologies are also provided in most cases, and since tools for converting OBO to OWL ontologies do exist, such as OboInOwl[4], it is always easier to adapt OBO ontologies to OWL-based development environments and semantic web applications [19].

The high-expressive power of OWL - owed to its rich collection of constructs and its support for rule languages such as SWRL, offer developers greater flexibility in domain modeling and expressing declarative knowledge (using rules) over ontologies. Moreover, the integration of OWL ontologies with rules help in expressing implicit domain knowledge by utilizing existing rule-based reasoning supports. With OWL being the standard ontology language approved by the World Wide Web Consortium (W3C), efforts to provide Crop Ontologies in RDF and OWL format will undoubtedly boost crop knowledge-sharing and allows interoperability between participants of the platform and beyond. Also Semantic Web applications can be developed to utilize such ontologies.

The focus of this paper is to practically explore how rules can be used to increase the expressive powers of ontologies focusing on the SWRL rules. By so doing, we develop OWL ontology for the Underutilized Crops domain and further integrate the ontology with basic SWRL rules. One significant role of ontologies is that they facilitate knowledge reuse. As such, we utilize some domain-independent as well as crop ontologies in the underutilized crops ontology. FAOs geopolitical ontology and OWL-time ontology are some of the ontologies imported. We hope in the future, to see more complex representation of general crops knowledge other than concept hierarchies and the simple is-a and part-of relationships currently offered by the popular crop ontologies (section 3.2). In a similar gesture, authors of Crop Ontology: vocabulary for crop-related concepts in [32], have suggested the use of OWL-DL in the future works of for added expressiveness and complex domain modeling.

The remainder of the paper goes as follows: We present our motivation and the scope of the review in the next section and section 2 discusses the expressive powers of OWL and the need for integrating OWL ontologies with rules. This is followed by a brief introduction of the SWRL formalism. Section 3 presents the relevant works on using ontologies to model a knowledge domain with emphasis on the crops domain. Section 4, which introduces the UC-ONTO, describing the

[4] $http://www.bioontology.org/wiki/index.php/OboInOwl:Main_Page$

problem background, approaches, and development methodology. In section 5, we present an implementation of the SWRL rules extension for the UC-ONTO case study. We evaluate the ontology and SWRL rule assertions in section 6 and finally conclude in 7.

1.1 Motivation and Scope

Inspired by [32, 28], the work presented in this paper is part of a PhD project which among others, aims at using ontologies (and related formalisms) to standardize knowledge representation in the field of Underutilized Crops. The review part of our work focuses on extending ontologies with rules and is restricted to the literature that discusses the use of SWRL rules and its expressive extensions. However, evaluation of computational and reasoning capabilities of OWL + SWRL combination is not provided in this paper.

Our work can serve as an introduction to the rule-based formalisms and a guide to new researchers and non-logic experts that plan to utilize these formalisms for their problem domain. The complete ontology can be found online[5].

2 Background

In this setion, we discuss the expressive powers of OWL ontologies and the importance of integrating such ontologies with SWRL rules. The SWRL formalism is then briefly discussed highlighting its condition for decidability, the DL-safeness.

2.1 OWL Expressiveness and the need for rules

Description Logic (DL)-based OWL is the standard ontology language approved by W3C for modeling domain knowledge in the Semantic Web [55]. In the quest for a more expressive web ontology language, the OWL family [33, 50], evolves from OWL 1 that consists of three sub-languages namely: OWL-Lite, OWL-DL and OWL-Full, to the more recent OWL 2, which is also partitioned into OWL2EL, OWL2QL and OWL2RL [34]. These languages offer different expressiveness and computational desirability with the current version, OWL 2, able to provide a wider range of constructs such as transitive and inverse properties, cardinality restrictions, as well as inheritance among others.

However, despite its success in achieving hierarchical definition and efficient classification of domain concepts when compared to Resource Description Framework (RDF) its predecessor, OWL suffers from other expressive limitations, such as its lack of support for composite role definition between concepts. Hence, there is the need for a more expressive domain modeling language than OWL as established by various researchers citing both theoretical and practical example

[5] https://www.dropbox.com/s/4l4bbcdus0bv7zm/BG1BG2MergeFinalOWL2RL.owl?dl=0

[24, 13, 35, 29, 17, 26]. Rule formalisms were consequently adopted to provide the needed support for more expressive power to the OWL language both being fragments of the classical logic.

The expressive limitations of OWL and the choice for Rules are not just mere coincidences. While OWL-DL ontologies provides simple, reusable and easy to understand knowledge models, they lack the expressiveness offered by rules. Furthermore, the rule formalisms apart from being in common practice, provides an efficient reasoning support to ontologies with the added expressiveness.

The integration of OWL-DL and SWRL provides many advantages that cannot be achieved using either OWL DL or Horn rules alone. Moreover, extending ontologies with rules is favored due to the wide acceptance of rules in knowledge modeling and the success of Rule-based formalisms in commercial applications among others.

2.2 SWRL Formalism

In the literature, various formalisms exist to extend DL ontologies with rules and they are often classified into Hybrid (loosely coupled) and Homogenous (tightly coupled) approaches [1, 40]. Among the homogenous formalisms, SWRL has received a considerable attention from the Semantic Web community over the last few years [24, 23, 8, 43] and forms the basis of our survey. The classification of the rule languages into hybrid or homogenous can be in terms of syntax, semantics or both. we refer the interested reader to [28] for a more detailed list of the popular formalisms.

SWRL is a direct extension of OWL-DL that exploits its model theoretic semantics while combining the syntaxes of OWL-DL with that of Rule-ML. SWRL, originally called ORL (OWL Rule Language) [25], is a horn-like rule formalism having antecedent (body) as well as consequent (head) with both having conjunctions of rule atoms. Usually in the form:

$$atom_1, atom_2, atom_3, \cdots, atom_n \rightarrow atom_1, atom_2, atom_3, \cdots, atom_l$$

As initially defined in [24] and further discussed in [28, 8], SWRL extensions are bindings that provides a mapping between variables used in the rules to elements of a given domain. Ontology elements in SWRL are identified using their URI[6] references. For technical details on the syntax and semantics of SWRL, we refer the reader to [25] and for background theory and implementations of Description Logic, see [2].

Decidability of SWRL Formalism: DL safeness. SWRL rules added to OWL ontologies need to be DL-safe to retain the decidability offered by OWL and ensure sound and complete reasoning over their ontologies. A DL-safe SWRL rule [35], ensures that only named concepts are used in the rules to avoid generating anonymous individuals during inference. In other words, only those

[6] Uniform Resource Identifiers, strings similar to URLs, used to identify all objects on the semantic web

variables (or named individuals) already declared in the antecedent may be used in the inference no new concepts may be introduced.

3 The Context: Ontologies in the Crop domain

This section discusses the relevant works of using ontologies to model a knowledge domain with emphasis on the crops domain. Starting with classifying ontologies, we present the popular crop-related ontologies showing their inadequacy in representing underutilized crops knowledge. Finally, we point the benefits of ontologies in life-sciences.

3.1 Ontologies as Knowledge Repositories - Classification

Ranging from generic taxonomies to specific application-based knowledge models, ontologies have commonly been categorized into three levels [48, 27] namely: (i) The foundational ontologies, (ii) Domain ontologies and (iii) Application-level ontologies.

Foundational Ontologies also called top-level or reference ontologies, provide general taxonomies with multi-domain knowledge. The Unified Foundational Ontology (UFO) [20], Basic Formal Ontology (BFO) [47], General Formal Ontology (GFO) [21], and the GFO-Bio [22] among others, are common examples of foundational ontologies. Foundational ontology being a repository of general knowledge provides a means for semantic evaluation of lower ontologies such as the domain ontologies.

Domain ontologies on their part provide conceptual and more descriptive definition of terms within scoped domain boundaries, usually for an organization or knowledge community comprising of concepts, their relationships and individual instances. They offer a common vocabulary for sharing, reuse and standardizing knowledge of a specific community or domain of discourse. Larger domain ontologies are sometimes referred as upper-domain, such as BIOTOP [4], which is an upper-domain ontology for molecular biology linking smaller domain ontologies with the BFO, FAOs AGROVOC [10, 11, 39], which has in the past thirty years grown from simple multilingual agricultural index to a Linked-Open-Data (LOD) set. Other examples of domain ontologies include the Crop Ontology [46], Plant ontology [12], Gene Ontology [5], and the Underutilized Crops Ontology (UC-ONTO), which is currently under development by Crops For the Future Research Center (CFFRC)[7].

Application ontologies are developed to be used for specific applications and usually utilize the domain ontologies by restricting conceptualizations to model a specified application domain. For example, the Food Ontologies for nutritional applications in [42, 9, 30] and sensor ontologies for manufacturing application reviewed in [44].

[7] http://www.cropsforthefuture.org/

3.2 Domain Ontologies in Life Sciences

In this section, we review some of the popular crop-based domain ontologies with emphasis on the expressiveness provided by their development languages.

Gene Ontology. The Gene Ontology [5] is a popular biological upper-domain ontology developed by the Gene Ontology Consortium to establish standards in the representation of gene-related knowledge for various species of organisms. It is designed as a collaborative community-based ontology development effort providing gene ontologies with three components: molecular functions, biological processes and cellular components, their annotations as well as tools to access and process the ontologies [51]. Like many existing biological ontologies, the Gene Ontology is available mostly in the OBO format. Though, OWL versions of these ontologies are provided in some cases. However, OBO ontologies even when converted to OWL formats are less expressive. This is due to the well-defined semantics, interoperability with other ontologies and the various tools and services that facilitate development, maintenance and reuse of OWL ontologies,

Plant Ontology. Considering it as a comparative tool for plant anatomy and genomic analysis [12], the Plant Ontology is developed to provide formal specification of terms that describe plant anatomy, morphology and growth stages with the first and later developed as components of the whole ontology. Plant Ontology utilizes the data model available in the Gene Ontology (GO) [5], for annotating the plant anatomy and growth stage ontologies with gene expressions and phenotype data from the GO. Similar to the Gene Ontology, the Plant ontology is also guided by the OBO Foundry ontology for seamless collaboration with other biological ontologies [12] and most of the ontology is available in the OBO format . However, some parts of the ontology are available in the OWL format. For efficient comparison of disparate data with similar terms, such as that of genomics, the use of ontologies is necessary for data curation and analysis as it helps to provide common structured vocabulary that permits automated reasoning.

Crop Ontology. Citing data management, accessibility and retrieval challenges as the main motivation, Generation Challenge Program (GCP) [8] developed the Crop Ontology to facilitate community sharing of crop-related information by semantically characterizing and annotating historic generic crop data sets (traits, phenotype, germplasm, breeding, etc.) [7, 46]. With a simple web-based interface and the help of semantic experts as moderators of the ontologies, the Crop Ontology platform allows community-based collaborative ontology development, where users can create and add their own ontologies to the pool. Originally in

[8] http://www.pantheon.generationcp.org

Open Biomedical Ontology (OBO) formats, the Crop ontology has evolved to utilize more terminological standards such as RDF and OWL [32].

With OWL being the most widely used standard for developing ontologies, effort to provide crop ontologies in RDF and OWL format will no doubt improve knowledge-sharing among researchers. This is basically due to the high expressiveness, efficient reasoning support, and the added advantage integrating OWL ontologies with declarative rule languages such as SWRL. Moreover, Semantic Web applications can be developed to easily utilize OWL ontologies.

From the foregoing exploration, the ontologies are able to provide an efficient and comprehensive hierarchical representation of their domains with common roles between concepts being of the form is-a and part-of relationships, which simply put, denotes that a concept is either a subtype of the connecting concept or that of the root/ancestral concept (see Fig. 1 on the right panel). However, they seem to lack complex representation of roles or relationships between concepts, which is one of the major differences between ontologies and hierarchical taxonomies such as thesauri. In a similar gesture, authors of Crop Ontology: vocabulary for crop-related concepts in [32], have suggested the use of OWL-DL in their future work for added expressiveness and complex domain modeling.

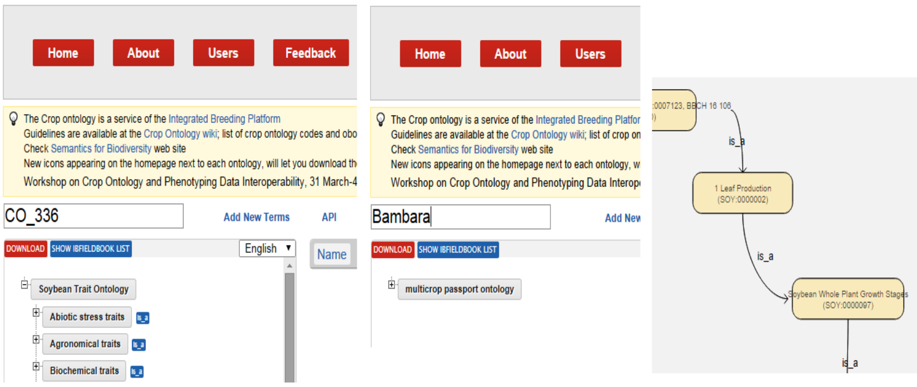

Fig. 1. Crop Ontology Curation Tool showing: general crops info. (left), lack of similar info. on underutilized crop (center) and the simple is-a relationship (right). Image source: http://www.cropontology.org/

Contribution of ontology to the crops domain (not exclusive though) can be summarized in the following points: i) For organization and sharing crop information ii) As integrative comparative tools iii) For standardization of domain knowledge and also iv) Useful for developing semantic web applications.

It should be noted however, that these contributions are not exclusive to the crops domain as they are simply benefits brought about by the use of ontologies. Though, the comparability advantage is more pronounced in the field of life sciences. Moreover, as ontologies are designed to be the knowledge modeling formalisms for an open-web [16, 6], their advantages may not be restricted to a particular domain. A review on the recent trends and applications of ontologies citing examples from various domain ontologies is presented in [14] and text book detail on uses of ontologies in bio-informatics is given in [49].

4 Case Study: Underutilized Crops Ontology (UC-ONTO)

This section introduces the motivation as a case study on the use of SWRL rules for integrating ontologies in the crops domain. The approach and specific design issues as related to our case, underutilized crops knowledge modeling are discussed.

Problem Background. With the United Nation's decade long efforts on biodiversity and food security, there is an awakening on the need to revitalize the cropping of neglected or underutilized crop species, many of which have the potential of providing food security as well as nutritional sustainability [15, 37, 53]. The Crops for the Future Research Center (CFFRC) , is one of the research bodies dedicated for research and development on Underutilized Crops. With many researchers working on different underutilized-crops related projects, there is a need for domain-level ontology to provide explicit specification of terms, the relationships between those terms and how they are related across the various research fields and outside partners.

Reusability Approach. As stated earlier, one of the benefits of developing ontology for a domain is knowledge reuse. Considering the available crop-domain collaborative ontologies (see section 3.2) and in line with the principle of ontology reuse, we ought not develop a new ontology but simply tailor these ontologies to present relevant information on underutilized crops species. Similar approach has been proposed in [3], where AGROVOC is used as a base vocabulary to develop the *CropOnt* a framework for relevant knowledge on crop production life cycle for individual farmers.

However, despite their nutritional, dietary-diversity, and economic importance [15], basic concepts definition on underutilized crop species are very rare and in some cases non-existent [53]. Consequently, most of the general crop ontologies do not have information on the underutilized or neglected crops due to the lack of available information on underutilized crops in general. In their book Global research on Underutilized Crops [37], the authors cited the lack of technical knowledge as one of the constraints to research and development on Underutilized crops.

Knowledge Gathering Approach. Two approaches have been considered in the early stages of our project: either to develop a complete Underutilized-crops ontology from scratch, or to utilize existing crop ontologies by importing relevant and shared concepts. The latter, which support knowledge-reuse and favored in the field of ontology engineering, was thus accepted.

To do this however, there is a need to analyze some of these general crop ontologies and critically evaluate them for possible integration, while considering compatibility issues. Specifically, how the Underutilized Crops ontologies, which share much of the concepts of the general crops, can fit together with proper source formats and linkages with the imported ontologies. Furthermore, this will not only support the reusability spirit of ontologies but will also save a great amount of development time on the part of CFFRC knowledge engineers. The choice will also ensure conformity of our ontology to the existing standards in crop-domain modeling.

4.1 UC-ONTO development methodology.

We employ the collaborative ontology development methodology, which is necessary to enable knowledge engineers work closely with the domain experts (underutilized-crops researchers in our context).

To achieve a comprehensive modeling, the general guidelines advised in the work of Noy and Mcguinnes [36], the METHONTOLOGY [18], DILIGENT [41], and the Onto-Knowledge methodology, were utilized. These guidelines help to structure the ontology engineering process by identifying important but non-obvious aspects, such as the target users of the ontology, supporting tools, and specifying what values can be allowed for properties. Other aspects that are apparent and also common to all methodologies - such as defining domain terms and roles, asserting their hierarchy, and filling the concept slots with individual instances - are performed iteratively for each source of data to populate the underutilized crops ontology. Similarly, our user-defined SWRL rules are added iteratively while ensuring the consistency of the ontology by invoking the *Pellet* reasoner. The major steps for UC-ONTO development can be summarized as follows: i) Ontology requirement specification ii) Domain knowledge gathering and conceptualization iii) Model implementation and iv) Evaluation of the model.

These steps were performed repeatedly for each component version of the UC-ONTO, leading to the final complete version. The two final stages were termed *versioning* and *assembly*. In 'versioning', we assign a label to represent each ontology fragment, specifying where it fits to the larger ontology. While in the 'assembly' stage, smaller ontology modules are put together and a reasoner is invoked to assert the overall classification and check for consistency. A common problem with the assembly stage however, is that for each module added to the main ontology, inconsistencies are bound to arise. As such, to minimize such inconsistencies, the assembly is carried out with the ontology Reasoner in active mode. Moreover, for each smallest ontology module assembled, the reasoner need to be invoked to check for the consistency. This way, it is easier to keep track of what causes the inconsistencies and where to correct them.

Other common issues in ontology development include, the failure to reuse existing ontologies in the beginning of development and also the failure to familiarize with basic domain concepts (by ontology engineers) - leading to the problems of modeling roles as classes and vice versa. Advisably, a domain expert should be available at all times to continuously check the progress of ontology modeling. This is because, while a Reasoner can cross-check inconsistencies arising from hierarchical representation and incorrect assertions, it is incapable of highlighting domain-related inconsistencies, among others.

4.2 The UC-ONTO

We have developed the first version of the underutilized crops ontology (UC-ONTO) using the Protégé 4.2 ontology editor. The ontology currently consists of SWRL built-ins, OWL-time ontology, and FAO geopolitical ontology as direct imports. This is because these ontologies being domain-independent and available in the OWL format can stand-alone without posing compatibility problems and inconsistencies.

While details on the development methodology and the aspects of the UC-ONTO (such as agronomic, physiological traits) are beyond the scope of this paper, we give a brief account of the composition of the ontology with a glimpse on the naming convention and structure. In the ontology, all crops related concepts are grouped together under the $DomainConcepts$ as super class and all other concepts such as $DaysOfWeek$, $TimeZone$, etc. offered by imported ontologies, are composed as siblings. The $UnderutilizedCrops$ class contains four sub classes with $Taro$, Tef, $Millet$ and $BambaraGroundnut$ class, which dominates most of the object properties such data-type property modeling in this version.

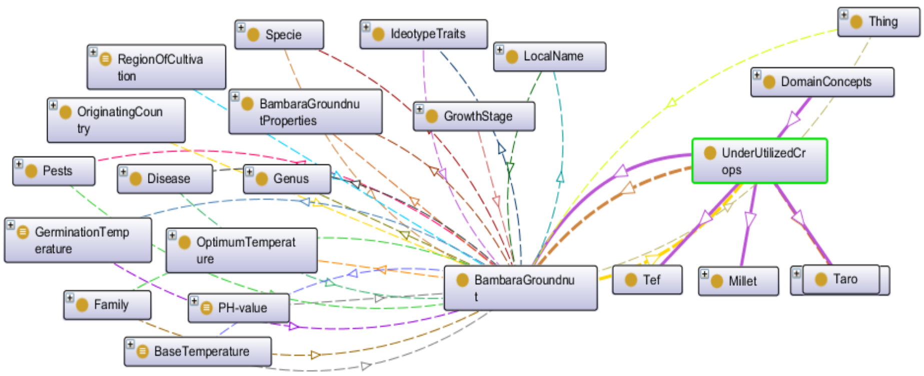

Fig. 2. Graphic view of concepts in UC-ONTO

The SWRL rules are written using the SWRL tab to specify more relationships between concepts on top of our ontology giving more flexibility to declarative property assertions in the UC-ONTO. Fig. 2 gives a partial graphic overview of the concepts and roles specified in the UC-ONTO.

5 Extending UC-ONTO with SWRL rules

This section presents an implementation of the SWRL rules extension for the UC-ONTO case study. The rules are intended to allow modeling declarative knowledge and for expressing complex roles (such as composite relations between concepts) that are not easily expressible with OWL alone.

The addition of our user-defined, DL-Safe, SWRL rules was delayed until the final version of the ontology was checked for consistency using *Pellet* reasoner and found to be consistent. Moreover, considering the main reason of using SWRL rules in our ontology, which is to express complex relations between domain concepts and utilize the SWRL built-ins to define and assert domain-specific concepts, it will still do no harm to our ontology if we express the OWL axioms using SWRL.

In the rules interface depicted in Fig. 3, we begin with a simple assertion in rule 8 that asserts a relationship between members of *BambaraGroundnut* and those of *BambaraGroundnutProperties* class using *hasProperty* relation. We then continue to assert more roles that are easily expressed with declaration, thereby extending the expressive power of the ontology. For example, the fourth rule:

$$\mathit{BambaraGroundnut}(?y), \mathit{Leaf}(?z), \mathit{isFeatureOf}(?z, ?y) \rightarrow$$
$$\mathit{hasLeafType}(?y, \text{"}\mathit{Trifoliate}\text{"})$$

States in simple terms, that if *BambaraGroundnut* class has a feature leaf, then it will be asserted that the leaf type is 'trifoliate'. However, since features such as leaf are not exclusive to *BambaraGroundnut* class, then unless the leaf individual is related to *BambaraGroundnut*, the leaf type trifoliate, cannot be asserted. Rules of these types that are based on certain conditions being true or otherwise, are hard to be expressed with OWL syntax alone.

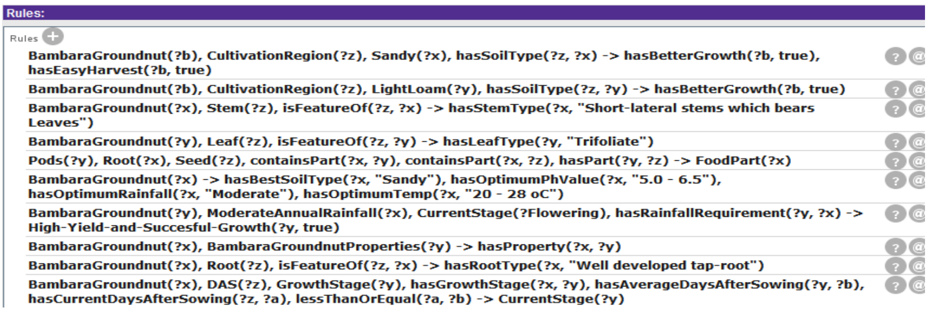

Fig. 3. Rules interface showing some user-defined SWRL rules

6 Evaluation

In this section, we evaluate the ontology and SWRL rules assertions by invoking the Pellet reasoner to classify and check for the consistency of the ontology.

Additional knowledge implicit in the crop ontology can then be inferred by this reasoner. Also to verify the conceptual facts and individual assertions, DL queries are used to probe the ontologies. We evaluate 2 to 3 queries for each SWRL rule,making a total of 46 DL queries. Results of frequent queries are saved and added as part of the ontology thereby evaluated automatically once the reasoner is invoked.

6.1 Reasoning and query processing

Using ontologies allow measuring performance at the design as well as run-time via a reasoner to compute the ontology classification and ensure consistency. As such a reasoner needs to be active and the ontology classified before writing any DL Queries. Our user-defined SWRL rules are validated by writing DL queries to check their inference or otherwise by the reasoner. For example, the query result of the sixth rule, determines the current 'growth stage' of a $BambaraGroundnut$,. The rule uses a SWRL built-in *'swrlb:lessThanOrEqual'*, to compare the days an individual $BambaraGroundnut(BG)$ is planted with the number of days asserted for the different growth stages (e.g. the flowering stage $hasAverageDaysAfterSowing = 50$). If there is a match, the reasoner will then assert this growth stage as the current stage of the individual BG. Results for some of the rules, which assert Datatype properties to $BambaraGroundnut$ individual, can be seen from the Inference provided by the Pellet reasoner in Fig. 4 (right).

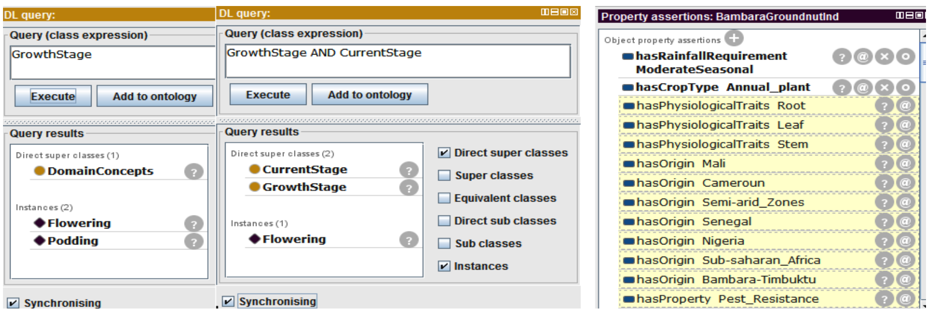

Fig. 4. Interface showing DL-Query Results (left) and Reasoner Inferences for Rules (right).

We would like to mention that the Underutilized crops ontology presented in this paper has: 24701 axioms, 111 classes, 397 individuals, with 94 object properties and 133 data properties. However, size and functionality of the ontology is expected to be continuously growing as more underutilized-crops data becomes available. The expressiveness of our ontology borders on SHOIN(D) algorithm and all SWRL rules added are DL safe, thereby decidable. The queries considered in the experiment were originated from the competency questions generated in our ontology engineering stage; due to space constraints we are unable to present those in details.

7 Conclusions and future work

In this paper, we propose a framework for representing knowledge using OWL ontologies and SWRL rules. Using the crops domain as a case study, we review and justify the need for integrating ontologies with rules. We present the SWRL-extended underutilized-crop ontology (UC-ONTO), highlighting our motivation, approach and development methodology. This is followed by an evaluation, which involves validation of the knowledge represented in the UC-ONTO through Reasoner inferences and writing appropriate DL queries. In the future, we aim to populate the ontology with more standard crop-related data from relevant Foundational Ontologies. Also we plan to publish the ontology and present the domain-knowledge to the public through a web-based, social-networking styled decision support system for underutilized crops.

For added expressiveness to our ontology, we intend to study and utilize the available SWRL extensions such as the first-order logic extension SWRL-FOL [40], the non-monotonic extensions for dealing with negation, exclusion and rule priority as in [8], and the X-SWRL [31], which allows for dealing with existential quantification of new individuals. Others extensions considered important includes the Fuzzy-SWRL [38], vague-SWRL [52], and SWRL-F [54] for modeling imprecise knowledge - a situation commonly encountered when dealing with domain experts, especially in the field of crops where informal and undocumented practices still hold sway.

Acknowledgement. This work is partially supported by a CFFRCPLUS CropBase PhD scholarship scheme.

References

1. Antoniou, G., Damásio, C.V., Grosof, B., Horrocks, I., Kifer, M., Maluszynski, J., Patel-Schneider, P.F.: Combining Rules and Ontologies. A survey. Tech. rep. (2005)
2. Baader, F., Calvanese, D., McGuinness, D.L., Nardi, D., Patel-Schneider, P.: The Description Logic Handbook: Theory, implementation, and applications. Cambridge University Press (2007)
3. Bansal, N., Malik, S.K.: A framework for agriculture ontology development in semantic web. In: Proceedings - 2011 International Conference on Communication Systems and Network Technologies, CSNT 2011. pp. 283–286 (2011)
4. Beisswanger, E., Schulz, S., Stenzhorn, H., Hahn, U.: BIOTOP : An Upper Domain Ontology for the Life Sciences. Applied Ontology 3(4), 205–212 (2008)
5. Berardini, T.Z.: The Gene Ontology in 2010: Extensions and refinements. Nucleic Acids Research 38 (2009)
6. Berners-Lee, T., Hendler, J., Lassila, O.: The semantic web. Scientific American pp. 29–37 (May 2001)

7. Bruskiewich, R., Davenport, G., Hazekamp, T., Metz, T., Ruiz, M., Simon, R., Takeya, M., Lee, J., Senger, M., McLaren, G., Van Hintum, T.: Generation Challenge Programme (GCP): standards for crop data. Omics : a journal of integrative biology 10(2), 215–9 (2006)
8. Calero, J.M.A., Ortega, A.M., Perez, G.M., Blaya, J.A.B., Skarmeta, A.F.G.: A non-monotonic expressiveness extension on the semantic web rule language. J. Web Eng. 11(2), 93–118 (2011), `http://dl.acm.org/citation.cfm?id=2230896.2230897`
9. Cantais, J., Dominguez, D., Gigante, V., Laera, L., Tamma, V.: An example of food ontology for diabetes control. In: Working notes of the ISWC 2005 Workshop on Ontology Patterns for the Semantic Web. p. 9 (2005)
10. Caracciolo, C., Morshed, A., Stellato, A., Johannsen, G., Jaques, Y., Keizer, J.: Thesaurus maintenance, alignment and publication as linked data: The AGROOVOC use case. In: Communications in Computer and Information Science. vol. 240 CCIS, pp. 489–499 (2011)
11. Caraccioloa, C., Stellatob, A., Morsheda, A., Johannsena, G., Rajbhandaria, S., Jaquesa, Y., Keizera, J.: The agrovoc linked dataset. Semantic Web 4(3), 341–348 (2013)
12. Cooper, L., Walls, R.L., Elser, J., Gandolfo, M.a., Stevenson, D.W., Smith, B., Preece, J., Athreya, B., Mungall, C.J., Rensing, S., Hiss, M., Lang, D., Reski, R., Berardini, T.Z., Li, D., Huala, E., Schaeffer, M., Menda, N., Arnaud, E., Shrestha, R., Yamazaki, Y., Jaiswal, P.: The plant ontology as a tool for comparative plant anatomy and genomic analyses. Plant & cell physiology 54(2), e1 (Feb 2013)
13. Cregan, A., Mochol, M., Vrandecic, D., Bechhofer, S.: Pushing the limits of owl, rules and protg – a simple example
14. Deshpande, N.J., Kumbhar, R.: Construction and applications of ontology: Recent trends. DESIDOC Journal of Library Information Technology 31, 84–89 (2011)
15. Ebert, A.: Potential of Underutilized Traditional Vegetables and Legume Crops to Contribute to Food and Nutritional Security, Income and More Sustainable Production Systems. Sustainability 6(1), 319–335 (Jan 2014)
16. Eiter, T., Ianni, G., Krennwallner, T., Polleres, A.: Reasoning web. chap. Rules and Ontologies for the Semantic Web, pp. 1–53. Springer-Verlag, Berlin, Heidelberg (2008)
17. Eiter, T., Lukasiewicz, T., Schindlauer, R., Tompits, H.: Well-founded semantics for description logic programs in the semantic web. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). vol. 3323 LNCS, pp. 81–97 (2004)
18. Fernández-López, M., Gómez-Pérez, A., Juristo, N.: Methontology: from ontological art towards ontological engineering. In: Proc. Symposium on Ontological Engineering of AAAI (1997)
19. Golbreich, C., Horridge, M., Horrocks, I., Motik, B., Shearer, R.: OBO and OWL: Leveraging semantic Web technologies for the life sciences. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). vol. 4825 LNCS, pp. 169–182 (2007)
20. Guizzardi, G., Wagner, G.: Using the Unified Foundational Ontology (UFO) as a foundation for general conceptual modeling languages. In: Theory and Applications of Ontology: Computer Applications, pp. 175–196. Springer Netherlands (2010)
21. Herre, H.: General Formal Ontology (GFO): A foundational ontology for conceptual modelling. In: Theory and Applications of Ontology: Computer Applications, pp. 297–345. Springer Netherlands (2010)

22. Hoehndorf, R., Loebe, F., Poli, R., Kelso, J., Herre, H.: GFO-Bio: A biomedical core ontology. Applied Ontology 3(4), 219–227 (2008)
23. Holford, M.E., Khurana, E., Cheung, K.H., Gerstein, M.: Using semantic web rules to reason on an ontology of pseudogenes. Bioinformatics (Oxford, England) 26(12) (Jun 2010)
24. HORROCKS, I., PATELSCHNEIDER, P., BECHHOFER, S., TSARKOV, D.: OWL rules: A proposal and prototype implementation. Web Semantics: Science, Services and Agents on the World Wide Web 3(1), 23–40 (Jul 2005)
25. Horrocks, I., Patel-schneider, P.F., Boley, H., Tabet, S., Grosof, B., Dean, M.: SWRL : A Semantic Web Rule Language Combining OWL and RuleML (2004), `http://www.w3.org/Submission/SWRL/`
26. Hustadt, U., Motik, B., Sattler, U.: Data Complexity of Reasoning in Very Expressive Description Logics. In: Kaelbling, L.P., Saffiotti, A. (eds.) Proc. of the 19th Int. Joint Conference on Artificial Intelligence (IJCAI 2005). pp. 466–471. Morgan Kaufmann Publishers (2005)
27. Keß ler, C., Raubal, M., Wosniok, C.: Semantic rules for context-aware geographical information retrieval. Smart Sensing and Context 5741 LNCS, 77–92 (2009)
28. Krisnadhi, A., Maier, F., Hitzler, P.: Owl and rules. In: Proceedings of the 7th International Conference on Reasoning Web: Semantic Technologies for the Web of Data. pp. 382–415. RW'11, Springer-Verlag (2011), `http://dl.acm.org/citation.cfm?id=2033313.2033320`
29. Krötzsch, M., Maier, F., Krisnadhi, A.A., Hitzler, P.: A Better Uncle for OWL: Nominal Schemas for Integrating Rules and Ontologies. In: 20th International World Wide Web Conference (WWW2011). p. 645 (2011)
30. Li, H.C., Ko, W.M.: Automated food ontology construction mechanism for diabetes diet care. In: Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, ICMLC 2007. vol. 5, pp. 2953–2958 (2007)
31. Li, W., Tian, S.: XSWRL, an Extended Semantic Web Rule Language and prototype implementation. Expert Systems with Applications 38(3), 2040–2045 (Mar 2011)
32. Matteis, L., Chibon, P., Espinosa, H., Skofic, M., Finkers, R., Bruskiewich, R., Hyman, G., Arnaud, E.: Crop Ontology: Vocabulary For Crop-related Concepts. In: Larmande, P., Arnaud, E., Mougenot, I., Jonquet, C., Libourel, T., Ruiz, M. (eds.) Proceedings of the first international Workshop on Semantics for Biodiversity. vol. 979 (2013)
33. McGuinness, D.L., Van Harmelen, F.: OWL Web Ontology Language Overview. W3C recommendation 10, 1–22 (2004), `http://www.w3.org/TR/owl-features/`
34. Motik, B., Grau, B.C., Horrocks, I., Wu, Z., Fokoue, A., Lutz, C.: OWL 2 Web Ontology Language Profiles (2009), `http://www.w3.org/TR/owl2-profiles/`
35. Motik, B., Sattler, U., Studer, R.: Query answering for owl-dl with rules. Web Semant. 3, 41–60 (Jul 2005)
36. Noy, N.F., McGuinness, D.L.: Ontology development 101: A guide to creating your first ontology. Tech. rep., Stanford (2001)
37. Padulosi, S., Hodgkin, T., Williams, J.T. and Haq, N.: Underutilized Crops: Trends, Challenges and Opportunities in the 21st Century. In: Jackson, J.E., Rao, V., M. (eds.) Managing plant genetic diversity, pp. 323–338. CAB International (2002)
38. Pan, J.Z., Stoilos, G., Stamou, G., Tzouvaras, V., Horrocks, I.: f-swrl: A fuzzy extension of swrl. Journal on Data Semantics, special issue on Emergent Semantics (2006)

39. Pazienza, M.T., Stellato, A., Tudorache, A.G., Turbati, A., Vagnoni, F.: An architecture for data and knowledge acquisition for the semantic web: The AGROVOC use case. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). vol. 7567 LNCS, pp. 426–433 (2012)
40. Peter F. Patel-Schneider, (Bell Labs Research, L.T.: A Proposal for a SWRL Extension towards First-Order Logic (2005), http://www.w3.org/Submission/SWRL-FOL/
41. Pinto, H.S., Tempich, C., Staab, S.: Diligent: Towards a fine-grained methodology for distributed, loosely-controlled and evolving engingeering of ontologies. In: Proceedings of the 16th European Conference on Artificial Intelligence. pp. 393–397. IOS Press (2004)
42. Pizzuti, T., Mirabelli, G., Sanz-Bobi, M.A., Goméz-Gonzaléz, F.: Food Track & Trace ontology for helping the food traceability control. Journal of Food Engineering 120(1), 17–30 (2014)
43. ROSATI, R.: On the decidability and complexity of integrating ontologies and rules. Web Semantics: Science, Services and Agents on the World Wide Web 3(1), 61–73 (2005)
44. Schlenoff, C., Hong, T., Liu, C., Eastman, R., Foufou, S.: A literature review of sensor ontologies for manufacturing applications. 2013 IEEE International Symposium on Robotic and Sensors Environments (ROSE) pp. 96–101 (2013)
45. Schuurman, N., Leszczynski, A.: Ontologies for bioinformatics. Bioinformatics and biology insights 2, 187–200 (2008)
46. Shrestha, R., Mauleon, R., Simon, R., Balaji, J., Channelière, S., Alercia, A., Senger, M., Manansala, K., Metz, T., Davenport, G., Bruskiewich, R., McLaren, G., Arnaud, E.: Development of GCP Ontology for Sharing Crop Information. Nature Precedings p. 6 (Apr 2009)
47. Smith, B., Grenon, P.: Basic formal ontology (bfo). INFOMIS Reports (2006)
48. Staab, S., Studer, R.: Handbook on Ontologies. Springer Science & Business Media (2010)
49. Stevens, R., Lord, P.: Application of Ontologies in Bioinformatics. In: Handbook on Ontologies, pp. 735–756 (2009)
50. W3C OWL Working Group: OWL 2 Web Ontology Language Document Overview (2012), http://www.w3.org/TR/owl2-overview/
51. Walls, R.L., Athreya, B., Cooper, L., Elser, J., Gandolfo, M.a., Jaiswal, P., Mungall, C.J., Preece, J., Rensing, S., Smith, B., Stevenson, D.W.: Ontologies as integrative tools for plant science. American journal of botany 99(8), 1263–75 (Aug 2012)
52. Wang, X., Ma, Z.M., Yan, L., Meng, X.: Vague-swrl: A fuzzy extension of swrl. In: Proceedings of the 2Nd International Conference on Web Reasoning and Rule Systems. pp. 232–233. RR '08, Springer-Verlag, Berlin, Heidelberg (2008)
53. Williams, J., Haq, N.: Global research on underutilized crops: An assessment of current activities and proposals for enhanced cooperation. Bioversity International (2002)
54. Wlodarczyk, T.W., Rong, C., O'Connor, M., Musen, M.: Swrl-f: A fuzzy logic extension of the semantic web rule language. In: Proceedings of the International Conference on Web Intelligence, Mining and Semantics. WIMS '11, ACM, New York, NY, USA (2011)
55. World Wide Web Consortium: Semantic Web "Layer Cake", =http://www.w3.org/2004/Talks/0412-RDF-functions/slide4-0.html