

ПРЕДСТАВЛЕНИЕ ЗНАНИЙ И ПОЛУЧЕНИЕ НОВЫХ ЗНАНИЙ НА ОСНОВЕ ОНТОЛОГИЧЕСКОГО ПОДХОДА: ОБЗОР ТЕХНОЛОГИИ SWRL

Сагайда П. И., Винницкая Я. А.

Рассмотрена необходимость и возможность объединения онтологии и правил для решения проблем представления знаний и получения новых знаний на основе онтологического подхода. Приведена схема объединения онтологии OWL DL и правил SWRL. Описаны основные особенности использования технологии SWRL, а также рассмотрены составляющие данной технологии (атомы). Проведен анализ атомов технологии SWRL, выделены и рассмотрены основные их характеристики. На примере описана возможность исключения основных характеристик, которые отличают технологию SWRL от DLP, таких как: конъюнктивные последовательности, дескрипции классов, равенства и неравенства. Определены платформы реализации рассмотренной технологии.

Розглянута необхідність та можливість об'єднання онтології та правил для рішення проблем представлення знань та отримання нових знань на основі онтологічного підходу. Наведена схема об'єднання онтології OWL DL та правил SWRL. Описані основні особливості використання технології SWRL, а також розглянуто складові даної технології (атоми). Проведено аналіз атомів технології SWRL, виділено та розглянуто основні їх характеристики. На прикладі описана можливість виключення основних характеристик, які відрізняють технологію SWRL від DLP, таких як: кон'юнктивні послідовності, дескрипції класів, рівності та нерівності. Визначено платформи реалізації розглянутої технології.

In this research, the necessity and possibility of merging ontologies and rules to solve the problems of knowledge representation and acquisition of new knowledge based on the ontological approach. Shows a diagram of combining OWL DL ontologies and rules SWRL. Describes the basic features of the use of technology SWRL, and discusses the technology components (atoms). The analysis of atomic technology SWRL, isolated and examined their basic characteristics. On the example described the possibility of excluding the main characteristics that distinguish technology SWRL from DLP, such as conjunctive sequence descriptions of classes, equality and inequality. Defined platform implementation technologies considered.

Сагайда П. И.

канд. техн. наук, доц. ДГМА
paulsagayda@ukr.net

Винницкая Я. А.

студент каф. КИТ ДГМА
yana9209@bk.ru

ДГМА – Донбасская государственная машиностроительная академия, г. Краматорск.

УДК 004.853

Сагайда П. И., Винницкая Я. А.

ПРЕДСТАВЛЕНИЕ ЗНАНИЙ И ПОЛУЧЕНИЕ НОВЫХ ЗНАНИЙ НА ОСНОВЕ ОНТОЛОГИЧЕСКОГО ПОДХОДА: ОБЗОР ТЕХНОЛОГИИ SWRL

Актуальным вопросом на данном этапе развития интеллектуальных систем является получение новых знаний. Существует много разработанных систем представления знаний, но проблема обучения систем до сих пор открыта.

Основной задачей представления знаний является хранение знаний таким образом, чтобы программные продукты могли осмысленно обрабатывать их и достигнуть тем подобия человеческого интеллекта.

Одним из наиболее перспективных подходов к решению этой проблемы являются разработки в области семантической паутины (Semantic Web). Основной логической базой данных разработок есть логический вывод, который использует технологии, объединяющие в себе онтологии и правила [1, 2].

Целью статьи является обзор технологии SWRL, как инструмента объединения онтологии OWL DL и правил, определение особенностей использования и анализ основных составляющих данной технологии.

Необходимо рассмотреть необходимость, определение, схему и особенности технологии SWRL, как одного из вариантов решения проблемы представления знаний и получения новых знаний на основе онтологического подхода.

Онтологии – богатые средства описания предметной области, ограниченные средствами вывода следствий из имеющихся знаний. Они основаны на дескрипционных логиках и являются частью классической логики.

Языки правил – специально построены для вывода следствий, для эффективности поиска вывода обычно в основе языка лежат продукции или Хорновские правила, что сильно ограничивает описательные возможности языка. Правила основаны на подмножестве логики первого порядка и возможных расширениях. С помощью правил достигается большая выразительность.

Комбинация онтологий и правил позволяет объединить достоинства подходов и обойти их недостатки.

Существуют разные подходы: SWRL, DLP (Grosz), dl-programs (Eiter), DL-safe rules, Conceptual Logic Programs (CLP), AL-Log, DL+log [1].

Выделены две главные стратегии объединения онтологий и правил:

- семантическая интеграция, известная также как однородный или гомогенный подход;
- строгое семантическое разделение или гибридный подход.

Основная характеристика однородного подхода состоит в следующем:

- как онтологии, так и правила содержатся в логическом языке;
- нет различия между предикатами правил и предикатами онтологий;
- взаимодействие с жесткой семантической интеграцией;
- правила можно использовать для определения классов и свойств онтологии;
- примерами данного подхода являются SWRL и DLP.

Гибридный подход обладает следующими преимуществами:

- интеграция со строгим семантическим разделением между двумя слоями;
- онтология используется как концептуализация ПрО;
- правила не могут определять классы и свойства онтологии за исключением некоторых специальных отношений;
- примерами данного подхода является Answer Set Programming (ASP).

Технология однородного подхода SWRL.

A Semantic Web Rule Language (SWRL) – это технология, которая основана на объединении OWL и RuleML, то есть объединены онтологии OWL-DL и правила.

RuleML (Rule Markup Language) [2] представляет собой язык разметки для описания правил. Данный язык позволяет публиковать и обмениваться правилами, созданными в рамках разных систем и задач. Ядром RuleML является XML-вариант языка Datalog, который позволяет записать факты и правила в форме продукций.

Схема объединения онтологии OWL-DL и правил SWRL представлена на рис. 1. Результатом концептуализации предметной области являются концепты и свойства, которые получает онтология. Зависимости предметной области описывает технология SWRL. Выполнение запросов SQWRL основано на базе правил SWRL, а запросы на языке SPARQL обращаются к самой онтологии.



Рис. 1. Схема объединения онтологии OWL-DL и правил SWRL

Особенности использования SWRL для описания логических правил заключается в следующем:

- правила SWRL не содержат конкретных объектов, а только ссылаются на них, что дает возможность применять одно и то же правило к нескольким группам объектов;
- правила SWRL могут быть добавлены к OWL-описанию, то есть, включены в онтологию;
- составление и чтение правил удобнее выполнять, если для этого существует специальный язык [1].

SWRL правила имеют форму импликации между предшественником (body) и последователем (header). Неформальное значение правила может читаться как: всякий раз, когда выполняются условия в предшественнике, условия в последователе тоже должны выполняться. Множество атомов в body обрабатывается как конъюнкция. Пустой body, обрабатывается как тривиальная истина, то есть удовлетворяется каждой интерпретацией. Таким образом, header тоже должен удовлетворяться каждой интерпретацией. Множество атомов в header обрабатываются как отдельные последовательности, то есть они все должны удовлетворяться. Пустой header обрабатывается как тривиальная ложь. Такие правила удовлетворяются \leftrightarrow предшественник не удовлетворяется ни одной расширенной интерпретацией. Правила с множественными атомами в header могут быть просто преобразованы в множественные правила, каждое с атомным последователем [3].

В некоторых литературных источниках под body понимают антецедент, а под header – консеквент, где оба являются конъюнкциями атомов (рис. 1).



Рис. 1. Основные составляющие SWRL правила

Общим использованием для правил является перенесение значений свойства от одного индивида к другому, связанному с ним[4].

Атомы SWRL технологии разделяют на унарные и бинарные. Унарные атомы представлены в формуле 1:

$$C(arg_1), D(dataArg_1), \quad (1)$$

где C – произвольное выражение класса OWL DL; D – тип данных URI или диапазон перечисленных значений; arg_1, arg_2 – либо индивиды, обозначающие URI, либо переменные, задающие диапазон индивидов; $dataArg_1$ – либо литералы данных, либо переменные, задающие диапазон значений данных.

Бинарные атомы представлены в формуле 2:

$$P(arg_1, arg_2), Q(arg_1, dataArg_1), arg_1 = arg_1, arg_2 \neq arg_2, \quad (2)$$

где P – свойство объекта; Q – свойство типа данных; $arg_1 = arg_1$ – равенство, или «sameAs»; $arg_2 \neq arg_2$ – неравенство, или «differentFrom».

Существует трансляция атомов и аксиом в SWRL, правила могут полностью заменить аксиомы ролей SH. При этом с учетом ограничения безопасности вводится специальный предикат ДЛ, O , который истинный для всех именованных индивидов (т. е. существует факт $O(a)$ для всех имен индивидов, a , встречающихся в онтологии) и не встречается в каком-либо выражении концепта онтологии (и, таким образом, не попадает в область квантора существования). Затем, для каждой переменной в правиле ДЛ-безопасности, к телу добавляется O -атом с этой переменной [5].

Встроенные модули – это атомы с фиксированной, предопределенной интерпретацией. SWRL утверждение включает встроенные модули для сравнения значений, математических и строковых манипуляций. Основная проблема – как интерпретировать встроенные модули, если их переменные находятся под экземплярами, при оценивании.

Правило может утверждать более сложные отношения между свойствами. Повышенная выразительная мощность приводит к неразрешимости ключевых задач вывода, в частности, согласованности онтологии. Расширенные языки требуют возможность представления бесконечный 2-мерный грид, чтобы согласованность стала неразрешимой. С добавлением правил, такое кодирование тривиально. Правило вместе с утверждением, что каждый узел грида связан точно с одним другим узлом посредством каждого x -succ и y -succ, позволяет представить такой грид (формула 3).

$$x - succ(?a, ?b) \wedge y - succ(?b, ?c) \wedge y - succ(?a, ?d) \wedge x - succ(?d, ?e) \rightarrow sameAs(?c, ?e), \quad (3)$$

где x -succ и y -succ – свойства.

SWRL правила выходят за рамки базовых клозов Хорна в следующих моментах:

- конъюнктивные последовательности;
- дескрипции классов также как и имена классов, и как предикаты в атомах класса;
- равенства и неравенства.

Однако большинство из них не добавляет мощности языка.

Конъюнктивные последовательности можно исключить, используя стандартное преобразование Lloyd-Торор: Например, правило вида $A \rightarrow C1 \wedge C2$ можно преобразовать в семантически эквивалентную пару: $A \rightarrow C1, A \rightarrow C2$.

Дескрипция d может быть исключена из правила просто добавлением *OWL* аксиомы, которая вводит новое имя класса и утверждает, что оно эквивалентно d , например, $\text{EquivalentClasses}(D\ d) - d$ заменяется именем класса D .

Эквивалентность атомов – свойство *sameAs* можно заменить подстановкой «определенное пользователем» *owl* свойство – например, Eq . Оно может быть задано подходящим значением, используя правило вида $\text{Thing}(?x) \rightarrow \text{Eq}(?x, ?x)$ и утверждая, что это функционал. Интерпретация Eq соответствует равенству элементов в $\text{EC}(\text{owl} : \text{Thing})$, показано в формуле 4:

$$\forall x, y \in \text{EC}(\text{owl} : \text{Thing}). \langle x, y \rangle \in \text{ER}(\text{Eq}) \Leftrightarrow x = y. \quad (4)$$

Поэтому Eq может использоваться вместо *sameAs* без изменения значения онтологии. Неравенства можно ввести *owl* свойство Neq , чтобы охватить некоторые из значений свойства *differentFrom* путем добавления правила вида (формула 5):

$$\text{Eq}(?x, ?y) \wedge \text{Neq}(?x, ?y) \rightarrow \text{Nothing}(?x) \quad (5)$$

Интерпретация Neq не пересекается с интерпретацией Eq (формула 6):

$$\forall x, y \in \text{EC}(\text{owl} : \text{Thing}). \langle x, y \rangle \in \text{ER}(\text{Neq}) \Rightarrow x \neq y, \quad (6)$$

Это приводит к правилу импликации (формула 7):

$$\text{Neq}(?x, ?y) \rightarrow \text{differentFrom}(?x, ?y) \quad (7)$$

Значит можно исключить *differentFrom* в *header* правила, путем подстановки Neq . Но Neq не полностью охватывает значение неравенства, так как это могут быть пары элементов в $\text{EC}(\text{owl} : \text{Thing})$, которые ни в расширении Eq ни Neq , т. е., *differentFrom* не означает Neq . Значит, нельзя использовать Neq , чтобы исключить вхождения *differentFrom* в *body* правила: чтобы это сделать, потребуется чтобы Neq было равно отрицанию Eq .

Возможность работы с SWRL реализована во многих редакторах онтологий, таких как Protégé и Ontolingua. Она поддерживается такими фреймверками, как Jess, Sesame и Jena [6].

ВЫВОДЫ

Технология SWRL является мощным инструментом для создания базы правил на онтологию предметной области. Данный подход является основным этапом для решения проблемы представления знаний и получения новых знаний из онтологии, так как дает возможность машинной обработки информации. Задачей для последующего анализа является рассмотрение средств для непосредственного получения новых знаний на основе SWRL технологии.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Фазлиев А. З. Языки запросов для OWL / А. З. Фазлиев // Второй симпозиум Казань – Онтологическое моделирование: состояние, направления исследований и применения, 2010.
2. Benedikt Linse. Data Integration on the (Semantic) Web with Rules and Rich Unification, Dissertation zur Erlangung des akademischen Grades des Doktors der Naturwissenschaften, June 21, 2010.
3. Проект электронной библиотеки методик и результатов исследований текстовых коллекций для системы «Источник»/ Н. В. Каргинова, И. В. Кравцов, Н. Д. Москин, А. Г. Варфоломеев // статья, 2008.
4. Heymans S. Semantic web reasoning with conceptual logic programs. In Grigoris Antoniou and Harold Boley, editors, RuleML/ Stijn Heymans, Davy Van Nieuwenborgh, Dirk Vermeir // volume 3323 of Lecture Notes in Computer Science, p.113-127. Springer, 2004.
5. Eiter T. Combining Answer Set Programming with Description Logics for the Semantic Web. In Didier Dubois, Christopher Welty, and Mary-Anne Williams/ Thomas Eiter, Thomas Lukasiewicz, Roman Schindlauer, Hans Tompits // Proceedings Ninth International Conference on Principles of Knowledge Representation and Reasoning (KR 2004), June 2-5, Whistler, British Columbia, Canada, pages 141–151. Morgan Kaufmann, 2004.
6. Horrocks I. SWRL: A semantic web rule language combining owl and ruleml/ Ian Horrocks, Peter F. Patel-Schneider, Harold Boley, Said Tabet, Benjamin Groszof, Mike Dean // W3C Note, 21 May 2004. Available at <http://www.w3.org/Submission/SWRL/>

Статья поступила в редакцию 14.01.2015 г.