УДК 004.822:657.1 DOI: 10.18698/2541-8009-2018-01-233

СПОСОБЫ ПОСТРОЕНИЯ ОНТОЛОГИИ ФИНАНСОВЫХ ОТЧЕТНОСТЕЙ В ФОРМАТЕ IFRS

А.И. Барышников

thief2008@yandex.ru
SPIN-код: 7947-5873

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

Аннотация

Работа посвящена вопросу создания онтологии понятий финансовых отчетностей различных организаций в формате IFRS. В ее рамках выполнены обзор и формализация предметной области и проанализированы профили, синтаксис и машины логического вывода онтологии для решения рассматриваемой задачи. Особое внимание уделено языку веб-онтологий OWL. В результате анализа установлено, что язык OWL является самым популярным и наиболее универсальным. Кроме того, в работе проведено сравнение машин логического вывода, поддерживающих OWL, по набору качественных критериев (выразительная мощность, диалект, поддержка правил и пр.) и вычислительной сложности их работы. Выводы, полученные в статье, позволяют выявить дальнейшие направления использования онтологий.

Ключевые слова

Онтология, OWL, машина логического вывода, ontology, reasoner, RDF, IFRS

Поступила в редакцию 23.11.2017
© МГТУ им. Н.Э. Баумана, 2017

Введение. В настоящее время для определения дальнейшей финансовой и инвестиционной политики анализируют финансовые отчетности компаний или предприятий. Обычно для этого используют отчеты о финансовых результатах. Кроме того, для финансового анализа сторонней компании изучают отчеты об изменениях капитала и о движении денежных средств. В таких отчетах фигурируют финансовые термины, сходные по своей сути, но различающиеся от компании к компании. Эти отчетности публикуют в печатном виде, который не приспособлен к автоматической обработке. Это вызывает необходимость создания средства автоматизированной обработки этих отчетностей и разработки механизма, который позволит определять значения одного и того же показателя у разных компаний. Одним из вариантов создания такого средства является построение обобщенной онтологии отчетов, описывающей типовые взаимоотношения понятий.

Обзор аналогов. В области финансов онтологии используются не впервые. Пабло Кастеллс разработал семантическую сеть технологий в управлении финансовой и экономической информацией в домене TIF (Tecnología, Información y Finanzas — технология, информация и финансы). Данная сеть используется

для поиска и визуализации информации в сети и интранете компаний [1]. Деян Лавбич и Марко Бажек построили семантическую сеть обмена финансовыми инструментами [2]. Маахре Салах и Туария Мохамед разработали онтологию финансовых вложений [3], Сандип Банерджи — онтологию взаимных фондов [4].

Все разработанные онтологии применяются в области финансов, но покрывают не всю необходимую область, а лишь ее малые или косвенные части [5]. Поставленная задача требует разработки и построения другого вида онтологии над иной частью предметной области. В рамках данной работы проанализировано понятие онтологии и выполнено предварительное рассмотрение типов онтологий, их характеристик работ и способов построения. После этого проведен выбор наиболее подходящего способа построения онтологии для выполнения поставленной задачи.

Постановка задачи. Задачей данной работы является формализация набора понятий документов финансовой отчетности семантической сетью онтологий и выбор наиболее подходящего синтаксиса, профиля и машины логического вывода для выполнения этих операций. Формализация переменной области будет проведена путем создания онтологии.

Онтология. Онтология — это артефакт, структура, описывающая значения элементов некоторой системы.

Согласно Томасу Груберу [6], также можно дать следующее определение: онтология — это формальная спецификация согласованной концептуализации (под согласованной концептуализацией подразумевается, что данная концептуализация не есть частное мнение, а является общей для некоторой группы людей).

Основными компонентами онтологии могут являться [7, 8]:

- классы (или понятия);
- отношения (или свойства, атрибуты);
- функции;
- аксиомы;
- экземпляры (или индивиды).

Существует множество языков описания семантических сетей, позволяющих реализовать онтологии. В табл. 1 показывает сравнительные характеристики языков KIF, OWL, RDF и DAML+OIL [9].

Анализ данной таблицы показывает, что язык OWL является самым популярным языком разработки. Кроме того, OWL обладает наиболее широкими возможностями описания, поэтому в рамках данной работы для разработки будет использоваться язык построения онтологий OWL.

OWL. Язык веб-онтологий OWL — это язык для определения и представления веб-онтологий [11]. Веб-онтология может включать описания классов, свойств и их примеры. Формальная семантика OWL описывает, как получить логические следствия, имея такую онтологию, т. е. установить факты, которые не представлены в онтологии буквально, но следуют из ее семантики.

Таблица 1

Сравнение различных языков построения онтологий

Критерий		KIF	OWL			RDF + RDF(S)	DAML+OIL
			Lite	DL	Full		
Конструкции или концептуализации домена		Литералы, лексемы и выражения	Классы, свойства, индивиды, экземпляры классов			Триплеты	Классы и свойства
Спецификация		Структурная	Структурная	Структурная	Структурная	Структурная	Структурная
Выразительность		Высокая	Низкая	Средняя	Высокая	Средняя	Аксиомы и конструкторы классов
Формальность	Синтаксис	Формальный	Формальный	Формальный	Формальный	Формальный	Формальный
	Семантика	Формальный	Формальный	Формальный	Не всегда формальный	Формальный	Формальный
Движок логического вывода		Есть	Есть	Есть	Встречаются проблемы вывода [10]	Нет механизма связного вывода	Есть
Проверка ограничений		Слабая	Хорошая	Хорошая	Хорошая	Слабая	Поддерживает SWRL
Связь с другими языками		Нет	RDF	RDF	RDF	OWL, DAML+OIL	RDF
Используемая парадигма		Логика предикатов первого порядка	DL		Не может быть транслирована в DL	Объектно-ориентированная структура	DL
Используемый W3C стандарт		Нет	RDF			XML	RDF + XML
Стандарт		Есть	Есть			Есть	Нет
Популярность (число ссылок в Google Scholar)		29100	429000			137000	10500

Машина логического вывода. Машины логического вывода (англ. *reasoner*) — класс программного обеспечения, предназначенный для работы с онтологическими моделями. Они позволяют проверять правильность модели и вычислять результаты логических выражений. В настоящее время реализовано множество различных машин логического вывода. Сравнительная таблица их характеристик представлена в табл. 2 [12].

Таблица 2

Сравнение различных машин логического вывода

Критерий		Pellet	Racer	Fact++	Snorocket	SWRL-IQ	HERMIT	CEL	TrOWL	ELK
Методология		Таб-личная	Таблич-ная	Таб-личная	Правила дополне-ния	Пра-вила SWRL	Гипертаб-личная	Прави-ла до-полне-ния	Правила дополне-ния	След-ственная
Непротиворечи-вость		Да	Да	Да	Да	Да	Да	Да	Да	Да
Полнота		Да	Да	Да	Да	Нет	Да	Да	Да	Да
Выразительная мощность		SROIQ (D)	SHIQ	SROIQ (D)	EL+	–	SROIQ(D)	EL+	SROIQ	EL
Диалект		DL; EL	DL	DL	EL	–	DL	EL	DL; EL;	EL
Инкремен-тальная классифи-кация	Сло-жение	Да	Нет	Нет	Да	Да/ Нет	Нет	Да	Нет	Да
	Удале-ние	Да	Нет	Нет	Нет	Да/ Нет	Нет	Нет	Нет	Да
Поддержка правил		Да (SWRL)	Да (SWRL)	Нет	Нет	Да (SWRL)	Да (SWRL)	Нет	Нет	Да (Соб-ственный формат правил)
Платформы ОС		Все	Все	Все	Все	Все	Все	Linux	Все	Все
Обоснование вывода		Да	Да	Нет	Нет	Да	Нет	Да	Нет	Нет
ABOX-модели		Да	Да	Да	Нет	Да	Да	Да	Да	Нет
OWL API		Да	Да	Да	Да	Нет	Да	Да	Да	Да
OWL Link API		Да	Да	Да	Нет	Нет	Да	Да	Нет	Да/Нет
Лицензия		DULI; AGPL;	Соб-ственная	GLGPL	Соб-ственная	Да/Нет	GLGPL	Apache License 2.0	DULI; AGPL;	Apache License 2.0
Поддержка Jena		Да	Нет	Нет	Нет	Нет	Нет	Нет	Да	Да/Нет
Язык реализации		Java	LISP	C++	Java	Prolog	Java	LISP	Java	Java
Доступность		От-кры-тый доступ	Коммер-ческая	От-кры-тый доступ	Коммер-ческая	Да/Нет	Открытый доступ	Откры-тый доступ	Коммер-ческая	Откры-тый до-ступ

Анализ данной таблицы говорит о том, что из свободных решений наиболее полным набором опций обладают машины логического вывода Pellet и Hermit. Однако при этом Pellet также поддерживает фреймворк Jenna и обладает под-

держкой инкрементальной классификации. Поэтому в рамках данной работы для вычисления результатов логических высказываний используется на машине Pellet. Однако для более полного исследования стоит оценить и скорость работы каждой из машин логического вывода.

Для сравнения скорости обработки запросов на подъязыке QL использованы три онтологии: онтология генеалогии (GO), онтология из Национального института онкологии США (NCI) и систематизированной номенклатуры медицины (SNOMED CT). Сравнение выполняли по трем признакам: скорость проверки когерентности класса (табл. 3), скорость проверки согласованности (табл. 4) и времени классификации (табл. 5) [5].

Таблица 3

Сравнение скорости классификации онтологий различными машинами логического вывода

Онтология	Машина логического вывода							
	CB	CEL	FACT++	Hermit	Pellet	RP	SR	TR
GO	0,,34	3,19	20,75	6,48	3,41	10,61	1,54	2,43
NBI	0,65	7,52	11,10	11,75	14,84	52,87	4,31	1,83
SCT	28,08	1 112,23	700,87	6 793,76	1 345,65	3 652,03	101,16	344,93

Таблица 4

Сравнение скорости проверки согласованности различными машинами логического вывода

Онтология	Машина логического вывода							
	CB	CEL	FACT++	Hermit	Pellet	RP	SR	TR
GO	—	2,17	0,36	0	0,27	—	0	0
NBI	—	0,65	0,71	0	0,38	—	0	0
SCT	—	0,88	15,3-	0	16,78	—	0	0

Таблица 5

Сравнение скорости проверки когерентности машинами логического вывода (BC — до классификации, AC — после классификации)

Онтология	Машина логического вывода							
	CB	CEL	FACT++	Hermit	Pellet	RP	SR	TR
GO BC	—	5,28	0,63	6,23	2,12	5,58	0,01	0,23
GO AC	—	02,08	0,07	0,06	0,12	0	0,01	0,03
NBI BC	—	4,46	1,26	11,73	3,47	37,73	0,01	0,06
NBI AC	—	3,19	0,14	0,09	0,18	0	0,01	0,04
SCT BC	—	38,42	22,37	5 276,85	56,91	273,45	0,07	5,17
SCT AC	—	34,59	1,76	1,36	6,07	0	0,06	0,46

Согласно анализу данных таблиц, по показателю скорости Pellet никогда не демонстрирует худшего результата, а потому его можно использовать в рамках поставленной задачи.

Выводы. В данной работе был проведен подробный анализ понятия онтологии и возможностей ее построения. Результаты исследования показывают, что OWL является самым популярным языком разработки, а также обладает наиболее полными возможностями описаний. Сравнение машин логического вывода также показало, что Pellet не является самой быстродействующей машиной, но имеет самый полный набор опций вывода и поддержку фреймворка Jenna, что лучше всего подходит для исполнения поставленной задачи.

В качестве дальнейшего направления данного исследования следует рассмотреть альтернативы. Онтологии являются не единственным способом представления предметной области, поэтому разумно также рассмотреть организацию и классификацию отчетов в формате IRFS с помощью нейронных сетей. Кроме того, в данной работе рассмотрены лишь семантические машины логического вывода, но не проведено их исследование и сравнение с вероятностными логическими сетями. Возможности использования этих сетей для логического вывода внутри онтологии также представляют и дальнейшее направление исследования [9].

Литература

[1] Castells P., Foncillas B., Lara R., Ricol M., Alonso J.L. Semantic Web technologies for economic and financial information management // *ESWS*. 2004. С. 473–487. URL: http://nets.ii.uam.es/aniceto/publications/esws04.pdf.

[2] Lavbič D., Bajec M. Employing semantic Web technologies in financial instruments trading // *International Journal on New Computer Architectures and Their Applications*. 2012. Vol. 2. P. 167–182. URL: https://www.researchgate.net/publication/221753930_Employing_Semantic_Web_technologies_in_financial_instruments_trading.

[3] Salah M., Mohamed T. Developing ontology for financial investment “Algeria case study” // *International Journal of Computer Applications (0975 — 8887)*. 2011. Vol. 24. No. 4. P. 1–6. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.259.1111&rep=rep1&type=pdf.

[4] Banerjee S.A Semantic Web based ontology in the financial domain. // *Proc. World Academy of Science, Engineering and Technology (WASET)*. 2013. Vol. 7.No. 6. P. 807–810. URL: http://www.waset.org/publications/10376.

[5] Dentler K. et al. Comparison of reasoners for large ontologies in the OWL 2 EL profile // *Semantic Web 2.2*. 2011. Vol. 2. No. 2. P. 71–87. URL: https://content.iospress.com/articles/semantic-web/sw034.

[6] Ontology (Computer science). URL: http://tomgruber.org/writing/ontology-definition-2007.htm (дата обращения 31.10.2016).

[7] Добров Б.В., Иванов В.В., Лукашевич Н.В., Соловьев В.Д. *Онтологии и тезаурусы: модели, инструменты, приложения*. М.: Бином. Лаборатория знаний, 2009. 173 с.

[8] Язык web-онтологий OWL 2: начальное руководство. URL: http://shcherbak.net/translations/ru_owl2primer_shcherbak_net.html (дата обращения 23.10.2016).

[9] Kalibatiene D., Vasilecas O. Survey on ontology languages // *Int. Conf. on Business Informatics Research. Springer. Berlin – Heidelberg, 2011*, pp. 124–141. URL: https://www.researchgate.net/profile/Diana_Kalibatiene/publication/221292703_Survey_on_Ontology_Languages/links/5616418708ae37cfe4085ae3/Survey-on-Ontology-Languages.pdf.

[10] Horrocks I. OWL: A description logic based ontology language. ICLP. 2005. Vol. 3668. P. 1–4. *Springer*, Heidelberg. URL: http://www.cs.ox.ac.uk/ian.horrocks/Publications/ download/2003/HPMW07.pdf.

[11] OWL Web ontology language overview. URL: https://www.w3.org/TR/owl-features (дата обращения 23.10.2016).

[12] Zhang Z., Miller J.A. Ontology query languages for the semantic web: A performance evaluation: Master's thesis. 2005. University of Georgia. URL: https://pdfs.semanticscholar.org/1bad/b19737cd278e4eb12f77ec78f6e35afeb157.pdf.

Барышников Андрей Игоревич — студент кафедры «Программное обеспечение ЭВМ и информационные технологии», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

Научный руководитель — Бекасов Денис Евгеньевич, магистр программной инженерии, преподаватель кафедры «Программное обеспечение ЭВМ и информационные технологии», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

METHODS OF AUTHORING A FINANCIAL REPORT ONTOLOGY IN THE IFRS FORMAT

A.I. Baryshnikov

thief2008@yandex.ru
SPIN-код: 7947-5873

Bauman Moscow State Technical University, Moscow, Russian Federation

Abstract

The study deals with the problem of authoring an ontology of financial reporting concepts for various organisations in the IFRS format. The study included a review and formalisation of the universe of discourse, and analysis of profiles, syntax and inference engines of the ontology that can solve the problem under consideration. We focus on the OWL Web Ontology Language. Our analysis established that the OWL language is the most popular and the most universal. In addition, the study compares inference engines that support OWL, in terms of qualitative criteria (expressive power, sublanguage, rule support and so on) and the computational complexity of running them. Conclusions derived in the article make it possible to pinpoint further potential applications of ontologies.

Keywords

Ontology, OWL, inference engine, ontology, reasoner, RDF, IFRS

© Bauman Moscow State Technical University, 2017

References

[1] Castells P., Foncillas B., Lara R., Ricol M., Alonso J.L. Semantic Web technologies for economic and financial information management. *ESWS*, 2004, pp. 473–487. Available at: http://nets.ii.uam.es/aniceto/publications/esws04.pdf.

[2] Lavbič D., Bajec M. Employing Semantic Web technologies in financial instruments trading. *International Journal on New Computer Architectures and Their Applications*, 2012, vol. 2, no. 1, pp. 167–182. Available at: https://www.researchgate.net/publication/221753930_Employing_Semantic_Web_technologies_in_financial_instruments_trading.

[3] Salah M., Mohamed T. Developing ontology for financial investment "Algeria case study". *International Journal of Computer Applications* (0975 — 8887), 2011, vol. 24, no. 4, pp. 1–6. Available at: http://citeseerx.ist.psu.edu/viewdoc/ download?doi=10.1.1.259.1111&rep=rep1&type=pdf.

[4] Banerjee S.A Semantic Web based ontology in the financial domain. *Proc. World Academy of Science, Engineering and Technology (WASET)*, 2013, vol. 7, no. 6, pp. 807–810. Available at: http://www.waset.org/publications/10376.

[5] Dentler K. et al. Comparison of reasoners for large ontologies in the OWL 2 EL profile. *Semantic Web 2.2.*, 2011, vol. 2, no. 2, pp. 71–87. Available at: https://content.iospress.com/articles/semantic-web/sw034.

[6] Ontology (Computer science). Available at: http://tomgruber.org/writing/ontology-definition-2007.htm (accessed 31.10.2016).

[7] Dobrov B.V., Ivanov V.V., Lukashevich N.V., Solov'ev V.D. Ontologii i tezaurusy: modeli, instrumenty, prilozheniya [Ontologies and thesauri models, tools and applications]. Moscow, Binom. Laboratoriya znaniy Publ., 2009. 173 p.

[8] Yazyk web-ontologiy OWL 2: nachal'noe rukovodstvo [*OWL Web ontology* language. Initial management]. Available at: http://shcherbak.net/translations/ru_owl2primer_ shcherbak_net.html (accessed 23.10.2016).

[9] Kalibatiene D., Vasilecas O. Survey on ontology languages. *Int. Conf. on Business Informatics Research*. Springer. Berlin – Heidelberg, 2011, pp. 124–141. Available at: https://www.researchgate.net/profile/Diana_Kalibatiene/publication/221292703_Survey_on_Ontology_Languages/links/5616418708ae37cfe4085ae3/Survey-on-Ontology-Languages.pdf.

[10] Horrocks I. OWL: A description logic based ontology language. ICLP, 2005, vol. 3668, pp. 1–4. Springer, Heidelberg. Available at: http://www.cs.ox.ac.uk/ian.horrocks/Publications/download/2003/HPMW07.pdf.

[11] OWL Web ontology language overview. Available at: https://www.w3.org/TR/owl-features https://www.researchgate.net/profile/Diana_Kalibatiene/publication/221292703_Survey_on_Ontology_Languages/links/5616418708ae37cfe4085ae3/Survey-on-Ontology-Languages.pdf (accessed 23.10.2016).

[12] Zhang Z., Miller J.A. Ontology query languages for the semantic web: A performance evaluation: Master's thesis, 2005. University of Georgia. Available at: https://pdfs.semanticscholar.org/1bad/b19737cd278e4eb12f77ec78f6e35afeb157.pdf.

Baryshnikov A.I. — student, Department of Computer Software and Information Technologies, Bauman Moscow State Technical University, Moscow, Russian Federation.

Scientific advisor — Bekasov D.E., Master of Software Engineering, Lecturer, Department of Computer Software and Information Technologies, Bauman Moscow State Technical University, Moscow, Russian Federation.