

# Автоматическое порождение фраз естественного языка по OWL-модели, семантике и прагматике

Полина Сазонова

Новосибирский Государственный Университет, Новосибирск, Россия  
psazonova@gmail.com

**Аннотация** Статья посвящена решению проблемы согласования пользовательской и программной онтологии. В ходе работы отработаны алгоритмы порождения фраз естественного языка по OWL-модели, семантике и прагматике. Данное решение можно использовать для повышения эффективности общения пользователей с виртуальными консультантами на сайтах, специализирующихся на продаже товаров и услуг.

**Ключевые слова:** онтология, OWL, virtual assistant, прагматика, семантика, NLP.

## 1 Введение

В настоящее время стремительно растет количество задач, связанных с обработкой данных, решаемых с помощью онтологий. Онтологии используются в теоретических исследованиях и практических разработках. По определению, формальная онтология предметной области представляет собой пару  $\langle S, \sigma \rangle$ , где  $\sigma$  - это множество ключевых понятий, а  $S$  - множество аналитических предложений, описывающих смысл данных ключевых понятий [1]. Предложение называется аналитическим, если его значение истинности зависит только от смысла понятий, содержащихся в этом утверждении [2][3].

Онтологии применяются для построения интеллектуальных и экспертных систем. В настоящее время среди компаний, реализующих свои товары через Интернет, стало популярным использование виртуальных помощников (virtual assistant). Их задачей является консультирование клиентов по вопросам, связанным с продаваемыми товарами, оказание помощи в выборе продукта, а в некоторых случаях и поддержание разговора на общие темы.

Виртуальный помощник ежедневно общается с разными людьми. Каждый потребитель (точнее, группа потребителей) имеют свое представление о предметной области и часто используют свою специфичную лексику. Поймут ли друг друга клиент и виртуальный помощник? Что следует делать, чтобы достигать наибольшего понимания? Здесь мы имеем проблему согласования программной и пользовательской онтологий [4]. В данной работе решается задача согласования этих двух онтологий для случая общения клиента и виртуального помощника.

Было установлено, что для решения обозначенной проблемы, необходимо найти "правила соответствия" между понятиями онтологии программной системы и онтологии пользовательских задач [4]. То есть виртуальному помощнику необходимо говорить на языке пользователя, а именно, осуществлять перевод информации с онтологии программной системы на онтологию конкретного пользователя (групп пользователей).

Сходным решением является технология, применяемая в чат-ботах (робот-собеседник), которая называется отзеркаливанием. Она заключается в перефразировании тех выражений, которые употребляет пользователь. Чат-бот отвечает на фразу собеседника подобной фразой, несколько измененной в соответствии с контекстом диалога. Таким образом, у пользователя создается ощущение, что его собеседник мыслит и ощущает так же, как он сам, что создает эмпатию. Однако, употребление только перефразированных выражений не может нести новую информацию. А в процессе диалога с виртуальным собеседником пользователю должна быть предоставлена новая информация, которая могла бы сподвигнуть его к совершению действий, например, покупке чего-либо. Причем новая информация должна быть представлена в том же стиле и при помощи такой же лексики, которую употребляет пользователь.

## **2 Разработка программной системы, осуществляющей порождение фразы естественного языка по модели, написанной на языке OWL**

Решение проблемы согласования онтологий заключается в разработке такого виртуального помощника, который сможет использовать "правила соответствия" между понятиями онтологии программной системы и онтологии пользовательских задач.

Такая программная система порождает фразы естественного языка в полуавтоматическом режиме на основе семантики и прагматики. Понятие прагматики определяется с использованием подхода Фреге [5], где прагматика - это соответствие между синтаксисом и семантикой. В соответствии с этим подходом одна и та же семантика может выражаться при помощи разного синтаксиса. Вид используемого синтаксиса для заданной семантики зависит от пользователя, с которым происходит общение. Например, при общении с пожилыми людьми и с подростками семантически одинаковое выражение будет иметь различное синтаксическое воплощение.

В данной работе семантика для порождения фразы задается OWL-моделью.

Прагматика задается следующим образом. Мы делим пользователей на группы по половому признаку, возрастным категориям (молодежь, взрослый и пожилой) и по роду деятельности. Для каждой группы пользователей лексика задается при помощи онтологии, соответствующей именно этой группе, и словаря сленговых выражений, характерных для их профессиональной деятельности. При порождении фразы учитывается тип пользователя, тем самым создается контекст общения, комфортный для пользова-

теля. Это происходит за счет того, что лексика и способ ее употребления подбираются так, чтобы они были привычны пользователю. Проанализировав профиль человека, с которым происходит общение, можно получить информацию о его возрасте, деятельности и интересах, и далее сделать вывод о необходимости использования той или иной лексики и подобрать характеристики будущей фразы (её тональность и стиль). Программная система в процессе общения относит пользователя, с которым она ведет диалог, к определенной группе, и начинает использовать онтологии, соответствующие данной группе. Фраза строится на основе характеристик тональности фразы и её стиля.

1. Тональность фразы задается наличием эмоционально окрашенной лексики. При порождении фразы, имеющей заданную тональность, подбираются соответствующие синонимы.
2. Стиль фразы определяется классом пользователей, к которому принадлежит собеседник. В соответствии с стилем выбираются нужная лексика, задаваемая соответствующей онтологией.

Построение фразы происходит на основе шаблонов с использованием принципов порождающей грамматики Хомского [6]. Каждое предложение состоит из именной группы и группы сказуемого. Именная группа состоит из определителя и определяемого, где определитель может опускаться. Группы могут содержать в себе прилагательные, наречия и т.д.

Согласование слов в предложении осуществляется за счет использования сторонних морфологических библиотек.

### 3 Заключение

В рамках данной работы были разработаны методы порождения фраз естественного языка по заданной OWL-модели, определяющей семантику фразы, и заданной прагматике фразы. Разработана программная система, реализующая данные методы.

Используя вышеописанный подход, можно решить проблему согласования пользовательской и программной онтологии и значительно повысить эффективность обмена знаниями между клиентом и виртуальным помощником. Виртуальный помощник при построении фразы учитывает не только модель ситуации, но и набор характеристик пользователя, с которым происходит общение, что гарантирует понимание сторон и увеличивает доверие к помощнику.

### Список литературы

1. *Пальчинов, Д. Е.* Моделирование мышления и формализация рефлексии I: Теоретико-модельная формализация онтологии и рефлексии. // *Философия науки.* 2006. № 4 (31). С. 86-114.

2. *Carnap, R.* Meaning and Necessity. // A Study in Semantics and Modal Logic. 1956.
3. *Пальчинов, Д. Е.* Решение задачи поиска информации на основе онтологий. // Бизнес-информатика. 2008. № 1. С. 3-13.
4. *Пальчинов, Д. Е., Целищев, В. В.* Проблема извлечения знаний в системе взаимодействия человека и компьютера (онтологии и пресуппозиции) // Философия науки. 2012. № 4. С. 20-35.
5. *Frege, G.* Über Sinn und Bedeutung / Zeitschrift für Philosophie und philosophische Kritik. s. 25–50. 1892.
6. *Chomsky, N.* Syntactic Structures — Berlin: Mouton de Gruyter, 2002, P. 119.
7. *Пальчинов, Д. Е.* Моделирование мышления и формализация рефлексии. Ч.2. Онтологии и формализация понятий. // Философия науки. 2008. № 2(37). С. 62–99.

# Automatic Natural Language Generation Using an OWL Model, Semantics and Pragmatics

Polina Sazonova

Novosibirsk State University, Novosibirsk, Russia  
psazonova@gmail.com

**Abstract.** This paper is focused on the problem of agreement between the ontology of a user and of a computer program. The natural language generation algorithms which use an OWL model, semantics and pragmatics have been studied. The present solution can be used to increase the interaction efficiency between users and virtual assistants on websites specialized on selling goods and services.

**Keywords:** ontology, OWL, virtual assistant, pragmatics, semantics, NLP.