

Е. Д. Махина¹, Д. Е. Пальчунов^{1,2}

¹ *Новосибирский государственный университет
ул. Пирогова, 2, Новосибирск, 630090, Россия*

² *Институт математики им. С. Л. Соболева СО РАН
пр. Академика Коптюга, 4, Новосибирск, 630090, Россия*

ekatermakhina96@gmail.com, palch@math.nsc.ru

ПРОГРАММНАЯ СИСТЕМА ДЛЯ ОПРЕДЕЛЕНИЯ РЕЧЕВЫХ ДЕЙСТВИЙ В ТЕКСТАХ ЕСТЕСТВЕННОГО ЯЗЫКА *

Статья посвящена разработке методов извлечения речевых действий из текстов естественного языка. В настоящее время одной из наиболее важных проблем является проблема создания человеко-машинных интерфейсов. В данной работе рассматривается задача построения подобных интерфейсов путем автоматизированного извлечения речевых действий из текстов естественного языка. Использование речевых действий позволяет задавать контекст диалога и уточнять смысл употребляемых понятий. По результатам исследования была разработана программная система, которая способна определять типы речевых действий, представленных в текстах естественного языка. Одно из возможных применений результатов данного исследования заключается в использовании разработанной программной системы при создании чат-ботов.

Ключевые слова: речевое действие, иллокутивная сила, теоретико-модельные методы, алгебра Линденбаума – Тарского, фрагменты атомарных диаграмм, онтология, естественно-языковые интерфейсы.

Введение

В настоящее время проектирование, построение и пополнение семантических и онтологических моделей различных предметных областей являются актуальными задачами, решаемыми как при автоматизации бизнес-процессов предприятий, так и при разработке программных продуктов для конечных пользователей. Одной из наиболее острых проблем является проблема предоставления интерфейсов на естественном языке для коммуникации человека и компьютера. В частности, такие проблемы ставятся и решаются при создании чат-ботов.

Компьютер не всегда способен «понять», какую цель хотел достичь человек той или иной фразой, исходя только из синтаксиса текста диалога с пользователем. Интерфейсы на естественном языке призваны помочь определять то, чего хочет человек, посредством не только синтаксического, но и семантического анализа текстов естественного языка.

Существуют разные подходы к разработке интерфейсов на естественном языке [1–5]. В рамках данной работы мы используем теоретико-модельные методы [6–11] и теорию речевых действий для создания естественно-языковых интерфейсов.

* Исследование выполнено при частичной финансовой поддержке Президиума СО РАН (проект «Инженерия интенциональных онтологий в дедуктивных и информационных системах» Комплексной программы ФНИ СО РАН П.1).

Целью настоящего исследования является разработка и программная реализация алгоритма определения типов речевых действий, представленных в текстах естественного языка. Достижение данной цели необходимо для решения более общей задачи – извлечения речевых действий из текстов естественного языка и их формального (теоретико-модельного) представления. Извлечённые речевые действия представляются в виде элементов куба алгебры Линденбаума – Тарского теории рассматриваемой предметной области [12; 13].

Одной из важных областей применения результатов данной работы является их использование при разработке чат-ботов нового поколения, которые по своим возможностям будут существенно превосходить существующие. Рассмотрим, например, чат-бот «Алиса» – виртуальный голосовой помощник, созданный компанией Яндекс. Алиса не создает контекста общения с собеседником, и по этой причине не всегда корректно отвечает на вопросы человека. Мы же хотим строить и далее использовать модель диалога между пользователем и чат-ботом.

На основе знаний о диалоге, представленных в модели, будут формироваться дальнейшие реплики чат-бота. Модель формально представляется в виде фрагмента атомарной диаграммы алгебраической системы, а речевые действия – в виде троек формул: <сообщение, побуждение, заявление>. Каждая такая тройка является элементом куба алгебры Линденбаума – Тарского.

Необходимо отметить, что целью данной работы является создание не полностью автоматической, а автоматизированной системы распознавания речевых действий, так как существуют проблемы, которые на настоящем этапе не представляется возможным решать полностью автоматически. К таким проблемам можно отнести полное описание контекста для определения пропозициональной составляющей и иллокутивной силы речевого действия (в частности, обработку пресуппозиций, содержащихся в тексте естественного языка [14; 15]).

Теория речевых действий

Разработка теории речевых действий была начата Дж. Л. Остином [16; 17] и детально разработана Дж. Р. Серлем и Д. Вандервекенем [18–28]. Нами была предложена определенная ревизия классической теории речевых действий [12; 13].

В теории речевых действий диалог между собеседниками рассматривается как совокупность речевых действий. В речевом действии выделяют две основных составляющих: высказывание (пропозициональная составляющая) и иллокутивная сила. Мы расширяем этот подход и рассматриваем текст естественного языка как последовательность речевых действий. Методы определения речевых действий в текстах естественного языка в рамках классической теории речевых действий исследовались в [29].

Высказывание (пропозициональная составляющая речевого действия) отражает общее содержание предложения. Мы формально описываем пропозициональные составляющие речевых действий при помощи фрагментов атомарных диаграмм (т. е. конъюнкций атомарных предложений логики предикатов первого порядка).

В классической теории речевых действий иллокутивная сила раскладывается на 7 компонентов:

- иллокутивная цель;
- мера стремления к иллокутивной цели;
- способ достижения иллокутивной цели;
- ограничения на пропозицию;
- предварительные условия;
- условие искренности;
- сила выражения условия искренности.

В данной работе мы рассматриваем две составляющие иллокутивной силы: иллокутивную цель и интенсивность выражения иллокутивной цели (т. е. меру стремления к иллокутивной цели).

Классическая теория речевых действий предлагает следующую классификацию речевых действий:

- 1) репрезентативы – речевые действия, целью которых является сообщение об общем положении дел;
- 2) директивы – речевые действия, побуждающие совершать или не совершать какие-либо действия;
- 3) комиссивы – речевые действия, направленные на возложение ответственности на говорящего за то, чтобы сделать что-то в будущем;
- 4) экспрессивы – речевые действия, служащие для выражения эмоционального состояния говорящего;
- 5) декларативы – речевые действия, результатом которых является осуществление главной цели высказывания, т. е. после их озвучивания положение дел в реальном мире соответствует сказанному.

Стоит заметить, что данная классификация не представляет собой разбиение на непересекающиеся подмножества, т. е. одно речевое действие может быть представителем нескольких классов.

Разработанная нами ранее классификация речевых действий [12; 13] выделяет три типа речевых действий (более точно – три типа компонентов речевых действий): сообщение, заявление, побуждение. Каждое речевое действие может иметь один, два или три компонента. Указанные типы речевых действий сопоставимы со следующими аспектами сознания собеседника: знания (прошлое, настоящее или будущее), представление о текущей ситуации (настоящее) и намерения (будущее). Например, заявление может менять контекст диалога, в частности изменять интерпретацию, смысл употребляемых понятий.

Таким образом, с одной стороны, речевое действие представляется как пара $\langle F, \varphi \rangle$ (иногда пишут $F\langle\varphi\rangle$), где F – иллокутивная сила, а φ – пропозициональная составляющая, а с другой стороны, как тройка предложений логики предикатов первого порядка $(\varphi_1, \varphi_2, \varphi_3)$, где φ_1 – сообщение, φ_2 – заявление, а φ_3 – побуждение. При этом формула φ_i может быть тождественно истинной. Это означает, что i -й компонент речевого действия отсутствует.

Заметим, что классический подход к теории речевых действий обладает рядом существенных недостатков [12; 13].

1. Переход от синтаксиса к семантике, определяемый только через иллокутивные глаголы.

Однако следующие примеры являются разными речевыми действиями с точки зрения классической теории речевых действий, но обозначают они одно и то же желание говорящего, поэтому с точки зрения семантики не имеет смысла разделять их:

- a. Закрой окно.
- b. Я прошу тебя закрыть окно.
- c. Я заявляю свое требование закрыть окно.
- d. Я сообщаю тебе свою просьбу – закрой окно.

2. Наличие в каждом речевом действии не более одной иллокутивной цели. Из-за этого классификация иллокутивных сил становится очень громоздкой.

3. Отсутствие точного определения классов речевых действий. Например, отсутствует четкий критерий, позволяющий ответить на вопрос, что представляет собой заявление. Подобного рода определения крайне важны, так как в настоящей работе для программной реализации необходимы четкие правила для определения типа речевого действия.

4. Изолированность речевого действия от диалога. Речевое действие рассматривается отдельно, как бы «в вакууме».

5. «Юридическая парадигма»: на самом деле меняет мир не слово, а его фиксация в рамках определенного контекста – протокола, регламента и т. д.

Основными преимуществами предлагаемого подхода к классификации речевых действий [12; 13] по сравнению с подходом Остина и Серля является следующее.

1. Речевое действие является композицией компонентов трех типов: сообщение, заявление и побуждение.

2. Каждый тип речевого действия имеет четкое определение.

3. Речевое действие рассматривается в контексте диалога. Это означает, что мы имеем представление о текущей ситуации (например, разговаривают ли два друга либо работода-

тель и подчиненный), и в зависимости от этого мы можем делать те или иные выводы о принадлежности речевого действия к тем или иным классам.

4. Меняется не мир, а собеседник: его знания, представления о текущей ситуации, намерения.

Определение типа речевого действия

Алгоритм автоматизированного извлечения речевых действий из текстов естественного языка можно представить в виде последовательности шагов:

- 1) распознавание речевого действия в предложении;
- 2) извлечение его пропозициональной части;
- 3) извлечение иллокутивной цели и интенсивности речевого действия.

Одной из задач данной работы является определение типа речевого действия, представленного в предложении естественного языка. Для наглядности рассмотрим следующие примеры:

- 1) Я объявляю войну (заявление)
- 2) Делай уроки! (побуждение)
- 3) В пятницу будет контрольная работа (сообщение)
- 4) В пятницу будет контрольная работа, все должны на нее прийти (сообщение + заявление)
- 5) Можно я оставлю у себя твою книгу еще на неделю, пожалуйста? (побуждение + заявление)
- 6) Пообещай, что больше не будешь так делать (побуждение)
- 7) Какое красивое небо! (заявление)

Разберем подробно пример предложения, содержащего все компоненты речевого действия: «Завтра в 9 часов утра вы должны прийти на сдачу долгов». Во-первых, в данном предложении присутствует сообщение, семантически выраженное в следующем: «Завтра в 9 часов утра состоится сдача долгов». Во-вторых, заявление является частью речевого действия, так как отдается приказ быть всем завтра на сдаче долгов. В-третьих, говорящий (предположим, что это преподаватель) побуждает всех прийти, т. е. он заинтересован в том, чтобы завтра все присутствовали на сдаче долгов, а иначе самому преподавателю не имеет смысла приходить. Перечисленные составляющие представляют 3 разных пропозиции, выделенные из основного предложения.

Для определения того или иного типа речевого действия можно задать определенные вопросы. Если ответы на эти вопросы могут быть получены из предложения, то речевое действие относится к тому или иному типу; в противном случае речевое действие не содержит искомого компонента.

Ниже перечислены вопросы, являющиеся содержательными критериями проверки наличия в речевом действии каждой из трех составляющих. Другими словами, на какие вопросы необходимо ответить, чтобы убедиться, что речевое действие имеет тот или иной компонент. Предполагается, что целью является определение типа речевого действия фраз субъекта *a* в диалоге с субъектом *b*.

1. Ставил ли целью говорящий *a* передать собеседнику *b* какие-то знания (об уже совершившихся, текущих или будущих событиях)?
2. Было ли целью говорящего установить, что имела место определенная речевая ситуация: вопрос, просьба, приказ, мольба и т. д.?
3. Хочет ли говорящий *a*, чтобы у собеседника *b* после разговора было намерение что-то сделать (при этом не имеет значения, было ли такое намерение у *b* до разговора; важно, чтобы после диалога это намерение появилось)?

Данные вопросы можно задавать пользователю в интерактивном режиме в том случае, если нельзя однозначно определить тип речевого действия. Также в случае невозможности однозначной классификации речевого действия можно выдавать пользователю несколько вариантов, отличающихся в зависимости от контекста.

Помимо этого, мы можем задавать вопросы, помогающие выяснить содержание каждой конкретной составляющей речевого действия:

1. $\varphi_1(a, b) = \text{“что } b \text{ узнал нового?”}$
2. $\varphi_2(a, b) = \text{“что произошло во время разговора } a \text{ и } b \text{ (какая речевая ситуация имела место)?”}$
3. $\varphi_3(a, b) = \text{“что намеревается сделать } b\text{?”}$.

В ходе работы был составлен следующий список правил определения типа речевого действия.

1. Побуждение:

- a. повелительное наклонение глагола
- b. условное наклонение глагола;
- c. будущее время глагола + 2-е или 3-е лицо;
- d. глагол «давай»;
- e. слово «пожалуйста»;
- f. слова «следует / должен / обязан / необходимо / нужно / надо» и т. д.;
- g. можешь ли ты / не мог бы ты / почему бы тебе не и т. д.

2. Сообщение:

- a. настоящее время глагола;
- b. прошедшее время глагола;
- c. 1-е или 3-е лицо.

3. Заявление:

- a. «я хотел бы / мне хотелось бы»;
- b. настоящее время глагола.

4. Заявление + побуждение:

- a. было бы хорошо / чудесно / здорово и т. д., если бы...;
- b. любой изолированный вопрос.

Видно, что некоторые правила совпадают для разных типов речевых действий. Это означает, что в исходном предложении речевое действие представляет собой комбинацию типов, например сообщение и побуждение, а возможно, и все три типа одновременно.

Предложенный список правил не является финальной версией алгоритма, используемого в программе, т. е. данный список является расширяемым и пополняемым в процессе использования программной системы для разных текстов русского языка.

Этапы анализа предложения

Алгоритм разработанной программной системы состоит из следующих шагов. На вход программной системе подается предложение на русском языке. В данном исходном предложении осуществляется морфологический анализ слов. После этого предложение проверяется на наличие каждого из типов речевого действия (сообщение, заявление, побуждение). На выходе программа выдает найденные типы компонентов речевого действия, представленного в данном предложении.

Опишем этапы работы программной системы более подробно. Итоговой идентификации типа речевого действия в предложении предшествует этап морфологического анализа предложения. Несмотря на то что данный этап не играет определяющей роли в идентификации типа речевого действия, он очень важен для предобработки предложения, так как именно на этом этапе происходит первоначальный выбор того или иного типа речевого действия.

Для морфологического анализа предложения естественного языка используется программа MyStem, предоставленная компанией Яндекс для открытого использования. Программа делает следующее: на вход подается предложение русского языка, а на выходе мы получаем

всю информацию о каждом слове – часть речи, время и наклонение для глаголов, падеж для существительных и т. п. К системе MyStem можно подключить свой словарь; также она позволяет строить гипотетические разборы для слов, не входящих в словарь.

Рассмотрим пример работы программы: для входного файла `sapgir.txt`, содержащего текст “В мурели шлепают пельсиски. В стакелках светится мычай.”, вывод будет следующим:

```
$ mystem -nig sapgir.txt

В{в=PR=|в=S, сокр=(им, ед|им, мн|род, ед|род, мн|дат, ед|дат, мн|вин, ед|вин, мн|твор, ед|твор, мн|п
р, ед|пр, мн)}
мурелки{мурелка?=S, жен, од=(им, мн|род, ед)|мурелка?=S, жен, неод=(им, мн|род, ед|вин, мн)}
шлепают{шлепать=V, несов, пе=непрош, мн, изъяв, 3-л}
пельсиски{пельсиск?=S, имя, муж, од=им, мн|пельсиска?=S, жен, неод=(им, мн|род, ед|вин, мн)}
В{в=PR=|в=S, сокр=(им, ед|им, мн|род, ед|род, мн|дат, ед|дат, мн|вин, ед|вин, мн|твор, ед|твор, мн|п
р, ед|пр, мн)}
стакелках{стакелка?=S, жен, неод=пр, мн|стакелка?=S, жен, од=пр, мн|стакелок?=S, муж, неод=пр, мн|
стакелка?=S, имя, жен, од=пр, мн|стакелк?=S, имя, муж, од=пр, мн|стакелк?=S, муж, неод=пр, мн}
светится{светиться=V, несов, нп=непрош, ед, изъяв, 3-л}
мычай{мычай?=S, муж, неод=(им, ед|вин, ед)}
```

Описанная программа встроена в разработанную программную систему SpeechActs и используется для предварительной обработки предложений.

После проведения морфологического разбора предложения мы переходим к его семантическому анализу. Этот этап играет большую роль в определении типа (комбинации типов) речевого действия, так как первоначально мы выбираем тип в соответствии с определенными морфологическими формами слов, а затем корректируем выбранный тип посредством семантического анализа. Так, к примеру, предложение “Какая прекрасная погода сегодня!” на этапе морфологического анализа мы определим как сообщение, а после семантического анализа поймем, что в действительности это заявление, так как данным предложением говорящий хотел выразить свое отношение к погоде, а в другом контексте это может быть побуждением отдохнуть на природе.

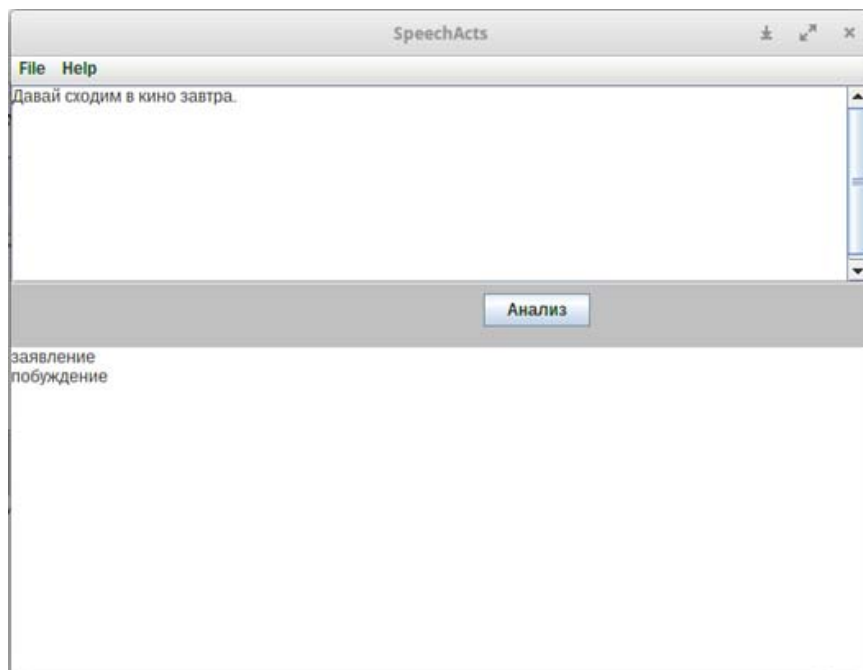
На этапе семантического анализа мы ищем употребленные в предложении глаголы в словаре иллокутивных глаголов, встроенном в программу. Нахождение глагола в одном из словарей определяет принадлежность речевого действия тому или иному типу.

Таким образом, на данном этапе мы подтверждаем выбор типа речевого действия, осуществленный на предыдущем этапе (этапе морфологического анализа), либо добавляем к уже выбранному типу еще один тип речевого действия, определенный в ходе семантического анализа.

Интерфейс программной системы

Алгоритм работы программной системы заключается в следующем: на вход подается предложение на русском языке. В исходном предложении осуществляется морфологический анализ каждого слова в предложении, после чего предложение проверяется на наличие каждого из типов речевого действия (сообщение, заявление, побуждение). На выходе программа выдает найденные типы речевых действий в предложении из входных данных.

Ниже приведен скриншот пользовательского интерфейса первой версии программной системы.



Архитектура программной системы

Разработанная программная система состоит из следующих компонентов.

1. Controller

Контроллер отвечает за инициализацию всех частей приложения. Он запускает остальные компоненты, а также обеспечивает передачу данных от графического интерфейса к модели и обратно.

2. LogicText

Модуль, запускающий программную систему LogicText, разработанную О. Г. Махасоевой и Д. Е. Пальчуновым [8; 9]. Программа делает следующее: каждое слово в предложении приводит к нормальной форме, выбирает из них субъект и объект; само действие является предикатом. В программу встроен словарь с некоторыми известными предикатами, где с каждым глаголом сопоставлены вопросы (кто, что, где, когда и т. д.). Алгоритм работы программной системы LogicText основан на применении теории «Смысл-Текст» И. А. Мельчука [30].

3. MyStem

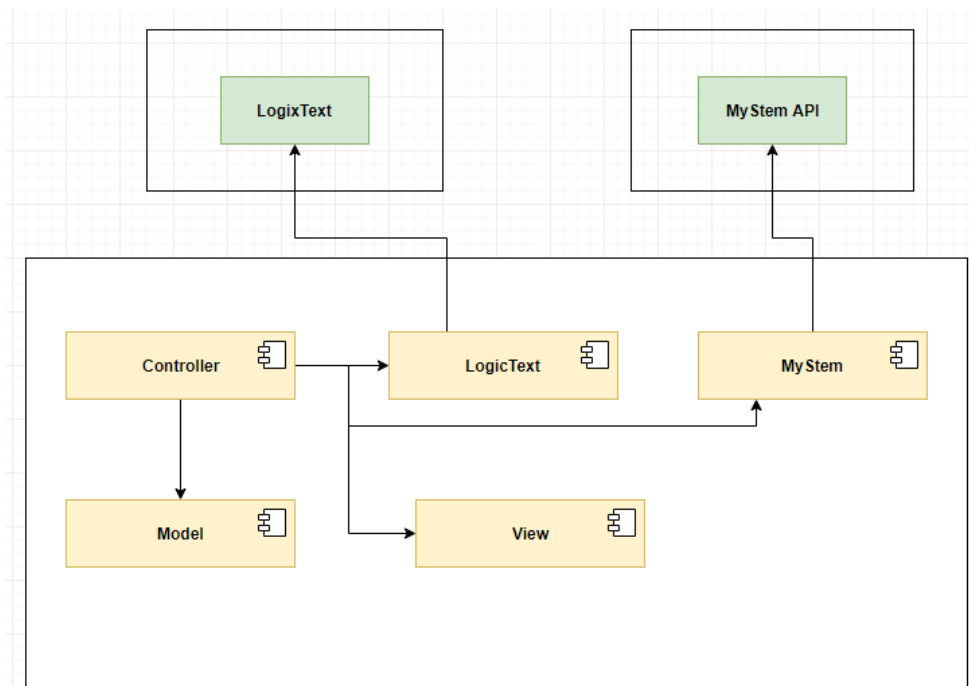
Модуль, который обрабатывает каждое слово в предложении и возвращает полную характеристику каждого из слов.

4. Model

Модель представляет собой главный компонент, в котором происходит обработка предложения, его анализ и определение типа речевого действия.

5. View

Данный модуль отвечает за отображение введенного пользователем предложения, а также за визуализацию ответа программной системы.



Речевое действие представлено в программной системе в виде композиции трёх компонентов: (сообщение, заявление, побуждение). Если компонент определённого типа присутствует в речевом действии, то приводится соответствующее правило, которое позволило определить этот тип.

Заключение

В данной работе были проведены исследования речевых действий, представленных в текстах естественного языка. В результате сделана классификация иллокутивных глаголов, соответствующих различным типам речевых действий, а также разработан алгоритм по определению типа (или комбинации типов) речевого действия.

На основе разработанного алгоритма была реализована программная система SpeechActs, которая позволяет определять тип речевого действия по предложению на естественном языке. Программная система также позволяет выводить информацию о том, по какому признаку (правилу) был определен тип речевого действия. Программная система SpeechActs интегрирована с программной системой LogicText, на основе которой она производит формальное представление информации, содержащейся в предложении естественного языка.

Разработанная программная система может быть применена для построения модели диалога, а также для построения модели пользователя при его общении с чат-ботом. Извлечение речевых действий дает возможность формального представления контекста диалога. В частности, речевые действия могут определять и переопределять смысл (определения) используемых понятий в контексте данного диалога.

Для дальнейшего совершенствования разработанной программной системы необходимо добавлять новые правила для более точного определения типа речевого действия. Помимо этого, планируется использовать методы определения интенсивности речевого действия, а также предполагается разработать методы построения фрагментов атомарных диаграмм для полного извлечения речевых действий из текстов естественного языка.

Список литературы

1. Елисеева О. Е. Естественнo-языковой интерфейс интеллектуальных систем: Учеб. пособие. Минск: Изд-во БГУИР, 2009. С. 84–85.
2. Dezhao S., Schilder F., Smiley C., Brew C., Zielund T., Bretz H., Martin R., Dale C., Pomerville S., Duprey J., Miller T., Harrison J. TR Discover: a natural language interface for querying and analyzing interlinked datasets // Proc. 14th Intern. conf. on the Semantic Web: ISWC 2015. Springer Intern. Publ., 2015. P. 21–37.
3. Крайванова В. А. Модель естественнo-языкового интерфейса для систем управления сложными техническими объектами и оценка эффективности алгоритмов на ее основе // Управление большими системами. М.: Изд-во ИПУ РАН, 2009. Вып. 26. С. 158–178.
4. Deshpande A. K., Devale P. R. Natural language query processing using probabilistic context free grammar // Intern. Journ. of Advances in Engineering & Technology. 2012. Vol. 3. No. 2. P. 568–573.
5. Li F., Jagadish H. V. NaLIR: An interactive natural language interface for querying relational databases // Proc. 2014 ACM SIGMOD Intern. conf. on Management of data, ACM. 2014. P. 709–712.
6. Пальчунов Д. Е. Решение задачи поиска информации на основе онтологий // Бизнес-информатика. 2008. № 1 (3). С. 3–13.
7. Dmitry E. Palchunov. Virtual catalog: the ontology-based technology for information retrieval // Knowledge Processing and Data Analysis. LNAI 6581. Berlin; Heidelberg: Springer-Verlag, 2011. P. 164–183.
8. Махасоева О. Г., Пальчунов Д. Е. Автоматизированные методы построения атомарной диаграммы модели по тексту естественного языка // Вестн. НГУ. Серия: Информационные технологии. 2014. Т. 12, № 2. С. 64–73.
9. Махасоева О. Г., Пальчунов Д. Е. Программная система построения атомарной диаграммы модели по тексту естественного языка. Св-во о гос. рег. программы для ЭВМ № 2014619198 от 10.09.2014.
10. Деревянко Д. В., Пальчунов Д. Е. Формальные методы разработки вопросно-ответной системы на естественном языке // Вестн. НГУ. Серия: Информационные технологии. 2014. Т. 12, № 3. С. 34–47.
11. Корсун И. А., Пальчунов Д. Е. Теоретико-модельные методы извлечения знаний о смысле понятий из текстов естественного языка // Вестн. НГУ. Серия: Информационные технологии. 2016. Т. 14, № 3. С. 34–48.
12. Пальчунов Д. Е. О логическом анализе естественного языка // Теория вычислений и языка спецификаций (вычислительные системы). 1995. Вып. 152. С. 61–75.
13. Dmitri E. Pal'chunov. Algebraische Beschreibung der Bedeutung von Aeusserungen der natuerlichen Sprache // Zelger, Josef; Maier, Martin (Hrsg.). GABEK. Verarbeitung und Darstellung von Wissen. Innsbruck; Wien: STUDIEN Verlag, 1999. S. 310–326.
14. Пальчунов Д. Е. Поиск и извлечение знаний: порождение новых знаний на основе анализа текстов естественного языка // Философия науки. 2009. № 4. С. 70–90.
15. Пальчунов Д. Е., Целищев В. В. Проблема извлечения знаний в системе взаимодействия человека и компьютера (онтологии и пресуппозиции) // Философия науки. 2012. № 4 (55). С. 20–35.
16. Остин Дж. Слово как действие // Новое в зарубежной лингвистике. М., 1986. Вып. 17: Теория речевых актов. С. 22–129.
17. Austin J. L. How to do things with words. Oxford Uni. Press, 1975.
18. Серль Дж. Р. Классификация иллокутивных актов // Новое в зарубежной лингвистике. М., 1986. Вып. 17: Теория речевых актов. С. 170–194.
19. Searle J. R. Indirect speech acts. 1975.
20. Searle J. R. A classification of illocutionary acts // Language in society. 1976. Т. 5, № 1. С. 1–23.
21. Searle J. R., Vanderveken D. Foundations of illocutionary logic. CUP Archive, 1985.
22. Searle J. R. Speech acts: An essay in the philosophy of language. CUP, 1969. Vol. 626.
23. Searle J. R. Expression and meaning: Studies in the theory of speech acts. CUP, 1985.

24. *Searle J. R. Austin on Locutionary and Illocutionary Acts // Philosophical Review. 1968. Vol. 77. No. 4. P. 405–424.*
25. *Vanderveken D. Meaning and speech acts: Volume 1, principles of language use. CUP, 1990. Vol. 1.*
26. *Friggieri J. et al. Actions and Speech Actions in the Philosophy of JL Austin. 1993.*
27. *Серль Дж. Р., Вандервекен Д. Основные понятия исчисления речевых актов // Новое в зарубежной лингвистике. М., 1986. Вып. 18: Логический анализ естественного языка. С. 242–264.*
28. *Vanderveken D. Fondements die la logique des attitudes // Manuscripto. 2006. Vol. 29. No. 2.*
29. *Бехтерева С. О. Анализ текстов на естественном языке на основании Теории речевых действий. Выпускная квалификационная работа бакалавра. Новосибирск, 2012.*
30. *Смыслы, тексты и другие захватывающие сюжеты: Сб. ст. в честь 80-летия И. А. Мельчука / Под ред. Ю. Д. Апресяна, И. М. Богуславского, Л. Ваннера, Л. Л. Иомдина, Я. Миличевич, М.-К. Л'Омм, А. Польгера. М.: Языки славянской культуры, 2012. 704 с.*

Материал поступил в редколлегию 07.08.2018

E. D. Makhina¹, D. E. Palchunov^{1,2}

¹*Novosibirsk State University
1 Pirogov Str., Novosibirsk, 630090, Russian Federation*

²*S. L. Sobolev Institute of Mathematics SB RAS
4 Academician Koptuyug Ave., Novosibirsk, 630090, Russian Federation*

ekatermakhina96@gmail.com, palch@math.nsc.ru

SOFTWARE SYSTEM FOR THE DETERMINATION OF SPEECH ACTS IN NATURAL LANGUAGE TEXTS

The paper is devoted to the development of methods of speech act extraction from natural language texts. Currently, one of the most significant problems is the problem of creating human-computer interfaces. The article discusses the task of development such interfaces by automated extraction of speech acts from natural language texts. The use of speech acts allows us to set the context of the dialogue and clarify the meaning of the concepts used. According to the results of the study, we develop a software system that is able to determine the types of speech acts represented in the natural language texts. One of the possible applications of the results of this study is to use the developed software system when creating chat bots.

Keywords: speech act, illocutionary force, model-theoretical methods, Lindenbaum – Tarski algebra, atomic diagram fragment, ontology, natural language interface.

References

1. *Eliseeva O. E. Estestvenno-yazykovoy interfeys intellektualnykh sistem [Natural Language Interface of Intellectual Systems]. Study guide. Minsk, BSUIR Publ., 2009, p. 84–85. (in Russ.)*
2. *Dezhao S., Schilder F., Smiley C., Brew C., Zielund T., Bretz H., Martin R., Dale C., Pomerville S., Duprey J., Miller T., and Harrison J. TR Discover: a natural language interface for querying and analyzing interlinked datasets. Proc. 14th Intern. conf. on the Semantic Web: ISWC 2015. Springer Intern. Publ., 2015, p. 21–37.*

3. Krayvanova V. A. Model of natural-language interface for complex technical objects control systems and efficiency evaluation of model-based algorithms. *Upravlenie bolshimi sistemami [Large-scale Systems Control]*. Moscow, Institute of Control Sciences Publ., 2009, vol. 26, p. 158–178 (in Russ.)
4. Deshpande A. K., Devale P. R. Natural language query processing using probabilistic context free grammar. *Intern. Journ. of Advances in Engineering & Technology*, 2012, vol. 3, no. 2, p. 568–573.
5. Li F., Jagadish H. V. NaLIR: An interactive natural language interface for querying relational databases. *Proc. 2014 ACM SIGMOD Intern. conf. on Management of data, ACM*, 2014, p. 709–712.
6. Palchunov D. E. The solution of the problem of information retrieval based on ontologies. *Bisnes-informatika*, 2008, no. 1, p. 3–13. (in Russ.)
7. Dmitry E. Palchunov. Virtual catalog: the ontology-based technology for information retrieval. *Knowledge Processing and Data Analysis. LNAI 6581*. Berlin, Heidelberg, Springer Verlag, 2011, p. 164–183.
8. Makhasoeva O. G., Palchunov D. E. Semi-automatic methods of a construction of the atomic diagrams from natural language texts. *Vestnik NSU. Series: Information Technologies*, 2014, vol. 12, no. 2, p. 64–73. ISSN 1818-7900 (in Russ.)
9. Makhasoeva O. G., Palchunov D. E. Program system for the construction of the atomic diagram of a model from natural language texts. Certificate of the State Registration of the computer program. No. 2014619198, registered 10.09.2014. (in Russ.)
10. Derevyanko D. V., Palchunov D. E. Formal methods of development of the question-answering system on natural language. *Vestnik NSU. Series: Information Technologies*, 2014, vol. 12, no. 3, p. 34–47. (in Russ.)
11. Korsun I. A., Palchunov D. E. Model-Theoretic Methods of Extraction of Knowledge on the Meaning of Concepts from the Natural Language Texts. *Vestnik NSU. Series: Information Technologies*, 2016, vol. 14, no. 3, p. 34–48. (in Russ.)
12. Palchunov D. E. On the logical analysis of natural language. *Theory of calculations and specification languages*. Novosibirsk, 1995, vol. 152: Computing systems, p. 61–75. (in Russ.)
13. Dmitri E. Pal'chunov. Algebraische Beschreibung der Bedeutung von Aeusserungen der natuerlichen Sprache. In: Zelger, Josef; Maier, Martin (Hrsg.). GABEK. Verarbeitung und Darstellung von Wissen. Innsbruck, Wien, STUDIEN Verlag, 1999, S. 310–326.
14. Palchunov D. E. Knowledge retrieval and elicitation: generation of new knowledge based on analysis of natural language texts. *Filosofiya nauki*, 2009, no. 4 (43), p. 70–90. (in Russ.)
15. Palchunov D. E., Tselishchev V. V. The problem of knowledge retrieval in interaction between a man and a computer: ontologies and presuppositions. *Filosofiya nauki*, 2012, no. 4 (55), p. 20–35. (in Russ.)
16. Austin J. L. Word as action. *New in foreign linguistics*. Moscow, 1986, iss. 17: The theory of speech acts, p. 22–129.
17. Austin J. L. How to do things with words. Oxford Uni. Press, 1975.
18. Searle J. R. Classification of illocutionary acts. *New in foreign linguistics*. Moscow, 1986, iss. 17: The theory of speech acts, p. 170–194.
19. Searle J. R. Indirect speech acts. 1975.
20. Searle J. R. A classification of illocutionary acts. *Language in society*, 1976, vol. 5, no. 1, p. 1–23.
21. Searle J. R., Vanderveken D. Foundations of illocutionary logic. CUP Archive, 1985.
22. Searle J. R. Speech acts: An essay in the philosophy of language. CUP, 1969, vol. 626.
23. Searle J. R. Expression and meaning: Studies in the theory of speech acts. CUP, 1985.
24. Searle J. R. Austin on Locutionary and Illocutionary Acts. *Philosophical Review*, 1968, vol. 77, no. 4, p. 405–424.
25. Vanderveken D. Meaning and speech acts. CUP, 1990, vol. 1: Principles of language use.
26. Friggieri J. et al. Actions and Speech Actions in the Philosophy of JL Austin. 1993.
27. Searle J. R., Vanderveken D. Basic concepts of calculus of speech acts. *New in foreign linguistics*. Moscow, 1986, iss. 18: Logical analysis of natural language, p. 242–264.
28. Vanderveken D. Fondements die la logique des attitudes. Manuscrito, 2006, vol. 29, no. 2.

29. Bekhtereva S. O. Natural language analysis based on the theory of speech acts. Bachelor final qualification work. Novosibirsk, 2012. (in Russ.)

30. Meanings, Texts, and Other Exciting Things: a Festschrift to Commemorate the 80th Anniversary of Professor Igor Alexandrovich Melchuk. Eds. Yu. Apresyan, I. Boguslavsky, M.-C. L'Homme, L. Iomdin, J. Miliüeviü, A. Polguère, L. Wanner. Moscow, Yazyki slavyanskoy kul'tury, 2012, 704 p. (in Russ.)

Received 07.08.2018

For citation:

Makhina E. D., Palchunov D. E. Software System for the Determination of Speech Acts in Natural Language Texts. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, p. 95–106. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-95-106