

УДК 004.5: 004.89

ЕСТЕСТВЕННО-ЯЗЫКОВЫЕ ИНТЕРФЕЙСЫ ИНТЕЛЛЕКТУАЛЬНЫХ ВОПРОСНО-ОТВЕТНЫХ СИСТЕМ

В.А. ЖИТКО¹, В.Н. ВЯЛЬЦЕВ², Ю.С. ГЕЦЕВИЧ², А.А. КУЗЬМИН³

¹Белорусский государственный университет информатики и радиоэлектроники
П. Бровки, 6, Минск, 220013, Беларусь

²Лаборатория распознавания и синтеза речи ОИПИ НАН
Сурганова, 6, Минск, 220012, Беларусь

³Белорусский государственный университет
пр. Независимости, 4, Минск, 220030, Беларусь

Поступила в редакцию 10 декабря 2011

Дано описание технологии проектирования естественно-языковых пользовательских интерфейсов, включающих в себя взаимодействие, используя естественные тексты, формальные тексты, синтез текста и речевой ввод сообщений. Статья описывает созданную вопросно-ответную систему с максимально комфортной для пользователя естественной формой коммуникации с помощью речи. Используемая технология развивается в рамках открытого проекта OSTIS [1]. Рассматривается библиотека компонентов проектирования естественно-языкового интерфейса: как ее пополнение сторонними компонентами, так и создание новых компонентов.

Ключевые слова: анализ естественно-языковых текстов, анализ речевого ввода, естественно-языковой интерфейс, речевой синтез, семантическая технология, технология проектирования.

Введение

В связи с динамичным развитием и распространением компьютерных систем возникает необходимость в снижении расходов на подготовку новых пользователей. Привлекательно в этом случае выглядит использование привычного для пользователя языка для организации его диалога с компьютерной системой. Такая возможность реализуется средствами естественно-языкового пользовательского интерфейса, обладающего рядом преимуществ: минимальной подготовкой пользователя, необходимой для работы с системой, простотой и высокой скоростью задания произвольных запросов к пользовательскому интерфейсу и высоким уровнем модели предметной области.

Для естественно-языкового пользовательского интерфейса вопросно-ответных систем возможно использование ограниченного набора лексики и грамматики, без серьезного ущерба функциональности вопросно-ответной системы. Ограниченный естественный язык – это подмножество естественного языка, текст на котором, без каких-либо усилий, воспринимается носителем исходного естественного языка, а также не требует длительного изучения для приобретения навыков составления текстов на этом языке, однако обладает сокращенным набором лексики и грамматики. Это позволит снизить время обработки естественно-языковых конструкций, а также устранить часть лингвистической неоднозначности.

Общая модель естественно-языкового интерфейса

Естественно-языковой пользовательский интерфейс, в рамках разрабатываемой технологии, рассматривается как специализированная интеллектуальная система, обеспечивающая диалог между прикладной вопросно-ответной системой и пользователем. Являясь интеллектуальной системой, естественно-языковой интерфейс включает в себя базу знаний, машину обработки знаний и пользовательский интерфейс.

Данный функционал естественно-языкового интерфейса интеллектуальной справочной системы определяет структуру интеллектуальной системы естественно-языкового интерфейса. Общая структура естественно-языкового пользовательского интерфейса имеет следующий вид:

- 1) база знаний:
 - лингвистическая база знаний;
 - предметная база знаний;
 - база знаний естественно-языкового интерфейса;
- 2) машина обработки знаний:
 - транслятор sc-конструкций в текстовое представление;
 - транслятор вопросов на семантический язык вопросов;
- 3) пользовательский интерфейс:
 - редакторы и просмотрщики текстов;
 - анализаторы и синтезаторы речи.

База знаний естественно-языкового интерфейса

База знаний естественно-языкового пользовательского интерфейса включает в себя лингвистическую и предметную базы знаний. Предметная база знаний используется вопросно-ответной системой для поиска и генерации ответов по запросам пользователя. Естественно-языковой интерфейс также использует эту базу знаний для решения ряда задач, связанных с разрешением неоднозначности естественного языка, т.к. предметная база знаний однозначно задает контекст диалога пользователя с вопросно-ответной системой [1].

Лингвистическая база знаний включает в себя формальное описание используемого естественного языка, привязку лексики к предметной базе знаний, спецификации семантических языков. Для описания знаний по каждому приведенному пункту существует специализированный семантический язык представления лингвистических знаний. Спецификация данного языка входит в состав лингвистической базы знаний [2].

Семантический язык вопросов

Семантический язык вопросов является одним из важнейших компонентов любой интеллектуальной системы, основанной на технологии OSTIS. Язык вопросов является семантическим языком представления знаний, текстами которого являются формальные записи вопросов и ответов. В естественно-языковом интерфейсе семантический язык вопросов выступает посредником между естественно-языковым пользовательским интерфейсом и интеллектуальной предметной системой.

На сегодняшний день существует множество различных языков запросов (языки запросов к базам данных, к данным, представленным по модели RDF, информационным системам, особенно к информационно-поисковым системам и др.):

- язык запросов XQuery для обработки данных в формате XML;
- язык запросов SQL к реляционным базам данных;
- язык запросов SPARQL к данным, представленным по модели RDF и др.

Для примера рассмотрим пример на языке SPARQL: поиск всех примеров каждого класса понятия треугольника.

Запрос на SPARQL будет выглядеть следующим образом:

- *SELECT \$x WHERE { <треугольник> rel:decomp [rel:example \$x] }*

Запрос, используя изоморфный поиск, представлен на рис. 1.

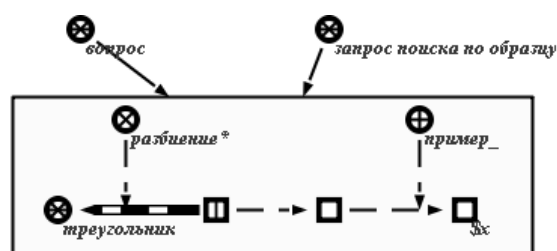


Рис. 1. Запрос изоморфного поиска

Из приведенного примера видно, что семантический язык вопросов обладает всеми функциональными возможностями языка запросов SPARQL, но в то же время обладает следующими достоинствами:

- позволяет в «краткой» форме формулировать известные системе вопросы (вопросы, описанные в базе знаний системы);
- расширять список возможных вопросов (семантический язык вопросов включает в себя средства описания и классификации вопросов);
- указывать дополнительную информацию о вопросе (автор вопроса, предметная область вопроса и пр.).

В то время как использование языка запросов SPARQL требует от пользователя знания внутренней организации интеллектуальной системы и как следствие, более высокого начального уровня подготовки пользователя, наличие упрощенной формы задания вопросов в семантическом языке вопросов позволяет пользователю общаться с интеллектуальной системой без знания внутренней структуры базы знаний.

Семантическая мощность языка вопросов в первую очередь зависит от многообразия типов вопросов. Т.к. для каждой предметной области могут существовать специфические только для нее типы вопросов, то в языке вопросов существует возможность расширять типологию и классификацию вопросов. Это позволяет создавать отдельные вопросы для специфических задач в предметных областях. Для этого достаточно описать новый вопрос и включить его в общую классификацию вопросов.

Машина обработки знаний

Машина обработки знаний естественно-языкового интерфейса включает в себя операции, обеспечивающие обработку различных лингвистических конструкций, перевод внешних языков на семантические языки интеллектуальной системы и обратно.

Все компоненты машины обработки знаний естественно-языкового интерфейса можно разделить на трансляторы и анализаторы [3]. Задачей трансляторов является перевод знаний из одного языка представления знаний в другой, к примеру, транслятор фактографических знаний по предметной области в текст на естественном языке. Задачей анализаторов является анализ фрагментов знаний и выявления ранее неизвестных фактов, к примеру, анализатор запроса пользователя, направленный на поиск цели и задачи вопроса.

Первыми этапами анализа запроса пользователя является морфологический и лексический анализ. Эти этапы на сегодняшний день хорошо проработаны и имеются их реализации. В системе такой анализ протекает поэтапно, «наращивая» формальное описание структуры предложения, используя информацию, полученную на этапе морфологического анализа. На этом этапе используются лексические и синтаксические правила используемого языка [4].

На следующем этапе – семантического анализа – строятся отношения эквивалентности элементов предложения и узлов в памяти системы. На этом этапе происходит соотношение лингвистической конструкции и конструкций в памяти системы, для выявления отношения соответствия, эквивалентности и т.д.

Результатом семантического анализа является конструкция запроса к системе. Сложность в получении такого результата заключается в семантической неполноте исходного естественно-языкового вопроса.

Вопросы, сгенерированные естественно-языковым интерфейсом, обрабатывает универсальный решатель вопросно-ответных систем, являющийся частью вопросно-ответной системы

по соответствующей предметной области. Стоит заметить, что универсальный решатель использует знания и лингвистической базы знаний таким образом система может отвечать на вопросы, связанные с используемым естественным языком.

Библиотека совместимых ip-компонентов естественно-языкового пользовательского интерфейса

Главным элементом семантической технологии компонентного проектирования естественно-языковых пользовательских интерфейсов является библиотека совместимых ip-компонентов. Такая библиотека включает в себя лингвистические базы знаний по разным языкам, различные трансляторы и анализаторы естественных текстов, элементы пользовательского интерфейса. Это позволяет проектировать естественно-языковые интерфейсы, комбинируя уже существующие компоненты, выбирая нужные лингвистические базы знаний, анализаторы и синтезаторы. Задачей разработчика в данном случае является привязка естественно-языкового интерфейса к предметной области интеллектуальной системы. Для этого необходимо добавить лингвистические знания о специфических для этой предметной области понятиях.

Разработка компонента речевого ввода информации на основе пакета НТК

Проблемы создания речевого дополнения к вопросно-ответному интерфейсу традиционны для систем распознавания речи: человеческая речь представляет из себя нестационарный сигнал, весьма сложный для анализа, физические характеристики сигнала в большой степени зависят от диктора (состояния голоса, тембра голоса, гендерной принадлежности диктора, языка ввода и т.д.).

К способам, призванным решить эти проблемы, относится исторически первый подход к распознаванию речи, основанный на технологии динамического искривления времени (Dynamic time warping (DTW)), а также практически заменившая его в последствии технология, основанная на скрытых марковских моделях (СММ).

Задача обработки сигналов. Для успешного решения проблемы распознавания речевых сигналов большую роль играет выделение их информативных признаков. Они должны соответствовать нескольким критериям [5]:

- быть легко вычисляемыми;
- сохранять всю необходимую информацию, которая содержится в сигнале;
- в режиме обучения на основании обучающего множества векторов признаков создавать модель, которая будет более общей, чем обучающий материал.

Алгоритм создания набора СММ для распознавания. Полный цикл создания хорошо обученных СММ включает два основных этапа: подготовка данных для обучения и непосредственно самообучение (рис. 2). Во многом успех распознавания зависит именно от качества обучающей информации. Поэтому этой части работы уделялось особенно много внимания.

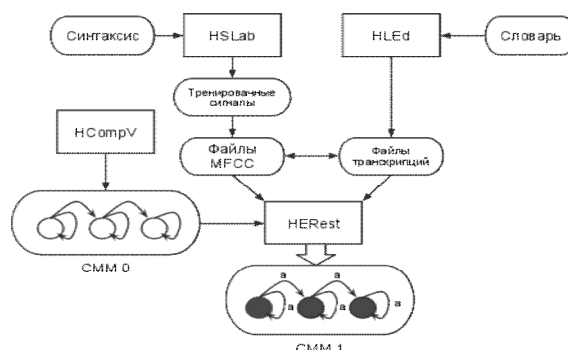


Рис. 2. Алгоритм создания набора СММ для распознавания

Подготовка данных.

Шаг 1 – грамматика. Первым шагом было создание грамматики. Так как система должна распознавать запросы к базе знаний, то распознаватель обязан идентифицировать последо-

вательность ключевых слов, выстроенных согласно грамматической последовательности в запросах.

Шаг 2 – словарь. Здесь первым делом создавался отсортированный список всех необходимых для распознавания слов. Далее запись каждого слова расширялась фонетической транскрипцией.

Шаг 3 – запись данных. На данном этапе осуществлялось создание набора файлов формата .wav, содержащих несколько вариантов (не менее трех) записей базовых слов, произнесенных диктором.

Шаг 4 – генерация файла транскрипций. Для обучения наборов СММ каждому файлу параметров тренировочных данных должна быть поставлена в соответствие фонетическая транскрипция.

Шаг 5 – кодирование данных. Финальным шагом в подготовке данных является обработка речевых сигналов и преобразование их в последовательности векторов признаков. Как и было упомянуто выше в данной работе, в качестве таких векторов были использованы кепстральные коэффициенты шкалы мел-частот.

Обучение. С этого момента начинается создание набора хорошо обученных СММ, все вероятности в которых описываются единичными Гауссианами.

Шаг 6 – создание начальных монофонов. Первым шагом в создании СММ является определение модели прототипа. Для системы, основанной на звуках, хорошей топологией является лево-правая с тремя состояниями.

Шаг 7 – пересчет коэффициентов. Параметры модели затем пересчитывались по методу Баума-Уэлча (алгоритм прямого-обратного хода или алгоритм максимизации правдоподобия).

В итоге были созданы три комплекта моделей: монофонные, трифонные, а также трифонные со связанными состояниями. Каждый из комплектов применяется для распознавания в прототипе системы.

Разработка компонента речевого вывода информации на основе синтезатора русского и белорусского языков

Для перевода текстовой информации в речевую из вопросно-ответной системы используется ip-компонент, основанный на синтезаторе речи по тексту. Данный ip-компонент реализован в качестве стороннего подключаемого модуля. Отметим, что внутренняя работа компонента не связана с семантическими сетями и взаимодействие с системой происходит с помощью трансляторов, но в результате такого подключения к системе синтеза речи по тексту пользователь получит озвученный ответ от системы. Таким образом, система естественно-языкового интерфейса станет еще более «естественной» для пользователя.

Синтез устной речи по тексту осуществляется на основе лексико-грамматического анализа входного текста путем моделирования процессов речеобразования с учетом правил произношения звуков и интонирования. Орфографический текст поступает на вход синтезатора и далее подвергается последовательной обработке рядом специализированных процессоров в соответствии с общей структурой синтезатора речи по тексту. Синтезатор включает модули: текстовый процессор, просодический процессор текста и сигнала, фонетический процессор и акустический процессор. Каждый из этих модулей поддерживается наборами соответствующих баз данных и правил.

Входной орфографический текст подвергается ряду последовательных обработок в каждом из процессоров. Текстовый процессор обрабатывает входной орфографический текст в следующей последовательности: очистка текста, преобразование знаков, расстановка словесных ударений и грамматических признаков словоформ.

Разработанная архитектура системы синтеза речи по тексту содержит в себе пятнадцать модулей, описание которых приводится ниже.

Главный модуль системы синтеза речи по тексту управляет работой всех остальных модулей, определением списка и загрузкой плагинов к остальным модулям, передачей данных между модулями.

Процессор нормализации текста перед лингвистической обработкой производит удаление из текста символов, не нужных для синтеза речи, удаляет случайное дублирование знаков препинания, заменяет похожие символы на один из них.

Главный модуль лингвистического процессора управляет другими модулями лингвистического процессора и контролирует преобразование этими модулями текста в последовательность синтагм. Процессор слов определяет возможные лексико-грамматические характеристики слова. Лексико-грамматический процессор определяет лексико-грамматические характеристики слова на основе вариантов, предложенных предыдущим процессором и лексико-грамматических характеристик других слов в тексте. Плагины дополнительных выражений производят обработку специальных выражений (например, чисел или сокращений) и преобразование их в словесное выражение. Процессор выражений расставляет ударения в словах, присоединяет к словам предлоги и частицы, заменяет конкретные выражения результатами обработки из плагинов. Процессор сборки словосочетаний соединяет отдельные слова в словосочетания, исходя из лексико-грамматических характеристик этих слов для недопущения разделения слов границей синтагмы. Процессор деления на синтагмы разделяет поток словосочетаний из предыдущего процессора на синтагмы, с указанием интонационного типа синтагмы в зависимости от знаков препинания и лексико-грамматических характеристик слов в словосочетаниях.

Фонетический процессор производит преобразование последовательности букв, из которых состоит синтагма в последовательность фонем.

Просодический процессор производит определение просодических характеристик (частоты основного тона, длительности, амплитуды сигнала) для каждой фонемы в последовательности, исходя из интонационного контура, определяемого типом синтагмы.

Акустический процессор соединяет аллофоны, определяемые фонемами, изменяет просодические параметры аллофонов, формирует звуковой сигнал. Контроллер преобразования звуковых форматов управляет плагинами, преобразующими звук.

Таким образом, спроектированная архитектура позволит разработать качественно новый синтезатор речи по тексту с высокой степенью «лингвистического понимания» входного текста и генерацией речи для самого широкого круга потребителей.

Практические и научные результаты

Результат данной работы представлен в качестве open source проекта на ресурсе <http://ostis.net>, где размещена вся документация по проекту а так же исходные коды реализации прототипа интеллектуальной справочной системы по геометрии с естественно-языковым интерфейсом. Результаты работы обсуждались на ряде конференций, в том числе и международных (OSTIS-2011).

Научные результаты проекта:

- предложена семантическая технология компонентного проектирования естественно-языковых интерфейсов;
- спроектирована семантическая модель естественно-языковых интерфейсов:
 - спроектирована база знаний;
 - спроектирован базовый набор компонентов машины обработки знаний;
 - специфицирован семантический язык взаимодействия естественно-языкового интерфейса с интеллектуальной справочной системой (язык вопросов);
 - специфицирован специализированный семантический язык представления лингвистических знаний.

Практические результаты проекта:

- описаны методы и средства проектирования естественно-языковых интерфейсов;
- спроектированы унифицированные интерфейсы реализации компонентов предложенной технологии;
- реализованы модули синтеза и анализа речи в качестве компонентов предложенной технологии;
- проведено обучение системы анализа речи (компонент может распознавать более 20 различных вопросов);

- настроены словари синтезатора речи (словари включают более 100 специфических терминов из предметной области прототипа системы);
- реализованы компоненты анализа и синтеза текста на основе семантического подхода к анализу тестов;
- реализован ряд (порядка 30) компонентов поиска ответов в семантическом пространстве;
- реализована и внедрена подсистема контроля компонентов, основанная на много-агентном подходе;
- реализован графический пользовательский интерфейс для естественно-языкового общения пользователя с интеллектуальной справочной системой по геометрии.

Заключение

Семантическая технология компонентного проектирования естественно-языковых интерфейсов вопросно-ответных систем предоставляет единую основу для анализа естественно-языковых текстов, начиная с ввода текста, морфологического, лексического, семантического анализа и генерации текста. Позволяет устанавливать отношения между эквивалентными лингвистическими конструкциями и структурой внутренней памяти системы.

Возможность интеграции сторонних разработок и проектов в качестве внешних ир-компонентов позволяет производить интеграцию различных подходов и методов в рамках одного проекта, что позволяет использовать их лучшие стороны.

NATURAL LANGUAGE USER INTERFACES FOR QUESTION ANSWERING INTELLIGENT SYSTEMS

V.A. ZHITKO, V.N. VALCEV, YU.S. GECEVICH, A.A. KUZMIN

Abstract

Description of technology to design a natural language interfaces for intelligent systems which is based on semantic networks (sc-technology) [1], applying to a prototype of natural language interface for question answering intelligent system on the geometry, is presented. Model includes input and output voice subsystems, the generation of responses in the form of natural language and formal text. Also article includes the description methods for linking various components of the system and description methods for development of new components.

Литература

1. Open Semantic Technology for Intelligent Systems. [Электронный ресурс]. Режим доступа: <http://www.ostis.net/>.
2. Сулейманов Д.Ш. // Сборник трудов Международной научно-технической конференции OSTIS-2011, Минск. 2011.
3. Byron Long // Dynamic Graphics Project Department of Computer Science University of Toronto, 1994.
4. Апресян Ю.Д. Избранные труды. Москва, 1995.
5. Житко В.А., Барбук С.Г. // Информационные системы и технологии. Сборник статей международной конференции-форума IST'2009. Минск. 2009.
6. Бовбель Е.И., Хейдоров И.Э. // Зарубежная радиоэлектроника. Успехи современной электроники. 1998. №3. С. 36–54.