

ПРИНЦИПЫ ПОСТРОЕНИЯ ЕСТЕСТВЕННО-ЯЗЫКОВЫХ ИНТЕРФЕЙСОВ ДЛЯ ИНТЕЛЛЕКТУАЛЬНЫХ СПРАВОЧНЫХ СИСТЕМ

ВВЕДЕНИЕ

Интеллектуальный интерфейс является важным компонентом интеллектуальной справочной системы для взаимодействия между интеллектуальной справочной системой и учащимися. Информация в интеллектуальной справочной системе хранится на формальном языке (например, в виде базы знаний), который не обязательно должны понимать обычные пользователи.

I. АНАЛИЗ СУЩЕСТВУЮЩИХ МЕТОДОВ ГЕНЕРАЦИИ ЕСТЕСТВЕННОГО ЯЗЫКА

Методы генерации естественного языка условно подразделяются на три категории: методы на основе шаблонов, методы на основе правил и методы на основе статистики. Методы на основе шаблонов используют предопределенный шаблон. Естественный язык генерируется путем заполнения данных в шаблоне. Методы на основе правил используют набор лингвистических правил для генерации естественного языка. В методах на основе статистики в большинстве случаев используются алгоритмы машинного обучения или глубокого обучения. Естественный язык генерируется путем обучения системы особенностям языка на основе некоторого корпуса.

Хотя некоторые методы в определенной степени реализуют процесс генерации естественного языка, но они не полностью применимы в интеллектуальных справочных системах. Интеллектуальные справочные системы в разных областях предъявляют разные требования к технологии генерации естественного языка. Разработка интерфейса интеллектуальных справочных систем требует применения гибридных методов.

II. ПРЕДЛАГАЕМЫЙ ПОДХОД

Предлагаемый подход строится на основе Технологии OSTIS, одним из ключевых принципов которой является использование онтологического подхода [1]. Интерфейс на китайском языке является компонентом, разработанным для интеллектуальных справочных систем (ИСС), разработанных на основе Технологии OSTIS.

Цянь Лунвэй, Ли Вэньцзу, аспиранты, кафедра интеллектуальных информационных технологий БГУИР, qianlw1226@gmail.com, wzzggml@gmail.com

Научный руководитель: Голенков Владимир Васильевич, профессор кафедры интеллектуальных информационных технологий БГУИР, доктор технических наук, golen@bsuir.by

Реализация интерфейса на китайском языке требует построения онтологии китайского языка. Структура онтологии китайского языка включает онтологию синтаксиса китайского языка и онтологию семантики китайского языка. В Технологии OSTIS формализация онтологии может рассматриваться как формализация и спецификация предметной области, включающая структурную спецификацию предметной области, терминологическую онтологию и др. Онтология китайского языка содержит лингвистическую информацию, такую как грамматическая информация и неглубокая семантическая информация о текстах на китайском языке.

При помощи понятий онтологии определяются конкретные лингвистические типы идентификаторов sc-элементов в базах знаний. Кроме того, могут определяться грамматические функции каждого конкретного фрагмента в предложении на естественном языке. После определения грамматических функций идентификаторов sc-элементов, в соответствии с грамматическими различиями разных языков, в процессе генерации естественного языка требуется разная обработка. Например, в английском и русском языках необходимо учитывать морфологические изменения слов, а в китайском языке необходимо добавлять функциональные слова, чтобы получить законченные предложения.

III. ЗАКЛЮЧЕНИЕ

Предложенный подход основан на Технологии OSTIS и реализует генерацию текстов. Онтология китайского языка предоставляет лингвистические знания для реализации генерации текстов. Генерация китайского языка в интерфейсе – это только часть разработки интерфейса на китайском языке. Интерфейс на естественном языке имеет большое значение для эффективности и применения ИСС.

1. Голенков, В. В. Гулякина, Н. А. Проект открытой семантической технологии компонентного проектирования интеллектуальных систем. Часть 1: Принципы создания. / В. В. Голенков, Н. А. Гулякина // Онтология проектирования. — 2014. — Vol. 1, P. 42–64.