

На правах рукописи

КЛЕВЦОВ ВАДИМ АЛЕКСАНДРОВИЧ

ОПРЕДЕЛЕНИЕ ПЕРЕЧИСЛЕНИЙ В ПРЕДЛОЖЕНИЯХ НА
АНГЛИЙСКОМ ЯЗЫКЕ НА ОСНОВЕ ИХ СИНТАКСИЧЕСКОГО
АНАЛИЗА

09.04.04 Программная инженерия

АВТОРЕФЕРАТ

диссертации на соискание степени магистра

Волгоград, 2016

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. В наше время набирает популярность дистанционное обучение. Дистанционное обучение использует различные подходы и методы, от видео лекций до систем управления обучением.

Одной из самых важных частей обучения, в особенности дистанционного, является тестирование полученных знаний.

Одной из самых развитых систем тестирования предоставляется системой управления обучением Moodle. Помимо тестов с вариантами ответа, тестов с развернутым ответом(эссе), существует возможность создания собственных типов вопросов за счет создания плагинов.

Однако не одна из систем обучения не может предоставить полноценного инструмента обработки ответов содержащих однородные члены или сложносочиненные предложения. Частичная их поддержка содержится в каждой подсистеме тестирования, то есть полный перебор вариантов перестановок однородных членов и сложносочиненных предложений, создание всех вариантов правильного ответа при создании вопроса. Так же СУО Moodle, есть типы вопросов, ответы для которых задаются с помощью шаблонов или регулярных выражений. Но в случае нескольких перечислений, особенно если одно из них является частью элемента другого перечисления, использование таких типов вопросов становится сложным из-за сложности шаблона и регулярного выражения, задающих правильный ответ, пропорционально количеству элементов перечислений и иерархии перечислений в предложениях.

Данные ограничения не позволяют в полной мере пользоваться вопросами с открытым ответом в рамках дистанционного обучения. Вот почему задача разработки метода и программного средства оценки ответа студента на естественном языке содержащего однородные члены и сложносочиненные предложения является актуальной.

Целью работы является снижение трудоемкости создания вопросов типа Correct Writing для СУО Moodle за счет автоматического определения

перечислений в английском языке на основе синтаксического анализа предложения.

Для достижения данной цели были поставлены и решены следующие задачи:

а) исследовать подходы, методы и средства обработки естественных языков

б) исследовать правила английского языка, выбрать структуры языка, изменение порядка записи которых, не изменяют семантику предложения, выбрать программное средство для осуществления синтаксического анализа;

в) разработать метод выделения перечислений в английском языке на основе синтаксического анализа;

г) реализовать и интегрировать метод в плагин Correct Writing, провести тестирование и эксперимент, оценить достижение цели.

Объектом исследования в диссертационной работе является процесс проверки ответов на естественном языке.

Предмет исследования: автоматическая проверка ответов на естественном языке с учетом наличия в них однородных членов и сложносочиненных предложений.

Методы исследования Для решения поставленных задач были использованы методы системного анализа, методы синтаксического анализа естественного языка - методы статистического анализа естественного языка.

Научная новизна работы.

Разработан метод определения перечислений, отличающийся от известных использованием синтаксического анализа.

Практическая ценность работы.

Разработанное метод позволяет с высокой точностью определить перечисления в предложении на естественном языке. Реализованное программное средство позволяет проводить оценку ответа студента на естественном языке, с учетом содержания в ответе однородных членов и сложносочиненных предложений.

Разработано программное средство использующее метод определения перечислений на основе синтаксического и семантического анализа, в процессе проверки ответа студента на естественном языке.

Положения выносимые на защиту:

а) Выработан метод определения перечислений в предложениях на естественном языке на основе синтаксического анализа.

б) Разработано программное средство проверки ответа студента на естественном языке, с учетом содержания в ответе однородных членов и сложносочиненных предложений.

в) Разработана схема интеграции программного средства с системой управления обучением Moodle.

Апробация работы. Основные положения и материалы диссертации докладывались на Внутривузовской научной конференции ВолгГТУ, Волгоград, 2014-2016 гг.

Публикации.

Структура и содержание диссертационной работы.

СОДЕРЖАНИЕ РАБОТЫ

Первая глава содержит анализ современного состояния вопроса поддержки перечислений в вопросах с открытым ответом систем дистанционного обучения. Также в глава содержится анализ существующих подходов к синтаксическому анализу естественного языка, сравнение средств синтаксического анализа, а также выбор средства для выполнения магистерской работы.

Вторая глава содержит описание процесса определения структур языка являющихся перечислениями, а также построение на их основе модели предложения, по которой возможно определить перечисления и их элементы.