

ЛАБОРАТОРНАЯ РАБОТА №8

ОРГАНИЗАЦИЯ ПОЛЬЗОВАТЕЛЬСКОГО ИНТЕРФЕЙСА НА ЕСТЕСТВЕННОМ ЯЗЫКЕ К ДИНАМИЧЕСКОЙ БАЗЕ ДАННЫХ.

1. Цель работы.

Целью работы является изучение лингвистических возможностей muLISP'a (newLISP-tk) и ознакомление с применением функционального программирования для решения современных задач теории и практики искусственного интеллекта.

2. Краткие теоретические сведения.

В настоящей работе рассматривается подход к организации интерфейса на Естественном Языке (ЕЯ), основанный на использовании ключевых слов.

Анализ ключевых слов - метод анализа ЕЯ-высказываний на предмет наличия ключевых слов, которые становятся значениями объектов предикатов. При этом компьютер одинаково реагирует на различные варианты входного текста, наличие грамматической правильности предложений не является обязательным, роль играет лишь наличие ключевых слов.

При разработке ЕЯ-интерфейса к Базе Данных (БД) с ключевым словом связывается либо характер действия программы (что хочет пользователь), либо значение некоторого атрибута искомого объекта. Например, при поиске по БД предприятий таким атрибутом может быть торговая марка (Старорусприбор, Новтрак, Квант, Планета), а ключевыми словами, определяющими действия - "Найти", "Выдать", "Вывести", "Выход". Процедура анализа входного текста с целью выделения ключей состоит из трех шагов :

- Ввод команды как строки символов.
- Преобразование введенной строки в список слов.
- Идентификация ключевых слов.

Для преобразования строки в список слов требуется написание функции, которая преобразует строку в список слов, причем под словом здесь понимается любая строка последовательность символов от одного разделителя до другого. В качестве разделителя может выступать пробел, символ возврата каретки, перевода строки, !, ", #, \$. Указанная функция может быть построена с применением встроенных функций обработки строк, функций ввода и вывода информации (о реализации данных функций в muLISP'e см. *Лекции 9 и 10*, а также файл *help.lsp*).

Один из наиболее распространенных и очевидных способов идентификации ключевых слов основан на частоте применения определенных грамматических конструкций : ключевые слова определяются позицией, которую они занимают в предложении. В настоящей работе рассматривается случай, когда ключевыми являются первое и последнее слово предложения. Первое слово определяет выполняемое программой действие, последнее – интересующий нас объект, информацию о котором мы хотим найти в БД. Работа с первым (командным) ключевым словом может быть организована по аналогии с обработкой пунктов

меню, см. *Лабораторную работу №6*. Здесь следует особо отметить необходимость выделения префикса, общего для разных поверхностных форм выражения одной и той же команды во входном ЕЯ-запросе. К примеру, применительно к поиску информации в качестве таких префиксов можно выделить "Най", "Выда", "Выве". Для выхода из программы это будут "Вых", "Вый", "exit", "quit". Выделение префиксов в muLISP'e можно организовать с применением встроенной функции SUBSTRING, используя заранее известную длину каждого префикса. Функция SUBSTRING имеет следующий синтаксис :

(substring <входная строка> <позиция начала подстроки> <позиция конца подстроки>), где номер позиции отсчитывается с нуля.

Аналогом функции substring в newLISP-tk является функция slice. Ее применение к строкам имеет следующий вид :

(slice <входная строка> <позиция начала подстроки> <длина подстроки>)

В отличие от substring отсчет позиций здесь производится с 1.

2. Задание на работу.

Изучить работу программы преобразования строки в список слов (файл key_wrds.lsp).

Дополнить реализованную в лабораторной работе №7 Систему Управления Базой Данных пользовательским интерфейсом на естественном (русском) языке. Подсистема понимания ЕЯ реализуется на основе описанной выше стратегии с использованием ключевых слов. Для формирования пользователем запроса на русском языке в произвольной форме в состав системы должен быть включен текстовый редактор. При его реализации рекомендуется использовать управляющие клавиши и списки свойств (см. *Лекцию 8*). Для запоминания таких текущих параметров редактирования, как координаты текущего положения курсора, параметры окна, рекомендуется использовать глобальные переменные.

3. Отчет по работе

Отчет по работе должен содержать :

- 1) Формулировку цели и постановку задач проводимых исследований;
- 2) Анализ задания, выбор метода решения с обоснованием, описание процесса разработки программ, полученные результаты и их анализ (обоснование);
- 3) Текст программы с комментариями и обоснованиями.
- 4) Выводы по проведенным машинным экспериментам.

Рекомендуемая литература.

Основная

- 1) Информатика и программирование шаг за шагом : Язык программирования LISP. // <http://it.kgsu.ru/Lisp/oglav.html>
- 2) Lutz Mueller newLISP For BSDs, Linux, Mac OS X, Solaris and Win32. Users Manual and Reference // http://www.newlisp.org/downloads/manual_frame.html

Дополнительная

- 3) Бородич Ю.С. и др. Паскаль для персональных компьютеров : Справ. пособие. – Мн.: Выш. шк., 1991.
- 4) Ин Ц., Соломон Д. Использование Турбо-Пролога : Пер. с англ. – М.: Мир, 1993. С. 452-473