

## ПРИМЕНЕНИЕ СРЕДСТВ РАСПОЗНАВАНИЯ РЕЧИ ПРИ РАЗРАБОТКЕ ОБУЧАЮЩИХ ПРОГРАММНЫХ МОДУЛЕЙ

Широкое применение автоматизированных голосовых порталов, справочных систем, автоинформаторов, автоответчиков, специальных программных средств, призванных упростить работу пользователя, вызвало сильный интерес к технологии распознавания речи человека (естественного языка).

Одной из тенденций современного образования является дистанционное обучение и применение компьютерных технологий. При работе с обучающей программой необходимо, чтобы компьютер понимал запросы пользователя. Наиболее простой способ выразить свой запрос: сказать. Следовательно, необходимо программное обеспечение, которое сможет распознавать речь человека. Такая программа намного упростит взаимодействие обучаемого с обучающим программным модулем.

В реализованном программном модуле предлагается сначала представить устную речь в виде массива слов, который затем будет обработан определённым образом. Именно обработка письменной речи определила первый этап исследования.

Сначала проводится лексический анализ фразы. Деление предложения на слова с использованием знаков препинания (при написании) или пауз (в разговоре). Далее в составе слов можно выделить корни, приставки и окончания.

Разработанный программный модуль делит запрос пользователя на отдельные слова и составляет массив из этих слов. Первый вариант программы производит сравнение по буквам (начиная с левой) каждого слова созданного массива с ключевой фразой, заранее определённой в каждой теме.

Например, на запрос пользователя: «сети» (рис. 1)

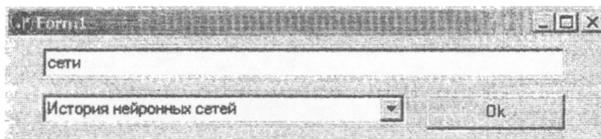


Рис. 1. Форма запроса.

получим результат (рис. 2):

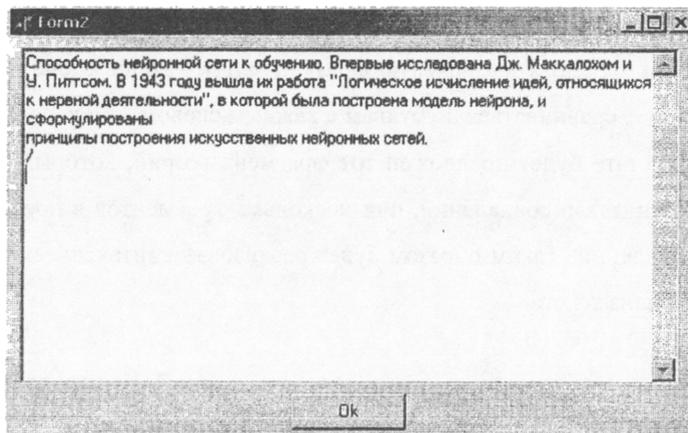


Рис. 2. Форма ответа на запрос.

Как видно из примера, первая буква в слове «сети» - «с», именно с неё начинается ключевая фраза «Способность нейронной сети к обучению».

Явный недостаток программы – хотелось бы получить несколько больше фрагментов теории, связанных со словом «сеть».

Тем не менее, при точных формулировках вопросов такой метод поиска оказывается допустимым, так как позволяет избежать лишних действий программы. Ответ на запрос обычно не содержит лишних теоретических фрагментов. Например, на запрос «Искусственные нейронные сети» нам предлагается следующий фрагмент теории (рис. 3):

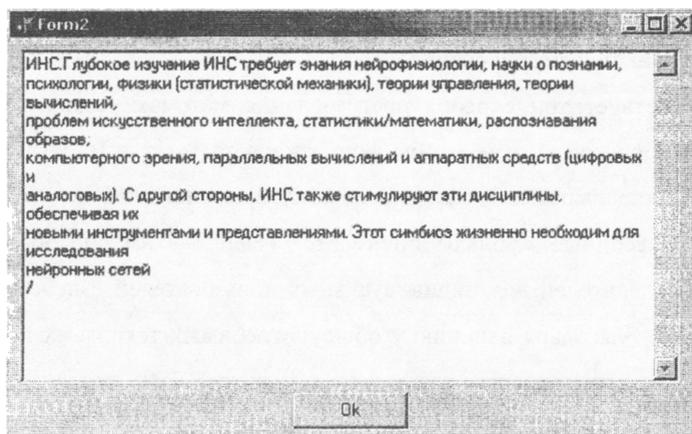


Рис. 3. Форма ответа на запрос.

Другой вариант программы, который сейчас находится в разработке, будет разбивать на слова и ключевую фразу, а затем, каждое слово из «массива запроса» будет сравниваться по буквам с каждым словом из «ключевого массива». В результате будет предложен тот фрагмент теории, который имеет наибольшее количество совпадений, или несколько фрагментов в порядке убывания их совпадений. Таким образом будет реализован синтаксический и семантический анализ текста.

**Ю. В. Угринова, гр. КТ-501**

## **ПРЕДСТАВЛЕНИЕ МАТЕМАТИЧЕСКИХ ДОКУМЕНТОВ В СЕТИ ИНТЕРНЕТ**

Сеть - это огромная доска объявлений. Все в большей и большей степени она становится средством для коммерции, и одновременно наиболее предпочтительным способом публикации материалов всех видов, включая и технические документы, содержащие математику. Наряду со вспомогательными программами, основные математические пакеты ПО играют важную роль в сетевых публикациях технического содержания.

Подавляющее большинство документов в Сети представлено в формате HTML (Hypertext Markup Language - язык разметки гипертекста), который не подходит для публикации многих технических документов, т.к. он не способен форматировать математические выражения с их специальными обозначениями, наподобие знаков интегрирования.

Существует три способа представления математических документов. Один из них состоит в том, чтобы форматировать текст в HTML и вставлять математические выражения как файлы с изображениями в документ HTML. Многие электронные журналы пользуются именно этим способом. Данный подход привлекателен для индивидуальных пользователей с небольшими потребностями. Благодаря наличию у общеупотребимых текстовых процессоров возможности экспортировать документы в формат HTML, создание необходимых для этого файлов не представляет труда.