

АВТОМАТИЗИРОВАННЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ

Автор программы проф. Ю.Н.Марчук

(Программа рассчитана на 32 часа). Целью данного курса является введение учащегося в проблематику современного информационного поиска. Объясняются основные понятия поиска, такие как ключевое слово, дескриптор, координатное индексирование, научные классификации знания и пр., информационно-поисковые языки, поисковый образ, поисковое предписание и т.д. Современные поисковые системы имеют широкую сферу применения. Распространяются экспертные системы, системы автоматизированного поиска информации, управления, машинного перевода, различного рода вспомогательные системы помощи человеку в выполнении интеллектуальных функций.

Основным содержанием данного курса являются лингвистические проблемы информационных систем. Общение человека с компьютером предполагает удобный естественно-языковой интерфейс, под которым понимается набор средств, позволяющих человеку взаимодействовать с компьютером без помощи программиста. Создание естественно-языкового интерфейса невозможно без решения лингвистических задач алгоритмизации работы с единицами главных языковых уровней.

Даются также и определенные сведения об организации сети научно-технического и управленческого информирования. Деловая проза существует в определенных рамках: в том, что касается обмена научной и управленческой информацией, эти рамки задаются системой производителей и потребителей такой информации. Организация информационных потоков внутри страны и вне ее определяет характер и объем деловой прозы, способы преодоления информационных и языковых барьеров, характер и возможности автоматизации обработки потоков информации на разных естественных языках и в различных информационных форматах.

Перспективы развития и совершенствования автоматизированных информационных систем зависят от успехов в фундаментальной области моделирования человеческого интеллекта. Воспроизведение роботом человеческих функций, таких, например, как распознавание речи, связано с разработкой сложных алгоритмов, лингвистических по своей природе и базирующихся на формализации обработки единиц практически всех языковых уровней – от уровня фонем и морфем до семантического, содержательного уровня с его сложными единицами. Особенность моделирования в практических целях заключается в том, что для распознавания единиц начального цикла, например, фонем или просодического содержания речи, требуется привлекать единицы других, более сложных уровней, в результате чего возникает комплексная система распознавания и машинного “понимания”.

Особую роль играют системы высокого интеллектуального уровня, такие, например, как системы машинного перевода. В курсе рассматриваются проблемы теории и лингвистической практики, затрагиваемые автоматизацией перевода с одного естественного языка на другой.

В ходе чтения курса предполагается проведение семинаров и практических занятий. Показ работы действующих систем машинного перевода предполагается осуществить в компьютерных классах. Анализ результатов работы систем позволит наметить линию взаимодействия человека с компьютером в решении интеллектуальных задач.

Программа называет основные разделы курса. В список литературы включена также только основная литература. Более широкий список литературы будет дан конкретно по каждой теме в ходе ее проработки.

ПРОГРАММА

Основные понятия современного автоматизированного информационного поиска. Информатизация мирового сообщества. Информационные языки. Ключевые слова, дескрипторы, координатное индексирование. Основные проблемы хранения, поиска и выдачи информации. Классификации наук и предметных областей. Словарный состав информационно-поисковых языков.

Информационно-поисковая система. Компоненты информационно-поисковой системы. Информационно-поисковые языки с грамматикой и без грамматики. Роль словаря-тезауруса в системе. Программная и техническая реализация ИПС.

Информационно-логические и экспертные системы. База данных и база знаний. Терминологические банки данных в информационных системах. Диалоговые системы. Проблемы разработки естественно-языковых интерфейсов.

Системы современного машинного перевода. История машинного перевода и развитие идей МП. Современное состояние теории и практики МП. Значение машинного перевода для развития кардинальных представлений лингвистики. Действующие системы машинного перевода и автоматические словари.

Организация системы научно-технической информации в стране. Потребители и производители информации. Информационные институты и центры документации. Барьеры на пути распространения информации.

Перспективы развития систем обработки информации и искусственного интеллекта.

ЛИТЕРАТУРА

(основная)

1. Атертон П. Справочное пособие по информационным системам и службам. – Париж: ЮНЕСКО / М.: ВЦП, 1980. 299 с.
2. Шемакин Ю.И. Начала компьютерной лингвистики. – М., МГОУ, 1992, 112 с.
3. Агеев В.Н., Узилиевский Г.Я. Человеко-компьютерное взаимодействие. – М., Мир книги, 1995, 288 стр
4. Марчук Ю.Н. Проблемы машинного перевода. – М.: Наука, 1983, 232 с.
5. Королев Э. И. Промышленные системы машинного перевода. – М.: ВЦП, 1991, 104 с.
6. Потапова Р.К. Речь: коммуникация, информация, кибернетика. – М.: Радио и связь, 1997, 522 ср.
7. Марчук Ю.Н. Теория и практика машинного перевода. Русский филологический вестник. – М., Московский Лицей, 1996, № 2, с. 123–135.
8. Диалог 95, Диалог 96, Диалог 97. Труды Международного семинара по компьютерной лингвистике и ее приложениям. – Казань 1995, 362 с.; М., 1996, 305 с.; М., 1997, 315 с.
9. Прикладное языкознание. Учебник / Отв. ред. А.С.Герд. – СПб.: СПб. Университет, 1996, 528 с.