

Общение с компьютером на естественном языке. Проблемы и задачи

Основы теории интеллектуального управления
Лекция 4.1

Бенджамин Уорф:

"Было установлено, что основа языковой системы любого языка (грамматика) не есть просто инструмент для воспроизведения мыслей. Напротив, грамматика сама формирует мысль, является программой и руководством мыслительной деятельности индивидуума."

Язык- это система взаимосвязанных категорий, которая, с одной стороны, отражает, с другой- фиксирует определенный взгляд на мир.

Прикладные системы, использующие ЕЯ

- Электронные переводчики,
- Распознаватели текста,
- Поисковые системы (в WEB, на рабочей станции – текстов, почты, в БД),
- Аннотаторы и рубрикаторы,
- Управление с помощью ЕЯ,
- Вопросно-ответные системы,
- Виртуальные личности.

Задачи при работе с ЕЯ:

- Анализ (понимание) ЕЯ,
 - Морфологический анализ,
 - Синтаксический анализ,
 - Семантический анализ,
 - Прагматический анализ,
- Синтез предложений ЕЯ,
- Классификация текстов,
 - Частотные методы,
 - Нейросетевые методы,
- Распознавание речи,
- Синтез речи.

Системы,
основанные на
использовании ЕЯ

Вопросно-
ответные

Управление
командами
на ЕЯ

"Читалки"

Электронные
переводчики

Поиск
информации по
запросу на ЕЯ

Понимание
ЕЯ

Генерация
ЕЯ

Анализ и синтез
речи

НГТУ, кафедра АППМ
А.В.Гаврилов

Может ли компьютер думать?

Вопрос связан с диалогом на ЕЯ

- Тест Тьюринга (Алан Тьюринг 1950)
- Китайская комната Серля (Джон Серль 1980)

Первые опыты

- **SHRDLU**

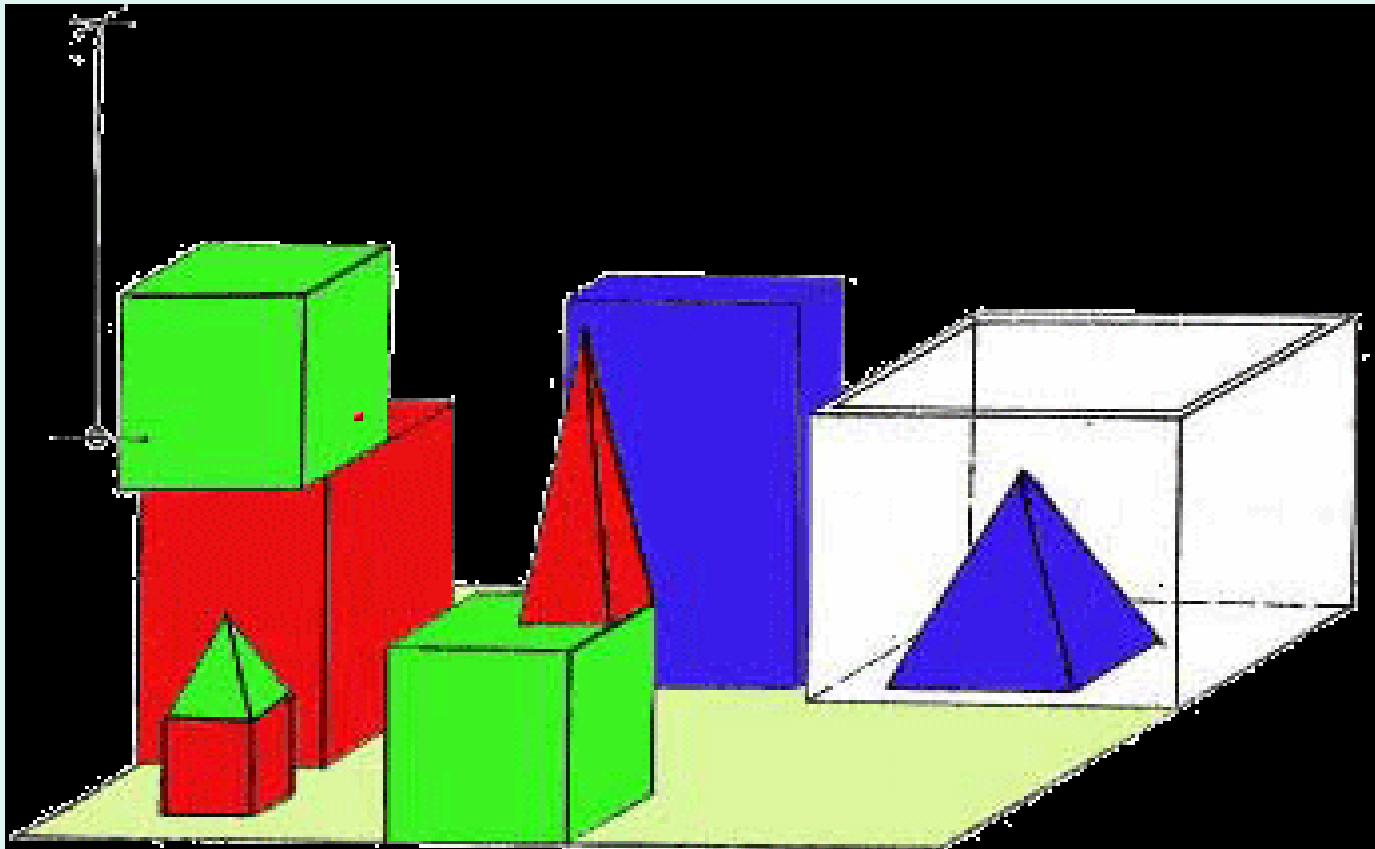
программа, разработанная Терри Виноградом в 1968-1970 г. в MIT: моделирует понимание компьютером ЕЯ в узкой области. Название происходит из сочетания самых частотных букв английского языка -

ETAOIN SHRDLU

- **Eliza**

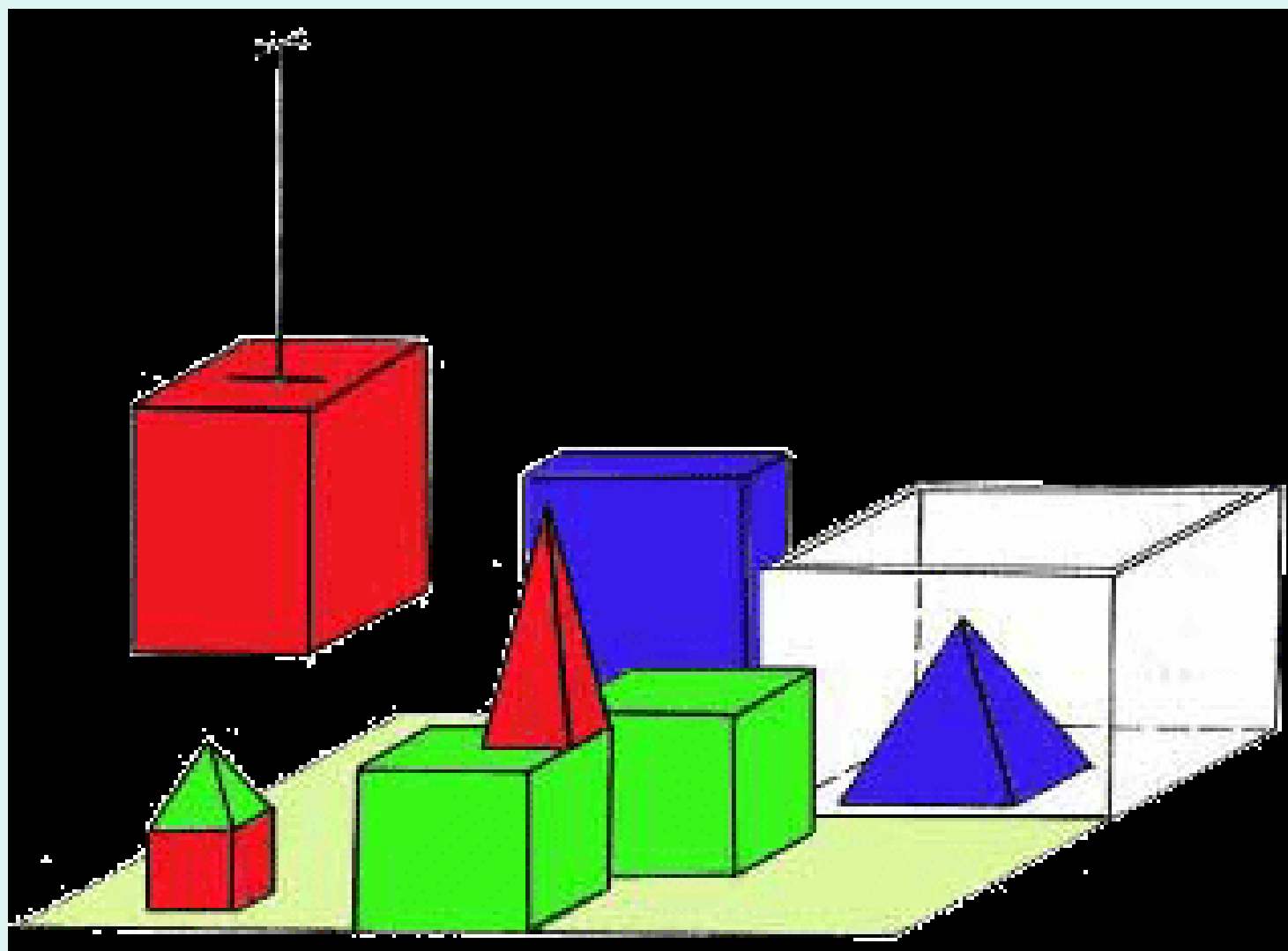
Виртуальный собеседник, созданный Джозефом Вейзенбаумом в 1966 г., пародирует понимание компьютером ЕЯ

SHRDLU



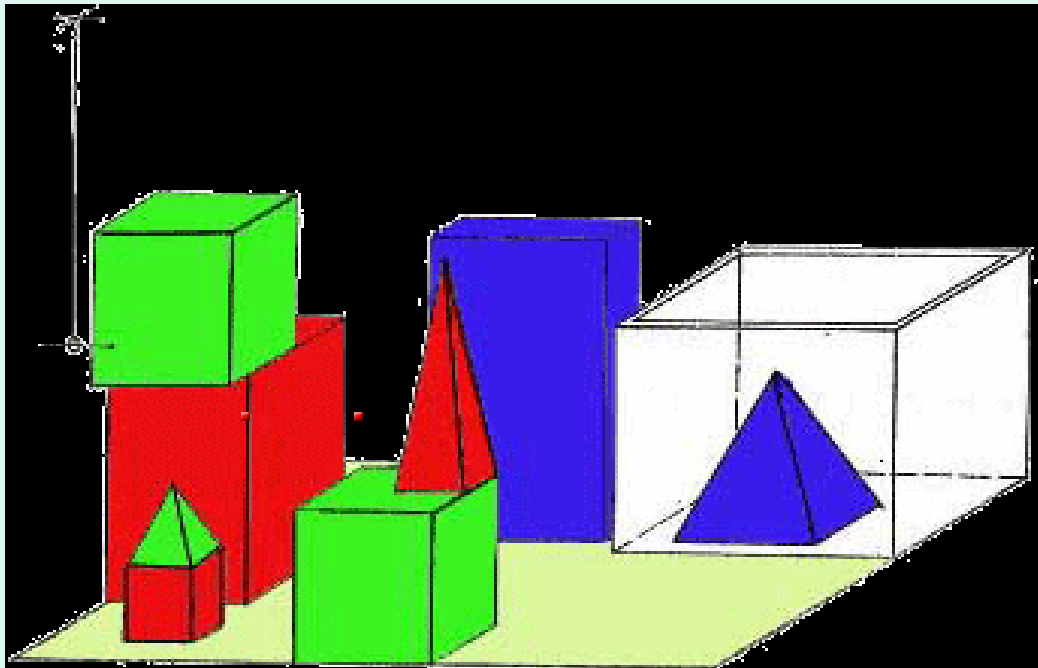
Человек: Возьми большой красный блок.

Компьютер: ОК.
ИГТУ, кафедра АППМ
А.В.Гаврилов



НГТУ, кафедра АППМ
А.В.Гаврилов

SHRDLU



Человек: Что находится в коробке?

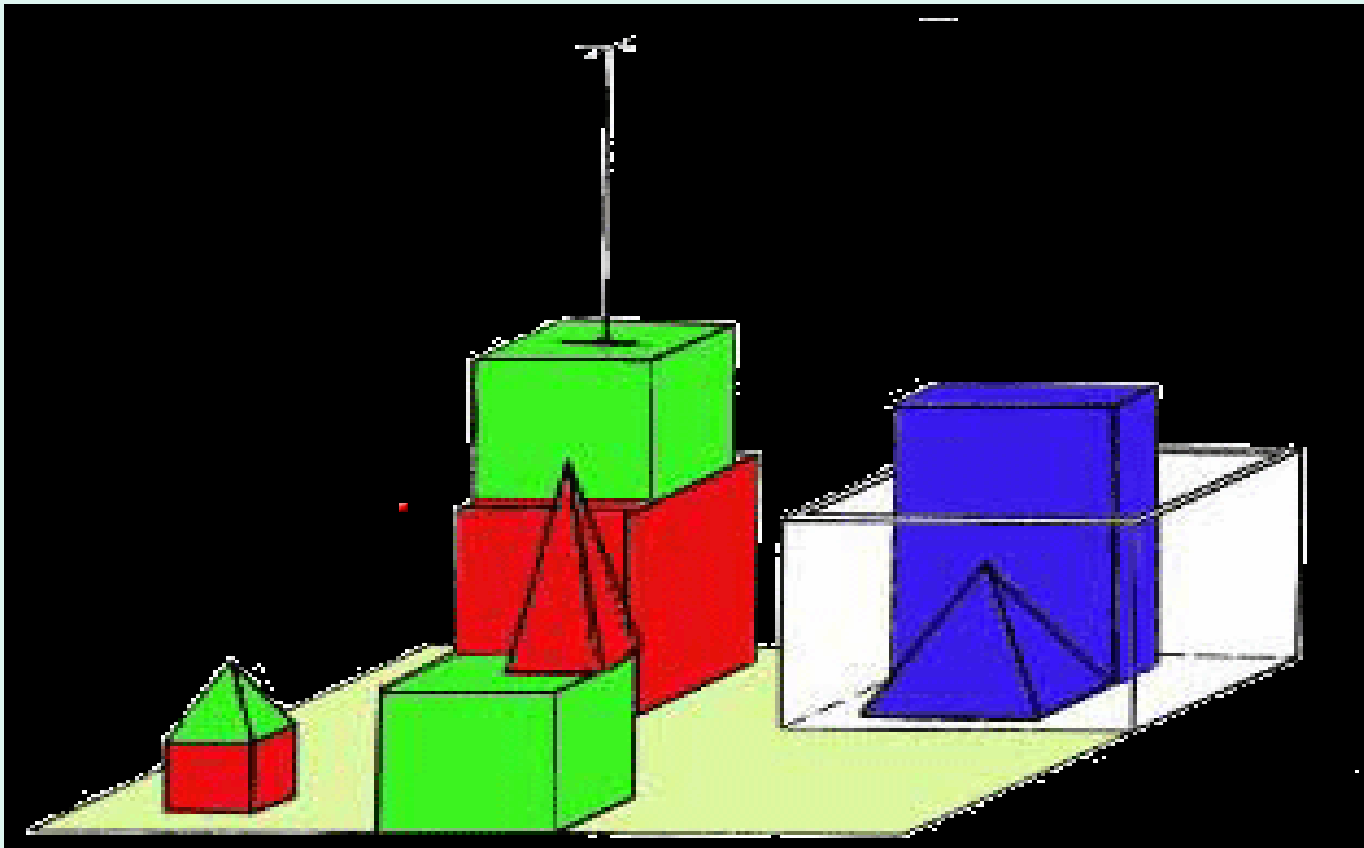
Компьютер: СИНЯЯ ПИРАМИДА И СИНИЙ БЛОК.

Человек: Что служит опорой для (этой) пирамиды?

Компьютер: КОРОБКА,

ИТХ, кафедра АППМ
А.В.Гаврилов

SHRDLU



Человек: Есть ли за пирамидой какой-нибудь большой блок?
Компьютер: ДА, ТАКИХ ТРИ: БОЛЬШОЙ КРАСНЫЙ, БОЛЬШОЙ
ЗЕЛЕНый КУБ И СИНИЙ БЛОК.

НГТУ, кафедра АППМ

А.В.Гаврилов

Человек: Помести малый (блок) на зеленый куб, на котором находится пирамида.

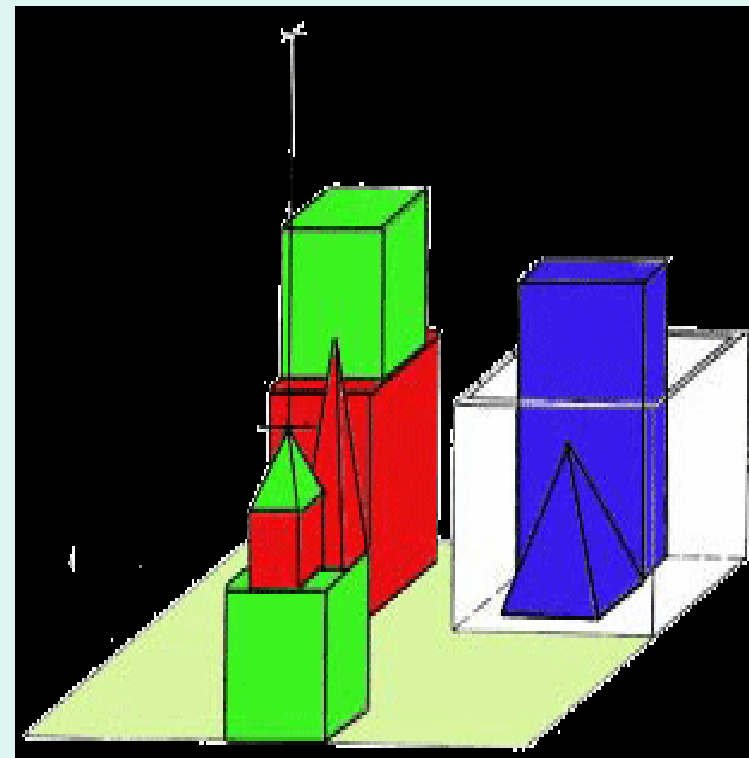
Компьютер: ОК.

Человек: Поставь самую маленькую пирамиду на него

Компьютер: ОК.

Человек: Служит ли самая короткая вещь, которая опирается на опору самой высокой пирамиды, опорой для чего-то еще?

Компьютер: ДА, ДЛЯ ЗЕЛеноЙ ПИРАМИДЫ.



Eliza и аналоги

- Классическая Eliza:

http://www-ai.ijs.si/eliza-cgi-bin/eliza_script

- Русский вариант: Киса

<http://hottabych.net/kisa/>

«Понимающий» компьютер должен осуществлять обработку всех уровней языка

- Фонетика
- Морфология
- Синтаксис
- Прагматика
- Дискурс

Решается только одна проблема:
НЕОДНОЗНАЧНОСТЬ

Компьютерная лингвистика

- направление в прикладной лингвистике, ориентированное на использование компьютерных инструментов – программ, компьютерных технологий организации и обработки данных – для моделирования функционирования языка в тех или иных условиях, ситуациях, проблемных сферах и т.д., а также вся сфера применения компьютерных моделей языка в лингвистике и смежных дисциплинах.
- область Искусственного Интеллекта, занимающаяся компьютерным моделированием владения языком с целью передачи информации, а также решением прикладных задач автоматической обработки текстов и звучащей речи
- Computational linguistics is a discipline between linguistics and computer science which is concerned with the computational aspects of the human language faculty. It belongs to the cognitive sciences and overlaps with the field of artificial intelligence (AI), a branch of computer science aiming at computational models of human cognition. Computational linguistics has applied and theoretical components.

Лингвистика

- Для моделирования языковой системы используются инструменты моделирования, пришедшие из физики, из экономики (и/или социологии), из физиологии, из философии и семиотики (теории о знаках).
- Лингвистика – хорошая лингвистика – должна уметь оценить рассматриваемый объект во всех этих плоскостях (быть междисциплинарной), конечно, если лингвистика – действительно наука о языке. Вернее сказать, это наука о языке (языковой системе) и тексте, формах и способах функционирования этой системы.

Основные понятия языка

- Дискурс (диалог) или текст
 - Контекст, активный диалог, роли участников диалога, прагматика (цели, мотивация)
- Предложение
 - Тип предложения, эллипсис (неполнота),
- Словосочетание
- Слово или словосочетание
 - Падеж (роль в предложении), часть речи, время
- Слог или морфема
- Символ (знак) или звук
- Синтаксис – структура предложения
- Семантика – смысл текста, предложения или словосочетания
- Прагматика – цели или мотивация

Два подхода к моделированию понимания смысла ЕЯ

- Подход основан на детальном синтаксическом разборе (**анализе**) предложения. Средствами синтаксического анализа вычленяются связанные понятия, которые объединяются в так называемые атомы смысла (АС). Создание АС идет только на основе данного предложения, определение связанных понятий идет только на основе синтаксических правил.
- На основе **распознавания** семантики. Разбор предложения идет путем вычленения связанных понятий с помощью базы знаний. База знаний хранит АС и определенным образом представленные связи между этими АС. На втором плане стоит синтаксический анализ с помощью которого вычленяются дополнительные АС, те которые не были сгенерированы с помощью БЗ.

Виды языковой неоднозначности:

- Синтаксическая (структурная) неоднозначность: во фразе *Time flies like an arrow* для ЭВМ неясно, идет ли речь о времени, которое летит, или о насекомых, т.е. является ли слово *flies* глаголом или существительным.
- Смысловая неоднозначность: во фразе *The man went to the bank to get some money and jumped in* слово *bank* может означать как банк, так и берег.
- Падежная неоднозначность: предлог *in* в предложениях *He ran the mile in four minutes/He ran the mile in the Olympics* обозначает либо время, либо место, т.е. представлены совершенно различные отношения.
- Референциальная неоднозначность: для системы, не обладающей знаниями о реальном мире, будет затруднительно определить, с каким словом - *table* или *cake* - соотносится местоимение *it* во фразе *I took the cake from the table and ate it*.
- Литерация (Literalness): в диалоге *Can you open the door? — I feel cold* ни просьба, ни ответ выражены нестандартным способом. В других обстоятельствах на вопрос может быть получен прямой ответ *yes/no*, но в данном случае в вопросе имплицитно выражена просьба открыть дверь.

Дополнительные трудности при диалоге на ЕЯ

- анафора (т.е. использование местоимений вместо их анафорических antecedентов - самостоятельных частей речи);
- Эллипсис (неполнота);
- экстраграмматические предложения (пропуск артиклей, опечатки, употребления междометий и т.п.);
- металингвистические предложения (т.е. попытка исправления введенного ранее).

Пример сложности распознавания смысла предложения

- Предложения «Мы отдали бананы обезьянам, потому что они были голодные» и «Мы отдали бананы обезьянам, потому что они были перезрелыми» похожи по синтаксической структуре.
- В одном из них местоимение *они* относится к обезьянам, а в другом — к бананам. Правильное понимание зависит от знаний компьютера, какими могут быть бананы и обезьяны.
- По нормам русского языка второе предложение некорректно, потому что в нем местоимение ссылается не на последнее подходящее слово, однако в *живой речи* такое предложение очень даже может встретиться.

Другие примеры трудностей

- **Свободный порядок слов** может привести к совершенно иному толкованию фразы: «Бытие определяет сознание» — кто кого определяет?
- В русском языке свободный порядок компенсируется развитой **морфологией**, служебными словами и знаками препинания, но в большинстве случаев для компьютера это представляет дополнительную проблему.
- В речи могут встретиться **неологизмы**, например, глагол «Пятидесятирублируй» — то есть высылай 50 рублей. Система должна уметь отличать такие случаи от опечаток и правильно их понимать.
- Правильное понимание **омонимов** — ещё одна проблема. При распознавании речи, помимо прочих, возникает проблема фонетических омонимов. Во фразе «Серый волк в глухом *лесу* встретил рыжую *лису*» выделенные слова слышатся одинаково, и без знания, кто глухой, а кто рыжий, не обойтись (Кроме того, что лиса может быть рыжей, а лес — глухим, лес также может быть рыжим (характеристика, в данном случае обозначающая преобладающий цвет листвы в лесу), в то время как лиса может быть глухой, что порождает дополнительную проблему, вытекающую из предыдущей.)

Применение речевых технологий

- управление базами данных;
- управление и регулирование технических систем;
- службы резервирования и заказов;
- диктовка и редактирование;
- услуги электронного секретаря;
- роботы;
- автоматический перевод;
- телефония;
- обеспечение конфиденциальности;
- помощь инвалидам (чтение для глухих, произнесение для немых).

Задачи, решаемые в речевых технологиях

- Анализ и распознавание речи
- Синтез речи

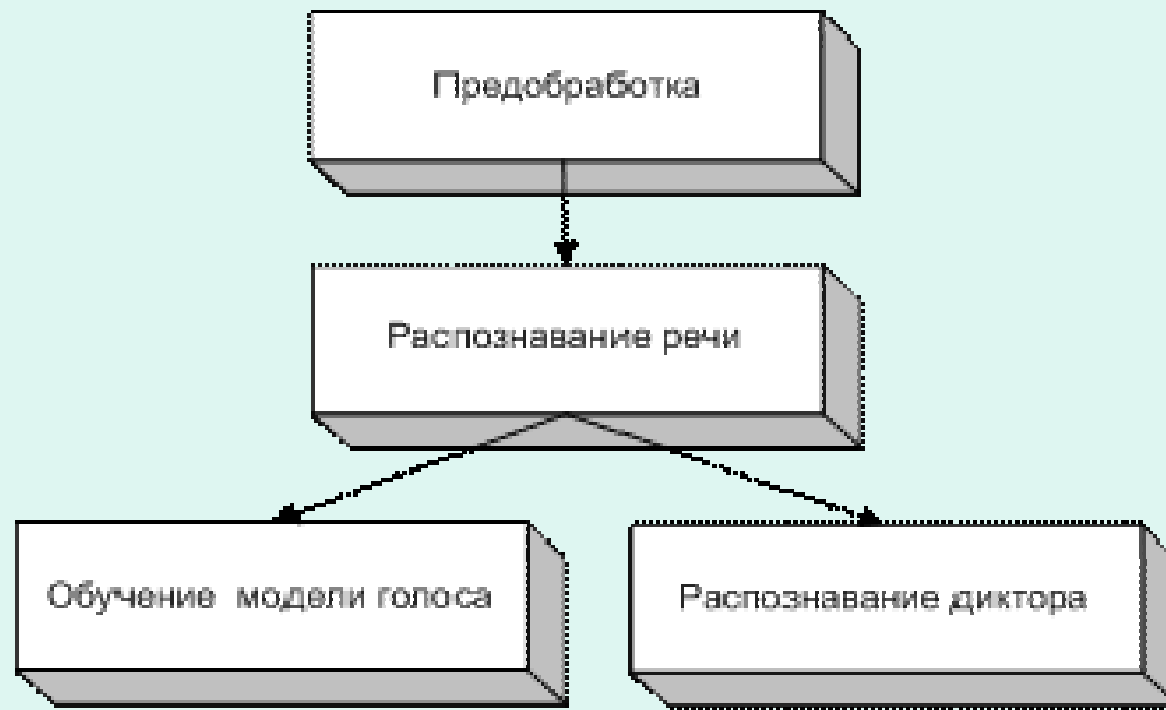
Специфические проблемы при распознавании речи

- Слитность речи (необходимо распознавать промежутки между словами для последующего синтаксического и семантического анализа)
- Существенные различия речевого сигнала для разных дикторов (дикторо-зависимое и дикторо-независимое распознавание речи), для разных эмоциональных состояний
- Звуковые помехи

Пример структуры системы распознавания речи

(Мурынин А. Б., В.Я. Чучупал, А. Десятчиков.

Вычислительный центр РАН)



Предобработка



Мел-кентральные коэффициенты и их производные по времени

$$\log S(i,t) = 1/2 (\log(\text{Re}[S(i,t)]^2 + \text{Im}[S(i,t)]^2))$$

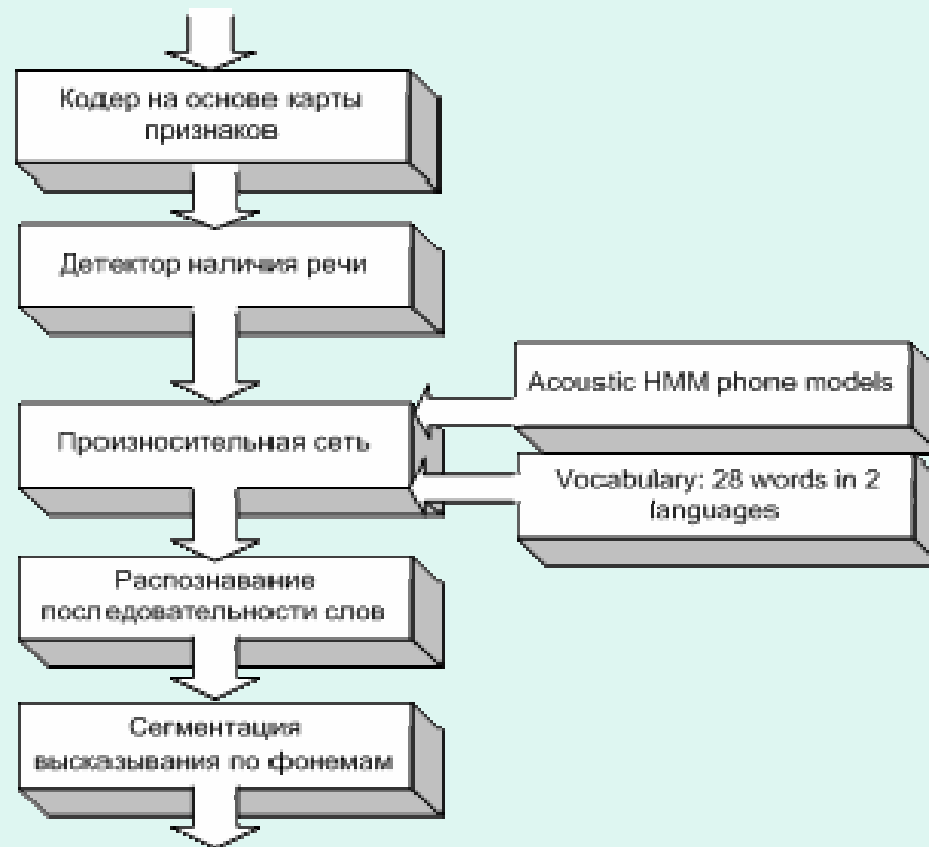
$$\{x(i)\}, i = 1 \dots 16 : f = 2595 \log_{10}(1 + |f|z/700)$$

$$\text{for } (k = 1; k \leq 16; k++) c[i] = x[k] * \cos((k-0.5) * (PI * i) / 16)$$

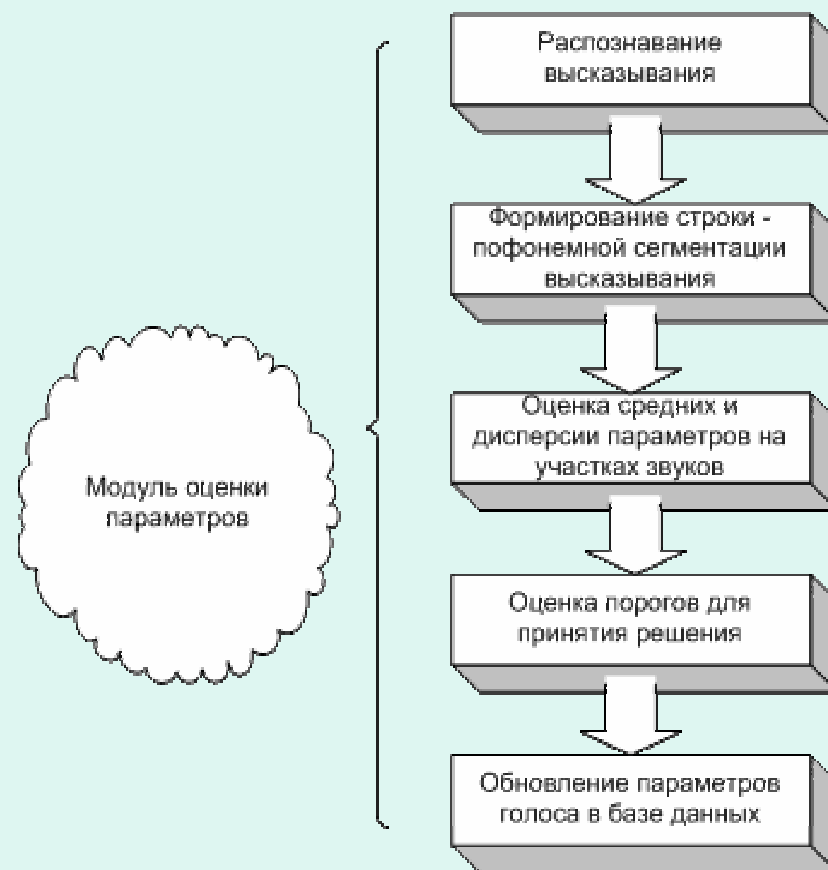
$$c[i] = c[i] * (1 + L/2 * \sin(PI * i / L))$$

$$dc(t) = -2 * c(t-2) - c(t-1) + c(t-1) + 2 * c(t-2)$$

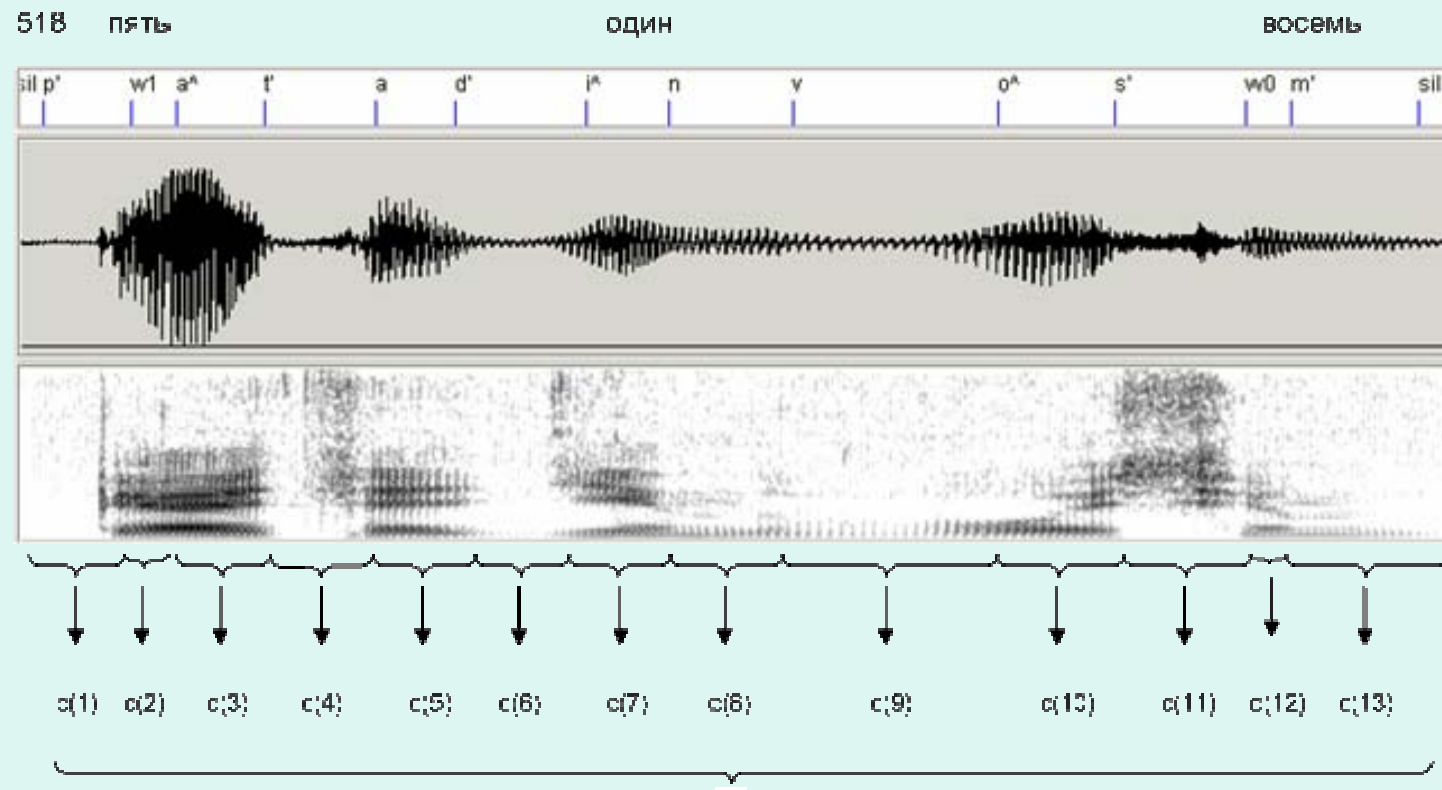
Распознавание речи



Обучение модели голоса



Сегментация и формирование шаблона



Шаблон голоса диктора для "один"

Распознавание диктора

