

Введение в робототехнику

Лекция 10.

Общение с роботом на естественном языке.
Проблемы и задачи

НГТУ, кафедра ПТМ
А.В.Гаврилов

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.

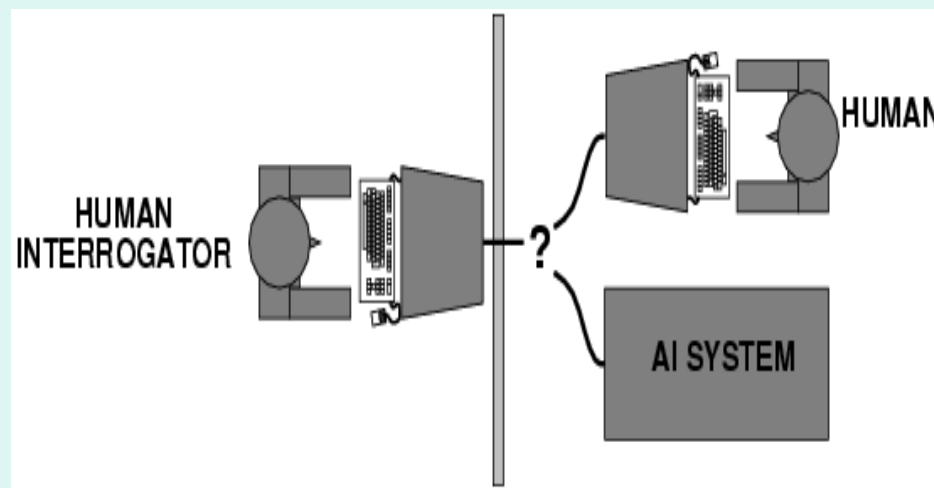
Может ли компьютер думать?

Вопрос связан с диалогом на ЕЯ

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

Тест Тьюринга

- Turing (1950) "Computing machinery and intelligence": «Может ли машина думать?» → «Способна ли машина к разумному поведению?»



Китайская комната

- Мысленный эксперимент, описанный Джоном Сёрлем, цель которого состоит в опровержении утверждения о том, что цифровая машина, наделённая искусственным интеллектом путём её программирования определённым образом, способна обладать сознанием в том же смысле, в котором им обладает человек. По сути, является критикой теста Тьюринга
- Эксперимент был представлен в 1980 году в статье «Разум, мозг и программы» («Minds, Brains, and Programs»).

Китайская комната (2)

- Возьмём, например, какой-нибудь язык, которого вы не понимаете. Для меня таким языком является китайский. Текст, написанный по-китайски, я воспринимаю как набор бессмысленных каракулей. Теперь предположим, что меня поместили в комнату, в которой расставлены корзинки, полные китайских иероглифов. Предположим также, что мне дали учебник на английском языке, в котором приводятся правила сочетания символов китайского языка, причём правила эти можно применять, зная лишь форму символов, понимать значение символов совсем необязательно. Например, правила могут гласить: «Возьмите такой-то иероглиф из корзинки номер один и поместите его рядом с таким-то иероглифом из корзинки номер два».

Китайская комната (3)

- Представим себе, что находящиеся за дверью комнаты люди, понимающие китайский язык, передают в комнату наборы символов и что в ответ я манипулирую символами согласно правилам и передаю обратно другие наборы символов. В данном случае книга правил есть не что иное, как «компьютерная программа». Люди, написавшие её, — «программисты», а я играю роль «компьютера». Корзинки, наполненные символами, — это «база данных»; наборы символов, передаваемых в комнату, это «вопросы», а наборы, выходящие из комнаты, это «ответы».

Китайская комната (4)

- Предположим далее, что книга правил написана так, что мои «ответы» на «вопросы» не отличаются от ответов человека, свободно владеющего китайским языком. Например, люди, находящиеся снаружи, могут передать непонятные мне символы, означающие; «Какой цвет вам больше всего нравится?» В ответ, выполнив предписанные правилами манипуляции, я выдам символы мне также непонятные и означающие, что мой любимый цвет синий, но мне также очень нравится зелёный. Таким образом, я выдержу тест Тьюринга на понимание китайского языка. Но все же на самом деле я не понимаю ни слова по-китайски. К тому же я никак не могу научиться этому языку в рассматриваемой системе, поскольку не существует никакого способа, с помощью которого я мог бы узнать смысл хотя бы одного символа. Подобно компьютеру, я манипулирую символами, но не могу придать им какого бы то ни было смысла.

Первые опыты

- **SHRDLU**

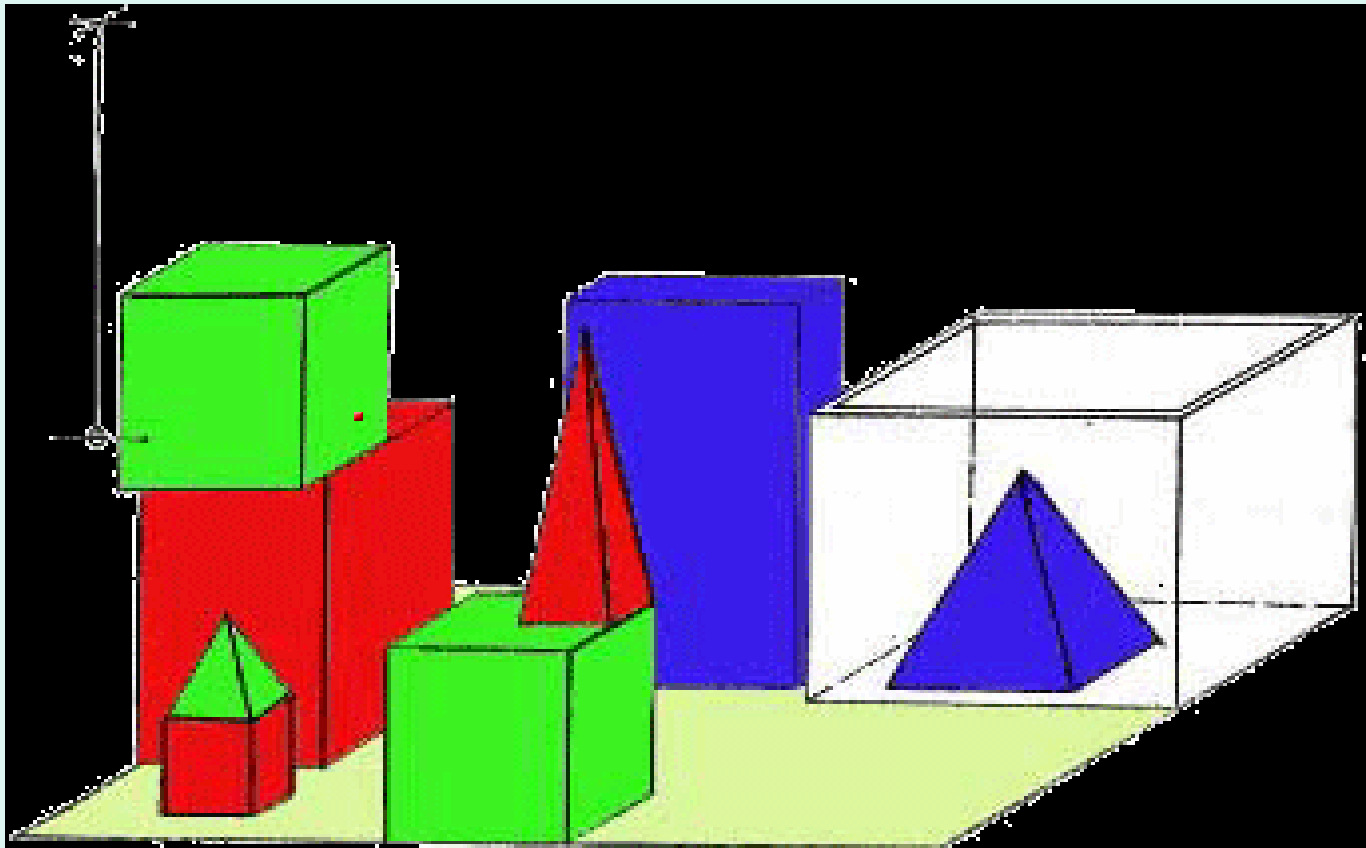
программа, разработанная Терри Виноградом в 1968-1970 г. в MIT: моделирует понимание компьютером ЕЯ в узкой области. Название происходит из сочетания самых частотных букв английского языка -

ETAOIN SHRDLU

- **Eliza**

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

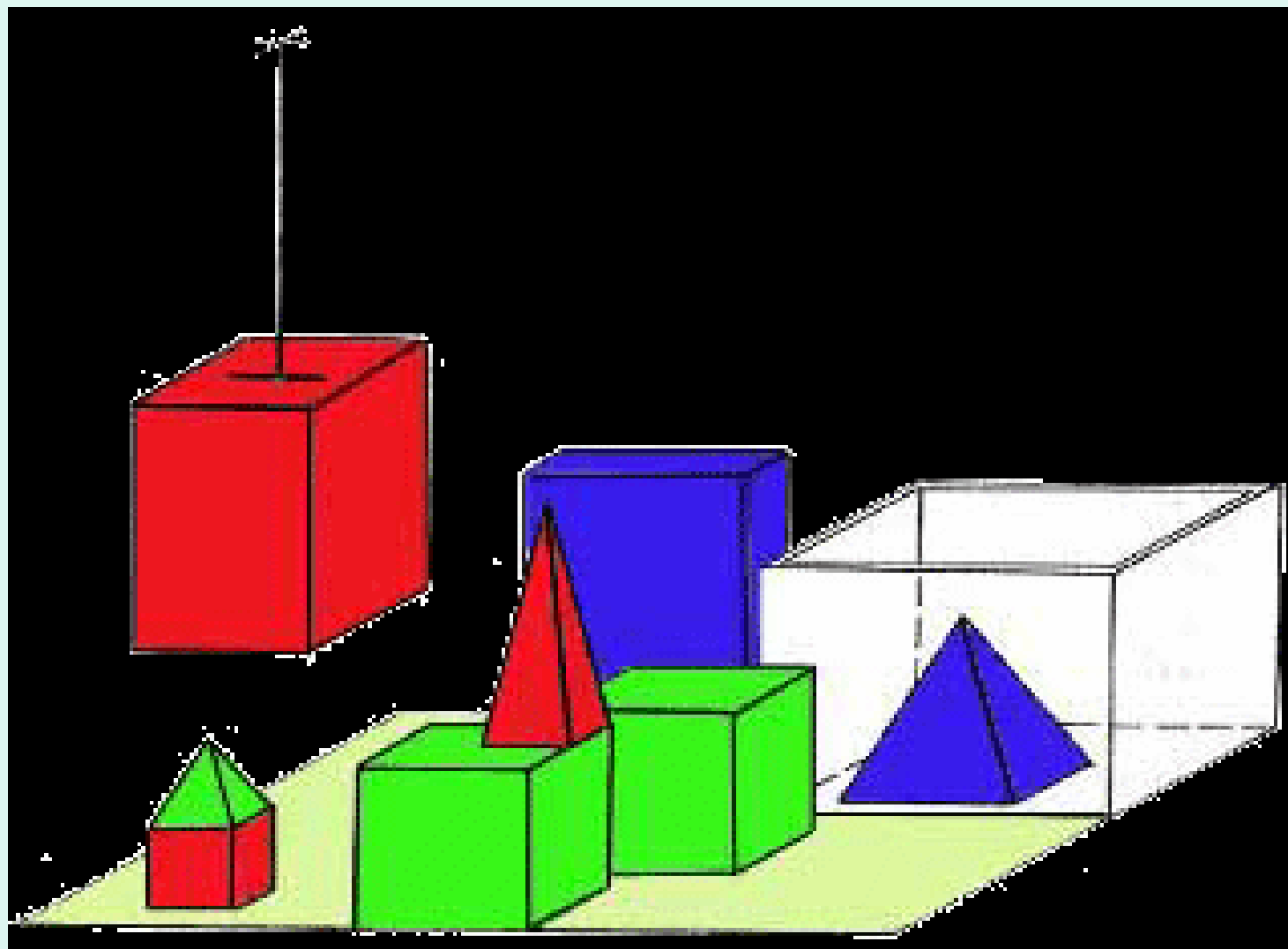
SHRDLU



Человек: Возьми большой красный блок.

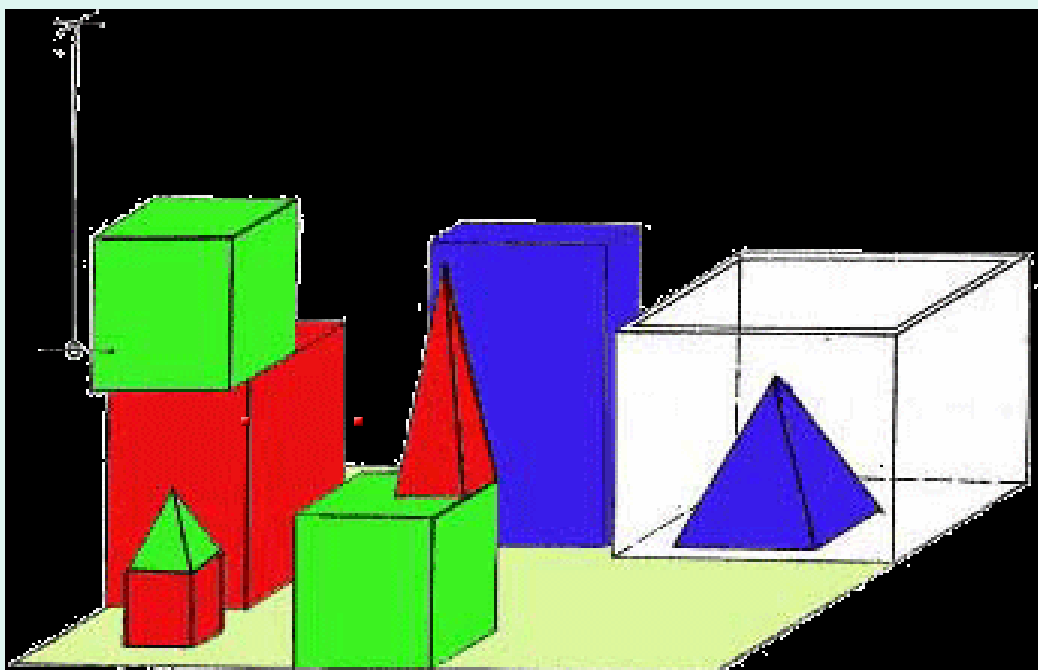
Компьютер: ОК. нгту, кафедра ПТМ

А.В.Гаврилов



НГТУ, кафедра ПТМ
А.В.Гаврилов

SHRDLU



Человек: Что находится в коробке?

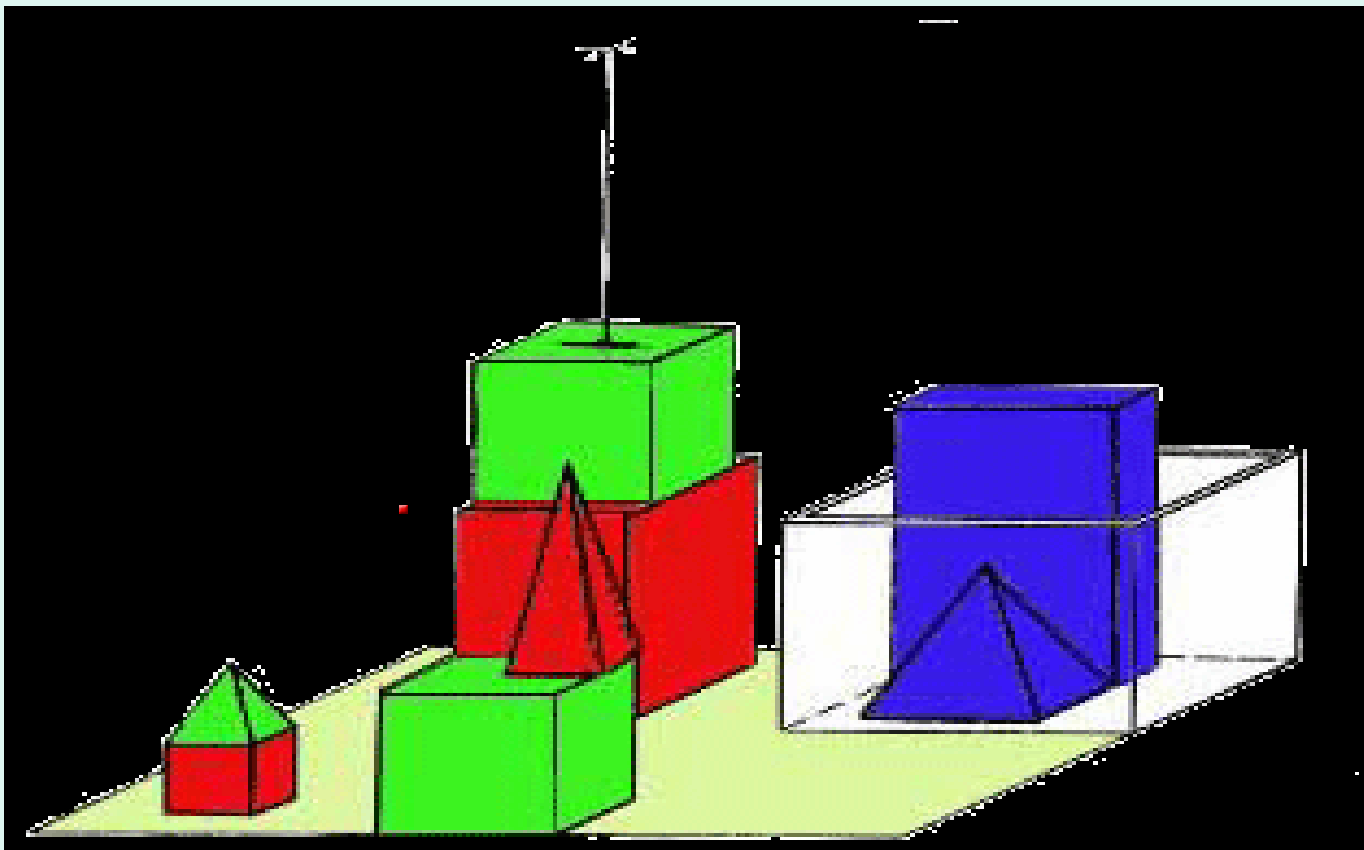
Компьютер: СИНЯЯ ПИРАМИДА И СИНИЙ БЛОК.

Человек: Что служит опорой для (этой) пирамиды?

Компьютер: КОРОБКА

ИКИИ, кафедра ПТМ
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 ПТМ

А.В.Гаврилов

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

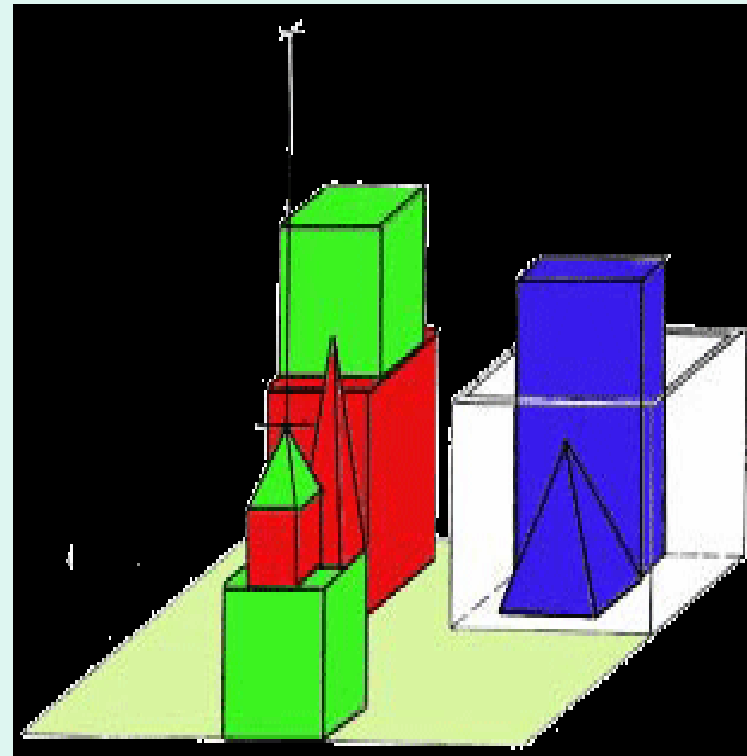
Компьютер: ОК.

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.



«Понимающий» компьютер должен осуществлять обработку всех уровней языка

- Фонетика
- Морфология
- Синтаксис
- Прагматика
- Дискурс

Решается только одна проблема:
НЕОДНОЗНАЧНОСТЬ

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 фонема
- 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

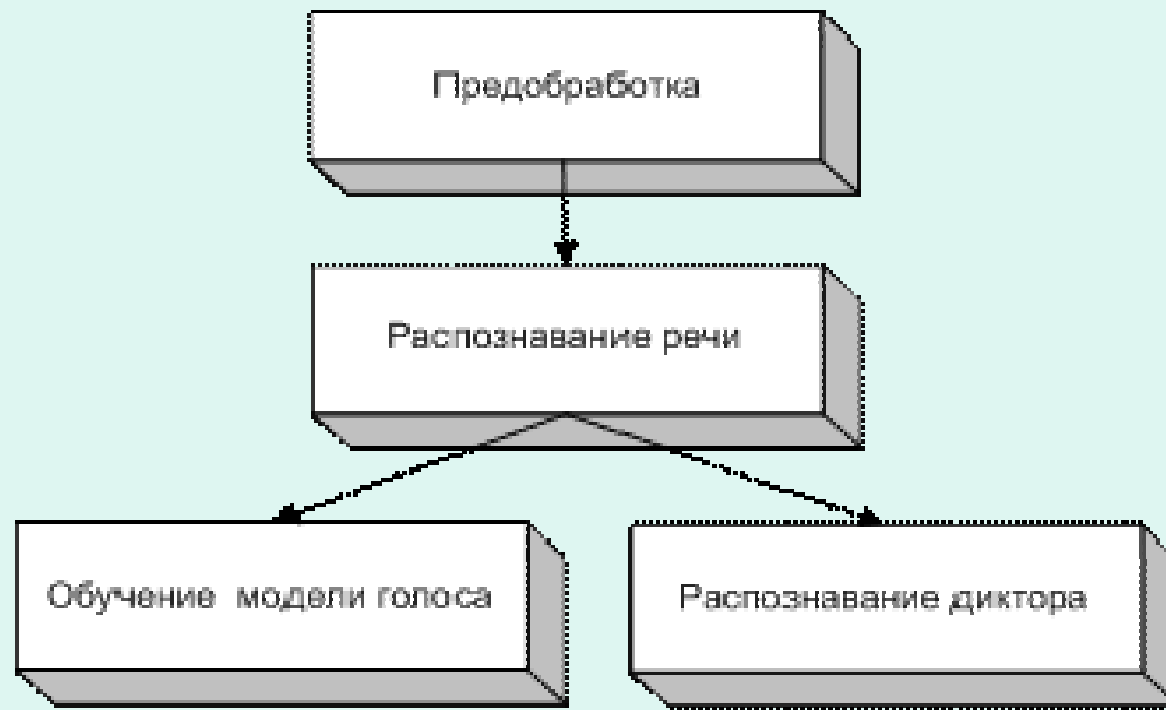
Специфические проблемы при распознавании речи

- 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

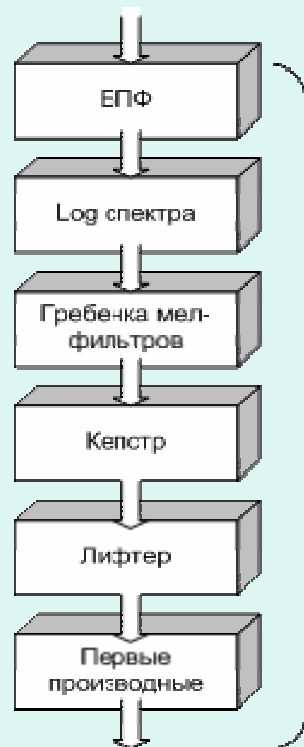
Пример структуры системы распознавания речи

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

Вычислительный центр РАН)



Предобработка



Мел-кентральные коэффициенты и их производные по времени

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

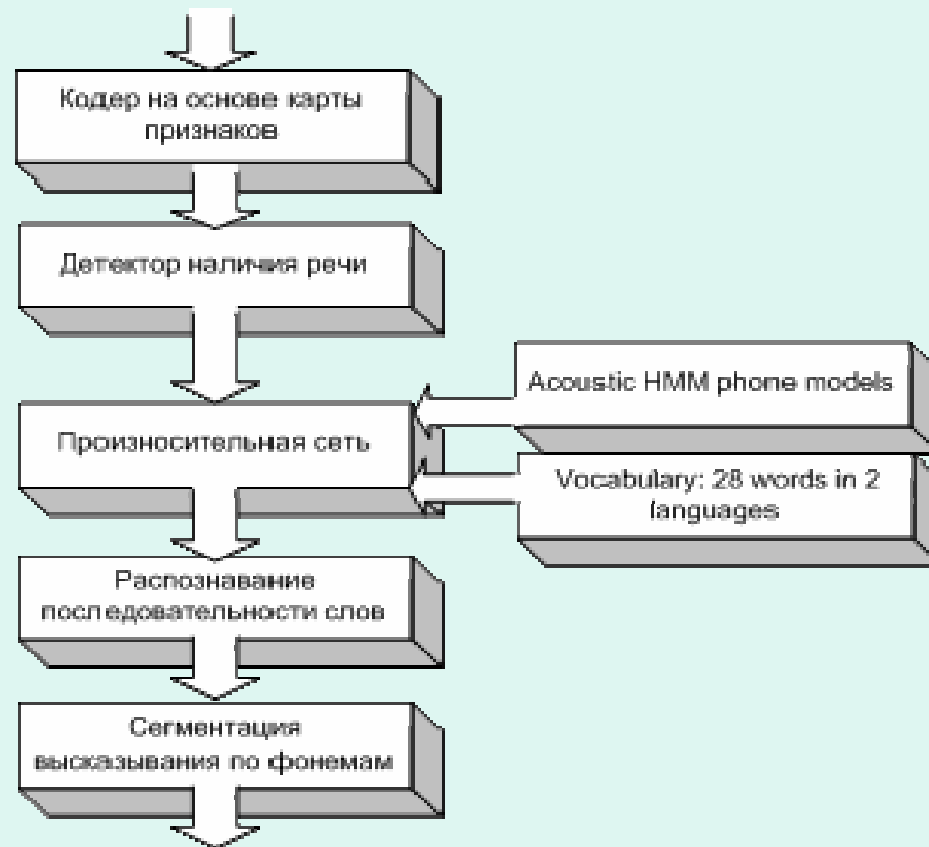
$$\{x(i)\}, i = 1 \dots 16 : f = 2595 \log_{10}(1 + |f|z/700)$$

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

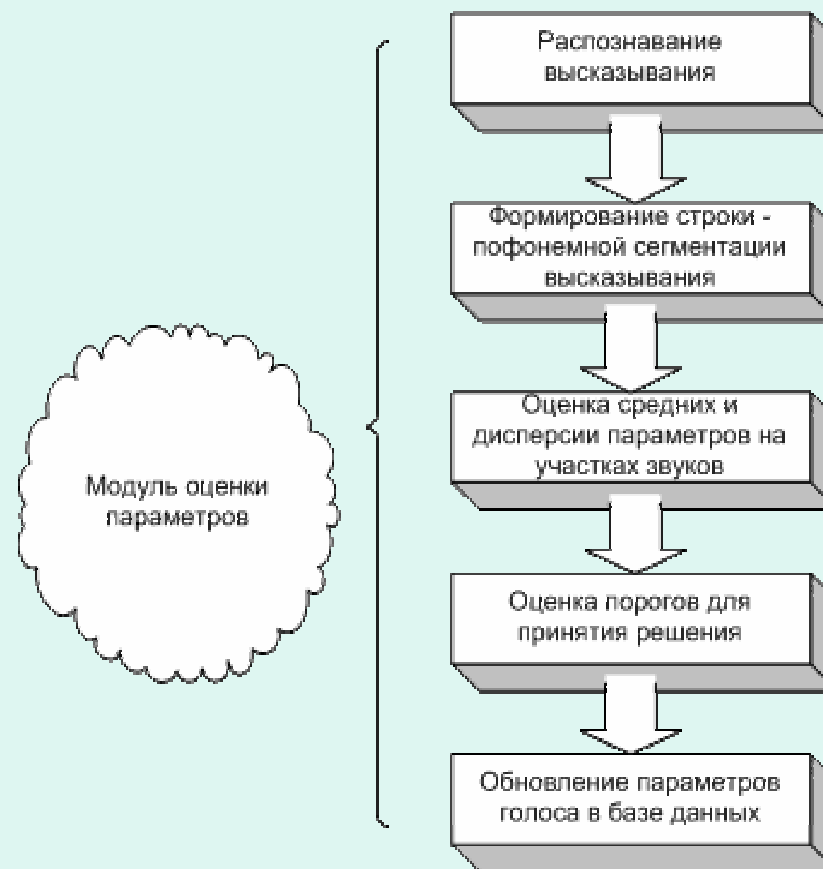
$$c[i] = c[i] * (1 + L/2 * \sin(PI * i / L))$$

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

Распознавание речи



Обучение модели голоса



Распознавание диктора

