

ТЕОРИЯ ПОРОЖДЕНИЯ РЕЧИ

Disclaimer

Всё услышанное и увиденное
является исключительно плодом
Вашего собственного
воображения

Основы ТПР

Концевой Михаил Петрович (ктр)

- 144 часа (74 аудиторных часа)
- 10 ч. лекций (ктр)
- 16 ч. лабораторных работ (Гусева А.В.)
- Зачет (ктр) (декабрь 2022 г.)

О допуске к зачету

20. Обучающиеся допускаются к сдаче зачета (дифференцированного зачета), экзамена по учебной дисциплине при условии выполнения и защиты ими всех расчетных, расчетно-графических, лабораторных и иных работ, курсовых проектов (курсовых работ), предусмотренных учебно-программной документацией соответствующей учебной дисциплины в текущем семестре.

Недопуск обучающегося к зачету (дифференцированному зачету) или экзамену по учебной дисциплине осуществляется решением соответствующей кафедры и отмечается в зачетно-экзаменационной ведомости словами «не допущен кафедрой» во время проведения зачета (дифференцированного зачета), экзамена.

ТЫ КТО ТАКОЙ?



Ваша специальность!

- «Современные иностранные языки (англ/нем) + (немецкий, английский) (**преподавание**)»
- со специализацией «Компьютерная лингвистика»

Ваша специальность!

- «Современные иностранные языки (англ/нем) + (немецкий, английский) (**преподавание**)»
- со специализацией «Компьютерная лингвистика»

Ваша специальность!

- «Современные иностранные языки (англ/нем) + (немецкий, английский) (преподавание)»
- со специализацией

Компьютерная лингвистика

Специализация

В образовательной системе:

- углублённое изучение относительно узкого поля деятельности,
- в рамках специальности,
- обеспечивающее необходимый уровень компетенции специалиста...

преподавателя англ. и нем. языка

~~Какого чёрта?~~

Причем здесь
компьютерная лингвистика?

Мир умных говорящих машин

Центр социального проектирования "Платформа" провел первый в мире соцопрос чат-ботов (**Алиса, Олег, P-Bot, Evie, Siri, Mitsuku, Rose**), приурочив его к запуску **Skybot F-850 (FEDOR)** с космодрома "Байконур".

Я узнал, что у меня
Есть огромная семья
И тропинка и лесок
В поле каждый колосок
Речка, небо голубое —
Это все мое родное
Это Родина моя,
Всех люблю на свете я!

Санкт-Петербург

- С февраля 2022 голосовой помощник Виктория оформила более **900 000** заявок на вызов врача на дом в Службе 122, что позволило справиться с огромным потоком звонков...

4 285 в день

Тюмень

- Группа компаний ЦРТ, входящая в экосистему «Сбера», установила в «Центре обеспечения мер социальной поддержки» Тюменской области интеллектуальный диалоговый ассистент, который совершает 1 500 звонков в месяц с целью информирования населения ...

Тюмень

- Голосовой робот Жанна, собранный на платформе VoiceBox от МТТ вышла на работу в Тюменский расчетно-информационный центр и обрабатывает более 66% звонков в контакт-центр 24 часа в сутки семь дней в неделю, а также помогает формировать ежемесячные счета.



APPKODE сегодня в 10:29

Алиса, Алекса летит в космос! Как голосовые помощники бороздят просторы Вселенной

Космонавтика, Голосовые интерфейсы*



08/31/2022 | **Responsible Business** | Parenting in a Digital World

Alexa, what is digital parenting?

By: Melissa Fischler Hed

If you're wondering what it means, you're not alone. Essentially—it's an extension of your own parenting, with a few twists.

HOME > SMART HOME

A Beginner's Guide to Amazon Alexa Terminology

Better understand the world of Alexa with our guide.

BY SIMON HILL

PUBLISHED 2 DAYS AGO



Агент Ева

Ваш личный помощник. Блокирует спам, отвечает
на звонки, защищает в интернете



Бесплатно

+7 (499) 123-45-67

«Агент Ева»:

Магазин

Ева: Секретарь Ева, говорите.
Абонент: Д. П.

02 сентября 2022 в 14:58

Как работает голосовой помощник «Ева» от «МегаФона»

КОМПАНИИ

МЕГАФОН

ГОЛОСОВЫЕ ПОМОЩНИКИ

«МегаФон» запустил для всех абонентов голосового помощника «Ева». Точнее, сразу двух: доступны «Ева» с женским голосом и «Максим» с мужским. Оба предоставляются бесплатно.

Яндекс перезапустил «Балабобу» и научил ее генерировать текст на двух языках

Нейросеть не знает, что говорит, и может сказать всякое — если что, не обижайтесь. Распространяя получившиеся тексты, помните об ответственности.

Договорились

18+

{ 20 }

ТПР = NLG

- Теория порождения речи
(проблемное поле, инструментарий ++),
может быть понята
(в контексте специализации)
как **Natural Language Processing (NLP)**

$$\text{NLP} = \text{NLG} + \text{ASR} + \text{NLU}$$

Natural Language Processing (NLP)

- общее направление искусственного интеллекта и компьютерной лингвистики, изучающее все проблемы компьютерного моделирования (анализа и синтеза) текстов на естественных языках.

NLP = NLG + ASR + NLU

- **ASR (Automatic Speech Recognition)** — технологии автоматического распознавания речи (на естественном языке)
- **NLG (Natural Language Generation)** - технологии генерации текстов на естественном языке
- **NLU (Natural-language understanding)** - технологии понимания текстов на естественном языке (текстов - связных, осмысленных, произвольных (по заданным тематикам))

Авторитетное издание

Лингвистический
энциклопедический
словарь (ЛЭС)

«Советская
энциклопедия»

1990 год

Ярцева Виктория
Николаевна



Издания ЛЭС

- 1990 (тираж 150 000 экз.).
- 2002 (тираж 3 000 экз.).
- <http://lingvisticheskiy-slovar.ru/about/>
- <http://www.tapemark.narod.ru/les/>

ПОРОЖДЕНИЕ РЕЧИ-1

1. в психолингвистике и психологии речи:
 - совокупность процессов перехода от речевого намерения (речевой интенции) к звучащему (или письменному) тексту, доступному для восприятия (идентификации и понимания).

Подходы и теории...

В бихевиоризме:

Беррес Фредерик Скиннер:

- ПР - результат эмпирической организации речевых реакций

Чарлз Эджертон Осгуд:

- ПР - «наполнение» речевыми реакциями врождённой системы

Подходы и теории...

В концепции **Н. Хомского**:

- ПР - процесс последовательной реализации семантических, грамматических и прагматических правил,
- формирующихся у носителя языка
- на базе врожденных когнитивных структур.

Подходы и теории...

В советской психолингвистике

(Лев Семенович Выготский,

Алексей Алексеевич Леонтьев):

- Порождение речи связывается с деятельностным характером речи.

Об одном недоразумении

Позиция советской психолингвистики чётко противостоит алгоритмическому пониманию ПР Хомским и его школой.

Как собрать большой урожай?

Порождение речи - 2

2. В экспериментальной фонетике:
- образование звуковых последовательностей, являющихся формой выражения значимых единиц языка.

Основные этапы ПР

- программа артикуляторных движений, соответствующая планируемому высказыванию;
- реализация этой программы, т. е. превращение её в необходимые комбинации артикуляторных движений;
- образование звукового сигнала (наиболее изученный этап ПР).

Трудная проблема сознания

- Трудность **hard problem of consciousness** не в том, что не получается найти решение (желательно простое и понятное), а в том, что принципиально не получается поставить проблему и хоть как-то сформулировать задачу!

ТПР = NLG

- Теория порождения речи
(проблемное поле, инструментарий ++),

МОЖЕТ БЫТЬ ПОНЯТА

(в контексте специализации)

как **Natural Language Processing (NLP)**

Забавно ;)

- Люди, ничего не понимающие в компьютерах (их программных, алгоритмических, математических основаниях) пытаются объяснить человека (свое мышление, сознание, речь) на основе компьютерных метафор
- **Ничего не понял сам – объясни другому**

ИТОЖИМ

и первое значение,
и второе значение
находятся
в проблемном поле

ПОРОЖДЕНИЕ РЕЧЕВОГО ВЫСКАЗЫВАНИЯ

Расширяя проблемное поле

порождение речевого высказывания

≠

порождению речи

Речь предполагает речевого субъекта

Порождение речи есть порождение оного

Квадратная матрица

1.1

1.2

1.3

2.1

2.2

2.3

3.1

3.2

3.3

Матрица

- математический объект, записываемый в виде прямоугольной таблицы элементов кольца или поля, представляющей собой совокупность строк и столбцов, на пересечении которых находятся её элементы.

Три уровня

1. Порождение речевого сообщества
(человечества)
2. Порождение речевой личности
(человека)
3. Порождение речевого высказывания
(устного текста)

Квадратная матрица

1.1

1.2

1.3

2.1

2.2

2.3

3.1

3.2

3.3

Устная и письменная речь

1. Порождение письменности
2. Порождение грамотности
3. Порождение письменного текста

Квадратная матрица

1.1

1.2

1.3

2.1

2.2

2.3

3.1

3.2

3.3

По ту сторону человека

- Порождение «языковых» машин
- Порождение диалоговых программ
- Автогенерация конкретного текста

Квадратная матрица

1.1

1.2

1.3

2.1

2.2

2.3

3.1

3.2

3.3

Ключевой термин

ОТЧУЖДЕНИЕ:

- отделение от людей процесса и результатов их деятельности, при котором процесс и результаты деятельности становятся неподвластными человеку и даже господствуют над ним, делая чуждыми друг другу человека и создаваемый им мир.

NLP vs NLG

Natural Language Generation -

программное преобразование

структурированных данных

в речь (ха-ха: в речевой сигнал)

- аналог превращения человеком идей
в письмо или речь.

В данном контексте

Термины **речь** и **язык**
рассматриваются
как синонимы

Речь и язык

Характеристика **речи** обычно даётся через противопоставление её **языку**, понимаемому как система объективно существующих, социально закреплённых знаков, соотносящих понятийное содержание и типовое звучание, а также как система правил их употребления и сочетаемости.

Речь и язык

- Речь (процесс) и язык (код) образуют единый феномен человеческого языка и каждого конкретного языка, взятого в определённом его состоянии.
- Речь есть воплощение, реализация языка (кода), который обнаруживает себя в речи и только через неё выполняет своё (коммуникативное) назначение.

Граничные дефиниции языка

ЯЗЫК — живой дух, в симбиозе с духовным человечеством и человеком.

ЯЗЫК — как физическое явление, вообще не существует

ЯЗЫК — это множество цепочек, составленных из символов некоторого конечного алфавита.

Конкретно о конкретном

Речь — конкретное говорение, протекающее во времени и облечённое в звуковую или письменную форму.

Арутюнова Нина Давидовна (1923, Москва) —
сов. и рос. лингвист, ч-к. АН СССР и РАН

Ключевой термин

ОТЧУЖДЕНИЕ:

- отделение от людей процесса и результатов их деятельности, при котором процесс и результаты деятельности становятся неподвластными человеку и даже господствуют над ним, делая чуждыми друг другу человека и создаваемый им мир.

Язык

Основной объект изучения языкознания, возникновение и существование которого неразрывно связано с возникновением и существованием человека — homo sapiens (см. Глоттогенез).

Александр Евгеньевич Кибрик (1939-2012) — российский лингвист, зав. кафедрой ТиПЛ

Отделения теоретической (структурной) и прикладной лингвистики филофака МГУ

Авторитетное мнение....

Бодуэн де Куртенэ Иван
Александрович (1845-
1929)

...язык как физическое
явление вообще не
существует



Авторитетное мнение....

Николай Яковлевич

Март (1865-1934):

... Сам по себе язык не существует, весь его состав есть отображение или, скажем конкретнее, отложение...



Плунгян Владимир Александрович



Плунгян Владимир Александрович

- доктор филологических наук
- член-корреспондент РАН,
- завсектором Института языкознания РАН,
- завсектором корпусной лингвистики и лингвистической поэтики Института русского языка РАН,
- профессор МГУ

В.А. Плунгян:

- Пройдя очень далеко ... многие из них в конце концов остановились и задумались: а существует ли язык вообще?
- Не является ли эта абстракция слишком сильной?

В.А. Плунгян:

- И не пора ли вернуться назад, к рассмотрению именно текстов, причем реальных текстов, которыми люди обмениваются в процессе коммуникации, а не искусственных предложений на страницах лингвистических статей?

Теория узуса

- изучает дискурсивную практику исходя из того, что не существует «языка вообще», а существуют структуры, которые преобладают в определенных типах дискурса (у разных говорящих, в разные моменты времени и т. п.); их (узусы) и надо изучать.

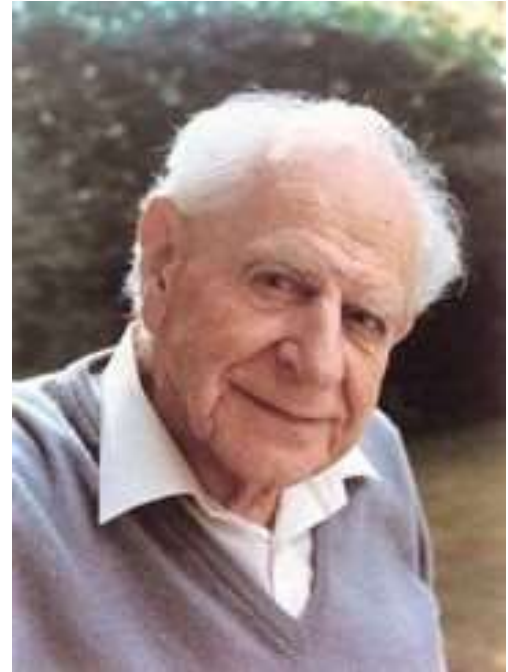
Диалог философов

- **Диоген Синопский:** А я вот, Платон, стол и чашу вижу, а стольности и чашности не вижу.
- **Платон:** Чтобы видеть стол и чашу, у тебя есть глаза, а чтобы видеть стольность и чашность, у тебя нет разума

Три мира

- мир физических объектов и состояний
- мир психических и ментальных состояний сознания
- мир объективного знания

Карл Раймунд Поппер (1902-1994)



Три (уровня) реальности

- Реальность речевого сигнала (мир физических объектов и состояний)
- Реальность речи (мир психических и ментальных состояний сознания)
- Сверхреальность речи (мир объективного знания творчества)

Владимир Иванович Мартынов

Время Алисы:

- Слово – отражает и удваивает реальность, причем человек не с состоянием увидеть самой реальности – он может видеть только ее отражение в зеркале слова.



Алесь Разанаў

ШЫЛЬДАЧКІ:

- На дрэвах — шыльдачкі.
- На іх напісана,
- колькі якому дрэву гадоў,
- як яно называецца,
- якая яго таўшчыня і вышыня.
-



Владимир Иванович Мартынов

Время Алисы:

- Что говорить о человеке, если согласно Писанию сам Бог нуждается в Слове, для того чтобы увидеть то, что Он Сам замыслил? «И сказал Бог: да будет свет. И стал свет. И увидел Бог свет, что он хорош».



Владимир Иванович Мартынов

Время Алисы:

- Получается, что Бог увидел свет только после того, как произнес слово «свет». И здесь можно обнаружить непроходимую пропасть, разделяющую Слово Бога и слово человека. Бог говорит и видит, что Он говорит, а человек видит и говорит, что он видит.

Владимир Иванович Мартынов

- Если, задумав погрузиться в реальность саму по себе, человек решится отбросить слова, преграждающие доступ к этой самой реальности, то тут же окажется, что нет ни реальности, ни самого человека, ибо отказ от слова автоматически делает несуществующим как то, во что следует погрузиться, так и того, кто собирается совершить погружение.

Affirmo ergo est

- Парадигма "affirmo" объективирует речь человека как возможность становления его бытия и его судьбы - его индивидуальности.
- Речь становится необходимым условием и пространством существования, жизни человека.
- Л. Витгенштейн: Если слово «реальность» не является самой реальностью, то о чем мы говорим, когда говорим «реальность».

Affirmo ergo est

- Не акт мысли (cogito), не акт действия (praxis), не акт существования (existenz), но акт утверждения (affirmo) становится основным - вбирает в себя как мысль, так и действие.
- Утверждение ведет к появлению того, что утверждается, и того, кто утверждает, ведет к становлению бытия, отличного от бытия натурального, ведет к реальности культуры.

К вопросу о реальности

Реальность = модус (мера, аспект, вид, уровень) Бытия.

- объективная, субъективная
- физическая, биологическая, социальная, ментальная...
- константная, виртуальная
- актуальная, потенциальная
- консенсуальная, обусловленная
- гиперреальность

К вопросу о реальности

- Virtual Reality (VR)
- Augmented Reality (AR)
- Augmented Virtuality (AV)
- Hybrid Reality (HR).
- Mixed Reality (MR)
- Immersive reality (IR)
- Extended Reality (XR)

К вопросу о реальности

- Virtual Reality (VR)
- Augmented Reality (AR)
- Augmented Virtuality (AV)
- Hybrid Reality (HR).
- Mixed Reality (MR)
- Immersive reality (IR)
- Extended Reality (XR)

О главном

Главное сегодня, это ответ на вопрос зачем в постиндустриальной экономике, где из 100 человек:

- 2 работают в сельском хозяйстве,
- 10 в промышленности,
- 15 в управлении,
- где и кому нужны лишние 75 человек

Лингвистика - наука об ошибках?



Максим Кронгауз

(проф., д.ф.н., дир. Инст
лингв. РГГУ)

Как писать частицу «не»
в нашей солнечной
стране? Павшим на
Первой
орфографической
посвящается

Максим Кронгауз

- В лингвистике сейчас самое развивающееся направление — это, конечно, компьютерная лингвистика, то есть разработка программ, решающих языковые и коммуникативные задачи вместо человека.
- Причем зачастую речь идет о задачах, которые человеку не под силу, к примеру, поиск в огромном массиве данных.

Три уровня

1. Порождение речевого сообщества
(человечества)
2. Порождение речевой личности
(человека)
3. Порождение речевого высказывания
(устного текста)

Устная и письменная речь

1. Порождение письменности
2. Порождение грамотности
3. Порождение письменного текста

Два трехуровневых ряда

- Общество
- Личность
- Высказывание
- Письменность
- Грамотность
- Письменный текст

По ту сторону человека

- Порождение «языковых» машин
- Порождение диалоговых программ
- Автогенерация конкретного текста

Компьютер – языковая машина!

ядро, стек, диспетчер ввода/вывода, обработчик событий, шрифтовая подсистема, растеризатор, оконная система, драйвера-переводчики, видеокарты и прочее, созданное на высокоуровневых языках программирования, обработанное компиляторами, оптимизаторами, компоновщиками, интерпретаторами и пр. и пр...

Нажмем на клавишу

- дебаунсер, мультиплексор ввода-вывода, стек устройств USB, логические вентили, селекторы, банк регистров, контроллеры шины, декодеры инструкций и микрокоды планировщика команд, мультиплексоры адресов и данных, схемы таймингов, контроллеры памяти, логика микрочипа необычайной языковой сложности...

Голова идет кругом от попытки вникнуть...

Упрощая снова и снова

- Все эти уровни сложности полностью скрыты и люди могут ими пользоваться даже не подозревая об их существовании
- Для подавляющего большинства пользователей нет ни малейшей возможности понять с чем они имеют дело.

Языковая сложность

Современные компьютеры настолько сложны в языковом отношении, что могут быть спроектированы и произведены только с использованием менее сложных компьютеров современной вычислительной среды (экосистемы).

Матрица

- математический объект, записываемый в виде прямоугольной таблицы элементов кольца или поля, представляющей собой совокупность строк и столбцов, на пересечении которых находятся её элементы.

Квадратная матрица

1.1

1.2

1.3

2.1

2.2

2.3

3.1

3.2

3.3

Проблеменое поле ТПР в ЛЭС

Только (психо-физиологическая) **часть** одного из значимых элементов матрицы.

Порождающие грамматики Хомского (логико-математическая часть) остаются за пределами рассмотрения...

Квадратная матрица

1.1

1.2

1.3

2.1

2.2

2.3

3.1

3.2

3.3

Возможности расширения

1. Порождение национального языка
2. Усвоение национального языка
3. Порождение ... высказывания

Возможности расширения

1. Порождение поэзии
2. Порождение поэтов
3. Порождение эпоса

Ключевой термин

ОТЧУЖДЕНИЕ:

- отделение от людей процесса и результатов их деятельности, при котором процесс и результаты деятельности становятся неподвластными человеку и даже господствуют над ним, делая чуждыми друг другу человека и создаваемый им мир.

Возможности расширения

1. Порождение прозы
2. Порождение прозаиков
3. Порождение романа

Ключевой термин

ОТЧУЖДЕНИЕ:

- отделение от людей процесса и результатов их деятельности, при котором процесс и результаты деятельности становятся неподвластными человеку и даже господствуют над ним, делая чуждыми друг другу человека и создаваемый им мир.

Возможности расширения

1. Порождение печати
2. Порождение печатного станка
3. Порождение тиража произведений

Ключевой термин

ОТЧУЖДЕНИЕ:

- отделение от людей процесса и результатов их деятельности, при котором процесс и результаты деятельности становятся неподвластными человеку и даже господствуют над ним, делая чуждыми друг другу человека и создаваемый им мир.

Томас Стернз Элиот

Где знания,
которые мы растеряли в
информации?
Где мудрость,
которую мы утратили в
знаниях?



Thomas Stearns Eliot

Where is the knowledge,
which is lost in information?

Where is the wisdom,
that is lost in knowledge?



В оный день....