

СЕКЦИЯ

НЕЙРОИНТЕЛЛЕКТ

Руководитель – д.ф.н., профессор **ШАПИРО Д.И.**

НЕКОТОРЫЕ ПЕРСПЕКТИВЫ РАЗРАБОТКИ НЕЙРОИНТЕЛЛЕКТА

Шапиро Д.И.

Институт проблем вычислительной техники и информатизации

Основным постулатом формирования "Общества знания" является развитие когнитивных и креативных способностей Человека и расширение его прагматических возможностей на основе компьютерного обобщения знаний о Мире и нем самом, и адекватных коммуникативных процедур. Подобные знания опираются на введенное В.И. Вернадским понятие "Ноосфера", которые по его представлению объединяет естественную среду, социальную среду и человека (с его особенностями и «делом рук его»). Развитие способностей Человека определяется, не в последнюю очередь, прогрессом культуры и науки. И хотя в настоящее время в этих сферах весьма широко распространены компьютерные технологии, важнейшим способом обмена научными знаниями традиционно являются личные контакты.

В Приложении к книге К. Popper, J. Ekkles "Self and its Brain"(1977) обсуждается проблема возможности сохранения неопубликованных идей и их влияния на прогресс науки. В развитие этой проблемы К. Popper сформулировал концепцию о "Трех Мирах": материальном (машины; приборы; книга, как материальный объект); человеческом (нейропсихолингвистические особенности); культурном (содержащим неопубликованные идеи, ненаписанные картины, не записанные нотными знаками симфонии).

В то время возможности и распространенность компьютерных технологий были весьма ограничены и этот аспект ими не обсуждался.

Одна из работоспособных гипотез может быть представлена следующим образом. Проблема обмена знаниями жидется (Шапиро 2004) на "трех китах": двух Участниках коммуникации и мультисенсорном дискурсе (применительно к технологии виртуальной реальности": " Пользователь (USER) - Среда обмена знаниями (SKC) - Интеллектуальный виртуальный коммуникант (IVC)"). К Среде обмена знаниями в соответствии с усложнением когнитивно-креативной стороны проблемы предъявляются все более глубокие требования (например, в проблемах образного или мультисенсорного обмена).

При этом применительно к проблеме обмена научными знаниями (да еще высказываемыми в неявной форме)существенную роль играют семантические (содержательные) особенности знаний.

Субъект в процессе реализации своих когнитивных возможностей расширяет индивидуальный имплицитный эталон (ИМ). И подобные имплицитные ценности эксплицируются по мере необходимости в соответствующее время (например, при необходимости реального порождения нового или вхождения в коммуникативную процедуру). Некий аналог подобного подхода можно наблюдать в определенный период развития маленьких детей, когда они, уже умея распознавать предметы, их взаимосвязи и назначение, еще не овладели речью. (См. также взрослый, попавший в страну со совершенно незнакомым языком, например, одним из дагестанских).В подобных классах проблем великого влияния "Уникального" субъекта Fo на субъект Fi с повышенной креативной активностью. Коммуникативная структура имеет вид:

$$(S \rightarrow T) \circ \implies \{ [P] \rightarrow [U] \rightarrow S' \rightarrow [F] \implies T \}$$

Проблема "Возбуждение творческой активности на основе "чужой мысли" может быть характеризована следующими факторами:

- спецификой входного метатекста To,
- умением «выуживать» из входной реализации метатекста To проблемно или личностно ориентированный смысл S,
- степенью корреляции смыслов S и S' (или T и T*), которая определяется как малыми значениями (например, воздействие на автора любых выступлений А.Н. Колмогорова, А. А. Ляпунова, Я.З. Цыпкина, В.В.Иванова независимо от тематики), так и высокими значениями (восстановление "ноги по пятке,...", Дон Жуан - А.С. Пушкина),

- внутренним настроем конкретного F_i на определенного F_o , который формируется на основе их априорных встреч, фиксации и запоминания аудио (интонации) и видео (лицо, манеры) образов, а также внутренне близким музыкальным сопровождением выступлениями.

Представления сэра Карла Поппера о "Трех Мирах", и, в частности, о "Мире Человеческом" и "Мире Культуры", как развитие его размышлений совместно с сэром Джоном Экклзом о способах передачи концептуальных научных идей без прямой вербальной или образной коммуникации являются фундаментом для исследования и разработки многих проблем нейроинтеллекта и нейрокомпьютинга.

В частности, приведенный в его примере "симфония, не записанная в виде нотных знаков", может быть зафиксирована (проиграна) на имплицитном уровне в Моделях "Orchestra" (McIntosh, 2000) и "Chorus" (Shapiro, 1994).

Эта симфония, с одной стороны, может быть реакцией на такие "факторы влияния", как:

- динамический входной паттерн ("партитура");
- иницирующий входной сигнал;

- самопроизвольное порождение, опирающееся на априорный опыт; а, с другой стороны, может сама явиться иницирующим входным сигналом или фрагментом опыта для определенного окружения.

В компьютерной реализации подобного подхода могут быть использованы расплывчатые фреймы, продукционные правила, нейрорасплывчатые алгоритмы и гетерогенные нейронные сети (HNN), функционирование которых основано на известных Структурно-функциональных принципах М.Л. Цетлина (1969). HNN способны при обработке образных знаний формировать определенную структуру для конкретного предъявления входного паттерна $\{IP\}$, определяя его смысл Sc с помощью проблемно-ориентированной шкалы Cs . Таким образом, предъявляя совокупность проблемно-ориентированных (лица, интонации) входных паттернов $\{IP\}_1, \dots, \{IP\}_p, \dots, \{IP\}_n$, можно сформировать соответствующую совокупность смыслов $\{S1, \dots, Sp, \dots, Sn\}$. Оттеночные оценки осуществляются человеком на основе его гено и фенотипических особенностей, сформированных в правополушарном ИМ. Совокупность смыслов $\{S1, \dots, Sp, \dots, Sn\}$ "имплицитно" хранится в компьютерном аналоге ИМ до востребованности, формируя "имплицитные" знания.

При появлении нового $\{IP\}_k$ в проблемно-ориентированном классе задач рассматриваются три варианта взаимосвязей $\{IP\}_p \iff \{IP\}_k$:

- | | | |
|--------------------|------|----------------------|
| - совпадение, | если | $C(Sp)=C(Sk)$, |
| - близость, | | $C(Sp)=C(Sk)+/-1$, |
| - рассогласование, | | $C(Sp)=C(Sk)>+/-1$, |

где $\iff$ - оператор взаимосвязи.

В последнем случае с помощью нечетких фреймов выявляются источники рассогласования (несовпадающие слоты нечетких фреймов).

В "человеческом" аспекте наиболее общим свойством, характеризующим отношения "Образ-Смысл", по-видимому, является Воображение (психическая деятельность, состоящая в создании представлений и мысленных ситуаций, никогда в целом не воспринимавшихся человеком в действительности и включающая воссоздающую и творческую составляющие). Проблемы порождения и представления знаний напрямую зависят от отношения «Образ - Смысл». Обсуждая проблему «невербальности мысли», Р. Пенроуз ссылается на Альберта Эйнштейна («Психические сущности, которые, по-видимому, и являются составляющими элементами мысли – это определённые знаки и более или менее отчётливые образы, которые могут «произвольно» воспроизводиться и комбинироваться по собственному желанию» [из письма Ж. Адамару]). Сам он, поддерживая эту точку зрения («Строит мысленные диаграммы»), отмечает возможность проявления в этой проблеме специфики ЛП и ПП функций. Подобная специфика, как известно, личностно ориентирована.

Проблемы описания и формального представления оттеночных характеристик образа и выявления его смысловых и эстетических особенностей уже давно активно совместно исследуются специалистами различных профессий ("Нейроэстетика", Цоллингер и др. 1985). Полученные результаты позволили подойти к компьютерным представлениям подобных процессов - "Компьютерная нейроэстетика" (Shapiro, 2004).

Компьютерная нейроэстетика обладает еще одним очень важным свойством. Р. Пенроуз (2003), анализируя процессы озарения и вдохновения в научной работе, подводит нас "к более, чем существенной роли, которую играют при формировании суждений эстетические критерии." И далее "красивая идея имеет гораздо больше шансов быть правильной, чем нескладная.

Таким образом, представляется возможным повысить креативный уровень участника коммуникации путём выявления и формирования оттеночных характеристик образа с учетом их эстетических и смысловых особенностей, расширяя сферу приложения подобных исследований.

Представляется целесообразным исследовать компьютерный аспект отношения "Образ-Смысл" с позиций его общего свойства

- "Нейровоображения" (способность нейроструктуры представить новый образ), включающее, в частности:

- введение нового элемента и соответствующих взаимосвязей, формирующих новый смысл или меняющих его;
- "достройку" образа на основе проблемно-ориентированных принципов (опыт, "шаблоны" ПМ);
- порождение или восстановление образов по известной оценке одного из элементов в данном классе объектов и др.

Подобные "имплицитные знания" используются ("эксплицируются") в ряде конкретных приложений, в частности:

- "Формирование образа" на основе эстетических и морально-этических оценок, "позитивного" образа $[..., \{IP\}k, ...,]_{pos} \implies [\{IP\}]_{pos}$ и "негативного" образа $[..., \{IP\}k, ...,]_{neg} \implies [\{IP\}]_{neg}$.

Эти образы $[\{IP\}]_{pos}$ и $[\{IP\}]_{neg}$ представляются как по видео (лицо), так и по аудио (интонации, музыка) каналам, на которые настроены соответствующие нейронные структуры W_{pos} и W_{neg} . В результате их функционирования формируются смыслы (S_{pos} и S_{neg}).

Появление нового $\{IP\}$ воспринимается и понимается компьютерным субъектом (соответствующей структурой Интеллектуального виртуального коммуниканта - IVC). Несоответствие входного паттерна сформированному эталону (ПМ) вызывает перестройку структуры W , формирование нового оттенка смысла и внешнюю реакцию.

Подобные процедуры, очевидно, основаны на устойчивой активации. Н.А. Бернштейн ещё в шестидесятые годы, рассматривая жизнедеятельность живых организмов, обратил внимание на существенные и несущественные признаки (по И.М. Гельфанду и М.Л.Цетлину), которые позволяют выделять основополагающие (например генетически обусловленные свойства на примере листьев дуба) и несущественные (обусловленные различным влиянием среды – освещение, питание и т.д.). «Организм реактивен по отношению к своим несущественным переменным, но в высшей степени нереактивен – активен по отношению к существенным» (стр.449, 1990г.) В данном классе проблем эти структуры связаны с существенными переменными.

Применительно к Человеко-машинному аналогу задачи влияния одного субъекта F_0 на другой F_i , коммуникативная структура имеет вид: .

$$(S \rightarrow T)_0 \implies \{ [P] \rightarrow [U] \rightarrow S' \rightarrow [F] \} \implies T',$$

$$\text{или } (S \rightarrow T)_0 \implies [W] \implies S' \rightarrow [F] \implies T',$$

где $[W]$ состоит из структуры восприятия $[W_p]$ и структуры понимания $[W_u]$.

Состояния нейроструктуры $[W_u]$ определяется по аналогии с «человеческим» следующими факторами:

а) спецификой внешнего воздействия T_0 , определяемого тремя типами:

- спокойное (отсутствие воздействия, Рис.)
- возбужденное (стандартное воздействия, Рис.)
- сильное возбуждение (нестандартное воздействия, Рис.)
- в) степенью корреляции S и S' (или T и T'),

с) внутренним настроем конкретного F_i на определённого F_0 , который формируется на основе их обучения и соответствующего возбуждения $[W_u]$.

Подобное возбуждение $[W_u]$ создаёт условия позволяющие породить новый смысл S' и новую реализацию метатекста T' .

«Семантическая близость» может быть достигнута:

- по одному объекту или признаку («красивая женская нога»);
- по всему образу («Софи Лорен»);
- по мультисенсорному паттерну (Рождественская Месса в Кафедральном Соборе, сопровождаемая музыкой И.С. Баха).

Выше был приведен один из подходов к решению проблемы "эксплицирования имплицитных знаний" в коммуникационной системе "Человек-компьютер". Приведенный подход, основанный на использовании особенностей "Нейровоображения", позволяет, по-видимому, подойти к изложенной выше концепции Карла Поппера о "Трёх Мирах" с учётом следующих факторов:

- условия порождения имплицитной (не высказанной) идеи,
- специфика контекста,
- совокупность объектов, включенных в эту систему,
- специфика процесса порождения,
- специфика сокоммуниканта.

Совокупность определяется посредством введенной А.А. Богдановым (2003) структуры взаимоотношений и оценки степени их подготовленности сокоммуниканта. Как отмечено (Л. Фейербах), "Искусство остроумно писать подозревает, между прочим, юмор в читателе".

Такой подход к проблеме использования имплицитных знаний может быть применён при реализации следующих функций:

- воображение,
- порождение нового (и достройка) образа,
- взаимопонимание,
- взаимовлияние,
- интуиция,
- ассоциативность.

Важная гносеологическая проблема, поставленная Сэром Карлом Поппером и Сэром Джоном Экклзом, многоаспектна и междисциплинарна. Ее решение имеет и серьезное прагматическое значение. Одним из путей решения этой проблемы является, по-видимому, расширение и углубление интеллектуальных возможностей компьютерного партнера-Интеллектуального виртуального коммуниканта и использование хорошо продуманных совместных компьютерных и нейропсихолингвистических экспериментов по обработке образных знаний.

Литература

1. Бернштейн Н.А. Формирование движения и активность. М., 1990, 450 с.
2. Богданов А.А. Тектология М., Экономика, 2003, 300 с.
3. Пенроуз Г. Новый ум короля, М., УРСС, 2004, 382 с.
4. Цетлин М.Л. Теория автоматов и моделирование биологических систем, М. Наука, 1969, 285 с.
5. Красота и мозг (ред. Цоллингер Г.) М. Мир, 1995, 295 с.
6. McIntosh A., Towards a Networks. The Theory of Cognition, N.N., v.13, N8, 2000, p. 1600-1621.
7. Popper K., Ekkles J., The Self and its Brain, Berlin, Schpringer International, 1972, 479 p.
8. Shapiro D.I., A Neurointelligence: From Virtual Cognition Systems to System of Virtual Reality, Pros. of АПСР'94, Bratislava, 1994, p.361-366.
9. Shapiro D.I., Multi-Sensorical Discourse and Problems of Estimation of Beauty, Proc. of EVA'04, London, 2004, p. 12.5 - 12.10.

Abstract: In this work the prospective problems of developping procedures of knowledge interchange in the Virtual Reality Technology on the basis of A.A. Bogdanov's and K. Popper's ideas, as well as M.L. Tsetlin's model concepts (Structural-Functional Specifics of Neural Systems) are considered.

The impact of the communicant features, their interrelations and conditions on the formation of such procedures is discussed.

The approaches to forming neuro-intellectual procedures of image generation and development, as well as image interchange are presented.

СЛУХОВОЕ ОБНАРУЖЕНИЕ СИГНАЛОВ И РАЗБОРЧИВОСТЬ РЕЧИ

Санников В.Г.

Московский технический университет связи и информатики

Передача речи по телефонным и радиотелефонным каналам связи сопровождается наличием различного вида шумов и помех, приводящих к снижению фонетической и смысловой информации, воспринимаемой слуховой системой человека. Одной из важных характеристик при слуховом восприятии речи в шумах является разборчивость фонетических речевых элементов (РЭ): звуков, слогов, слов, фраз. Известны результаты измерения разборчивости речи в шумах, полученные на основе проведения психофизических экспериментов с привлечением методов субъективно-статистической экспертизы [1]. Поскольку эти методы требуют длительной организационной работы и привлечения большого количества аудиторов, они не могут быть использованы в условиях производства и эксплуатации систем речевой передачи. В связи с этим актуальной становится проблема оперативной, адекватной слуховому восприятию, оценки качества передачи речи по разборчивости.

При решении указанной задачи важно учитывать преобразования речи на периферии слуховой системы как приемника речевой информации. Согласно современным представлениям слуховая система человека – это сложная биологическая нейросенсорная система, анализирующая поступающие на её вход звуковые стимулы и осуществляющая распознавание передаваемых РЭ с оценкой качества их передачи.

Основные закономерности преобразования речи на периферии слуха рассматривались автором в работах [2, 3]. Показано, что ощущения на уровне рецепторных потенциалов внутренних волосковых клеток (откли-

ков слуховых детекторов) пропорциональны среднему квадрату откликов (СКО) полосовых фильтров слуховых детекторов. Поскольку эти потенциалы – случайные процессы, то справедлива гипотеза о том, что аудитор для принятия решения о передаваемых РЭ использует некоторое правило, основанное на анализе статистики СКО.

В результате открытия наличия внутренних шумов в нейронных каналах слухового нерва и порога принятия решения, эффективным методом анализа работы аудитора оказалась теория статистических решений [4]. Эта теория может служить основой для понимания психофизических законов и объяснения экспериментальных результатов по оценке разборчивости РЭ.

При исследовании психофизических законов слухового восприятия звуковых стимулов в соответствии с теорией статистических решений принята гипотеза о том, что наблюдатель использует для принятия решения функционал отношения правдоподобия. В этом случае для описания процесса решения универсальное значение имеет рабочая характеристика, представляющая собой зависимость вероятности правильного решения $P_{ss} = P(s | s)$ - вероятность правильного распознавания РЭ при условии, что он в действительности передавался, от вероятности ложного обнаружения $P_{sn} = P(s | n)$ - вероятность решения о передаче РЭ при условии, что он в действительности не передавался. Здесь приняты две альтернативы s и n , где s - величина, характеризующая присутствие сигнала в шуме, и n - величина, характеризующая отсутствие полезного сигнала, т.е. на входе слуховой системы имеется один шум.

Методика получения рабочих характеристик слуховой системы представлена в работе [5]. В частности, там получены следующие соотношения для рабочих характеристик

$$P_{ss} = \left(1 + \frac{0.5 \cdot x}{1 + h^2}\right) \cdot \exp\left(-\frac{0.5 \cdot x}{1 + h^2}\right), \quad (1)$$

$$P_{sn} = (1 + 0.5 \cdot x) \cdot \exp(-0.5 \cdot x), \quad (2)$$

где x - нормированный порог принятия решения, изменяющийся в пределах от нуля до бесконечности, $h^2 = D_s/D_n$ - отношение сигнал/шум (ОСШ), D_s и D_n - дисперсии, соответственно, полезного сигнала, подлежащего распознаванию, и шума наблюдения.

На рис.1 представлены рабочие характеристики, полученные на основе соотношений (1) и (2). Здесь же приведена экспериментальная рабочая характеристика (***) , полученная в эксперименте по обнаружению звукового сигнала во внешнем белом шуме [4]. Эксперимент проводился по абсолютной схеме с одним интервалом стимулирования. Как видно при $h^2 = 1.7$ теоретическая характеристика достаточно хорошо аппроксимирует экспериментальную зависимость.

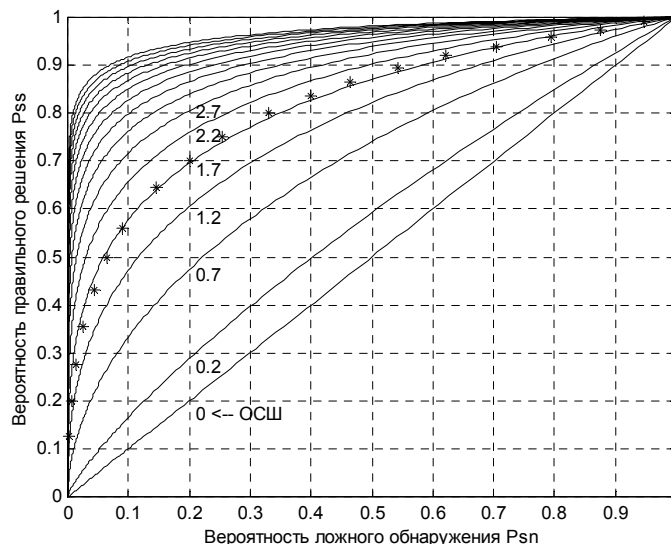


Рис. 1. Рабочие характеристики слуховой системы

Из данного исследования можно заключить, что развиваемый теоретический анализ преобразования звукового стимула на периферии слуха с достаточной точностью описывает статистические явления при проведении психофизических экспериментов. Распространим данный метод на нахождение теоретических кривых для экспериментального семейства слоговой разборчивости русской речи, приведенного в работе

[1]. Здесь на стр. 223 (рис. 9.14) представлено экспериментальное семейство зависимостей слоговой разборчивости от уровня речи при разных уровнях шума гладкого спектра.

В соотношениях (1) и (2) с уровнями речи L_s и шума L_n связан параметр ОСШ

$$h^2 = \frac{D_s}{D_n} = 10^{0.1L}, \quad (3)$$

где, после учета нелинейных свойств амплитудных характеристик слуха, величину L представим в следующем виде

$$L = k_1 \cdot [L_s - 10 + k_2 \cdot (L_n - 40)] / \sqrt{1 + \left[\frac{L_s - 10 + k_2 \cdot (L_n - 40)}{97.5 - k_3 \cdot (L_n - 40)} \right]^6}. \quad (4)$$

Подставляя в (4) параметры: $k_1 = 0.22$, $k_2 = 1.06$, $k_3 = 0.4$, фиксируя уровень шума L_n и изменяя уровень речи L_s , находим ОСШ по (3) и, подставляя его в соотношение (2), находим вероятность правильного обнаружения. Семейство теоретических характеристик, получаемых таким образом при различных уровнях шума, представлено на рис. 2. Здесь же приведены экспериментальные зависимости (***)

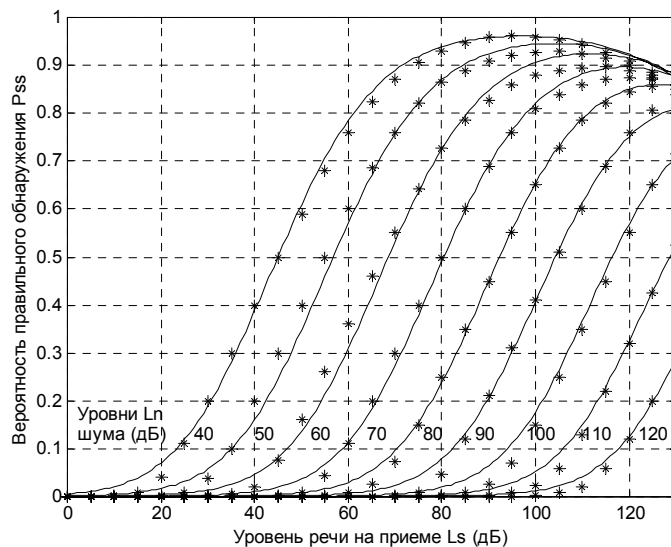


Рис. 2. Кривые слоговой разборчивости русской речи

Анализ приведенных теоретических и экспериментальных зависимостей показывает, что они достаточно близки. Следовательно, построение рабочих характеристик позволяет оперативно оценивать качество передачи речи по показателю её разборчивости без привлечения дорогостоящих психоакустических измерений.

Литература

1. Покровский Н.Б. Расчет и измерение разборчивости речи. - М.: Связьиздат, 1962. - 391 с.
2. Санников В.Г. Теоретический анализ заметности искажений речевых сигналов по громкости их слухового восприятия. "ЭЛЕКТРОСВЯЗЬ", № 12, 2002, с. 38-42.
3. Санников В.Г. Слуховой прием сигналов и оценка громкости их восприятия. Труды LVII научной сессии, посвященной дню радио. Том 1. - М.: РНТОРЭС им. А.С. Попова, 2002, с. 196-198.
4. Леонов Ю.П. Теория статистических решений и психофизика. -М.: Наука, 1977. - 223 с.
5. Санников В.Г. Рабочие характеристики слухового обнаружения звуковых сигналов. Материалы V международной НТК: "Перспективные технологии в средствах передачи информации". -Владимир.: Связьоценка, 2003, с. 129-132.

ACOUSTICAL DETECTION OF SIGNALS AND LEGIBILITY OF SPEECH

Abstract. In conditions of manufacture and operation of systems of speech communication the problem operative, adequate to acoustical perception, an estimation of quality of transfer of speech on a parameter of legibility is actual. In job on the basis of the theory of statistical decisions the design procedure of performance data of acoustical system develops. The comparative analysis of theoretical characteristics with experimental data of syllabic legibility of speech is given.

ИДЕНТИФИКАТОР СТРУКТУРЫ ДЛЯ АДАПТИВНОЙ СИСТЕМЫ КОДИРОВАНИЯ РЕЧЕВЫХ СИГНАЛОВ

Басов О.О, Рыжков А.П.

Академия Федеральной Службы Охраны Российской Федерации

Одной из важнейших проблем развития информационно-телекоммуникационной системы является создание высокоэффективных устройств аналого-цифрового и цифро-аналогового преобразования речевых сигналов.

При разработке систем кодирования речевых сигналов основные трудности в их функционировании связаны с необходимостью преодоления априорной неопределенности относительно свойств источника и нестационарности речевого сигнала. Для преодоления этих трудностей используется два подхода.

Первый связан с разработкой робастных алгоритмов обработки, которые обеспечивают нечувствительность или слабую чувствительность результатов обработки от изменения характеристик источника в значительных пределах.

Второй подход связан с разработкой адаптивных алгоритмов, основной особенностью которых является то, что на каждом этапе обработки параметры и структура таких систем изменяются так, чтобы качество сжатия сигналов было наилучшим. В существующих кодеках речевых сигналов реализована только параметрическая адаптация. Для улучшения адаптивных свойств кодеков речевых сигналов предлагается использовать систему с переменной (переключаемой) структурой [1]. На рисунке 1 показана структурная схема такой системы.

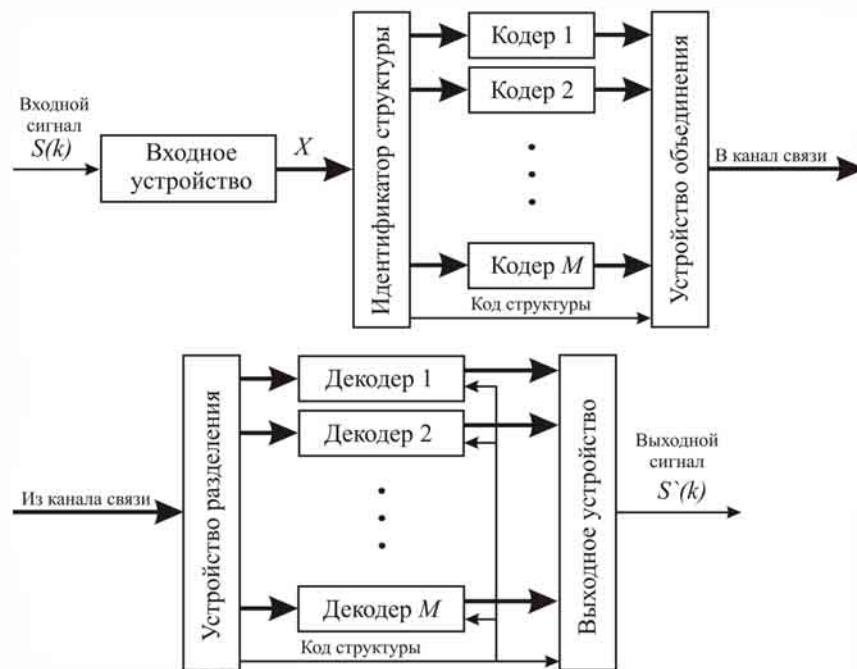


Рис.1 Структурная схема системы с переменной структурой.

На входное устройство поступают отсчеты речевого сигнала $S(k)$ взятые с частотой дискретизации 8 кГц. Здесь происходит формирование речевого вектора X , состоящего, как правило, из $N=40-160$ отсчетов, что соответствует длине кадра анализа равной 5-20 миллисекундам соответственно. В состав входного устройства может включаться фильтр, с помощью которого можно обеспечить лучшее функционирование идентификатора или/и лучшее субъективное восприятие синтезированной речи.

Идентификатор структуры обеспечивает анализ пришедшего вектора речевого сигнала. Результатом такого анализа является сигнал, указывающий номер кодека, который обеспечит наилучшее качество сжатия обрабатываемого кадра речевого сигнала.

Анализ кодеков речи, предназначенных для скоростей передачи $V=1,2-9,6$ кбит/с, показывает, что число типов их структур не должно превышать 128. При этом в качестве кодеков предлагается использовать линейные предсказатели с различными длинами фильтров и наборы векторных квантователей, отличающиеся размерами кодовых книг и характеристиками сигналов, используемых для их формирования [5].

Существенно уменьшить число необходимых типов структур для эффективного кодирования речевых сигналов на низких скоростях передачи, а соответственно, и число реализуемых кодеков в приведенной

адаптивной системе с переключаемой структурой, позволяет применение кодеков речи на основе нелинейного предсказания. Такие кодеки функционируют, основываясь на математическом аппарате нейронных сетей [4].

Устройство объединения обеспечивает формирование кадра передачи, состоящего из кода структуры, параметров кодируемого сигнала, сигнала синхронизации. В общем случае в канал связи будет подаваться сигнал с переменной скоростью, но на выходе устройства объединения можно обеспечить и постоянную скорость передачи. В приемной части кодера реализуются обратные процедуры, в соответствии с пришедшим из канала связи кодом структуры M_i , восстановление сигнала осуществляется декодером M_i .

Важнейшим этапом функционирования представленной системы кодирования с переменной структурой является анализ блока (вектора) речевого сигнала. В общем случае идентификатор структуры реализует разделение всего входного сигнального пространства на M субпространств. Возможные устройства анализа кадра могут применять:

- анализ статистических и параметрических характеристик речи;
- анализ функции распределения сигналов;
- процедуру векторного квантования;
- цепь Маркова [2].

Однако лучший результат для системы кодирования с кодерами на основе нелинейного предсказания был показан идентификатором структуры на основе нейронной сети Кохонена (рис.2).

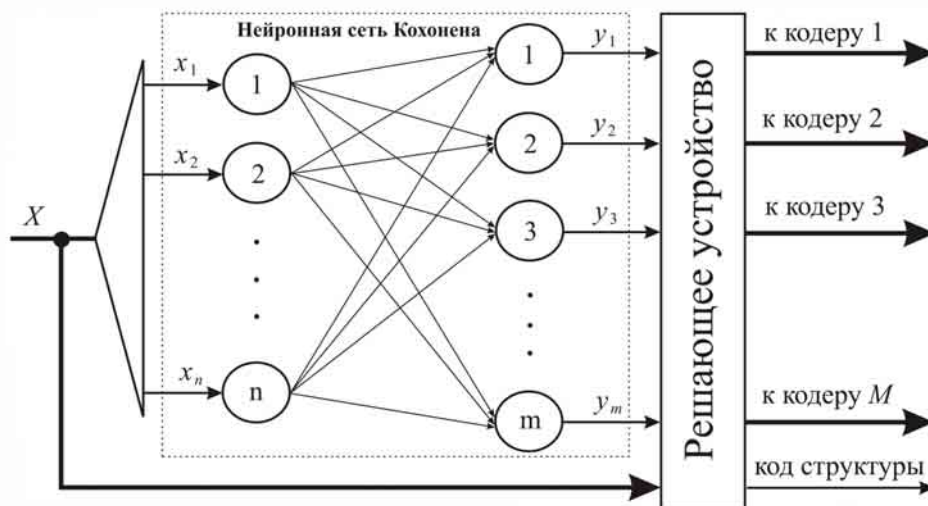


Рис.2 Структурная схема идентификатора структуры.

Данные сети позволяют в результате обучения осуществлять топологически непрерывное отображение F входного n -мерного пространства в выходное m -мерное пространство. При этом обучение здесь происходит без учителя на основе образов, поступающих на сеть. В качестве метода обучения используется конкурентное обучение.

Нейронные сети, использующие такой метод обучения, называются *конкурентными*. Конкурентная нейронная сеть в общем случае представляет собой двухслойную нейронную сеть с прямыми связями. Первый слой выполняет чисто распределительные функции, причем каждый нейрон его имеет соединения со всеми нейронными элементами выходного слоя. Второй слой осуществляет конкуренцию между нейронными элементами, в результате которой определяется нейрон-победитель.

Для нейрона-победителя синаптические связи усиливаются, а для остальных нейронов не изменяются или могут уменьшаться. В результате этого процесса осуществляется конкурентное обучение. Победителем в конкуренции является нейрон, который в результате подачи на вход сети определенного образа имеет максимальную взвешенную активность

$$S_j = \sum_i w_{ij} x_i = W_j X^T, \quad (1)$$

где $X = \{x_1, x_2, \dots, x_n\}$ - входной вектор (образ) речевого сигнала, $W_j = \{w_{1j}, w_{2j}, \dots, w_{nj}\}$ - вектор-столбец весовых коэффициентов j -го выходного нейрона.

Пусть $S_k = \max_j S_j$.

Тогда активность выходных нейронов

$$y_j = F(S_j) = \begin{cases} 1, & \text{если } j = k; \\ 0, & \text{если } j \neq k, \end{cases} \quad (2)$$

где $j = \overline{1, m}$.

Таким образом, после обучения нейронной сети при подаче входного речевого вектора активность нейрона-победителя принимается равной единице, а остальных нейронов – нулю. Это правило известно под названием «победитель берет все».

На практике чаще всего оказывается, что активности выходных нейронов принимают значения меньше 0 или больше 1. Поэтому, при определении нейрона-победителя необходимо нормализовать весовые и входные векторы для каждого нейрона. Нормализация осуществляется следующим образом:

$$|X| = \sqrt{\sum_i x_i^2} = 1, \quad |W_j| = \sqrt{\sum_i w_{ij}^2} = 1. \quad (3)$$

В результате можно записать следующее правило обучения для весового вектора нейрона-победителя:

$$W_k(t+1) = \frac{W_k(t) + \gamma(X(t) - W_k(t))}{|W_k(t) + \gamma(X(t) - W_k(t))|}, \quad (4)$$

где $0 < \gamma < 1$ характеризует скорость обучения.

В качестве нейрона-победителя выбирается такой нейрон, весовой вектор которого наиболее близок к входному вектору. В скалярной форме весовые коэффициенты k -го нейронного элемента

$$w_{ik}(t+1) = \frac{w_{ik}(t) + \gamma(x_i(t) - w_{ik}(t))}{|W_k(t) + \gamma(X(t) - W_k(t))|}, \quad (5)$$

где $i = \overline{1, n}$.

При применении данного правила к k -му нейрону усиливается его выходная активность. Таким образом, при конкурентном обучении все множество входных образов речевого сигнала разбивается на кластеры, каждому из которых соответствует свой нейрон. При поступлении на вход нейронной сети неизвестного образа она будет его относить к такому кластеру, на который он больше всего похож. В этом заключается обобщающая способность такого типа нейронных сетей [3].

Обучение нейронной сети Кохонена для идентификатора структуры производится на реальном речевом сигнале. Число нейронов входного слоя устанавливается равным величине входного вектора речевого сигнала, выходного – необходимому числу типов структур кодеров. При этом для каждой области решений, определяемой соответствующим выходным нейроном, определяется кодек, обеспечивающий наилучшее качество сжатия речевого сигнала при заданной скорости кодирования.

Алгоритм функционирования идентификатора структуры для адаптивной системы кодирования речевых сигналов состоит из следующих шагов:

1. На нейронную сеть подается входной вектор речевого сигнала X .
2. Определяется нейрон-победитель, обладающий наибольшей активностью.
3. В соответствии с номером нейрона-победителя, решающее устройство подает входной образ X на один из кодеров системы кодирования с переменной структурой.

Создание алгоритмов сжатия речевых сигналов на основе систем с переменной структурой позволяет повысить качество передачи речевых сигналов наряду с уменьшением ее скорости, а применение идентификаторов структуры на основе нейронных сетей Кохонена в таких системах, в свою очередь, позволяет наиболее точно изменять их структуру, выбирая наилучшую из заданного конечного множества вариантов.

При использовании предложенного варианта идентификатора структуры в адаптивных системах кодирования речи с переменной структурой в идеальных каналах обеспечивается выигрыш в качестве восстановления речевого сигнала на приемной стороне на 3-5 дБ. Однако в реальных каналах необходимо учесть тот факт, что ошибка при передаче служебной информации о структуре кодека, приведет к потере всего речевого кадра. Поэтому при разработке систем с переменной структурой рекомендуется использование помехоустойчивого кодирования такой информации [1].

Таким образом, применение идентификатора структуры на основе нейронной сети в системах с переменной структурой для построения кодеков речи является более эффективным по сравнению с традиционными процедурами анализа кадра речевого сигнала.

Литература

1. Афанасьев А.А., Басов О.О., Богачёв Г.В. Структурно – адаптивная система для кодирования речевых сигналов.//Академия ФАПСи. – Орёл, 2003. – 20 с.: ил.
2. Б.Уидроу, С.Стирнз. Адаптивная обработка сигналов. – М.: Радио и связь, 1989 – 450 с.

3. Головкин В.А. Нейронные сети: обучение, организация, применение. Кн.4.: Учебное пособие для вузов / В.А.Головкин; Под ред. А.И.Галушкина. – М.: ИПРЖР, 2001 – 256 с.: ил.- (Нейрокомпьютеры и их применение).

4. Нейроматематика: Учеб. пособие для вузов. Кн. 6/ Д.А.Агеев, А.Н. Болухто, С.А. Верещагин и др.: Под общ. ред. А.И.Галушкина. – М.: ИПРЖР, 2002. – 448 с.: ил.

5. Шелухин О.И., Лукьянцев Н.Ф. Цифровая обработка и передача речи / Под ред. Шелухина О.И. – М.: Радио и связь, 2000. – 456 с.: ил.

In the paper the opportunity of application of neural networks in adaptive systems low-speed speech coding is considered. Algorithms of training and functioning of the identifier of structure are resulted on the basis of Kohonen's neural network. Advantages of such identifier of structure are shown in comparison with traditional procedures of the speech analysis.

СИНТЕЗ СИСТЕМЫ УПРАВЛЕНИЯ КОММУНИКАЦИОННЫМ РЕСУРСОМ СПУТНИКОВОЙ СИСТЕМЫ СВЯЗИ С ПРЕДОСТАВЛЕНИЕМ КАНАЛОВ ПО ТРЕБОВАНИЮ НА ОСНОВЕ НЕЙРОСЕТЕЙ

Чистопрудов Д.А.

Серпуховской военной академии ВВС

Современное развитие телекоммуникаций обусловлено ростом потребностей в обеспечении связи как со стороны обычных пользователей, так и со стороны специфических потребителей. К последним относятся различные ведомства и корпорации, развертывание системы связи которыми решает специфические (обусловленные конкретной корпорацией) задачи. Как правило, данные системы связи строятся как самостоятельные, обособленные системы, охватывающие различные виды и типы связи. Подсистемами таких систем могут выступать: проводная подсистема, спутниковая подсистема, подсистема радиосвязи, причем подсистема спутниковой связи (ПСС) представляет собой наиболее дорогостоящий сегмент.

Система связи, построенная на базе спутниковой системы связи (ССС), является иерархической в рамках решаемых на узлах связи (УС) корпорации задач. Это обуславливает разнородность состава конечных средств связи на каждом узле связи – потребителей коммуникационного ресурса СССР. Кроме того, иерархичность УС обуславливает различие значений начального трафика, приходящегося на абонентскую аппаратуру спутниковой связи УС. Все это позволяет говорить о крайней неоднородности абонентов описываемой СССР.

При решении задачи обеспечения связью абонентов с низким исходящим трафиком применяется организация СССР с предоставлением каналов по требованию (ПКТ – DAMA). Известные способы описания процессов обслуживания запросов абонентов СССР с ПКТ базируются на моделировании протоколов случайного множественного доступа (СМД). Полученные при этом результаты говорят о нестабильности данных протоколов, в частности для систем с конечным числом абонентов характерно такое явление, как бистабильность и блокировка [1, 2]. Это в свою очередь актуализирует проблему управления параметрами протокола СМД для повышения его устойчивости[3]. В качестве таковых выделяют две группы параметров: параметры повторной передачи пакета и параметры структуры запросного канала.

В ряде работ, посвященных исследованию эффективности организации СССР с ПКТ [4], для оценивания реализованного качества обслуживания абонентов рассматривается критерий требуемого качества, выраженный как вероятность доведения сообщения за время, не превышающее заданного. В свою очередь, данная вероятность является вероятностью сложного события, состоящего в том, что сумма времени предоставления канала и времени передачи сообщения по каналу связи не превысит установленного предела:

$$P_{\text{дов}}(t) = P(t_{\text{предостав. канала}} + t_{\text{передачи по каналу}} \leq t_{\text{треб}})$$

Обе процедуры - и предоставления канала по требованию и передачи сообщения, являются нетривиальными. Дело в том, что при наблюдении режима бистабильности и блокировки, в запросном канале конфликтуют вызывающие станции, увеличивающие трафик повторного запроса. Конфликты происходят на фоне обслуживания принятых к обслуживанию заявок. При этом высвобождается каналный ресурс, который может быть перераспределен под нужды запросного канала. Кроме того, свое влияние накладывают и особенности СССР, коими являются:

1. большое время распространения при передаче на спутник и со спутника (от 0.24 до 0.27 с);
2. по стволу передаются широкополосные сигналы;

3. спутник транслирует на все наземные станции одновременно, за исключением направления на центральный узел спутниковой связи (ЦУСС);

4. все наземные станции прослушивают служебный канал управления, в котором передается запрос вызываемой земной станции на организацию канала.

Первая особенность вводит ограничение, связанное с тем, что время недоступности информации об исходе прохождения запроса сопоставимо со временем передачи нескольких запросных пакетов, что обуславливает необходимость упреждающего прогноза о данном исходе.

Вторая особенность вводит ограничение, связанное с предоставлением самого канала связи: земной станции может быть предоставлен только запрашиваемый канал со специфическими свойствами (радиоданные и скорость канала).

Третья особенность вводит возможность реализации централизованного управления параметрами протокола доступа к ССС, реализованного с задержкой (см. особенность №1);

Четвертая особенность обуславливает отсутствие необходимости квитирования поступления запроса в ЦУСС.

Так как число спутниковых каналов связи является конечным, то задача их распределения между конечным числом активных земных станций (см. особенность №2) – является комбинаторной задачей. Причем решение данной задачи должно быть:

- достаточно быстрым, для обеспечения приемлемого времени реакции системы управления на изменение наблюдаемых параметров;
- достаточно надёжным, для обеспечения устойчивости выработанного варианта управления к изменению внутренних и внешних параметров ССС, в течении времени запрета управления (связано со временем идентификации состояния ССС: время набора статистики входных параметров [4]);
- адекватным описываемой ССС.

Как правило, сегмент спутниковой системы связи разрабатывается под конкретные типы оконечной аппаратуры связи - потребителей коммуникационного ресурса. В свою очередь эти потребители характеризуются:

- количеством - m_i на i -ом УС, $i \in 1..N$, N -количество УС;
- типом k_m запрашиваемого коммуникационного ресурса для m -го средства связи (1 пакет на окно со скоростью v_1 бит/сек, 2 пакета на окно со скоростью v_2 бит/сек,..., выделенное направление со скоростью v на t -сек., ...);
- интенсивностью запросов на предоставление коммуникационного ресурса k -го типа для m -го средства связи - $\lambda_m(k)$;
- значением примитивного первичного трафика $p_{0,m}$ для m -го средства связи.

Вероятность повторной передачи запроса в очередном временном сегменте $p_{r,i}$ является общим показателем для i -ой земной станции в целом.

Структура запросного канала описывается двумя значениями: общим количеством каналов, выделенных под запросно-вызывной канал (ЗВК) - $\Omega_{ЗВК}$, и количеством каналов, выделенных под его подканал $\Omega_{ЗВК2}$ (подканал служит для обмена служебной информацией – тип запрашиваемого соединения и готовность к соединению [4]).

Для синтеза системы управления коммуникационным ресурсом спутниковой системы связи с предоставлением каналов связи по требованию необходимо разработать такую математическую модель, аргументами которой выступали бы: $N, \lambda_m(k), p_{0,1..N}, p_{r,1..N}, \Omega_{ЗВК}, \Omega_{ЗВК2}$. Результатом работы модели должны быть значения следующих показателей: среднее время ожидания обслуживания заявки m -ого средства связи $\bar{T}_{об, m, k}$, вероятность блокировки ЗВК - $P_{бл.ЗВК}$, и обобщенный показатель качества ССС $P_{оов}(k)$. Кроме того, в силу использования ШПС, необходимо в модели учитывать энергетические характеристики сигналов, эффект захвата сигнала [3] и процедуру распознавания исходных данных: из-за ШПС распознавание столкновения заявок на фоне помех представляет определенные трудности, что обуславливает применение косвенных методов.

Тогда, имея такую модель, можно составить поле возможных значений искомым характеристик ССС. Получаемая таким образом функциональная зависимость $P_{оов}(k)$ от входных параметров может служить

для нахождения обратной функциональной зависимости $p_{r,1..N}, \Omega_{ЗВК}, \Omega_{ЗВК2}$ от $P_{дог}^{треб}(k)$, при неизменных значениях $N, \lambda_m(k)$ и энергетических показателей.

Очевидно, что просчет подобной области решений $P_{дог}(k)$, является громоздкой задачей. Адекватность решения реальной ситуации накладывает ограничения на детальность проработки математической модели. В работах Шиманова и Ковалькова показано, что зависимости $P_{дог}(k)$ являются нелинейными зависимостями на наблюдаемом промежутке времени. Это в свою очередь обосновывает решение применения в качестве таковой математической модели реального времени – нейросети, обученной по обучающей выборке, полученной с помощью имитационной модели.

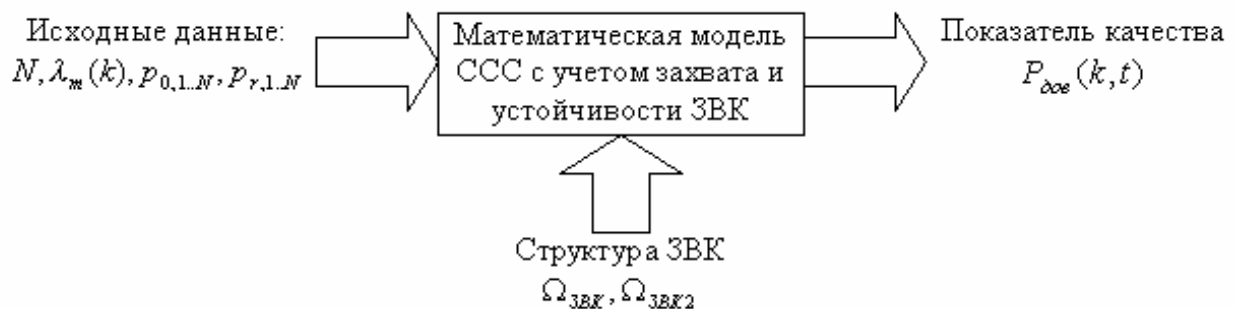
Тогда структура процедуры синтеза системы управления коммуникационным ресурсом может быть представлена в виде следующей схемы (см. рис. 1).

Следует отметить, что важным остается вопрос идентификации значений входных параметров математической модели. Одним из вариантов таковой идентификации может послужить применение опять же нейросетевого подхода, основанного на теории распознавания образов и нечеткой логики.

Литература

1. Назаров А.А., Юревич Н.М. Исследование явления бистабильности в сети с протоколом Алоха для конечного числа станций// Автоматика и телемеханика. 1996, №9, с.91-100;
2. Людоговский А.С., Цимбал В.А., Шиманов С.Н. Протоколы случайного множественного доступа в спутниковой сети связи с низкоорбитальными спутниками ретрансляторами//Технологии электронных коммуникаций. – Т.49, Спутниковые системы связи, 1994. – С.71-81;
3. Бунин С.Г., Войтер А.П. Вычислительные сети с пакетной радиосвязью. – К.:Тэхника, 1989. – 223с.
4. Ковальков Д.А. Особенности оценивания качества запросно-вызывного канала с двухфазовым обслуживанием и вероятностным захватом запросов абонентов// Проблемы обеспечения эффективности и устойчивости функционирования сложных технических систем. Сборник трудов. Ч.1. XXII Межведомственная НТК. – Серпухов, 2003. С.176-180.

1 Этап: просчет поля возможных значений $P_{дог}(k)$ с учетом устойчивости ЗВК:



2 Этап: замена математической модели на адекватную нейросеть

3 Этап: решение обратной задачи: нахождение $p_{r,1..N}, \Omega_{ЗВК}, \Omega_{ЗВК2}$ при известных $P_{дог}(k, t)$

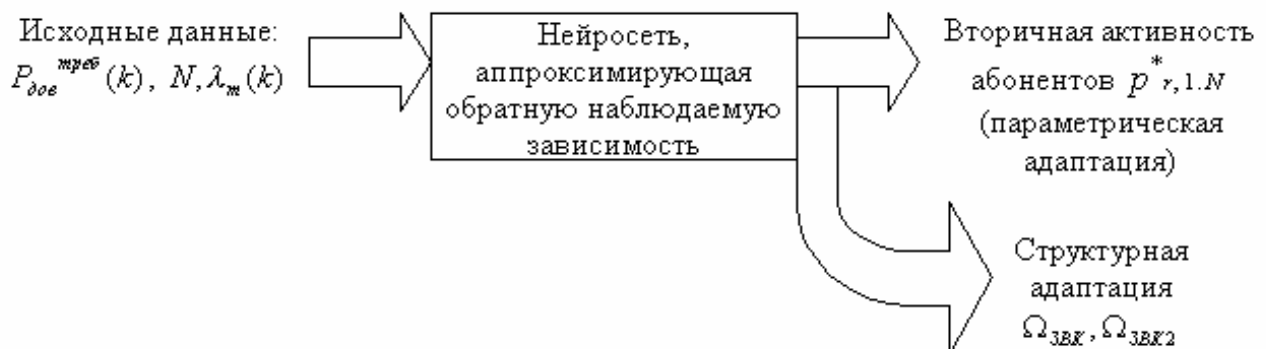


Рис. 1. Процедура синтеза системы управления коммуникационным ресурсом ССС с ПКТ.

SYNTHESIS OF A CONTROL SYSTEM BY A COMMUNICATION RESOURCE OF SATELLITE SYSTEM OF COMMUNICATION WITH GRANTING CHANNELS ON DEMAND ON A BASIS NEUREAL NETWORKS

Chistoprudov D.

Serpukhov Military Institute of Rocket Forces

In article is offered procedure of synthesis of a control system by a communication resource of satellite system of communication of DAMA.

ИСПОЛЬЗОВАНИЕ НЕЙРОСЕТЕЙ ДЛЯ ПОЛУЧЕНИЯ И ОБРАБОТКИ СИГНАЛОВ С ПЕРСПЕКТИВНЫМИ ВИДАМИ МОДУЛЯЦИИ

Ковальков Д.А., Репин М.В.

Серпуховской военный институт РВ

Известно [1], что максимальное значение пропускной способности канала связи описывается известным уравнением К. Шеннона и зависит от ширины полосы пропускания канала и отношения сигнал/шум, действующем в канале связи. При этом решающее значение имеет выбранный вид модуляции несущего сигнала. Приоритетным в данной области является направление, связанное с использованием квадратурной (QASK) модуляции. Применение QASK позволяет максимально приблизиться к пределу пропускной способности канала за счет максимального использования энергетики сигнала.

В настоящее время технология использования QASK проработана достаточно глубоко. Она базируется на том, что каждому значению дискретного символа C_k ставится в соответствие пара величин - амплитуды синфазной и квадратурной составляющих, что эквивалентно амплитуде и начальной фазе несущего колебания. Данное преобразование основано на известном в радиотехнике преобразовании Гильберта, которое позволяет выделять у произвольного сигнала $S(t)$ значения амплитуды и фазы. Однако данное преобразование является физически нереализуемым, что накладывает ограничения на практическое использование данного метода преобразования исходного сигнала. В силу этого на практике рассматривают квадратурный сигнал как сумму синусоидальной и косинусоидальной составляющих несущего колебания. На приемной стороне используют два умножителя (для синуса и косинуса, соответственно) и фильтры низких частот. Получаемый на передающей стороне сигнал обладает размытым спектром, что обусловлено мгновенными изменениями значений амплитуды передаваемого сигнала. Для уменьшения спектра получаемого сигнала используют так называемые фильтры с косинусным сглаживанием АЧХ (или SQRT-фильтры). При этом, при наличии согласованности SQRT-фильтров на передающей и на приемной сторонах можно получить довольно хорошие результаты по распознаванию принимаемого сигнала [2].

В качестве продолжения идей QASK используется способ многомерной многоуровневой модуляции [3]. Для этого необходимо отказаться от традиционного синусно-косинусного представления сигналов и рассматривать сигналы в базисе из других функций. При этом сигналы базиса должны обладать следующими свойствами:

- каждый сигнал должен занимать конечную полосу частот F ;
- сигналы должны быть ортогональны на конечном интервале T .

В качестве такового базиса рассматриваются смещенные вытянутые сфероидальные функции (см. рис. 1)

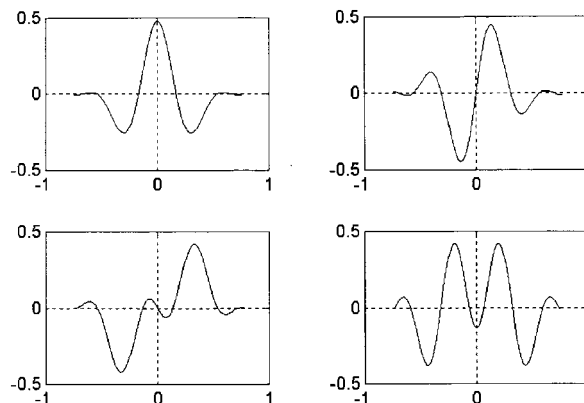


Рис.1 Базисные функции для четырехмерного пространства

Использование смещенных вытянутых сфероидальных функций в качестве базисных элементарных сигналов может представлять интерес как средство повышения спектральной эффективности беспроводных сетей передачи информации[3].

Таким образом для многомерной многоуровневой модуляции число точек созвездия (сигнально-кодовой конструкции), используемых для представления дискретного символа, определяется числом уровней амплитуды базисных сигналов L и количеством используемых ортогональных базисных функций N . На рисунке 2 представлена сигнально-кодовая конструкция для $L=4$ и $N=3$, при этом каждая ось ассоциируется с соответствующей базисной функцией.

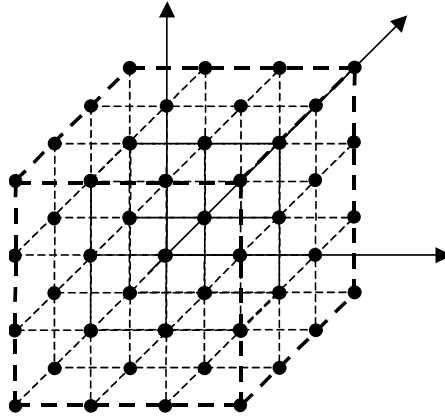


Рис. 2. – Сигнально-кодовая конструкция для трехмерного случая

Однако, открытым в настоящее время является вопрос получения таковых функций.

Одним из вариантов получения на передающей стороне базисных функций может быть вариант использования нейросетей, обученных по шаблону базиса. Дело в том, что обученная нейросеть достаточно быстра в процедуре формирования отклика. Кроме того, все процессы, происходящие в нейросети не связаны с процедурой коммутации (описываются дифференциальными уравнениями переходов), а следовательно, на выходе нейросети возможно получение слабоизменяющихся структур. Это свойство, в свою очередь, может быть использовано для формирования формы огибающей спектра передаваемого сигнала. Наличие же информационной составляющей, как входного параметра нейросети, делает возможным воспроизведение значения шаблона с необходимой точностью, заложенной в структуре сети.

Кроме того, использование базисных сигналов обуславливает необходимость применения оптимальных приемников, что влечет за собой отказ от использования SQRT-фильтров, вносящих искажения в получаемый сигнал.

Применение QASK для радиосигналов также накладывает ограничения, связанные с видом спектра получаемого сигнала (для повышения стойкости к помехам целесообразно использовать сигнал с широким спектром). Это также является основой для альтернативного SQRT- свертке способа в силу отказа от корректирующих АЧХ фильтров.

Отказ от SQRT-фильтров актуализирует задачу идентификации принимаемого сигнала. При этом механизм приема таковых сигналов должен быть достаточно гибким для реализации узнавания «размытых» образов сигналов. В данном контексте понятие «размытости» связано с неточностью соответствия образа и распознаваемого сигнала. Эта неточность обосновывает применение на приемной стороне аппарата нейросетей для решения задачи распознавания принимаемых сигналов.

Так как дискретность пространства амплитудных значений известна заранее, то необходимости в использовании аппарата нечеткой логики и гибридных нейросетей нет. Вместо этого целесообразно в качестве классификатора использовать так называемые вероятностные нейросети (PNN) [4]. Единственным управляемым параметром этих нейросетей является среднее квадратическое отклонение определяемого класса. В качестве входных параметров этого классификатора будем рассматривать мгновенные значения синфазной и квадратурной составляющих (для многомерной модуляции, соответственно, мгновенные значения составляющих базиса). Особенностью нейросети PNN является наличие области неопределенности, т.е. области равновероятности двух или более классов. Для интеграции PNN в конечное устройство предлагается (обход области неопределенности) использовать инверсный входной сигнал и классификатор инверсного сигнала. Поясним.

Для примера зададимся QASK-16 ($L=4$, $N=2$) в поле значений: для синфазной составляющей от -1 до 1 , для квадратурной составляющей от -1 до 1 (см. рис.3). Для построения PNN нейросети была использована среда Matlab. В качестве обучающей выборки использовались координаты центров классов. Впоследствии построенная нейросеть опробывалась на всей области допустимых значений. Полученные результаты распознавания приведены на рис. 2.

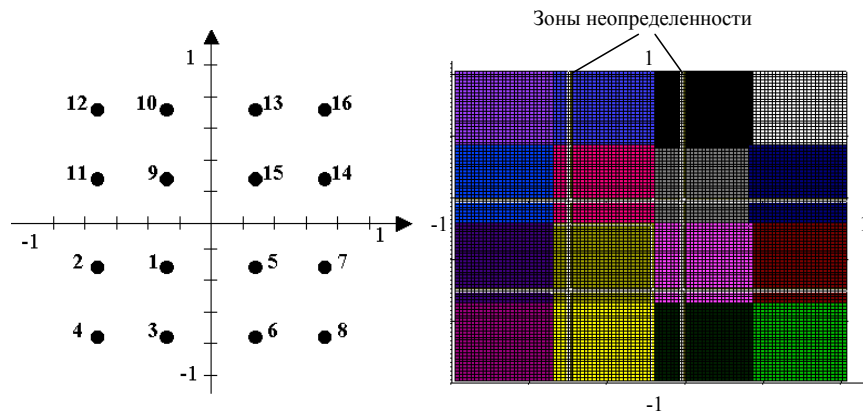


Рис. 3. Созвездие для QASK-16 и результат классификации сигнала на приемной стороне

Нетрудно заметить, что зоны неопределенности классифицируемого сигнала смещены относительно центра поля возможных значений синфазной и квадратурной составляющих. Вариация значений параметра среднеквадратического отклонения при обучении PNN нейросети к устойчивым результатам не привела, однако позволила сделать следующее наблюдение: при изменении обучающей матрицы на инверсную ($P'_{обуч} = -P_{обуч}$), зоны неопределенности остаются, сохраняя при этом смещение относительно центра влево и вниз. Если при этом использовать инверсное значение синфазной и квадратурной составляющих, то возможна правильная идентификация распознаваемого образа.

Признаком неопределенности распознаваемого класса служит «0» на выходе PNN нейросети (в других случаях на выходе номер класса). Это значение может быть использовано как разрешающий сигнал для работы инверсной PNN нейросети. При этом возникает задержка сигнала, связанная с повторным процессом идентификации сигнала (задержка работы нейросети). Однако, в силу того, что нейросеть работает гораздо быстрее, чем передаются информационные символы, от схемы компенсации временной задержки можно отказаться.

Все вышесказанное относительно PNN нейросетей подтверждает возможность их использования в системах с QASK модуляцией, где применение SQRT фильтров невозможно или нецелесообразно.

Литература

1. О.И. Лагутенко. Модемы. Справочник пользователя. - СПб.: «Лань», 1997 г.;
2. А.Б. Сергиенко. Цифровая обработка сигналов. - СПб.: «Питер», 2002 г. - 608 с.;
3. В.В. Крылов, С.С. Курганова. Многомерная многоуровневая модуляция. - ННГУ: Труды научной конференции по радиофизике, 2002 г.;
4. В.В. Круглов, М.И. Дли, Р.Ю. Голунов. Нечеткая логика и искусственные нейронные сети. - М.: Издательство физико-математической литературы, 2001 г. - 224 с.

USE NETWORKS FOR RECEPTION AND PROCESSING OF SIGNALS WITH PERSPECTIVE KINDS OF MODULATION

Kovalkov D., Repin M.

Serpukhov Military Institute of Rocket Forces

In article on the basis of modelling it is offered to use probabilistic neural networks for identification accepted quadrature amplitude shift keying (QASK) a signal.

ЭМОЦИОНАЛЬНЫЕ АСПЕКТЫ ДИСКУРСА В ТЕХНОЛОГИИ ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ В ПРОБЛЕМЕ ФОРМИРОВАНИЯ БАЗЫ ЗНАНИЙ

Шапиро Д.И.¹, Захарченко И.Н.²

¹Институт проблем вычислительной техники и информатизации

²Московский Гуманитарный Университет

В разработке технологий виртуальной реальности требуется, с одной стороны, адекватный подход к особенностям Пользователя, а, с другой стороны, глубокое и тонкое представление «человеко-подобных» свойств интеллектуального виртуального коммуниканта (IVC). К подобным свойствам относятся, в частности, понимание вопросов различной степени сложности и образных представлений, их обработка и порождение высказываний. По мере развития информационного общества («общества знаний») существенно расширяется сфера применения VRT и, следовательно, требований к возможностям IVC.

В проекте «Virtual Walks on Rinascimento Roma with Stendhal» (VIRSTENDHAL) индивидуальные особенности IVC (Стендаля) являются основой организации дискурса.

В докладах на конференциях EVA в Лондоне и Москве (Шапиро, 2003) были приведены четыре типа вопросно-ответных процедур: простые, стандартные, нестандартные, уникальные.

Порождение ответов на уникальные вопросы требует использования личностных особенностей прото-типа IVC (Стендаля).

Великий Стендаль (Анри Бейль) был незаурядной творческой, романтической личностью: офицер наполеоновской армии и убеждённый республиканец, замечательный яркий романист и жёсткий трезвый аналитик, великолепный знаток итальянской истории, искусства, музыки и участник движения карбонариев против австрийского владычества в Италии. И при этом он был глубоким ценителем женской красоты и тонким ранимым человеком. Стефан Цвейг написал о нём: «он жил одновременно в двух Мирах – мире рассудка и мире чувств»

Для изучения эмоциональной специфики Стендаля использовались некоторые фрагменты его эпистолярного наследия (в частности, переписка с дамами). В ходе анализа переписки Стендаля с дамами было выбрано чувство, определяющее характер его переживаний и поведения – чувство любви.

На основе писем Стендаля к дамам предпринята попытка структурировать эмоциональные оттенки чувств великого писателя. Анализ показал наличие двух типов чувств, переживаемых им как любовные – романтической, возвышенной любви к миланской даме г-же Дембовской, урожденной Метильде Висконти-ни, и более приземленной, основанной на плотских отношениях любви к светской красавице г-же Кюриаль. Переживаемое писателем чувство можно в обоих случаях назвать любовным, но анализ эмоциональных оттенков показывает различие их характеристик.

В ситуации возвышенной романтической любви диапазон эмоциональных оттенков оказывается весьма широким. Выстроенные по нарастанию (см. табл. 1), оттенки чувства Стендаля к г-же Дембовской описываются как

1. Возвышенная мечтательность («мечта, завладевшей моей душой»).
2. Идеализация любимого человека («Если бы у вас были недостатки, я говорил бы, что обожаю ваши недостатки, и говорил бы правду»).
3. Робость в поведении («В вашем присутствии я робок, как ребенок ...»).
4. Тревога, смятение чувств («В полном смятении, в которое повергает меня непреодолимая потребность видеть вас ...»).
5. Роковая, всепоглощающая страсть («Я охвачен пагубной страстью», «Отныне эта страсть стала важнее всего в моей жизни», «Когда человеком владеет всепоглощающая страсть...»).

Шкалирование оттенков демонстрирует сложную градацию переживаемого автором писем чувства, полностью его захватившего.

Менее дифференцированной является шкала оттенков чувства любви в отношениях с г-жой Кюриаль (см. табл. 2). Переживаемое писателем чувство возможно разложить на следующую шкалу эмоциональных оттенков:

1. Любовь как потребность в душевной дружбе («потребность в душевной дружбе», «люблю больше, потому что наша дружба становится еще тесней»).
2. Любовь как потребность в интимном общении («Если бы мы провели вместе целую неделю и наши сердца всегда бились бы так же горячо, мне кажется, мы кончили бы тем, что уже не расставались бы больше»).
3. Нежность как высшее выражение любви («Вчерашний обед был совершенной картиной счастья, интимности, нежности»).

Таблица 1. Письма к г-же Дембовской

Чувство	Эмоциональные оттенки	Цитаты
---------	-----------------------	--------

Романтическая любовь	1. Возвышенная мечтательность.	С. 182: «Я буду любить вас весь остаток жизни; что бы вы ни сделали, это ничего не изменит в мечте, завладевшей моей душой».
	2. Идеализация любимого человека.	С. 187: «Сударыня, в вашей душе так много благородства, что она не могла не понять моей».
	3. Робость в поведении.	С. 187: «Если бы у вас были недостатки, ... я говорил бы, что обожаю ваши недостатки, и говорил бы правду».
	4. Тревога, смятение чувств.	С. 185: «Я кажусь вам смешным, моя робость и молчаливость наскучили вам».
	5. Роковая, с трудом переносимая страсть.	С. 185: «В вашем присутствии я робок, как ребенок, слова замирают у меня на губах, я могу только смотреть и любоваться вами».
		С. 187: «Едва лишь я полюблю, как становлюсь робким, и вы можете судить об этом по растерянности, которую я проявляю каждый раз, как нахожусь подле вас».
		С. 186: «В полном смятении, в которое подвергает меня непреодолимая потребность видеть вас, я все же сохранил пока одно качество ...».
		С. 186: «Бывают мгновения в долгие одинокие вечера, когда я мог бы стать убийцей, если бы нужно было совершить убийство для того, чтобы увидеть вас».
		С. 187: «Тот миг, который должен был стать блаженнейшим в моей жизни, стал, напротив, одним из самых тревожных из-за боязни не угодить вам».
		С. 191: «Я чуть не отказался от этого счастья; я чувствовал, что не в силах перенести его».
		С. 193: «Я имею несчастье любить по-настоящему».
		С. 185-186: «Я охвачен пагубной страстью и уже не могу отвечать за свои поступки».
		С. 186: «Отныне эта страсть стала важнее всего в моей жизни. Все другие интересы, все соображения бледнеют перед нею. Эта роковая потребность видеть вас владеет мною, увлекает и вдохновляет меня».
		С. 186: «Всегда в моей жизни все, что отвлекало меня, что не имело отношения к моей страсти, для меня не существовало; счастливая или несчастная это страсть, она заполняет каждое мое мгновение».
		С. 188: «Когда человеком владеет всепоглощающая страсть, все, что он говорит, и все, что он делает в каких-либо обстоятельствах, еще не позволяет судить о нем самом».

Таблица 2. Письма к г-же Кюриаль.

Чувство	Эмоциональные оттенки	Цитаты
---------	-----------------------	--------

Любовь, основанная на близких отношениях	1.Потребность в задушевной дружбе	С.213: «Я чувствую такую потребность в задушевной дружбе, что уже, кажется, жалею о парижских туманах». С.215: «Когда я вижу с тобой, мой ангел, три дня подряд, мне всегда кажется, что я люблю тебя еще больше, люблю больше, потому что наша дружба становится еще тесней».
	2.Потребность в интимном общении.	С.215-216: «Я так люблю тебя сегодня, так тебе предан ... Если бы мы провели вместе целую неделю и наши сердца всегда бились бы так же горячо, мне кажется, мы кончили бы тем, что уже не расставались бы больше».
	3.Нежность как высшее выражение любви.	С.216: «Вчерашний обед был совершенной картиной счастья, интимности, нежности. Таких моментов, по-моему, ... никогда не испытываешь, если позволишь себе разыгрывать комедию с тем, кого любишь» С.216: «Мне хотелось бы, чтобы любовь моя казалась тебе самым нежным из всего, что существует на свете».

Таким образом, формирование специализированного фрагмента [ВК-ЕМ] базы знаний основано на совокупном представлении многоэлементных качественных шкал, определяющих конкретные эмоциональные особенности IVC.

Литература

1. Стендаль. Собрание сочинений в 15 томах. Т. 15. М., 1959.
2. Шапиро Д.И. (2003). Основы технологии виртуальной реальности. М., ИКФ, 2003. Каталог.
3. Шапиро Д.И., Захарченко И.Н. Проблема организации образного дискурса «Пользователь – Интеллектуальный Виртуальный Коммуникант» // LVIX научная сессия, посвященная дню радио. Труды. Т.2. М., 2004.
4. Шапиро Д.И., Захарченко И.Н. Интеллектуальный виртуальный коммуникант: эмоциональная специфика высказываний. EVA 2004. М., 2004.

The problems of the formation of Base Knowledge for the preparation of the statements with the emotional specifics are discussed. The Stendhal's epistolary cultural heritage is used.



ОСОБЕННОСТИ РАЗРАБОТКИ БАЗЫ ЗНАНИЙ ДЛЯ ИНТЕЛЛЕКТУАЛЬНОГО ДИСКУРСА В СИСТЕМАХ ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ. ВЛИЯНИЕ ЭТАПА РАСПОЗНАВАНИЯ РЕЧИ НА ФОРМИРОВАНИЕ БАЗЫ ЗНАНИЙ ВИРТУАЛЬНОГО КОММУНИКАНТА

Шапиро Д.И., Овчаренко Г.А

Всероссийский научно-исследовательский институт проблем вычислительной техники и информатизации

Этот доклад имеет целью раскрыть одну из основных проблем при построении дискурса между человеком и искусственным коммуникантом, а именно распознавание речи. Постараемся понять, каким образом этот этап дискурса влияет на реализацию базы знаний и какими путями можно пойти при его построении.

Постараемся вспомнить с какими особенностями приходится сталкиваться разработчикам при построении баз знаний для виртуального коммуниканта:

- 1) Языковые особенности (сложность русского языка, многообразие правил, форм спряжений, склонений и т.п.);
- 2) Особенности, связанные с поиском по базе знаний (установление коэффициента точности при котором распознанное слово будет сопоставляться слову из базы знаний);
- 3) Особенности формирования ответных реплик при работе с базой знаний по клишированным выражениям;
- 4) Временные особенности (попытка приближения к "realtime" диалогу, без задержек на поиск, распознавание и формирование ответной фразы).

В настоящее время распознавание речи сводится к решению трех типов задач:

1. распознавание отдельно произносимых слов;
2. распознавание слитной речи;
3. идентификация по образцу речи.

Распознавание отдельных слов по большей части используется для речевого управления вычислительной машиной, и является для нас частным случаем.

Распознавание слитной речи имеет целью преобразования в текст естественной речи человека. Идентификация по образцу речи используется для целей обеспечения безопасности, и не будет рассматриваться в этом докладе.

Остановимся подробнее на распознавании слитной речи.

Механизмы распознавания речи.

В большинстве существующих механизмов можно выделить четыре основных модуля (рис.2):

1. модуль сбора данных,
2. экстрактор,
3. компаратор,
4. интерпретатор.



Схема механизма распознавания речи.

Модуль сбора данных включает получение входного сигнала и его предварительную обработку, которая может включать автоматическую регулировку усиления, подавление эха, обнаружение присутствия/отсутствия речи и обнаружение интонационного конца фразы.

Этот модуль включает также выделение отрезка речи из входящего речевого сигнала. Существует несколько алгоритмов определения начала и конца речи.

В одном из них определяется некоторый пороговый уровень сигнала. Начальная точка речи в этом случае соответствует моменту, когда входящий сигнал начинает превышать пороговый уровень, а конечная точка – моменту, где амплитуда входящего сигнала меньше пороговой. Основным недостатком этого метода заключаются в невозможности точного определения речевого отрезка в случае сильного шума, или, наоборот, тихой речи. В нашем случае, необходимо отслеживать только конец поступления сигнала, так как начало будет ознаменовываться нажатой собеседником кнопкой. В любом случае система определения входного сигнала поможет избежать лишней информации, поэтому не будем торопиться и отказываться от нее.

Другой метод использует нормализацию амплитуды входного сигнала в соответствии с минимальной амплитудой. Полученные нормализованные значения сравниваются с пороговым значением.

Экстрактор выполняет частотный анализ сигнала. Акустическо-фонетический поток данных разбивается на короткие кадры, или векторы, продолжительностью, как правило, около 10 мс. Как правило, для каждого кадра определяется ряд параметров, используя быстрое преобразование Фурье. Кроме того, многие системы используют вместо, или вместе с этими характеристиками другие, например, спектральные характеристики, а также первую и вторую производную от спектральных характеристик.

Компаратор осуществляет акустическое сравнение: каждый кадр, или вектор, сравнивается с имеющимися акустическо-фонетическими образцами, хранящимися в специальной базе данных. При этом сравниваться могут как отдельные фонемы, так и слова, и даже фразы. При небольшом количестве слов, используемых диктором, более высокую надежность и скорость можно ожидать от распознавания целых слов, но при увеличении словаря скорость резко падает, и оптимальным становится распознавание отдельных фонем.

В основном используются три алгоритма для распознавания кадров:

- алгоритм динамической трансформации шкалы времени (dynamic time-warping)
- скрытое Марковское моделирование (Hidden Markov Modeling)
- нейронная сеть с временной задержкой (time-delay neural network)

Алгоритм динамической трансформации шкалы времени использует оптимизационный принцип для минимизации числа ошибок, возникающих при сравнении распознаваемого слова с эталонной моделью. Несмотря на то, что эта технология дает хорошие результаты, но очень трудоемка.

Скрытое Марковское моделирование использует вероятностные модели слов. При использовании этой технологии для каждого возможного варианта распознаваемого слова вычисляется вероятность, затем полученные вероятности сравниваются и выбирается слово с наибольшей вероятностью. Данное моделирование успешно используется в настоящее время во многих приложениях. Однако она содержит ряд недостатков: невозможность различия двух схожих по произношению слов, чувствительность к шуму и большое число вычислений.

Нейронная сеть с временной задержкой в случае распознавания ограниченного числа слов дает лучшие результаты, чем метод скрытого Марковского моделирования. Однако этот метод содержит недостаток: часто он сходится к локальному минимуму, не являющемуся глобально оптимальным решением.

Один из методов, основанный на сравнении фонем, использует понятие «контекстная фонема». В данном методе фонема рассматривается в сочетании с предшествующей и последующей фонемой. Далее в процессе распознавания определяется фонема, наиболее близко соответствующая распознаваемой.

Большинство систем распознавания речи при сравнении кадров с эталонами (в нашем случае словами из базы знаний) используют балльную систему, в том числе и скрытое Марковское моделирование, для определения наиболее подходящего эталона из всех возможных.

Интерпретатор решает задачу динамического программирования с целью найти наилучшее разбиение полученного от компаратора "алфавитного" потока на слова и фразы. В зависимости от объема используемого словаря и действующих синтаксических правил, применяются различные стратегии поиска и отсева.

В данном блоке из распознанных фонем формируются слова, а из слов фразы. При этом также часто используется балльная (или вероятностная) система сравнения результатов.

Проанализируем те требования, которым должна отвечать выбранная нами система распознавания речи:

1. произвольный пользователь
2. аппаратные требования
3. быстрота работы
4. работа с именами собственными и т.п.
5. возможность адаптации к новым базам знаний (обновляемость)

-произвольный, пользователь;

Все предельно ясно когда систему обучают распознавать речь конкретного человека. Система программируется с учетом речевых особенностей индивидуума, при этом страдая ограниченностью применения, но обладая высоким процентом распознавания. Обычно процент распознавания может достигать 90-95%.

Таким образом, при создании системы распознавания необходимо понимать, что система будет либо дикторнезависимой для разбора отдельных слов и выражений с небольшим словарем, но с хорошим качеством распознавания даже на фоне шумов, либо дикторозависимой для распознавания слитной речи, с практически неограниченным словарем, но с труднопредсказуемым качеством распознавания, зависящим от положения микрофона, состояния говорящего (радость, смех, грусть), фоновых шумов и так далее.

Для того, чтобы попытаться сделать систему дикторозависимой, следует постараться по первым фразам идентифицировать собеседника и отнести его к какой либо группе (по произношению, акценту, эмоциональности и т.п.).

Для попытки определения первичных признаков собеседника и его речевых особенностей следует создать на этапе проектирования базу знаний по определенным типам собеседников, чтобы в дальнейшем сократить количество ошибок связанное с речевыми особенностями.

- аппаратные требования

Немаловажной остается техническая сторона процесса, а именно чувствительность микрофона, шумы, помехи и т.п. Речь в виде акустических волн попадает на микрофон, где преобразовывается в электрический сигнал. От качества микрофона, линии передачи и наличия акустических шумов в большой мере зависит правильность опознавания слов, численно выраженная как качество распознавания (отношение числа правильно распознанных слов к общему числу слов, произнесенных перед микрофоном). В случае невозможности повлиять на качество микрофона (наш случай, когда собеседник использует микрофон, который имеет характеристики, далекие от идеальных), необходимо хотя бы учитывать возможные особенности частотных характеристик такого микрофона. То же относится и к линиям связи.

И наконец еще одна нерешенная проблема, которое в теории восприятия речи носит название “cocktail-party” эффект или более расширенно – анализ акустических сцен. Этот эффект основан на способности слушающего сосредоточить внимание на выделенном источнике звука в условиях сильной зашумленности. В настоящий момент в автоматическом распознавании речи разрабатываются в основном методы подавления сравнительно гладких помех и искажений. Хотя помехоустойчивость систем распознавания считается одним из основных направлений в области создания систем практического назначения, заметных прорывов в этой области не наблюдается.

- быстрота работы

В нашей системе как нигде важна скорость работы системы распознавания речи. После распознавание речи и разбиения фразы на составляющие, проходит этап выделения и поиска по базе знаний ключевого слова. Первый этап диалога, а именно получение первой реплики окажется самым долгим. Какие есть возможности для сокращения времени работы системы и одновременного повышения ее эффективности. Вот

некоторые моменты для размышления:

Проверки распознавания (“+” повышение качества распознавания, “-” увеличение времени)

За и против:

А) Некоторые проверки не дают результатов лучших, чем 5% ошибок.

Б) Так как слова, включенные даже в хорошо и аккуратно произносимый текст, оказываются как бы плавающими в океане омонимии, то количество ошибок (словесных) возрастает приблизительно в 5 раз. Беглое отслеживание таких ошибок, кроме случаев возникновения нелепых текстов, уже затруднительно. Аппарат коррекции ошибок в большинстве демонстрационных систем слабо отлажен.

В) Были упоминания, что даже для хорошо организованных спонтанно произнесенных текстов вероятность правильного распознавания слов не превышает одной трети.

Г) Время обработки введенного отрезка речи в таких системах может занимать минуты.

Запараллеливание процесса работы

А) Использование нескольких параллельно работающих способов выделения одних и тех же элементов речевого сигнала на базе анализа акустического сигнала. Примером может служить параллельное использование формантных и полосных признаков для идентификации фонетических элементов речевой структуры

Б) Параллельное независимое использование сегментного (фонемного) и целостного восприятия слов в потоке речи.

- работа с именами собственными и т.п.

Для сокращения поиска ключевого слова, следует заведомо исключить все слова, так или иначе не подходящие под это определение. Следует производить их идентификацию прямо в Компараторе для последующего исключения из претендентов. К таковым словам могут относиться междометия, слова паразиты, звуковые особенности речи, названия, слова собственные и т.п.

- обновляемость

Постоянные наработки и новые методы в области распознавания речи обязывают нас постоянно дополнять и обновлять необходимые элементы, как программные, так и аппаратные. Доработка баз знаний, а так же их увеличение и дополнение могут реально улучшить конечный результат.

Литература

1. Ямов С.И., Кабак И.С., Курочкин С.Н., Бродин А.Г. Многоуровневая система распознавания речи. <http://art.bdk.com.ru/govor/infast49.htm>
2. В.И.Левин. Обработка речевых сигналов. <http://art.bdk.com.ru/govor/infast61.htm>

This article is aimed to reveal one of the basic issues in creating a virtual discourse between a human being and an artificial communicant, in particular, the issue of speech recognition. We will try to understand the way, this phase of discourse effects the realization of knowledge base and what method should be chosen for its realization. We will show the difficulties preventing developers from realizing knowledge bases for virtual communicant.



ПРОБЛЕМЫ ОРГАНИЗАЦИИ АНГЛО-РУССКОГО ДИСКУРСА «ПОЛЬЗОВАТЕЛЬ-ИНТЕЛЛЕКТУАЛЬНЫЙ КОММУНИКАНТ» В ТЕХНОЛОГИИ ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ

Шапиро Д.И., Орлова Е.С.

Всероссийский научно-исследовательский институт проблем вычислительной техники и информатизации,
113114, Москва, Второй Кожевнический переулок, дом 4/6

Одно из самых передовых направлений телекоммуникационных и компьютерных технологий сегодня-это технология виртуальной реальности, которая применяется практически во всех сферах жизни общества. Подавляющее большинство работ в этой области связаны с техническими приложениями, в которых общение пользователя с компьютерными системами (виртуальным миром) осуществляется операциональным способом, то есть через операцию ввода-вывода данных. Отдельный класс представляют задачи применения технологии виртуальной реальности для решения гуманитарных задач и задач межкультурного обмена, в которых пользователь общается в виртуальном мире с интеллектуальным субъектом (индивидуальным виртуальным коммуникантом). И здесь возникает проблема организации вербального интеллектуального дискурса, который представляет особый интерес для музейных технологий виртуальной реальности, где произвольный пользователь общается с выдающимся специалистом- интеллектуальным виртуальным коммуни-

кантом. Для нашей страны это является очень важным, так как межкультурный обмен должен быть адекватным для российских пользователей, а международные произведения культуры представлены на английском языке. На этом этапе возникает сложность из-за большого различия языков. Например, пользователь, владеющий языком романо-германской группы, не испытывает проблем при использовании интернета на английском языке, так как алфавиты и написание букв идентичны, а большинство слов пишется одинаково и имеют более или менее сходные значения. Русский пользователь не обладает такими преимуществами.

Реализация интеллектуального дискурса в технологии виртуальной реальности связана с реализацией диалога, то есть широкого круга вопросно-ответных процедур, обмена высказываниями и мультисенсорными образами, причем общение должно представлять собой двуязычный диалог по схеме: вопрос от пользователя на русском языке, компьютер переводит на свой язык, ищет нужный ответ и переводит на язык пользователя, то есть ответ дается на том языке, на котором был задан вопрос, при этом компьютер работает на английском языке, то есть речь идет об использовании двуязычной вопросно-ответной процедуры. Очень важно учитывать специфику пользователя, а именно, эстетическую, образовательную и поведенческую специфику, так как люди различного социального происхождения и различного образования используют разные слова и обращают внимание на совершенно различные детали. В общем, интеллектуальный дискурс имеет несколько особенностей, например, использование специфического множества слов, использование специфических интонационных характеристик, использование двуязычной вопросно-ответной процедуры. И здесь возникает вопрос перевода с одного языка на другой.

Хорошо известны и изучены технологии машинного перевода, устного перевода, аспекты общения человека и компьютера. В случае машинного перевода главная проблема заключается в искусственном построении фразы, нарушении семантики и стилистики языка. Перевод является искусственным и также нарушается естественное позиционирование слов в предложении, которое существует в устной речи. Также компьютер не воспринимает нетрадиционное построение фразы (например, инверсию) или все, что так или иначе отклоняется от заданной модели: подлежащее, сказуемое, дополнение, обстоятельство (в английском языке). В случае устного (синхронного или последовательного) перевода таких проблем не возникает из-за того, что фразы строит человек, который свободно владеет «живым» естественным языком, соответственно он владеет стилистикой и семантикой языка, причем все процессы обработки фраз происходят автоматически, неосознанно. Поэтому остается огромный пласт для изучения и применения на практике – это двуязычный диалог «пользователь-компьютер» без привлечения дополнительных программ машинного перевода и с минимальным ограничением информации. Еще один важный момент, который необходимо учитывать, то, что при общении на родном языке (или при свободном владении иностранным) собеседники всегда понимают больше, чем просто слова, то есть они воспринимают скрытую в языке информацию, эмоции, подтекст. Зачастую этот процесс происходит неосознанно, и здесь возникает проблема анализа неосознаваемого скрытого в языке знания. Это особенно важно, если речь идет о восприятии информации с голоса.

Работа по реализации двуязычного диалога должна идти по двум направлениям:

1. Учет особенностей английского и русского языков особенно при непосредственном общении:

Русский язык	Английский язык
1).грамматические особенности: различная структура фраз: -нефиксированный порядок слов; -пропуск членов предложения; -нет структуры вопросительных предложений, вопрос задается интонационно; -падежи и склонения и падежные окончания.	-Фиксированный порядок слов; -пропуск членов предложения допускается, но не является грамматической нормой; -четкая структура вопросительных предложений; -предлоги, артикли и суффиксы.

2) фонетические особенности (если восприятие идет с голоса), существующие в обоих языках такие как: ассимиляция(уподобление звуков), редукция(качественное сокращение звуков), элизия(выпадение звуков), появление связующих согласных, интонация различных типов вопросов и утверждения, ритмико-мелодический строй речи, тон и тембр голоса, даже дефекты речи. Учет интонационных особенностей является очень важным, так как интонации языков различны, а в английском языке интонация имеет решающее действие, и этот вопрос требует серьезной проработки (например, несовпадение интонации вопросов и вариативность английской интонации вопросов в зависимости от типа вопроса).

3) лексические и стилистические особенности: многозначность слов и фраз, а также наличие большого количества фразовых глаголов и идиоматических выражений в английском языке, официальный и неофициальный стили, формулы вежливости, все то, что называется лингвистическим соответствием. Особую проблему представляет постоянный процесс изменения языка, появление новых слов и новых значений слов, сленг. Здесь возникает проблема словарного запаса компьютера.

Необходимо учитывать и тот факт, что при естественном общении около 50% информации мы «считываем» с лица собеседника, а в случае с компьютером это невозможно.

2. Расширение базы знаний: создание адекватных ответов на различные типы вопросов. Вопросно-ответные процедуры характеризуются четырьмя типами вопросов: простые(что, где, когда), стандартные(сколько, как долго, какой), нестандартные(почему, какие особенности), уникальные(какая связь..., какое влияние). При этом важно учитывать особенности машинного перевода и построение фразы компьютером, о чем говорилось выше, а также тот факт, что вопрос об одном и том же может быть задан по-разному. Например: «Он был женат? У него была семья?»

Процесс создания двуязычных вопросно-ответных процедур может состоять из следующих этапов: отбор материала и составление словаря-минимума, подбор клишированных фраз и установление русско-английских соответствий (например, падежи русского языка – английские предлоги), разработка вопросных процедур с учетом различных типов вопросов и разработка соответствующих ответов, а также организация вопросно-ответных процедур для формирования алгоритмов на основе семантических карт, в которых слова со сходными значениями размещаются на одной и той же или соседних позициях, так что расстояние между ними показывает семантическое сходство. Слова определяются рядом совместно встречающихся слов, затем кодируются в векторы, принимая во внимание семантические корреляции между ними. Метод может быть расширен для создания двуязычных семантических карт, причем позиционирование слов определяется через Евклидово расстояние векторов, которые представляют слова. Вычисление сходства значений слов является важным методом, который может быть применен ко многим областям использования языка: увеличение количества вопросов в поиске информации, объяснение неясности смысла слов, изучение лексических баз данных и словарей синонимов. Семантические карты могут быть сконструированы с функцией самоорганизации. Данные для создания карты могут быть индивидуальными и интегрированными. Для создания практической семантической карты необходимо определить, какой тип информации важен, и в зависимости от этого создавать определенную семантическую карту или многоцелевую семантическую карту, чтобы можно было избежать нежелательных отклонений. Двуязычные семантические карты можно также применить в изучении иностранного языка или в письме на иностранном языке, используя двуязычный параллельный словарь. Важно то, что в этом случае перевод будет свободным и предложения будут стоять в семантическом порядке.

Особый интерес представляет разработка ответов на нестандартные вопросы. В этом случае необходимо выделить ключевые слова, определить степени сходства. Также необходимо учитывать психологические особенности человека: если задан вопрос «Была ли у него семья? / Был ли он женат?», то ожидаемый ответ не просто «Да/Нет», а хотя бы одно-два предложения о семье и детях. Рассмотрим несколько примеров простых, стандартных, нестандартных и уникальных вопросов:

1. простые вопросы.

Когда была написана картина?
Где была написана картина?
Кто художник?
Где жил художник?
Когда родился и когда умер художник?
Кто были родители художника?
Где родился художник?
Где сейчас находится картина?

2. стандартные вопросы.

Что изображено на картине?
Как долго художник писал эту картину?
Сколько было лет художнику, когда он написал картину?
В какой технике написана картина?
К какому течению принадлежал художник?
Можно ли выделить определенные периоды в творчестве художника?

3. нестандартные вопросы.

Существуют ли копии картины?
Где находятся копии?
Кто авторы копий?
Художник когда-либо продавал эту картину?
Ее вообще когда-нибудь продавали?
Сколько стоит картина?
Сколько стоила картина при первой продаже?
Какие цвета доминируют в палитре?
Какие особенности палитры художника?
Менялась ли палитра художника в течение его жизни?
Что стало с картинами этого художника после его смерти?

4. уникальные вопросы.

Почему художник нарисовал эту картину?

Типична ли эта картина для творчества художника?

Влияли на творчество художника другие художники или великие люди того времени?

Почему он стал художником?

Кем стали дети художника?

Эти процедуры и технологии могут найти широкое применение, но идеальным и необходимым является создание тонко-структурированной проблемно-ориентированной базы знаний, анализ ключевых слов и их определителей, а также анализ языковых особенностей и формирование клишированных вопросов и адекватных фраз-ответов.

Литература

1. Шапиро Д.И. Человек и виртуальный мир. Когнитивные, креативные и прикладные проблемы. Эдиториал УРСС, Москва, 2000

2. Шапиро Д. И. Основы технологии виртуальной реальности. ИКФ-каталог, 2003

3. Brown, Statistical Approach to Language Translation, COLING'S-88

4. Qing Ma and others, Self-organizing Semantic Maps and its Application to Word Alignment in Japanese-Chinese Parallel Corpora// Neural Networks, October N17-2004, pp. 1241-1253

5. Shapiro D.I., Abramson L.M., Andreeva I.A., Andreeva O.B., Organization of Multisensorial Discourse for Virtual Reality Technology: Methodology Problems. Proc. Of EVA'02, Beijing.

PROBLEMS OF ORGANIZATION OF ENGLISH-RUSSIAN DISCOURSE "USER-INTELLECTUAL VIRTUAL COMMUNICANT" IN VIRTUAL REALITY TECHNOLOGY

This report discusses some problems of realization of intellectual discourse in virtual reality technology connected with working out the bilingual question-answer procedures. Special attention is paid to the problematic spheres of realization of the problem, such as peculiarities of Russian and English languages and advantages and disadvantages of computer translation. There is also a review of the problem of Database enlargement using semantic maps.



ОБУЧЕНИЕ ПО ОШИБКЕ ПРОГНОЗА НАГРАДЫ (TD- LEARNING) - ТРЕТЬЯ ПАРАДИГМА ТЕОРЕТИЧЕСКОЙ НЕЙРОБИОЛОГИИ

Смирнитская И.А.

Институт Высшей Нервной Деятельности и Нейрофизиологии РАН

Несколько дисциплин ставят своей целью создание вычислительных устройств основанных на тех же принципах, по которым работает живой мозг. Сюда относят искусственный интеллект, теорию нейронных сетей, создание адаптивных роботов и т.д.. Эти области науки, возникли на основе достижений в понимании работы мозга, из попыток моделирования работы нервной системы. И по мере развития видно, что нейробиология и вышеупомянутые технические воплощения взаимно обогащают друг друга.

Можно сослаться на знаменитое положение Д. Марра о том, что недостаточно только исследовать строение и функционирование нервной системы, необходимо идти навстречу, пытаться строить технические аналоги живых управляющих систем, и на пути взаимодействия этих двух движений возникнет понимание того, как устроен мозг (Марр, 1987)

Три этапа есть результаты и вехи этого взаимодействия.

Этап первый – модель нейрона Мак-Каллока и Питтса (Мак-Каллок, Питтс, 1943)

Модель отразила удивительные пороговые свойства нейрона, и внедрила взгляд на нервную систему, как на составленную из элементарных унифицированных модулей - пороговых сумматоров, осуществляющих вычислительные операции.

Этап второй – синапс Хебба (Дунин-Барковский, 1978). Это построенная на одном нейроне модель Павловского условного рефлекса. Модель привлекла внимание к обучению, как определяющему свойству нервной системы, выявила способ обучения – путем запоминания сочетания условного и безусловного стимулов, и постулировала механизм обучения на клеточном уровне – изменение эффективности (модификация) межнейронных связей - синапсов.

После появления Хеббовской модели, в нервной системе начали искать, и нашли изменяемые синапсы. А устройства из пороговых сумматоров (нейронов) МакКаллока и Питтса с синапсами Хебба широко распространились. Из них строят ассоциативные сети, модели нейронных структур, устройства для распознавания образов и т.д. Но все, что до недавних пор создавалось в этой области, опиралось только на два описанных выше модельных обобщения.

Третьим этапом мы считаем модель обучения по ошибке прогноза награды (Sutton, Barto, 1998).

Под влиянием естественного отбора, живой организм в природе стремится строить свои действия так, чтобы максимизировать получаемую из окружающей среды награду. Действие, результатом которого явилось получение большей, чем раньше, награды запоминается,

По аналогии с поведением живого организма, модель обучения по ошибке прогноза награды рассматривает структуру, управляющую поведением технического устройства, как осуществляющую непрерывную коррекцию последовательности его действий в направлении получения большей суммарной награды, и запоминание этой последовательности.

Фактически это есть замыкание Павловского обучения на окружающую среду. Здесь включается в рассмотрение цель запоминания, которой не могло быть у Павловских животных, потому что они не были автономны - запоминали только то, чего хотел добиться от них экспериментатор.

ЧАСТЬ 1. МОДЕЛЬ ОБУЧЕНИЯ ПО ОШИБКЕ ПРОГНОЗА НАГРАДЫ.

Модель обучения по ошибке прогноза награды (TD-learning) является частным случаем теории обучения по подкреплению (reinforcement learning) Ричарда Сатона и Эндрю Барто (Sutton, Barto, 1998).

Ее авторы хотели сконструировать систему, которая в ходе взаимодействия с окружающей средой достигает некоторой цели. Хотя сама эта теория возникла в начале 80-тых, цепь ее предшественников восходит к началу 50-х. Среди них Р. Беллман (Беллман, 1957а,б), М. Л. Цетлин (Цетлин, 1969) и др..

Назовем автономное обучающееся устройство агентом. Целью, к которой стремится агент, является максимизация суммарной награды. Для того, чтобы добиться этой цели, обучающееся устройство вычисляет по ходу своих действий некоторую функцию от награды, которую это устройство рассчитывает получить в будущем. Эта функция называется значимостью текущего состояния или прогнозом. Она задается следующим выражением

$$V^{\Pi}(s) = E_{\Pi} \{r_{t+1} + \gamma V^{\Pi}(s_{t+1}) + \gamma^2 V^{\Pi}(s_{t+2}) + \dots / s_t = s\} \quad (1)$$

здесь s_t - состояние системы агент-окружающая среда в момент времени t ; s – конкретная реализация этого состояния; Π (политика) - распределение вероятностей выбранной последовательности действий; r_t - награда, которую получает агент из окружающей среды в момент времени t ; $V^{\Pi}(s)$ – значимость состояния, или средний прогноз суммарной награды, которую получит агент, если в момент времени t он находится в состоянии s , и в выборе действий придерживается политики Π ; E_{Π} - усреднение в рамках выбранной политики Π ; $\gamma < 1$. – коэффициент затухания, отражающий свойство дисконта - ослабления представления о будущей награде. Такое ослабление отражает реальный факт, что до использования отсроченной награды, агент должен израсходовать заметную часть энергии, которую он от награды получит.

В этих терминах основной постулат теории Сатона и Барто состоит в том, что автономное устройство стремится выбрать такую политику Π , при которой $V^{\Pi}(s)$ будет максимальным.

Для этого нужно выполнить выбранную последовательность действий от момента t до конца исследуемого им эпизода, запомнить ее, запомнить полученную суммарную награду, повторить всю последовательность действий такое число раз, которое необходимо, чтобы проверить все возможные политики Π , и все соответствующие каждой политике варианты, потом выбрать из них наилучшую Π , запомнить ее, и придерживаться в следующий раз, попав в то же состояние s .

Если мы хотим применить теорию обучения по подкреплению к поведению животных, то представляется нереальным, чтобы реальное животное в самом деле следовало такой тактике для нахождения оптимального поведения.

Теория обучения по подкреплению предлагает вариант: обучение по текущей ошибке прогноза - “temporal difference learning (TD learning)”.

Чтобы найти ошибку прогноза, перепишем уравнение (1)

$$V^{\Pi}(s) = E_{\Pi} \{r_{t+1} + \gamma (V^{\Pi}(s_{t+1}) - V^{\Pi}(s_t)) + V^{\Pi}(s_t) / s_t = s\} = \\ = E_{\Pi} \{r_{t+1} + \gamma V^{\Pi}(s_{t+1}) - \gamma V^{\Pi}(s_t) + V^{\Pi}(s_t) / s_t = s\} \quad (2)$$

Т.е. прогноз награды в момент времени t складывается из награды, которая будет получена при совершении следующего шага и прогноза, который будет в следующий момент.

Уравнение (2) дает возможность оптимизировать не всю последовательность действий, а каждый шаг по отдельности, что гораздо больше соответствует условиям поведенческих опытов.

Второе важное замечание касается способа решения уравнения (2). В ходе повседневного опыта агент фактически на каждом шаге корректирует свое поведение в направлении получения большей награды, что

эквивалентно использованию этого уравнения, как итеративного соотношения, дающего последовательные приближения V_0, V_1, V_2 ,

Выберем начальное приближение любым, например $V_0 = 0$, и будем на каждом шаге заменять старое значение $V(s_t)$ на новое

$$V(s_t) \leftarrow V(s_t) + \alpha [r_{t+1} + \gamma V(s_{t+1}) - V(s_t)] \quad (3)$$

здесь α - параметр; $r_{t+1} + \gamma V(s_{t+1}) - V(s_t)$ - называют ошибкой прогноза награды (temporal difference error);

Правильней называть эту величину текущей ошибкой прогноза награды.

Эту ошибку используют для оценки только что выбранного действия. Если ошибка положительна, делается вывод, что тенденция выбирать это действие должна быть на будущее усилена, если же ошибка отрицательна, - то ослаблена.

Как правило, решение сходится достаточно быстро, часто за несколько итераций. Но поскольку в этом случае уравнение (2) решается не на всем множестве возможных значений, есть опасность попасть в локальный максимум прогноза. Теория обучения по подкреплению в качестве выхода из этого положения указывает на «исследовательское поведение». Это означает, что, найдя итеративным способом достаточно хорошую политику Π , агент должен время от времени поступать не придерживаясь этой политики, а, выбирая свои действия случайным образом. Если существует лучшая политика Π , то при такой тактике агент с большой вероятностью «нащупает» ее. Исследовательским это поведение будет по отношению к отдаленному результату, к суммарной награде.

В применении к животному это означает, что оно время от времени будет действовать спонтанно, в противовес выученному поведению.

Одним из источников теории обучения по подкреплению явилось изучение поведения животных. Поэтому кажется естественным попытаться применить формализм этой теории к функционированию тех структур животного, которые осуществляют выбор действия, и в которых, - удивительно, но в тех же самых структурах, - находятся области связанные с ощущением удовольствия. Отделение управляющих областей мозга от областей непосредственно регистрирующих избыток или нехватку пищи, отвечает постулату теории Сатона и Барто, что управляющим сигналом является не сама награда, а ее прогноз.

В последнее время получено множество доказательств того, что ключевую роль в выборе действия играют нейроны стриатума (Redgrave, Prescott, Gurney, 1999). А опыты В. Шульца показали, что источники вещества, связанного с удовольствием - дофамина, выделяющегося в стриатуме - дофаминовые нейроны в VTA и черном веществе физически меняют спайковую активность в связи с ошибкой прогноза награды (Shultz, Dayan, Montague, 1997, Shultz, 1998).

Появление пачки импульсов свидетельствует о положительной ошибке: «ситуация лучше, чем ожидалось». Пауза соответствует отрицательной ошибке – «дело идет хуже, чем ожидалось». Если же активность остается прежней, значит, все идет как ожидалось.

Открытие Шульца послужило весомым аргументом в пользу применения теории обучения по подкреплению для построения моделей целенаправленного поведения животных (Pennartz, 1996, Suri, Bargas, Arbib, 2001, Montague, Hyman, Cohen 2004).

Мы применили модель обучения по ошибке прогноза к поведенческой задаче отсроченной награды.

ЧАСТЬ 2. ТРАКТОВКА РЕЗУЛЬТАТОВ ПОВЕДЕНЧЕСКИХ ЭКСПЕРИМЕНТОВ В РАМКАХ МОДЕЛИ ОБУЧЕНИЯ ПО ОШИБКЕ ПРОГНОЗА НАГРАДЫ.

Г.Х. Мержанова, экспериментально изучает поведение животного в условиях задержанного подкрепления (Мержанова и др., 2003, Merzhanova, 2003). По условию опыта, кошка после включения сигнала должна нажать лапой на педаль. Если она нажмет через 1 секунду после включения светового сигнала, то получит сухарную смесь; если через 10 секунд – мясной корм. По типу поведения кошки четко разделились на группы.

В первую вошли те, которые обучились, и были в состоянии выдержать 10 секунд, и получить ценный мясной корм – их называют «самоконтрольными». Во вторую те, которые не сумели научиться контролировать свое поведение, и нажимали педаль раньше – таких животных называют «импульсивными». Кроме того, была выделена третья группа – эти кошки умели вести себя самоконтрольно, но достаточно часто срывались и нажимали педаль раньше.

По нашему предположению предпочтение близкой награды «импульсивными» животными связано с нехваткой дофамина, проявляемой в форме слишком сильного дисконта. Животное плохо прогнозирует отдаленную награду; это означает, что для него она кажется менее ценной. Формально это учитывается параметром γ .

МОДЕЛЬ

Мы считали каждый отдельный трайл вектором, размерность которого равна числу секунд эксперимента. После включения светового сигнала первая секунда соответствовала задержке. «Импульсивные кошки» в течение следующих двух секунд хватали пищу. Потом следовали 7 секунд, когда педаль была отключена, и в течение последних двух секунд разрешали себе брать корм «самоконтрольные животные».

Если кошка нажимала на педаль в течение 2-ой и 3-ей секунды, она получала награду R равную 1. Если нажимала на 10-ой, или 11-ой секунде, то R равнялось 2. Многие кошки во время периода самоконтрольного сдерживания совершают небольшие движения, слегка трогая педаль. За такие движения кошка получала наказание R равное -0.5 .

Предполагалось, что животное на каждой i -ой секунде совершает выбор: продолжить ожидание, или совершить немедленное действие. Совершение действия вело к получению соответствующей награды и завершению данного трайла. Чтобы смоделировать этот выбор на i -ой секунде каждого трайла производилась коррекция прогноза награды V_i и двух вариантов прогноза награды при условии выбранного варианта поведения Q_i^0 - при условии совершения действия и Q_i^1 - при условии ожидания

$$\begin{aligned} V_i &\leftarrow V_i + \alpha [R_i + \gamma V_{i+1} - V_i] \\ Q_i^0 &\leftarrow Q_i^0 + \alpha [\gamma V_{i+1} - Q_i^0] \\ Q_i^1 &\leftarrow Q_i^1 + \alpha [\gamma V_{i+1} - Q_i^1] \end{aligned} \quad (4)$$

Мы приняли, что в каждый i -тый момент животное выбирает: продолжать ожидание, или не продолжать, под влиянием двух факторов. Во-первых, в зависимости от переменной P , экспоненциально зависящей от величины Q_i^1 (Sutton Barto, 1998)

$$P_i = C_0 \exp(\eta Q_i^1) / (C_0 \exp(\eta Q_i^1) + \exp(\eta Q_i^0)) \quad (5)$$

здесь параметр η – аналог температуры Больцмановского распределения – отражает «жадность» животного; параметр C_0 возможно, связан с вниманием. Использование зависимости (5) приводит к быстрому предпочтению действий с большим Q_i^1 .

Второй фактор, от которого зависит выбор вариантов поведения – некоторая пороговая величина ϕ , значения которой выбираются случайно и равновероятно. Величина ϕ отражает исследовательский компонент поведения. Мы принимали, что животное продолжает ожидание на i -том шаге, если $P_i > \phi$.

РЕЗУЛЬТАТЫ

Моделирование показало, что при $\gamma < 0.7$ кошка ведет себя «импульсивно», и только более чем 20-кратное увеличение задержанной награды R по отношению к немедленной награде, может превратить ее в «самоконтрольную».

При $\gamma = 0.99$ кошка ведет себя самоконтрольно.

Кошки относятся к промежуточной третьей группе при γ порядка 0.9, при этом параметр η должен быть не меньше 3, но и тогда они достаточно часто допускают срывы.

В связи с этим стоит вспомнить что условием нахождения оптимальной политики Π является использование исследовательского поведения. В применении к задаче выбора из двух вариантов, это означает, что время от времени животное должно поступать наоборот по отношению к выученному поведению. Если строить предположения о том, как это устроено, то в основе поведения должны лежать два разных механизма: один, обеспечивающий выучивание оптимальной политики и следование ей; другой, генерирующий исследовательское поведение. На разных стадиях процесса обучения соотношение этих тенденций может быть разным. И также оно может быть разным у разных животных.

Отсюда вытекает возможная трактовка экспериментальных данных Г.Х.Мержановой. Достаточное или недостаточное количество дофамина отвечает за разделение кошек на две группы по параметру γ , а взаимодействие величин P и ϕ дает разделение кошек на «самоконтрольных» и кошек третьей группы, которых можно называть спонтанными. Это разделение определяется относительным соотношением «логики» и «спонтанности»

Такое разделение может иметь отношение к противопоставлению субъектов по признаку творчества и рутинности.

Литература

1. Дунин-Барковский В.Л. Информационные процессы в нейронных структурах. М., «Наука», 1978.

2. Мак-Каллок У.С., Питтс У. Логическое исчисление идей, относящихся к нервной активности. - в кн. Автоматы. Под ред. К.Э. Шеннона и Дж. Мак-Карти. М., ИЛ, 1956.
3. Марр Д. Зрение. Информационный подход к изучению представления и обработки зрительных образов. М., «Радио и связь», 1987
4. Мержанова Г.Х., Долбакян Э.Е., Хохлова В.Н. Межнейронные фронто-гиппокампальные взаимодействия у кошек, обученных выбору качества подкрепления. Ж. Высш. Нервн. Деят. И.П. Павлова, 2003,53(3):290-8.
5. Цетлин М.Л. Исследования по теории автоматов и моделированию биологических систем. М., «Наука», 1969.
6. Bellman R.E. (1957a) . Dynamic Programming. Princeton University Press, Princeton, NJ.
7. Bellman R.E. (1957b) . A Markov decision process/ Jurnal of Mathematical Mech., 6:679-684.
8. Logue A.W. Reseach on self-control: An integrative framework// Behavioral and brain sciences, 1988, II, p. 665.
9. Merzhanova G.Kh, Local and distributed neural network and individuality. Neurosci.Behav.Physiol.,2003, Feb;33(2):163-70.
10. Montague P.R., Hyman S.E., Cohen J.D. Computational role for dopamine in behavioral control. Nature, 2004, V.431, 760-767.
11. Pennartz C.M. The ascending neuromodulatory systems in learning by reinforcement: comparing computational conjectures with experimental findings. Brain Res. Rev. 1996, 21, 219-245.
12. Redgrave P., Prescott T.J., Gurney K., The basal ganglia: a vertebrate solution to the selection problem ? Neuroscience, 1999, V. 89, N 4, 1009-1023
13. Saton R., Barto A. Reinforcement learning – MIT Press, 1998
14. Shultz W., Predictive reward signal of dopamine neurons. J. Neurophysiol, 1998, 80, 1-27
15. Shultz W., Dayan P., Montague P.R. A neural substrate of prediction and reward. Science, 1997, v275, 1593-1599
16. Suri R.E., Bargas J., Arbib M.A. Modelling function of striatal dopamine modulation in learning and planning. Neuroscience, 2001, V. 103, №1, pp65-85.

TD-LEARNING – THE THIRD BASE OF COMPUTATIONAL NEUROSCIENCE. AN EXAMPLE OF IT'S APPLICATION FOR DOPAMINE IN SEL-CONTROL PROBLEM

Smirnitskaya I.A.

Institute for Higher Nervous Activities and Neuropsychology RAS

The authors formulate, that TD-learning is the third, after formal neuron of McCulloch-Pitts and Hebb synaps, framework of neural nets theory. In present paper, TD-learning is used for modelling self-control versus impulsiveness behavior in cats: for verifying the hypothesis that dopamine level in striatum plays the role of discount γ . Computer simulation had confirm this assumption: - small value for γ corresponding low dopamine level leads to impulsive behavior; large value of γ corresponding high dopamine level leads to self-control.

ПОНЯТИЕ «ОРГАНИЗАЦИЯ» И ВОЗНИКНОВЕНИЕ ЯЗЫКА

Парменов В.К.

ОАО «Протон – Пермские моторы»

Истоки общесистемного понятия «организация» тесно связаны с понятием «душа» и возникновением языка и мышления – одной из семи «Мировых загадок». Научный и практический интерес представляет установление непосредственной связи понятия «организация» и с остальными шестью «мировыми загадками»¹. отождествление понятий «душа» и «организация», понимание их взаимосвязи было всегда, и лишь по-разному выражалось в зависимости от уровня развития знания или преобладающей терминологии. Как бы взглядов на происхождение языка не придерживался исследователь, от догматов христианского учения о сотворении путем чуда до поисков причины в «импульсе», сообщенном сверхъестественной силой

¹ 1-сущность материи и силы, 2-происхождение движения, 3-происхождение жизни, 4-целесообразность природы, 5-возникновение ощущения и сознания, 6-возникновение мышления и речи, 7-свобода воли, о которых физиолог Дюбуа-Раймон в речи «О границах познания природы» (1872) говорил *ignorabimus* (не узнаем), на что немецкий естествоиспытатель Эрнст Геккель в книге «Мировые загадки» возразил - *Ignoramus* (не знаем)

живым формам, или эволюционной теории Ч.Дарвина, он неизбежно сталкивается с этой загадкой. А поскольку границы человеческого мышления определяются языком, теория познания должна опираться на изучение языка². Уже первые письменные источники содержат понятие «душа», а поскольку письмо при своем появлении зафиксировало то, что в мифах существовало до этого тысячелетиями, понятие «душа», с полным основанием можно отнести к «первопонятиям». Существует более десятка теорий происхождения языка, каждая из которых имеет своих приверженцев, полагающих, что существующие данные свидетельствуют о справедливости избранной ими гипотезы и несостоятельности остальных. Однако, ни богословская теория “божественного” происхождения человеческой речи, ни учение просветителей о происхождении языка путем “общественного договора”, ни теория Кондильяка, выводящая язык из криков животных, не приблизили к решению этой мировой загадки.

В таблице отражены хронология появления первых орудий труда, первых письменных источников и религиозных учений. Интервалы времени на пути к языку, если только он изначально не был заложен в сознании или подсознании человека, различны у разных авторов. Современные данные лингвистов и нейролингвистов, археологов и антропологов проводят временные границы формирования и дивергенции языка между средним и поздним палеолитом, между неандертальцами и кроманьонцами. Бунак В.В. (1951) и Кочеткова В.И.(1973) обнаружили морфологические следы асимметрии функций двух полушарий мозга на черепках ископаемых людей в находках периода 50 - 40 тыс. лет и связали их с развитием речевых зон Брока и Вернике. (Шапиро,2000, с.67) Моррис Сводеш в середине 20 века разработал метод глоттохронологии, позволивший сделать предположение о времени распада языкового единства в 15-12 тысячелетиях. Сильным аргументом в пользу дивергенции праязыков явился вывод генетиков (Л.Л.Кавелли-Скорца), показавших в ряде исследований корреляцию лингвистических и генетических эволюций, где при разной скорости дифференциации генов и языков сохраняется пропорциональность. Значительно приближают время происхождения языка исследования Филиппа Либермана и Эдмунда Крелинга (1971), показавших при реконструкции надгортанных структур черепов неандертальцев, их неспособность произносить все гласные современного английского языка. Несмотря на то, что образования второй, речевой сигнальной системы относят к стадии питекантропов, периферические органы речи по данным археологов сформировались у кроманьонца. Из таблицы видно, что два с половиной миллиона лет труда потребовалось для того, чтобы (в соответствии с трудовой теорией происхождения языка) сформировались органы речи, и появилась потенциальная возможность рождения языка, еще спустя 36 тыс. лет возникла письменность. Весьма вероятным представляется развитие языка по экспоненте.

Однако в появлении письменности от древнесемитского письма до письма древнеперсидского, как и в формировании религий, мы наблюдаем строгую цикличность.

Число лет до и после новой эры	Археологические находки, языки, письменность, религии
2000	Каодай, РЛжэдан, Сэйте-но (Япония). Каодай, Хоахао (Вьетнам). Церковь саентологии (США и еще 100 стран)
1900	Бахаизм (Иран). Чхондотё (Корея). Тэнри (Япония)
1500	Сикхизм
600	Ислам
500	Индуизм. Синтоизм
300	Письмо Майя. Христианство
Новая эра	
300	Арамейский язык. Древнеперсидское письмо
400	Классическое греческое письмо. Ветхий завет. Иудаизм
500	Латинское письмо. Трипитака – Буддизм. Веды. Даосизм. Джайнизм
600	Этрусский алфавит. Коран – Магометанство. Зороастризм. Конфуцианство
700	Греческое письмо. Все западные алфавиты. Упанишады
1000	Санскрит. Веды. Гаты – древнейшая часть Авесты
1100	Критское письмо. Китайское письмо
1300	Древнесемитское (финикийское) письмо
2000	Шумерское письмо. Клинопись Двуречья. Брахмы
2100	Список царей
2600	Псалмы. Вавилонские покаяния
2800	Древнее царство. Фараон. Чиновники-писцы
4000	Раннее царство. Древнеегипетское письмо.
12000	Кроманьонцы – появление периферических органов речи

² Разочарованный тем, что в фундаментальном труде его учителя И.Канта “Критика чистого разума”, нет столь важного понятия как язык, И.Г.Гердер в своей “Метакритике” впервые высказал это положение.

40000	Неандертальцы – образование второй, речевой сигнальной системы
50000	Питекантропы, синантропы и другие
2000000	Архантропы оловянного (Танзании) Homo habilis
2600000	Возраст древнейших каменных орудий

Иоганн Гердер в “Трактате о происхождении языка”, удостоенном премии Королевской Академии наук за 1770 год впервые соединяет в происхождении языка его естественные, животные истоки с мышлением и понятием организации. Сомневаясь в том, что физиология сумеет когда-нибудь создать учение о законах душевной жизни³, он считает, что анализ строения нервной системы если и прольет свет на это явление, то раздробит его при этом на ничего не говорящие части. Между тем своей работой он закладывает основы такого учения⁴. Его исторический анализ зарождения и развития языка создает последовательную, логически стройную картину от природных истоков⁵ до сознательной внутренней и внешней организации человека, взаимного формирующего влияния языка и мышления. Он показывает как за счет потери человеком инстинкта, этого результата слепого предопределения, *вытекающего из самой организации* чувств и сферы представлений животного, человек приобретал большую ясность восприятия.

Принципиальным является его вывод для последующего определения понятия “организация”: “Назовите, как хотите всю эту расстановку сил человека: разумом, рассудком, самосознанием и т.п. *Это единая, целостная организация всех человеческих сил вместе взятых*, единое хозяйство его чувственной и познающей, его познающей и волевой природы... *И различие здесь не в степени или в количестве* приложенной силы, а в совершенно особом, *своеобразном направлении и развитии* всех сил» (Гердер 1959.)

Остается, однако, важный круг вопросов, поднятый Гердером, ответ на которые, пытаются найти до сих пор: “Откуда же душа брала для них слова- приметы, которыми она могла бы их назвать? В какой связи находятся зрение, слух, цвет и слово, запах и звук? Да и кто же может сравнить звук и цвет, явление и чувство?”⁶. Первым сделал такое сравнение Герман Гессе в “Игре в бисер” и предложил обучать этому искусству элиту. Вторым, кто еще в 1955 году предсказал технологию виртуальной реальности и обратил внимание на возможности эффективного использования совокупной деятельности различных сенсорных каналов человека, был Станислав Лем в “Магеллановом облаке”, там же предупредивший о грозящей при этом опасности отчуждения сознания. Сегодня методы виртуальной реальности становятся распространенной технологией. В последнее пятилетие формируется междисциплинарное научно-технологическое направление под условным названием “Сенсорная биология и новые сенсорные технологии” с задачей “сравнить возможности сенсорных систем организмов (органов чувств) и биосенсоров, установить реально ли применение современных знаний о принципах организации и функционирования органов чувств для создания новых сенсорных технологий”⁷ (Варфоломеев и др., с. 99-108).

Исходя из положения Сеченова “мысль есть рефлекс, прерванный на двух третях”, А.А.Богданов уточняет его терминологию. Он расширяет понятие “мысль” до всякого представления и ощущения, всякого “внутреннего” психического акта, определяет какие именно рефлекс “прерываются на двух третях”, трактует недостающую треть мысли как акт речи - “мышечное сокращение, сопровождаемое звуком речи” (Бо-

³ Проблема здесь по мнению А.А.Богданова заключается в том, что “связь физического ряда и связь психического ряда принципиально различны, не сводимы одна на другую, и не допускают объединения в какой-нибудь третьей, высшей закономерности” (Эмпириомонизм. кн. I, с.21). Решение задачи устранения дуализма познания он ищет в способе определения связи, которая могла бы быть установлена строго объективно между рядом физического и рядом психического опыта, а также между категориями причинности и параллелизма, используя методы подбора, подстановки и другие.

⁴ В 1913 г. на собрании Общества московского научного института И.П.Павлов говорил: «...сознание представляется мне нервной деятельностью определенного участка больших полушарий... светлое пятно, окруженное всем остальным пространством полушарий более или менее значительной тенью». Сегодня с помощью методов компьютерного анализа электрической активности мозга, позитронно-эмиссионной томографии, функционального радиоманнитного резонанса, компьютерной термоэнцелоскопии - “светлое пятно”, которое мысленно прозрел Павлов, можно увидеть собственными глазами в виде карт внутрикоркового взаимодействия или корковых связей в диапазоне частот бета-ритмов при решении пространственных или вербальных задач. (Симонов П.В, с.785-789.)

⁵ “Последним материнским прикосновением творческой руки природы была заповедь всем ее созданиям: “чувствовать не только для себя, но в звуках выражать свое чувство!”... Эти вздохи, эти звуки - не что иное, как язык; существует язык чувств, как закон самой природы” (Гердер 1959, с.134)

⁶ Там же, с.150-151.

⁷ Уже сегодня сенсорные технологии обеспечивают определение и анализ сигналов, далеко выходящий за рамки возможности органов чувств человека. Потенциально достижимыми считают создание механо- и терморесепторов при визуализации изображений в различных электромагнитных полях, сопряженных с ресепторами биосенсоров, способных детектировать внешние химические сигналы и проводить мониторинг внутреннего метаболического состояния.

гданов, 1910, с.7). Основываясь на достижениях сравнительной филологии, “сравнительной антропологии”, “сравнительной лингвистики” Гумбольдта, “лингвистической археологии” Гейгера, работах И.Гердера, Макса Мюллера, Людвига Нуаре, он впервые подошел к происхождению речи с точки зрения трудовых рефлексов, связал их с двигательными реакциями организма, со звуком и совместным сокращением мышц дыхательных, гортанных, глоточных, ротовых. “Эта часть реакции, благодаря свойствам звука, доступна членам коллектива и тогда, когда трудовой акт выполняется вне поля их зрения; она, следовательно, может играть роль “сообщения” об этом акте, и вообще его “обозначения” (Богданов, 1925, с. 71).

Так, замещение действия социально выработанным символом, а его изображения или звукового оформления - мысленным представлением двигательного образа, привели к мышлению. А.А.Богданов обратил внимание на то, что в греческом языке “речь” и “мышление” обозначается одним словом “λογος”, что для Платона мышление есть “разговор, который душа ведет с самой собою по поводу различных предметов ее созерцания”. В одной из первых своих работ (Богданов, 1901) он высказывает точку зрения на психику вообще, как систему двигательных реакций организма; на “сознание” - как на область комбинирования и отбора неполных реакций; на трудовой процесс, как на цепь сознательно-целесообразных реакций. Первые опорные пункты - принцип голосовой координации действий и принцип “самообозначения” рефлексов действия междометиями (принцип Нуаре), он дополняет организационно - биологической точкой зрения. Более того, - рассматривает реакции восприятия и воздействия, входящие в содержание понятия не в их индивидуальных формах, а в социально-сложившихся, исторически выработанных жизнью коллективов, проявляя, тем самым, *организационную функцию языка в формировании не только мышления, но и трудового процесса в целом.*

В “Логико-философском трактате” (1921) Людвиг Витгенштейн считает язык состоящим из имен, *организованных* в предложения определенным способом, и именно *способы их организации определяют границы смысла, отражающего определенную организацию элементов мира.* Джон Дьюи в своей прагматической теории исследования утверждает, что люди отличаются от других чувствующих организмов тем, что они являются тонкими лингвистическими существами, что усвоенные языковые навыки, позволяют отличать безопасные ситуации от ситуаций, таящих угрозу. Ситуационный опыт возникает в природе как *про-дукт сложных взаимодействий организмов с их окружением.*

Ноам Хомский в 50-е годы разработал теорию «глубинных структур» синтаксиса, общих для всех языков, вывел законы преобразований, по которым эти глубинные структуры отбирают и переводят лексические единицы в идиоматический разговорный язык. Американский исследователь языка индейцев Бенджамен Уорф установил, что основа языковой системы любого языка не просто инструмент для воспроизведения мыслей. Напротив, грамматика сама формирует мысль, является программой и руководством мыслительной деятельности. Он считает, что мир предстает перед нами как калейдоскопический поток впечатлений, который должен быть *организован* нашим сознанием, а это значит в основном - языковой системой, хранящейся в нашем сознании. Физиолог Роджер Браун в 1970 году выдвинул предположение, что в первых высказываниях ребенка проявляется его сенсомоторный интеллект, и что характерные черты этих высказываний являются общими для всех людей. Подобный взгляд на существование неких эволюционных стадий в процессе умственного развития человека согласуется с биогенетическим законом Эрнста Геккеля, сформулированного в 1866 году: «Онтогенез повторяет филогенез» - то есть история каждого организма повторяет историю вида. Положение Брауна подразумевает, что ребенок на этой стадии развития в соответствии с доктриной Геккеля, проходит последний эволюционный этап, предшествующий стадии полного расцвета умственных способностей человека. Такое предположение вполне вероятно; известно, что формирование мозга заканчивается лишь через несколько недель после рождения ребенка. Браун полагал, что в простейших двухсловных предложениях ребенка на первом уровне развития находит отражение нечто большее, чем просто сенсомоторный интеллект, что уже в них проявляются зачатки врожденных лингвистических способностей человека (Линден, 1981, с.44).

Невролог Норман Гешвинд (доклад на симпозиуме «Общение животных и язык человека: качественные различия или эволюционная преемственность?») считает язык сложным явлением, контролируемым разными участками мозга. В мозге человека левое полушарие доминирует при выполнении обеих функций языка: высказывании и восприятии. Среди исследованных им людей с неполноценными языковыми способностями (афазия) у 97% было повреждено левое полушарие. Было установлено, что сходные повреждения определенных частей мозга влекут за собой неспособность пострадавшего, как к последовательному совершению некоторых действий, так и к соблюдению правильного порядка слов в предложении. На этом основании антрополог Гордон Хьюз сделал предположение о существовании взаимосвязи между языком и технологией (в статье «Однозначное формулирование взаимосвязи между использованием орудий труда, их изготовлением и возникновением языка»). Исследования антропологами языков жестов, которыми пользовались североамериканские индейцы, аборигены Австралии и некоторые другие народы, показали, что в этих языках используются те же синтаксические формы, что и других человеческих языках. Они пришли к выводу, что глубинные структуры современного языка тесно связаны с аналогичными структурами естественных языков

жестов. Эта идея перекликается с представлением Хьюза о том, что «глубинные структуры» представляют собой грамматику, которая при своем первом появлении позволила людям *организовывать последовательности двигательных актов*. В рамках этой же схемы «порождающая грамматика» Хомского в действительности может отражать переход в *процессе эволюции из руслу жестикуляторного языка к речи*. Можно также предположить, что в первичных структурах, кроме грамматики в ее расширительном толковании, генетически закреплены первичные организационные механизмы: преклонение перед божествами, уважение к старшим, нравственные принципы, моральные установки и т.п.

Известный лингвист В.А. Звягинцев считал, что язык выполняет две взаимосвязанные, но не идентичные функции: мышление и коммуникация. Из всего сказанного следует это определение дополнить третьей функцией - *организация*.

Просматривается интересная аналогия в том положении, которое заняли в научном сообществе Всеобщая организационная наука (тектология) А.А. Богданова и теория Ноама Хомского. Принципы и закономерности организации, выделенные в тектологии Богданова, широко использовались в методологии системных научных направлений в 50-80 годы прошлого века. Точно также фундаментальные положения, заложенные в основу модели теории Хомского, применялись во всем мире лингвистами и психологами, нейропсихологами и антропологами при разработке методик исследования. И наука Богданова, и теория Хомского (которая при своем появлении была названа революцией Хомского) в полной мере отвечают всем критериям концепции научных революций Томаса Куна. Но ни та, ни другая не стали новой парадигмой теории познания в XX веке по различным причинам. Сегодня эти причины и их инерция в основном изжиты, а бурные темпы развития системных научных направлений в последнее десятилетие заметно поутихли. Так что есть достаточно оснований надеяться, что в веке наступившем *организационная наука станет полноценной теорией организации и займет подобающее место в теории познания*.

Литература

- Богданов А.А. Познание с исторической точки зрения. СПб. 1901.
 Богданов А.А. Падение великого фетишизма. СПб., Дороватовский и Чарушников, 1910.
 Богданов А.А. Учение о рефlekсах и загадки первобытного мышления. В сб. Комакадемии, 1925.
 Варфоломеев С.Д., Евдокимов Ю.М., Островский М.А. Сенсорная биология, сенсорные технологии и создание новых органов чувств человека. М., Вестник РАН, том 70, № 2, 2000.
 Симонов П.В. Светлое пятно сознания. М., Вестник РАН, том. 69, № 9 – 2000.
 Шапиро Д.И. Человек и виртуальный мир. М., «Эдиториал УРСС» - 2000.
 Гердер И.Г. Избранные сочинения. М., Л., 1959.
 Гердер И.Г. Идеи к философии истории человечества. М., «Наука», 1977.
 Кун Т. Структура научных революций. М., «АСТ», 2001.
 Линден Ю. Обезьяны, человек, язык. М., «Мир», 1981.

3-Д МОДЕЛИРОВАНИЕ В ИСТОРИИ КУЛЬТУРЫ: О СОЗДАНИИ ВИРТУАЛЬНОЙ МОДЕЛИ АНТОНИЕВА-СИЙСКОГО МОНАСТЫРЯ

Трубочкин Д. В.

Российский комитет программы ЮНЕСКО «Информация для всех»

В своем докладе я расскажу о нашей работе над 3-д моделью Антониева-Сийского монастыря – памятника культуры и русской духовности. Этой работе предшествовал этап международного сотрудничества в сфере виртуального моделирования исторических зданий, с которого уместно начать рассказ.

В 2001 г. было объявлено о сотрудничестве Российской академии театрального искусства–ГИТИС с европейским консорциумом ТЕАТРОН (www.theatron.org); был сформулирован проект «История театра на компьютере», цель его – формирование виртуального музея исторических театров и использование его в преподавании искусствоведческих дисциплин студентам. 3д-лаборатория группы «Театрон» находится в Уорикском университете (Великобритания). В 1990-е годы эта группа одной из первых в мире занялась применением трехмерного моделирования в истории театральной архитектуры. В основе проекта виртуального музея ТЕАТРОН лежит намерение предоставить посетителю ресурса возможность самому избирать последовательность усвоения предлагаемой электронной информации, исходя из его собственных целей и предпочтений. Задуманный ими интернет-музей онлайн будет состоять из виртуальных моделей 20 важнейших европейских театров от античности до XX века.

В 2003 году к сотрудничеству присоединился Российский комитет программы ЮНЕСКО «Информация для всех» с намерением расширить деятельность по 3-д моделированию и в итоге разместить в российском интернете виртуальные модели зданий, имеющих культурную и историческую ценность.

Изучение опыта применения 3-д технологий в истории культуры показало ценность 3-д моделей для образования, но было сопряжено с дискуссиями об их ценности в исследовательской науке. В этой связи были сформулированы два различных подхода:

1) трехмерная реконструкция — это усовершенствованная форма иллюстрации, а функция иллюстрации — наглядно представить результат, достигнутый ученым ранее и независимо от художника; признав это, мы выносим трехмерные модели театров за рамки научного исследования и отводим им роль необязательного дополнения к нему, вступающего в действие только для того, чтобы придать результатам завершившегося уже исследования интересную и увлекательную форму;

2) трехмерная реконструкция – не просто иллюстрация, но новый и эффективный инструмент научного исследования; это значит, что с помощью трехмерной реконструкции становится возможным по-новому взаимодействовать с уже известным объектом исследования и достичь результатов, не достижимых никаким иным образом; при этом никакая 3-д модель, разумеется, не может заменять собою настоящий объект исследования, так как по определению принадлежит исключительно к сфере исследовательского инструментария.

Очевидно, что эти два подхода обозначают, так сказать, две степени серьезности в отношении к виртуальной реконструкции со стороны ученых. Среди приверженцев первого совета, наверное, найдутся те, кто будет воспринимать 3-д модели как своего рода интеллектуальную забаву, разновидность познавательной игры, умело сочетающей полезное с приятным. Признавая их правоту, я, тем не менее, отношу себя к тем, кто склонен воспринимать виртуальное моделирование прежде всего как научный метод. Приведу один подтверждающий пример.

В настоящее время проф. Р. Бичем, руководитель консорциума ТЕАТРОН, и археолог проф. Д. Пэкер (США) возглавляют проект, получивший название «Pompeii Project», «Помпеев проект». Цель его – комплексное изучение первого римского каменного театра – Театра Помпея: его строительство в 55 г. до н. э. и было подлинным историческим «Помпеевым проектом». Результатом современного «Помпеева проекта» будет создание специального электронного ресурса по этому театру, который будет включать в себя разнообразные данные: от древнейших литературных свидетельств до результатов современных акустических испытаний. Важнейшая особенность «Помпеева проекта» в сочетании традиционной археологии и археологии «компьютерной», основанной на 3-д моделировании. Что такое «компьютерная археология», станет ясно ниже после краткого исторического экскурса.

Театр Помпея вмещал до 40000 человек; диаметр внешней окружности аудитории достигал 150 м, а длина активного пространства просцениума — 90 м. Помпей построил не просто театр, но колоссальный религиозно-зрелищный комплекс. Огромный «портик за сценой» четырехугольной формы (135x180 м) одной своей колоннадой прилегал к обратной стороне сценического фасада и обрамлял собою сады с фонтанами; в портике были выставлены знаменитые на весь средиземноморский мир произведения искусства, а в стороне, противоположной зданию театра, размещалась *Curia Pompeia* – дом заседаний сената. Храм Венеры Победительницы был вознесен на самую вершину зрительской аудитории, и пик его крыши и достигал 43 метров над поверхностью оркестры, почти на 9 метров превышая крышу гигантского трехуровневого сценического фасада: выше этого храма в Риме не было ни одного здания вплоть до самого заката Римской империи.

В Средние века и позднее жители Рима использовали структуры Театра Помпея для постройки новых зданий, и сегодня на всей территории то здесь, то там встречаются элементы Театра Помпея, встроенные в потолки и стены. Значительная часть древних сооружений теперь скрыта под землей, а на поверхности видны лишь призрачные очертания, сохраненные в контурах современных зданий. Легко представить себе, что полномасштабные раскопки театра Помпея сегодня невозможны из-за плотной застройки и густонаселенности этих районов. Вот здесь и приходит на помощь «компьютерная археология».

Суть ее в том, чтобы в процессе накопления разрозненных археологических и литературных сведений о Театре Помпея постоянно модифицировать трехмерную модель этого театра, учитывая мельчайшие детали и приближая ее к наибольшей точности. Легко редактируемая и безопасная для экспериментов трехмерная модель выступает здесь как наиболее совершенное средство аккумулировать непрерывно поступающую разрозненную информацию и, так сказать, подводить под общий знаменатель разнородные данные посредством их визуализации. Эта трехмерная модель, наложенная на современные фотографии местности, позволит выявить наиболее желательные точки для т. н. «замочных скважин», через которые археологи смогут замерять и изучать объекты с минимальным вторжением в существующие структуры. Поэтому «компьютерную археологию» еще называют «археологией в замочную скважину».

В этом постоянном взаимообмене информацией между вещественными данными и интегрирующей их виртуальной моделью как раз и заключается смысл «компьютерной археологии», которая позволяет бережно обходиться с современными зданиями, достигая при этом хороших исследовательских результатов. Такой взаимообмен информацией и, как следствие его, непрерывное уточнение и редактирование виртуальной модели служит в то же время надежной гарантией тому, что создание красивых картинок на компьютере ни

Но, быть может, главное, чем был знаменит монастырь, – это своей библиотекой, крупнейшей на русском Севере; например, в собрании библиотеки хранился уникальный экземпляр рукописного Сийского Евангелия XIV века. Также в собрании монастыря – уникальные примеры иконописи, так как здесь была своя иконописная школа. Все это в совокупности объясняет название проекта: монастырь – хранитель подлинного сокровища русской духовности в течение многих веков.



Мы остановились на 3-мерной модели как интегральном ресурсе не случайно. Известно, насколько велико значение зримого пространственного образа русских церквей для русской национальной культуры: сама их архитектура призвана нести в мир дух православия. Наша современность очень чувствительна к зримым образам, и потому 3-мерные изображения смогут выполнить не только исследовательскую, но и просветительскую миссию в России и во всем мире. Таким образом естественно формулируется проект соз-

дания 3-мерного путеводителя по Антониеву-Сийскому монастырю, который будет выпущен на нескольких языках. Этот путеводитель расскажет о стадиях изменения архитектурного облика монастыря, об истории и предназначении каждого здания. В путеводитель будут помещены модели объектов, находящихся сейчас в разрушенном или ветхом состоянии. В него можно будет включить трехмерные изображения церковной утвари с рассказом о том, какой смысл несут православные символы и как они используются в богослужении. Таким образом, наш ресурс сможет выполнять роль введения в православную культуру. С путеводителем легко можно связать и все цифровые копии книг: например, войдя в помещение библиотеки в келейном корпусе, мы получим список электронных изображений и сможем загрузить любое из них, а почта – дальше двигаться по монастырю.

Почему мы предпочитаем 3-мерное изображение панорамным фотографиям, широко применяющимся в последнее время для оцифровки памятников культуры? – Мы полагаем, что свободное взаимодействие с трехмерными объектами, то есть свобода перемещения в виртуальном пространстве будет гораздо более увлекательным путем познания монастыря, чем обзор с нескольких фиксированных точек.

Мы хотим опубликовать виртуальный путеводитель в «облегченном» виде на сайте монастыря (в формате VRML) и издать более весомую версию на диске. Работа над VRML-моделью неожиданно стала поводом для широкой постановки вопроса об эстетических качествах современных трехмерных моделей памятников культуры.

До недавнего времени единственным критерием оценки моделей была точность очертаний и размеров, и перевод объекта в виртуальное пространство сам по себе считался достижением, сколь бы примитивны ни были внешние качества получающейся модели. Трехмерное изображение мыслилось как продукт технологического копирования, а не как произведение искусства: в результате современный интернет переполнился 3-д объектами, выполненными в абстрактной технократической эстетике с искусственными цветом и светотенью, ненатуральными поверхностями и совершенным отсутствием выражаемого смысла.

Авторы проекта иначе поняли свою задачу. Действующий монастырь – это памятник духовности, и именно этот смысл надо выразить при воспроизведении его на компьютере; отсюда ясно, что моделирование Антониева-Сийского монастыря не есть абстрактное копирование некоторого набора точек, линий и плоскостей – но создание целостного художественного произведения, чьи эстетические свойства должны соответствовать богатейшему выразительному языку оригинала. Именно поэтому в проекте «Явить миру Сийское сокровище...» над 3-д моделью работает художник в традиционном смысле слова – не «компьютерный дизайнер», но именно художник, который владеет 3-д моделированием как средством выразительности. Значительная часть его работы была посвящена тому, чтобы найти баланс между размером получающейся модели и ее художественными качествами. В результате общий объем модели Троицкого собора в формате VRML составил не более 1 Мб (рисунок 2).



Рис. 2: Первые результаты работы над 3-д моделью Антониева-Сийского монастыря (художник К. Еремин). Троицкий собор (на переднем плане) и Сергиевский надвратный храм. (Снимок с VRML-модели)

Итак, сегодня очевидно, что 3-д модели сочетают увлекательность игры и познавательность книги, и потому они могут привлечь внимание всех, кто интересуется русской культурой, в том числе туристов и спонсоров, которые сейчас весьма нужны для реставрационной деятельности. Хочется надеяться, что этот проект будет успешно доведен до конца, и тогда он может стать хорошим примером многостороннего сотрудничества разных организаций в деле формирования и сохранения национального цифрового наследия.

ON VR-MODELLING OF ANTONIEV-SIYSKY MONASTERY

Trubochkin D.

Russian Committee of the UNESCO Programme “Information for All”

This paper tells about the project “Disclose Siysky Treasure to the World” which involves the Russian Committee of the UNESCO Programme “Information for All”, Antoniev-Siysky Monastery, Arhangelskaya Regional State Library, Arhangelsky

Regional State Archive and Saint Petersburg National Library. The project is aimed at building a multilingual digital resource on the history and contemporary life of Antoniev-Siysky Monastery.

The Monastery was founded in XVI century in the North of Russia. It is famous for its unique stone architecture (survived from XVII century, partly destroyed), old library and workshop of cultic images. During XX century most of these treasures of national culture were removed from the Monastery to different locations all over Russia. Most of them face a threat of disappearance, and all of them can remain unknown to Russian people. Our project will found the collaboration of different organizations to collect all items together – in digital form – and provide a wide access to them through Internet and digital media (CD).

A virtual reality model of the Monastery will be put into the basis of the resource. All digital data – books, pictures, audio files, 3d-copies of the church items – will be associated with it. A precise virtual visualization of architecture and scenery will make possible a wide range of “virtual excursions” over the Monastery, and therefore add game-like opportunities to the resource. This will let considerably increase the number of potential users of the resource to include not only scholars, but also children, tourists etc.