

# Использование ассоциативных семантических сетей для классификации звукозаписей

Евгений Иванко, Денис Перевалов

## Введение

Считается, что человек отражает действительность не только и не столько непосредственно органами чувств и соответствующими образами. Большую роль в процессах восприятия и осознания человека играет вторая сигнальная система (согласно школе Павлова это система ориентировки человека на словесные сигналы-знаки, на основе которых возможно образование временных нервных связей). Замкнутая система взаимосвязанных знаков является важной абстрактной надстройкой над образным восприятием.

Исследование структуры второй сигнальной системы как основы для функционирования человеческого мышления и логического аппарата предвосхищалось Колмогоровым еще в середине XX века: “Быть может, для теоретической науки одно из самых интересных исследований (в котором могут естественно сочетаться идеи кибернетики, новый математический аппарат и современная логика) есть исследование процесса образования слов как второй сигнальной системы <...> Проследить этот механизм выкристаллизования слов как сигналов, несущих в себе комплекс образов, и создание на этой базе ранней логики – крайне благодарная область исследований” [2].

Следуя этим идеям, в предыдущих работах мы сформулировали и исследовали гипотезу о возможности обработки различных видов информации с помощью анализа знаков в некотором смысловом пространстве. Оказалось, что определенные задачи обработки текстов, рисунков и трехмерных объектов можно решать одними методами, переводя данные в универсальную структуру – ассоциативную семантическую сеть [1,3]. В данной работе мы развиваем идею использования метода для обработки звуковой информации.

Текст устроен следующим образом. В первой части статьи мы приводим основные положения метода, которые иллюстрируем на задаче обработки текста во второй части. Во второй части мы опишем построение семантической сети для текстов и приведем результаты сопоставления слов. Наконец, в третьей части изложена схема преобразования звука в семантическую сеть и пример результатов классификации коротких звукозаписей.

## Семантическая сеть

Любая информация, будь то текст, изображение или звук, может быть представлена как совокупность неких единиц – назовем их знаками. Выделять единицы, воспринимаемые человеком – задача сложная. Поэтому можно попытаться перенять сам принцип восприятия информации в виде множества связанных единиц. В этом случае мы моделируем двухфазное восприятие информации человеком. На вход поступает

цельный образ (что соответствует первой сигнальной системе человека), затем образ расчленяется и структурируется на множество связанных, осмысленных знаков (вторая сигнальная система). Это множество знаков со связями между ними мы будем называть семантической сетью.

Возьмем, к примеру, такой словарь, в котором каждому встречающемуся в его тексте слову есть определение. Примем за знак слово из этого словаря, причем каждое слово свяжем со всеми словами из соответствующей ему статьи. Такой словарь упрощенно отражает одну из самых крупных, развитых и важных семантических сетей человечества, а именно – сеть естественного языка.

В общем виде семантическую сеть можно представить как взвешенный граф, в вершинах которого находятся знаки, а ребра с весами отражают связи знаков между собой. В зависимости от того, что считать знаками и как определять меру близости между ними, можно получить различные семантические сети. Далее рассматриваются ассоциативные семантические сети. Для определения понятия близости двух знаков в таких сетях требуется для каждого из двух изучаемых знаков составить список часто встречающихся рядом с ним знаков. Таким образом, для каждого из двух знаков мы получим список «ассоциированных» с ним знаков. Сравнивая эти списки можно сделать количественные выводы о том, насколько близки исходные два знака.

Следуя вышесказанному можно предложить краткий алгоритм получения ассоциативной семантической сети по произвольному входному образу:

- 1) определить алфавит знаков,
- 2) выделить знаки алфавита из входного информационного образа,
- 3) определить понятие «рядом» – близость любых двух знаков, пользуясь их взаимным расположением в информационном образе,
- 4) пользуясь частотой появления рядом, для каждого знака составить списки близких к нему знаков (знаков с которыми он в некотором смысле ассоциирован),
- 5) сравнивая списки любых двух знаков, определить меру близости этих знаков в силу информационного образа,
- 6) по полученной информации построить взвешенный граф со знаками в качестве вершин и взвешенными ребрами в качестве связей между знаками.

Оказывается, что сравнивая между собой семантические сети группы входных образов, можно делать выводы о близости, в некотором смысле, самих образов. Примеры построения семантических сетей и их сопоставления приведены ниже.

## **Обработка текстов на русском языке**

### *Построение семантической сети*

В качестве алфавита знаков возьмем множество первых 5 букв всех возможных слов русского языка. Если слово имеет длину меньше 5 и больше 2, то рассматриваем первые три буквы слова. Если же длина слова 1 или 2 буквы, берем его без изменения.

Для заданного текста множеством знаков, входящих в него, будет множество соответствующих начал слов, входящих в него, длиной 1,2,3,5 букв. Два слова объявим близкими, если они встречаются в тексте на

расстоянии к слов. Знаки препинания, встречающиеся в тексте, игнорируем. Степень близости двух знаков – это число случаев, когда два слова, соответствующих данному знаку, встретились рядом.

Таким образом, для фразы «Богами вам еще даны золотые дни, золотые ночи» (А.С.Пушкин) знаками будут: (богам, вам, еще, дан, золоты, дни, золоты, ночь). При  $k = 2$  со знаком «дни» будут рядом знаки (дан, золоты, дни, ночь), с весами (1,2,1,1) соответственно.

#### *Сопоставление двух слов в тексте*

Рассмотрим применение метода расчета близости двух знаков семантической сети в случае текста к анализу близости двух слов в тексте Энциклопедического словаря (Интернет-редакция под псевдонимом «Готье Неимуший», 75 тыс. слов, 2003). Возьмем знаки, соответствующие этим словам, и упорядочим по убыванию списки близких к ним знаков. Например, для текста энциклопедического словаря начала таких списков для слов «животное» и «картофель» будут следующими (для того, чтобы по возможности не изучать часто встречаемые предлоги и союзы, мы не показываем знаки длиной 1 или 2 буквы):

Таблица 1, Упорядоченные списки знаков, близких к знакам «живот» и «карто»

<b>Животное</b> [живот]	<b>Картофель</b> [картоф]
2737 живот	123 карто
613 челов	45 культ
452 мир	30 овощи
308 объект	27 выраст
238 расте	24 кормо
202 испол	24 сахар
187 орган	23 свекл
160 польз	17 молоч
152 позво	17 плодо
142 или	17 ячмен

Если теперь построить новый список, в котором связи для каждого знака будут минимумы связей для знаков «живот» и «картоф», получится таблица, в которой указаны знаки, наиболее близкие к обоим рассмотренным:

Таблица 2, Упорядоченный список знаков, близких к обоим знакам «живот» и «картоф»

17 культ	(17, 45)
11 плодо	(11, 17)
9 кормо	(9, 24)

7 возде	(7, 14)
6 выращ	(6, 27)
6 расте	(238, 6)
5 длина	(53, 5)
5 живот	(2737, 5)
5 крупн	(36, 5)
5 сельс	(98, 5)

Видно, что начало этого списка соответствует смысловому наполнению двух рассматриваемых слов.

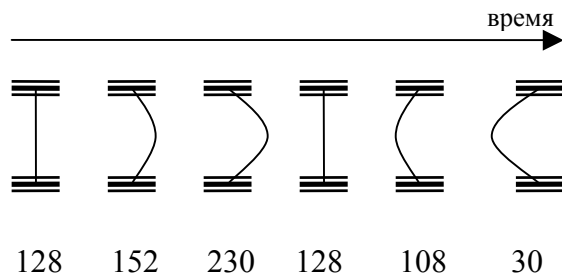
## Обработка звуков

### *Построение семантической сети*

Выше на примере текстов мы показали, как строятся семантические сети. Однако оказывается, что метод построения и последующего сравнения семантических сетей пригоден не только для анализа текстовой информации. Хорошие результаты были получены, в частности, при анализе звукозаписей слов русского языка.

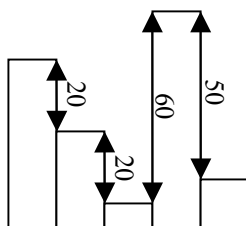
В качестве входных данных мы использовали звуковые WAV файлы. Упрощенно такой файл можно представить как список чисел от 0 до 255, которые отражают отклонение мембраны звукозаписывающего устройства от нейтрального положения в каждый момент времени:

Рис.1, Цифровое представление звука



Следуя схеме, описанной во втором разделе, мы должны определить, что будет являться элементарным знаком в случае аудио информации и задать меру близости между двумя знаками. Оказалось удобным принимать за знак модуль разности между соседними числами:

Рис. 2,  
Знак как разность  
амплитуд



Аудио информация, аналогично текстовой, имеет линейную структуру, поэтому после выбора подходящего знака процедура построения семантической сети полностью сводится к повторению аналогичной процедуры для текста. Рассчитав модули разностей между числами, характеризующими соседние отклонения мембраны, запишем получившуюся последовательность чисел через пробел. Мы получим текст, словами которого являются числа от 0 до 255. Этот текст далее можно преобразовать в семантическую сеть, по алгоритму, описанному в предыдущем пункте, с тем изменением, что знаками являются не слова русского языка, а числа от 0 до 255.

#### *Сопоставление аудиозаписей*

Для нескольких семантических сетей, соответствующих звуковым файлам, можно проанализировать расстояние между ними, чтобы сделать вывод о близости самих звукозаписей. Так как семантические сети являются взвешенными графами, для определения их близости имеет смысл использовать метрику, учитывающую разности весов соответствующих ребер графов. Результаты сравнения текстов между собой проще всего интерпретировать, рассматривая эти тексты как точки в некотором пространстве, расстояние между которыми равно расстоянию между их семантическими сетями. Найдя таблицу попарных расстояний, это множество точек разбиваем алгоритмом кластеризации на заданное число классов. В данной работе мы используем алгоритм кластеризации, рассчитанный на небольшое число классов  $K$ : полным перебором анализируем все наборы из  $K$  точек, объявляем их центрами классов, оставшиеся точки относим к ближайшему центру, и рассчитываем сумму квадратов расстояний между точками и ближайшими к ним центрами. Разбиение, дающее минимум этой суммы, и объявляем результатом кластеризации.

В качестве одного из результатов применения описываемой технологии к обработке аудиозаписей был проведен следующий эксперимент: два человека произнесли и записали каждый по 10 слов, 5 вариантов слова «машина» и 5 вариантов слова «паровоз».

Программе, реализующей описанный метод построения и сопоставления семантических сетей, было предложено разбить аудиозаписи на 2, 3 и 4 группы, каждая из которых будет содержать близкие, по мнению программы, звукозаписи:

Таблица 3, Результаты классификации звукозаписей

2 класса	3 класса	4 класса
«Машины» одного человека	«Машины» первого человека	«Машины» первого человека
Все остальное	«Машины» второго человека	«Машины» второго человека
	«Паровозы»	«Паровозы» первого человека
		«Паровозы» второго человека

## Заключение

Описанный подход к обработке информационных образов нашел применение не только в обработке звуков. Получены хорошие результаты в области распознавания 2D двуцветных контурных изображений и 3D воксельных объектов. Полученные результаты позволяют надеяться, что найденный метод отражает некоторые закономерности процесса обработки произвольной информации и заслуживает изучения с позиций наук, касающихся понятия информация.

По мнению авторов, важным моментом является единообразие алгоритмов обработки информации предлагаемым методом, независимо от природы самой информации. Практически любой информационный образ можно превратить в семантическую сеть, а далее действуют одинаковые алгоритмы сопоставления сетей. Этот факт позволяет сопоставлять между собой внутреннюю структуру таких непохожих видов информации как например, звук и изображение.

Авторы надеются, что сотрудничество со специалистами различных областей знаний позволит углубить теоретическое понимание знаковой природы информации и создать на этой основе несколько новых интересных программных систем.

## Литература

- 1) Иванко Е.Е., Первалов Д.С., Об одном знаковом методе обработки информации и его приложения к анализу текстов и контурных изображений// Труды 32-й регион. молодеж. конф. Проблемы теоретической и прикладной математики, Екатеринбург, 2004, с. 264-268.
- 2) Колмогоров А.Н., Автоматы и жизнь. Возможное и невозможное в кибернетике, 1963.
- 3) Ivanko E.E., Perevalov D.S., On using sign method for 3D images recognition and classification, //Int. Conf. on Computing, Communications and Control Technologies 2004, Texas, (отправлено).
- 4) Sharples M. et al Computers and Thought: A practical Introduction to Artificial Intelligence, <http://www.cogs.susx.ac.uk/local/books/computers-and-thought/index.html>, 1996.
- 5) Sowa J.F Semantic Networks, <http://www.jfsowa.com/pubs/semnet.htm>

# **The application of semantic associative nets for sound records classification.**

The work is devoted to the description of information processing method, worked out by authors, and its application for automatic analysis and classification of texts and sounds. The method is based on the idea that information can be presented as an associative semantic net. Usually such nets are used to solve the problems of computer text analysis. The authors' research touches upon much wider area of such nets application.

The main differences of the proposed nets are homogeneity and high degree of automation of different information types processing. In order to transform some information into such a net it is necessary to determine the notion of an elementary sign, to decompose the informative image into elements- signs and then to identify correlation degrees between them. For instance, words will be sign if a text is considered, while the correlation between words will be counted as the frequency of events where these two words are situated close to each other.

As a consequence of this unified technology application the authors could get results in the assessment of the word sense correlation in a encyclopedia dictionary as well as of the Russian language sound recognition. Earlier the authors got the results in classification of contour two and three dimensional images. Interesting is the fact that the experiments were based on the same analysis algorithms and on the same nets comparisons, independent of the nature of the information being processed.

## **Использование ассоциативных семантических сетей для классификации звукозаписей**

Статья посвящена описанию метода обработки информации, разработанного авторами, и его применению для автоматического анализа и классификации текстов и звуков. В основе метода лежит представление информации в виде структуры ассоциативной семантической сети. Обычно такие сети используются в задачах компьютерного анализа текстов. Авторами исследуется более широкая область применения таких сетей.

Основными отличиями предлагаемых сетей являются однородность и высокая степень автоматизации процессов обработки для различных типов информации. Для преобразования информации в такую сеть требуется определить понятие элементарного знака, произвести декомпозицию информационного образа на элементы-знаки, и, затем, установить связи близости между ними. К примеру, для текста знаками будут слова, а связь между двумя словами – частота встречаемости этих слов поблизости в тексте.

В результате применения единообразной технологии были получены результаты по оценке смысловой близости слов в энциклопедическом словаре и распознаванию звукозаписей слов русского языка.

Ранее авторами были получены результаты по классификации контурных двумерных и трехмерных изображений. Интересным представляется тот факт, что в экспериментах использовались одни и те же алгоритмы анализа и сопоставления сетей, не зависящих от природы обрабатываемой информации.