



УДК 004.822:514

СЕМАНТИЧЕСКАЯ СЕТЬ ПРЕДМЕТНОЙ ОБЛАСТИ КАК ОСНОВА ДЛЯ ФОРМИРОВАНИЯ СЕТИ ПЕРЕХОДОВ ПРИ АВТОМАТИЧЕСКОМ РАСПОЗНАВАНИИ СЛИТНОЙ РЕЧИ

Харламов А.А. *, Ермоленко Т.В.**

* *Институт высшей нервной деятельности и нейрофизиологии РАН, г. Москва*

kharlamov@analyst.ru

** *Институт проблем искусственного интеллекта, г. Донецк, Украина*

etv@iai.dn.ua

Задача распознавания слитной речи в настоящий момент не считается успешно решенной. Тем не менее, известны лицензионные продукты таких фирм как Nuance, Autonomy, которые используются в приложениях, в том числе, русскоязычных. Однако сеть допустимых переходов между словами в заданной предметной области строится в этих системах вручную, а это трудоемкий процесс. В статье описывается подход к автоматическому формированию сети переходов для заданной предметной области на основе использования ассоциативной сети корпуса текстов этой предметной области.

Ключевые слова: распознавание ключевых слов в потоке слитной речи, сеть переходов, автоматическое формирование ассоциативной сети текста.

ВВЕДЕНИЕ

Базовая модель распознавания слитной речи может быть реализована на основе разных подходов. Одним из наиболее известных и хорошо зарекомендовавших себя является подход на основе скрытых Марковских моделей, имеющий общие корни с алгоритмом динамического программирования, разработанным Т.К. Винцоком [Винцюк, 1987]. При этом контекст различных уровней учитывается сетью переходов: фонемы соединяются в слова, слова – в предложения, предложения принадлежат текстам, описывающим предметную область. Предполагается, что на вход системы поступает параметризованный речевой сигнал в виде последовательности векторов параметров в пространстве первичного описания. Параметризация сигнала производится препроцессором, который моделирует физическое восприятие речевого сигнала с учетом свойств акустического канала и фоновых помех. В ответ на произнесенное речевое сообщение на выходе системы получается результат распознавания в виде одной или нескольких гипотез о произнесенной последовательности слов. Эти последовательности, дополненные оценкой сопутствующих речевых параметров, передаются в блок смысловой интерпретации, где происходит принятие

окончательного решения в пользу одной из гипотез. Знания о предметной области, в которой используется система распознавания, позволяют существенно сузить и ускорить поиск.

В одном из подходов к распознаванию слитной речи при преобразовании речи в текст используется словарь, в котором каждое слово задано своим графемным представлением. По этому представлению слова формируются одна или несколькими фонемными транскрипциями слова. Далее, исходя из фонемного состава слова, формируется акустический прототип слова путем объединения различных прототипов акустических образов фонемного уровня (например, трифонов). При распознавании предъявленная реализация слова сравнивается с преобразованными сформированными акустическими прототипами слова. При преобразованиях прототипов слова сохраняется порядок следования акустических форм фонемного уровня прототипа и варьируется в разрешенных пределах длительность акустических форм. Процесс сравнения и поиск наилучшей меры сходства осуществляется методом динамического программирования. Распознаваемая реализация относится к тому слову, преобразованный прототип которого дал наибольшее интегральное сходство с распознаваемым сигналом.

Также как прототипы слов из прототипов графем, создаются прототипы словосочетаний и предложений из прототипов слов и рассматриваются их допустимые преобразования. И в этом случае процесс перебора допустимых фраз и возможных границ между словами достигается методами динамического программирования. Ответом на слитную речевую последовательность является та допустимая фраза или предложение, преобразованный прототип которой оказался самым похожим на предъявленный для распознавания речевой сигнал. Наиболее просто, но не с наилучшим качеством, решается задача распознавание слитной речи в случае свободного порядка следования слов. Если же учитывать не только лексику, но и синтаксис и семантику речи, то на порядок слов накладываются дополнительные тематические ограничения (ограничения на сочетаемость слов).

Фонемные эквиваленты графемных представлений фраз и предложений конкретной предметной области, также как и в случае формирования фонемных представлений слов, формируются вручную. При этом строится сеть переходов, в которой учитываются все возможные комбинации слов всех предложений всех текстов, описывающих предметную область. Так сеть просьбы «разбудить в определенное время» из предметной области «Гостиница» будет выглядеть следующим образом (см. рисунок 1).



Рисунок 1 – Сеть переходов для просьбы «разбудить в определенное время»

Здесь в круглых скобках указаны подсловари, которые можно менять местами, а в квадратных – которые нельзя переставлять. Переставлять подсловари можно только внутри старших скобок. Символ (*) означает пустое слово. Не все слова в этой фразе являются ключевыми. Только «разбудите», «часов» и «минут».

В отличие от трудоемкого метода построения моделей слов с учетом правил комбинаторики элементов, не представляет большой сложности автоматически построить Марковскую модель слова. Для этого берется речевой корпус, содержащий определенное количество произнесенных слов словаря (достаточное для формирования представительной модели), и в автоматическом режиме формируются модели слов. Объединение этих моделей в более крупные единицы представляет собой иногда непреодолимые трудности. Модель языка таким образом построить трудно (а для русского языка просто невозможно), так как требуемого для обучения модели количества произнесенных текстов просто не существует физически. Для модели предметной области это ограничение менее критично, но тоже достаточно серьезно. Именно поэтому известные системы

распознавания слитной речи работают более или менее прилично на нескольких предметных областях (для которых сформированы модели), и плохо для других предметных областей.

Менее сложна, по сравнению с распознаванием слитной речи, задача распознавания ключевых слов в потоке слитной речи. В этом случае трудоемкий процесс ручного формирования сети допустимых переходов между ключевыми словами в заданной предметной области может быть автоматизирован с использованием технологии автоматического построения семантической сети текста на основе корпуса текстов этой предметной области. Семантическая сеть предметной области (текста) является одним из способов представления в сетевом виде семантики предметной области (текста). Особенностью этого автоматического процесса формирования сети переходов является необходимость выявления ключевых, в заданной предметной области, слов, а также существенно важных для описания предметной области предложений. Предложения, не несущие информации о предметной области, в этом случае отбрасываются. Чем тщательнее будет произведен отбор, тем более точно будет работать система распознавания речи. После выбора ключевых слов, а также значимых предложений текста, на основе этих предложений строится однородная семантическая (ассоциативная) сеть. Далее, эта сеть используется для преимущественного выбора гипотез ключевых слов, которые входят в эту семантическую сеть и находятся на наименьших расстояниях от предыдущего распознанного слова.

1. Автоматическое формирование сети переходов

Идея автоматического формирования сети переходов заключается в следующем. В тексте выявляются ключевые понятия в их взаимосвязях. Сеть из ключевых понятий текста, учитывающая их связи (ассоциативная – однородная семантическая – сеть), может быть основой для построения сети переходов. Ключевые понятия выявляются на фоне других, второстепенных слов. Достигается это [Харламов, 2006] вычислением частоты встречаемости слов, а также попарной встречаемости слов в осмысленных фрагментах текста – предложениях. Далее, частотные характеристики слов – вершин сети – перенормируются в их смысловые веса, в результате чего появляется возможность ранжировать и ключевые слова, и предложения текста с точки зрения их важности в этой сети, и, следовательно, в тексте (или в корпусе текстов, описывающих предметную область). Другими словами, такой анализ позволяет выявить важные ключевые слова и существенные предложения текста на фоне второстепенных.

Построенная на основе этого множества предложений ассоциативная сеть является сетью переходов между ключевыми словами для этой

группы предложений. Сеть, построенная на всех предложениях корпуса текстов, описывающего предметную область, ранга выше порогового, является сетью переходов между ключевыми словами для всей предметной области.

2. Автоматическое формирование ассоциативной сети

Текст представляет собой внутренне структурированную последовательность символов, такую, что различные ее элементы имеют различную сложность и различную частоту встречаемости. В процессе автоматического анализа текста формируются словари этих элементов $\{B_i\}$. Это, например, словари флексивных морфем, корневых основ. Выявленные лингвистические единицы в дальнейшем можно использовать для обработки текстовой информации. Словарь флексивных морфем можно использовать для морфологического анализа, словарь корневых основ – для выявления ключевых понятий в тексте и формирования однородной (ассоциативной) семантической сети.

Ранее одним из авторов была реализована технология обработки текстовой информации TextAnalyst, позволяющая автоматически выявлять ключевые понятия в тексте на основе только информации о структуре самого текста (независимо от предметной области). Для этого формируется частотный портрет текста, содержащий информацию о частоте встречаемости понятий текста, представленных как корневые основы соответствующих слов, или их устойчивых сочетаний, встречающихся в тексте, а также об их совместной (попарной) встречаемости в смысловых фрагментах текста (например, в предложениях). Частотный портрет, таким образом, содержит информацию о частоте встречаемости понятий и их попарной (в терминах их ассоциативной связи) встречаемости в тексте. Использование хопфилдоподобного алгоритма [Hopfield, 1982] позволяет перейти от частоты встречаемости к смысловому весу (вес связей при этом остается неизменным).

Эта обработка включает несколько этапов. На первом этапе осуществляется первичная обработка: из текста удаляется нетекстовая информация, текст сегментируется на слова и предложения, из текста удаляются стоп-слова, рабочие и общеупотребимые слова, а оставшиеся слова подвергаются морфологической обработке. Для простоты анализа морфологическая обработка производится с использованием традиционного морфологического словаря – словаря первого уровня – $\{B_i\}_1$. Далее формируется словарь второго уровня – $\{B_i\}_2$ – словарь корневых основ (и устойчивых словосочетаний). На следующем этапе строится частотный портрет текста, то есть выявляются частоты p_i встречаемости корневых основ понятий

(полученных в результате морфологического анализа) и их устойчивых сочетаний, и частоты p_{ij} их попарной встречаемости в предложениях текста (то есть формируется словарь третьего уровня $\{B_i\}_3$). Эта сеть является исходной для формирования сети переходов.

3. Автоматическое выявление ключевых понятий текста

И, наконец, на третьем этапе обработки, частоты встречаемости перенормируются в смысловые веса с использованием итеративной процедуры, похожей на алгоритм искусственной нейронной сети, предложенной Хопфилдом [9]:

$$w_i(t+1) = \left(\sum_{i \neq j} w_i(t) w_{ij} \right) \sigma(\bar{E}), \quad (1)$$

здесь $w_i(0) = \ln p_i$; $w_{ij} = \ln p_{ij} / \ln p_j$ и $\sigma(\bar{E}) = 1/(1 + e^{-k\bar{E}})$ функция, нормирующая на среднее значение энергии всех вершин сети \bar{E} . В результате итеративной процедуры перенормировки наибольшие веса получают понятия, связанные с наибольшим числом других понятий с большим весом, то есть те понятия, которые стягивают на себя смысловую структуру текста.

Полученные таким образом смысловые веса ключевых понятий показывают значимость этих понятий в тексте. В дальнейшем эта информация используется для выявления предложений текста, содержащих наиболее важную информацию в тексте.

В результате получается так называемая ассоциативная (однородная) семантическая сеть N как совокупность несимметричных пар понятий $\langle c_i, c_j \rangle$, где c_i и c_j – понятия, связанные между собой отношением ассоциативности (совместной встречаемости в некотором фрагменте текста):

$$N \cong \{ \langle c_i, c_j \rangle \}. \quad (2)$$

В данном случае отношение ассоциативности несимметрично: $\langle c_i, c_j \rangle \neq \langle c_j, c_i \rangle$.

Семантическая сеть, описанная таким образом, может быть переописана как множество так называемых звездочек $z_i = \langle c_i < c_j \rangle$:

$$N \cong \{ z_i \} = \{ \langle c_i < c_j \rangle \}. \quad (3)$$

Под звездочкой $z_i = \langle c_i < c_j \rangle$ понимается конструкция, включающая главное понятие c_i , связанное с множеством понятий-ассоциантов $\langle c_j \rangle$, которые являются семантическими признаками главного понятия, отстоящими от

главного понятия в ассоциативной сети на одну связь. Ассоциативные связи направлены от главного понятия к понятиям-ассоциантам.

4. Сеть переходов в задаче распознавания ключевых слов в потоке слитной речи

Задача распознавания ключевых слов в потоке слитной речи проще, чем задача точного распознавания слитной речи: в этом случае распознаются не все произносимые слова, но только некоторые – ключевые – имеющие большое значение (заранее оговоренные) в тексте. Если в случае распознавания слитной речи для успешного распознавания требуется построение полной сети переходов (см. рисунок 1), которая позволяет существенно сократить оперативный словарь на каждом шаге распознавания от слова к слову, в случае распознавания ключевых слов сеть переходов упрощается. Из нее уходит второстепенная информация. В ней остаются только ключевые слова и переходы между ними.

Ассоциативная сеть, построенная так, как это описано в разделах 2 и 3 может быть использована для уменьшения числа гипотез при распознавании текущего ключевого слова. В этой сети имеются только ключевые слова. Причем после распознавания очередного слова оперативный словарь ограничивается только ближайшими ассоциантами распознанного ключевого слова, что существенно упрощает саму задачу распознавания, так как словарь системы при этом существенно редуцируется.

5. Пример построения ассоциативной сети

Для наглядности представления описанного подхода рассмотрим пример построения сети переходов, на основе текстов разговорников для раздела «Бронирование гостиницы». Сеть формируется на проходе с помощью персонального продукта, реализующего статистический нейросетевой подход к анализу неструктурированных текстов TextAnalyst. Исходный текст из единственного разговорника представлен ниже

Пример. «Я делал заказ. Места были зарезервированы для меня и моей семьи. Заказ был подтвержден в Париже. У вас есть свободные места? Мне нужна комната. Я хотел бы одноместный номер. Я хотел бы номер с ванной. Я хотел бы номер с двумя кроватями. Нам нужен двухместный номер с дополнительной кроватью. Есть что-нибудь подешевле? Не могли бы Вы показать мне комнату получше? Не могли бы Вы показать мне комнату побольше (поменьше)? Комнату, с видом на море. Какова плата за обслуживание и налог? Надбавка за обслуживание учтена? Сколько стоит номер, включая завтрак?»

Завтрак включен? Сколько это стоит в день? Вам нужен залог? Когда я должен освободить номер? Вам нужен мой паспорт? Можете порекомендовать другую гостиницу? Я сниму этот номер на неделю (месяц). Я пробуду два дня. Меня зовут Вася. Где мне расписаться? Можно одолжить Вашу ручку?»

Сформированная ассоциативная сеть этого текста представлена на таблице 1.

Таблица 1 – Ассоциативная сеть текста Примера 1

Родитель	Подчиненный
включая	
	завтрак
	включая завтрак
	завтрак включен
нужен	
	нужен номер
	нужен залог
	нужен мой паспорт
завтрак	
	включая
	включая завтрак
	завтрак включен
нужен залог	
	нужен
нужен номер	
	нужен
нужен мой паспорт	
	нужен
показать мне комнату	
	комнату
плата	
	обслуживание
делал заказ	
	заказ
включая завтрак	
	включая
	завтрак
надбавка за обслуживание учтена	
	обслуживание
завтрак включен	
	включен
	завтрак

зарезервированы	
	семьи
заказ	
	делал заказ
	подтвержден в Париже
подтвержден в Париже	
	заказ
комнату	
	показать мне комнату
	нужна комната
семьи	
	зарезервированы
обслуживание	
	плата за обслуживание и налог
	надбавка за обслуживание учтена
нужна комната	
	комната

Обработка большого по объему текста (собранный из разных разговорников) даст более совершенную сеть переходов.

Конечно, сеть выглядит несколько примитивно, кроме того, в ней есть информационный шум. Сеть построена на маленьких текстах, описывающих предметную область. Для более полного учета переходов, необходимо собрать более представительный корпус текстов. Необходима также чистка сети экспертом, и, возможно, ее дополнение. Однако если учесть, что ручное формирование сетей переходов, описывающих предметную область в системах распознавания речи – очень трудоемкое занятие, подспорье в построении сети имеет несомненные преимущества.

6. Дальнейшие исследования

В рамках данной работы сеть переходов строилась с использованием статистического подхода и для решения задачи распознавания ключевых слов в потоке слитной речи. Наряду со статистическими методами для автоматического анализа текстов используются лингвистические методы. Лингвистический подход к анализу текстов позволяет решить задачу семантического анализа более точно, но только для отдельных предложений текста. Объединение семантических представлений отдельных предложений в этом случае – до сих пор не решенная задача. Поэтому, совмещение двух подходов – статистического и лингвистического – позволит улучшить качество анализа при

возможности построения семантической сети всего текста. В этом случае семантическая сеть всего текста строится статистическим методом, но вместо пар слов в ней учитываются тройки, включающие пару слов и связь между ними. Эти тройки выявляются в процессе синтаксического анализа. Из предложений текста выявляются содержащиеся в них предикатные структуры, а также отношения сочинения, и атрибутивные отношения.

Под предикатной структурой понимается ядерная подлежащно-сказуемая конструкция, включающая в свой состав n актантов. В общем случае ядро – это глагольная конструкция, а актанты объединяются с ним системой отношений, заполняя его валентные гнезда [Харламов и др., 2012]. Таким образом, предикатная структура имеет вид (1):

$$Subj - R_0 - Pred - R_i - Obj_i, i = 1 \dots n, \quad (4)$$

где *Subj* – активный субъект (грамматическое подлежащее), который инициирует использование предиката *Pred* (грамматического сказуемого); R_0 – отношение «быть субъектом», R_i – предикативные отношения, Obj_i – актанты или узлы предикатной структуры (имя существительное, местоимение, числительное) в их атрибутивной форме, n – количество актантов.

Общую схему действий можно представить в виде последовательности шагов.

- Членение предложения по знакам пунктуации и сочинительным союзам на начальные сегменты (фрагментация). Определение вершин и типов начальных сегментов.
- На декартовом произведении омонимов внутри начальных сегментов построение множества однозначных морфологических интерпретаций каждого сегмента.
- Построение синтаксических групп для каждой интерпретации сегмента с помощью синтаксических правил, выявляющих синтаксические связи между словами. Оценка синтаксического покрытия каждой интерпретации.
- Установление иерархии между сегментами с помощью синтаксических правил: вложения контактно расположенных сегментов (причастных, деепричастных оборотов, обособленного определения); определение однородности между контактно расположенными сегментами; определение отношения импликации между сегментами по подчинительным союзам, в них входящим.

Предикатная структура простого предложения выявляется на основе анализа главных членов предложения путем выявления шаблона, соответствующего одной из минимальных структурных схем предложений (МСС), описывающих предикативный минимум предложения. Использование МСС в качестве формального образца позволяет выявить

предикативную основу (структурную схему) простого предложения, и в дальнейшем – его предикатную структуру. Смысловая связь между понятиями предложения (объектом/субъектом) в общем случае может быть описана предикатом, актантами которого выступают данные понятия. Установление таких синтактико-семантических связей позволяет сформировать схему ситуации, описываемой во фразе.

Обусловленный валентностью предиката семантико-синтаксический уровень анализа конструкций, не соответствующий узкому собственно формально-синтаксическому подходу, дает возможность даже из набора неправильных форм (посредством приведения их к начальным формам) с помощью заполнения валентных гнезд определить схему предложения.

Выявление ключевых предикатных структур в тексте возможно с использованием того же подхода, который был использован для выявления ключевых понятий и ключевых предложений в тексте в технологии TextAnalyst.

После выявления ключевых предикатных структур текста, в ассоциативной сети текста отношения между парами ключевых понятий размечаются их типами из соответствующей предикатной структуры таким образом, что звездочка ассоциативной сети заменяется предикатной структурой, причем, главное понятие ассоциативной сети совмещается с субъектом, а ближайшие ассоцианты – главным и второстепенными объектами и их атрибутами предикатной структуры. Ассоциативные связи заменяются соответствующими типами отношений предикатной структуры – собственно предикатным отношением между субъектом и главным объектом, а также другими отношениями – валентными связями актантов предиката – субъекта с, соответственно, второстепенными объектами и их атрибутами.

Заключение

Автоматический смысловой анализ текста с использованием технологии TextAnalyst позволяет выявить ключевые понятия в их взаимосвязях в тексте, а также ранжировать их по степени их смысловой значимости в данном тексте. Ранжирование ключевых понятий, в свою очередь, позволяет ранжировать предложения текста, и выбирать наиболее существенные из них для множества текстов, описывающих предметную область. Такая сеть может быть исходной для построения сети переходов между ключевыми словами в задаче распознавания ключевых слов в потоке слитной речи. В дальнейшем предполагается исследовать возможность объединения статистического и лингвистического подходов к автоматическому смысловому анализу текстов, что позволит строить более точную сеть переходов за счет разметки ассоциативных связей между

ключевыми словами типами их отношений в предикатных структурах соответствующих предложений.

Работа выполнена при поддержке гранта Минобрнауки 2012-1.4-07-514-0018 «Исследование и разработка программного обеспечения понимания неструктурированной текстовой информации на русском и английском языках на базе создания методов компьютерного полного лингвистического анализа».

Библиографический список

[Винцюк, 1987] Винцюк Т.К. Анализ, распознавание и смысловая интерпретация речевых сигналов. – Киев. Наукова думка, 1987. – 267 с.

[Харламов, 2006] Харламов А.А. Нейросетевая технология представления и обработки информации (естественное представление знаний). - М.: «Радиотехника», 2006. - 89 с.

[Hopfield, 1982] Hopfield, J.J. Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. 79, 1982. Pp. 2554 – 2558

[Харламов и др., 2012] Харламов А.А., Ермоленко Т.В., Дорохина Г.В., Гнистько Д.С. Метод выделения главных членов предложения в виде предикатных структур, использующий минимальные структурные схемы. Речевые технологии, № 2, 2012.

SUBJECT AREA SEMANTIC NETWORK AS A BASIS FOR COMPILING TRANSITION NETWORK IN AUTOMATIC SPEECH RECOGNITION

Kharlamov A.A. *, Yermolenko T.V. **

**Institute of Higher Nervous Activity of the Russian Academy of Science, Moscow, Russia*

kharlamov@analyst.ru

***Institute of Artificial Intelligence Problems, Donetsk, Ukraine*

etv@iai.dn.ua

The problem of speech recognition is not satisfactory fulfilled at the moment. But there are some products of the famous companies like Nuance and Autonomy which are used in applications. But the transition network between words of the subject area there compiled manually. Approach to automatical compilation of such a network is represented in the paper.