Оценка гипотез с использованием синтаксического анализа

Батальщиков А.А.
Зулкарнеев М.Ю.
Шамраев Н.Г.

ФГНУ НИИ «Спецвузавтоматика»
Россия, г. Ростов-на-Дону, Газетный пер., 51,
тел. (863) 297-50-84, факс (863) 297-50-84,
zulkarneev@mail.ru, ncam1977@yahoo.com, asni@asni.rsu.ru

В настоящее время в распознавании речи широко распространены декодеры, использующие скрытые марковские модели и алгоритмы Баум-Уолша и Витерби. В процессе распознавания производится оценка акустических и языковых вероятностей для различных возможных вариантов произнесённой фразы, и результатом работы декодера является набор из N наилучших гипотез. В случае использования биграммной или трёхграммной языковой модели учитываются вероятности сочетания пар и троек слов соответственно. Тем не менее, полученные на выходе гипотезы — фразы или предложения не обязательно соответствуют правильному построению предложения. В докладе предлагается улучшить качество распознавания речи, используя дополнительно синтаксический анализ речи. Для каждой из N лучших гипотез дополнительно выполняется анализ соответствия содержимого фразы синтаксическим правилам. При анализе каждой гипотезы выставляется оценка соответствия синтаксическим правилам, для полностью правильного предложения она полагается равной 1. В докладе рассматриваются различные способы и методы оценки синтаксического соответствия для предложений русской и английской речи, основанные на фрагментации синтаксических групп и алгоритме Коке-Янгера-Касами.

ВВЕДЕНИЕ

В настоящее время в распознавании речи широко распространены декодеры, использующие скрытые марковские модели, основанные на алгоритмах Баум-Уолша и Витерби. На вход декодеру поступает последовательность векторов признаков, которая затем восстанавливается как последовательность слов. Более подробно процесс декодирования описан в [1]. Конкретная восстановленная последовательность слов в дальнейшем будет называться гипотезой, поскольку она отражает одну из возможных последовательностей слов, произнесённых в анализируемом речевом сигнале. В процессе распознавания производится оценка акустических и языковых вероятностей для различных возможных вариантов произнесённой фразы, причём это происходит не поэтапно, а в совокупности по результатам вычислений алгоритма Витерби. Использование других типов языковой модели (например, использование вероятностных контекстно-свободных грамматик) может позволить улучшить качество распознавания. Кроме этого, возможно улучшение за счет модификации весов акустической и языковой

вероятностей, а также задания параметров глубины поиска и штрафующей функции на количество слов.

В случае использования биграммной или трёхграммной языковой модели учитываются вероятности сочетания пар и троёк слов соответственно. Тем не менее, полученные на выходе гипотезы — фразы или предложения не обязательно соответствуют правильному построению предложения. На основании этого факта в работе разрабатывается метод модификации вероятностей гипотез для того, чтобы повысить вероятность грамматически более правильных гипотез.

Далее даётся описание используемого метода и приводятся результаты экспериментов.

ОПИСАНИЕ МЕТОДА

Метод модификации вероятностей гипотез основывается на оценке синтаксической целостности предложения. Анализ правильности построения предложения проводится с использованием формальных грамматик Хомского [2,3].

В работе исследуется система распознавания речи для английского языка. Контекстно-свободная грамматика английского языка в упрощённой форме может быть задана с помощью приблизительно 40 правил, описанных в [4]. Правила имеют вид:

[S]->[NP-VP][PP],
[S]->[PP-NP][VP],
[S]->[NP-VP][PP-PP], ...,

где S — начальное правило, [NP-VP], [PP], [VP], [PP-NP] — правила, отвечающие грамматическим категориям (соответственно комбинированной группы существительного-глагола, предложной, глагольной группы, комбинированной группы предложной и существительного и др.)

Для проверки правильности гипотез каждое слово переводится в соответствующую грамматическую категорию. В случае омонимов или неоднозначного соответствия грамматическим категориям, создаются альтернативные гипотезы. Затем проводится анализ получившейся структуры данных, основанный на алгоритме Коке-Янгера-Касами [4]. В результате анализа получается дерево синтаксического разбора. Кроме этого на основании синтаксического анализа делается вывод, описывается ли предложение данной формальной грамматикой, и строится последовательность использованных правил.

В методе используется набор гипотез, сгенерированный декодером. При анализе каждой гипотезы выставляется коэффициент α соответствия синтаксическим правилам, для полностью правильного предложения он полагается равным 1. Для предложений, которые не соответствуют правильному грамматическому построению, коэффициент варьируется от минимального до близкого к 1 в зависимости от величины охвата слов предложения деревом правил с наибольшей глубиной.

Если предложение полностью не удовлетворяет грамматическим правилам, оно разбивается на фрагменты (фрагмент получается как последовательность слов-элементов поддерева, генерируемым некоторым правилом, не обязательно начальным S), и для каждого фрагмента i вводится величина α_i — оценка правильности грамматической структуры. Оценка правильности грамматической структуры предложения α вычисляется как

$$\alpha = \sum_i \alpha_i = \sum_i \frac{N_i}{N},$$

где N — общее количество слов в предложении, N_i — количество слов в -ом фрагменте, суммирование ведётся по фрагментам с глубиной дерева не менее 2.

Правильная гипотеза выбирается на основе скорректированной на α величины правдоподобия гипотезы, полученной от декодера.

ОПИСАНИЕ ЭКСПЕРИМЕНТОВ

Для проверки предложенного метода была проведена серия экспериментов по распознаванию слитной речи. Для экспериментов использовалась речевая база английского языка, полученная из новостных каналов. Для тестирования были использованы речевые сообщения общей длительностью 1 час. Для распознавания использовались контекстно-зависимые акустические трифонные модели (СММ), с 4500 различных состояний. В качестве языковой модели использовалась трёхграммная модель. При тестировании все слова из тестирующей выборки содержались в словаре.

Таблица

Конкурирующие предложения — гипотезы на английском языке

№ гипотезы	Слова гипотезы и их величины правдоподобия							
1	WAGNER	CAME	FROM	YOU	AS	AID	OF	INDIANA
	740.1657	449.9958	134.9487	580.8	518.4	349.6	335.8	1190.7481
2	WAGNER	CAME	FROM	YOU	AS	A	OF	INDIANA
	740.1657	449.9958	134.9487	580.8	518.4	359.09	325.5	1190.7489
3	WAGNER	CAME	FROM	YOU	AS	THEY	OF	INDIANA
	740.1657	449.9958	134.9487	580.8	467.6	416.1	316.5	1190.7489
4	WAGNER	CAME	FROM	YOU	AS	EIGHT	OF	INDIANA
	740.1657	449.9958	134.9487	580.8	518.4	339.0	342.7	1190.7489

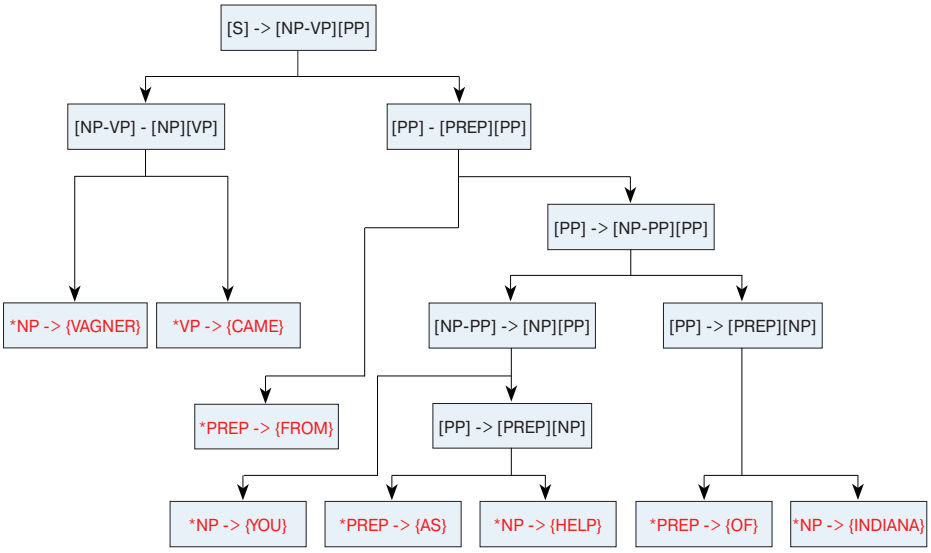

Рис. Дерево правил, полученное в результате работы алгоритма Коке-Янгера-Касами

В качестве примера в таблице приводится набор конкурирующих гипотез для одного предложения тестирующей выборки. В первой строке приводится правильная гипотеза. По результатам синтаксического анализа при помощи алгоритма Коке-Янгера-Касами эта гипотеза получила оценку правильности, равную 1. Остальные гипотезы получили оценку правильности меньше 1. На рисунке показано дерево правил для первой из гипотез.

По результатам проведённых экспериментов точность распознавания повысилась на 1,2%. Улучшение было достигнуто для относительно коротких предложений длительностью от 6 до 10 слов.

ЗАКЛЮЧЕНИЕ

Эксперименты показали, что синтаксический анализ позволяет повысить точность распознавания для речевых сообщений на английском языке, полученных из новостных каналов. Повышение точности объясняется тем, что метод повышает приоритет грамматически корректных гипотез, которые характерны для новостных сообщений. Планируется провести эксперименты по распознаванию речевых сообщений на русском языке. Грамматика Хомского для русского языка устроена гораздо сложнее, однако уже сейчас построены основные правила, позволяющие проверять правильность предложения. Используя правила, разработанные в проекте АОТ факультета лингвистики РГГУ [5], можно добиться достаточно компактного представления языковых правил в форме грамматики Хомского. Кроме того можно модифицировать метод для вероятностных контекстно-свободных грамматик. В этом случае синтаксическая модель будет выступать в качестве модели языка.

ЛИТЕРАТУРА

1. *L.R. Rabiner, B.H. Juang*. Fundamentals of Speech Recognition. Prentice Hall, Englewood Cliffs, NJ, 1994.

2. *Н. Хомский*, Аспекты теории синтаксиса. МГУ, 1972.

3. *Н. Хомский, Дж. Миллер*. Введение в формальный анализ естественных языков. Кибернетический сборник. Вып. 1. Мир, 1965. С. 229–292.

4. *S. E. Levinson*, Mathematical models for speech technology, Chippenham, John Wiley and Sons, 2005.

5. Ссылка в сети Internet: http://www.aot.ru.