

АЛФАВИТ С МИНИМАЛЬНОЙ ИЗБЫТОЧНОСТЬЮ

Н.Г.Загоруйко, В.Н.Елкина

При распознавании ограниченного набора устных команд по последовательности более мелких речевых элементов (слов, звукотипов и т.д.) обычно возникает следующая ситуация: в силу избыточности языка для распознавания команд требуется распознавать не все мелкие элементы, входящие в состав этих команд, а лишь некоторую их часть.

Задача устранения избыточных элементов решена в работе [1]. В этой статье даётся решение задачи минимизации количества распознаваемых элементов не за счет "вычеркивания лишних", как в [1], а за счет объединения мелких элементов в группы. При этом минимизируется как число таких групп, так и ожидаемое количество ошибок распознавания групп.

В качестве примера будем рассматривать модель распознающего устройства, в котором на первом этапе (в первом автомате) распознаются звукотипы, а на втором — по последовательности звукотипов — отдельные команды.

Алгоритм минимизации алфавита предполагает использование исходной матрицы (M_0) ошибок распознавания звукотипов. Эту матрицу можно получить в процессе предварительной эксплуатации макета (или модели на ЭВМ) такого распознающего устройства, в котором алфавитом объектов распознавания является пол-

ный (исходный) набор звуко типов данного языка. Матрица ошибок (M_0) будет, в общем случае, не симметричной (например, может оказаться, что звонкие взрывные будут распознаваться как глухие чаще, чем наоборот), т.е. что ошибки I-го и II-го родов не равны друг другу.

Среднюю вероятность ошибки (P_{ij}) распознавания между элементами i и j будем определять по формуле:

$$P_{ij} = q_i P_{i/j} + q_j P_{j/i}$$

где $P_{i/j}$ - ошибка I-го рода,

$P_{j/i}$ - ошибка II-го рода,

а q_i и q_j - априорная вероятность появления i и j элементов, соответственно.

Прежде чем приступить к минимизации алфавита мелких элементов нужно проверить его достаточность. Для этого составляется матрица (M_T), строками (S) которой являются команды (слова) данного набора, столбцами (K) - звуко типы, а на пересечении столбца и строки стоят номера следования данного звуко типа в данном слове. Достаточным будем считать такой набор звуко типов, при котором матрица M_T не содержит одинаковых строк. Если условие достаточности выполняется, можно переходить к проверке условия необходимости. Необходимым будем называть минимальный набор звуко типов или их групп, для которого выполняется условие достаточности.

В процессе минимизации набора распознаваемых элементов будем объединять в группы те звуко типы, для которых выполняется условие совместности. Совместными будем называть звуко типы, объединение которых в одну группу не приводит к слиянию разных команд (строк). Это условие эквивалентно принятому в фонетике понятию о смысловозначительной функции фонемы. При неограниченном словаре все звуко типы, очевидно, будут несовместимыми.

Назовем групповым весом (D_j^e) звуко типа k_j среднюю вероятность ошибочной классификации между ним и всеми остальными звуко типами данной e -ой группы:

$$D_j^e = \sum_{i=1, i \neq j}^{K_e} P_{i/j}$$

Здесь K_e - число звуко типов в e -ой группе.

Общим весом (D_j) звуко типа k_j будем называть среднюю вероятность ошибочной классификации между ним и всеми остальными звуко типами исходного базиса:

$$D_j = \sum_{i \neq j}^{K_0} P_{i/j}$$

где K_0 - общее число звуко типов.

Аналогично межгрупповым весом $D^{e,r}$ будем считать среднюю вероятность ошибки распознавания всех звуко типов e -ой группы со всеми звуко типами r -ой группы:

$$D^{e,r} = \sum_{i=1}^{K_e} q_i \sum_{j=1}^{K_r} P_{i/j} + \sum_{j=1}^{K_r} q_j \sum_{i=1}^{K_e} P_{j/i} =$$

$$= \sum_{i=1}^{K_e} \sum_{j=1}^{K_r} q_i P_{i/j} + \sum_{i=1}^{K_e} \sum_{j=1}^{K_r} q_j P_{j/i} =$$

$$= \sum_{i=1}^{K_e} \sum_{j=1}^{K_r} (q_i P_{i/j} + q_j P_{j/i}) = \sum_{i=1}^{K_e} \sum_{j=1}^{K_r} P_{ij}$$

здесь K_e и K_r - число звуко типов в e -ой и r -ой группах, соответственно.

Для сокращения перебора процедура группировки начинается с поиска ядра группы. Ядрами групп могут служить звуко типы из любой несовместимой пары или - если таковой пары нет - два звуко типа k_i и k_j , для которых величина

$$\Delta = D_i \cdot D_j (1 - P_{i/j})$$

имеет максимальное значение (т.е. наиболее "тяжелые" и наименее связанные между собой элементы).

Затем берётся любое слово словаря и для всех входящих в его состав звуко типов определяется групповой вес со всеми имеющимися к этому времени группами. Звуко тип присоединяется к той из совместимых групп, с которой он имеет наибольший групповой вес. Если некоторый звуко тип (k_i) имеет со всеми совместимыми группами H_e вес $D_i^e = 0$, то k_i может быть либо присоединен к любой из этих групп, либо выделен в каче-

стве ядра новой группы. В первом случае достигается минимум числа групп с минимальной для этого числа групп средней ошибкой распознавания (см. блок-схему 1). Во втором случае минимизируется средняя ошибка распознавания с наименьшим при этой ошибке числом групп (см. блок-схему 2).

После каждого шага (присоединения звукотипа к группе или формирования ядра новой группы) производится проверка группового веса всех элементов каждой группы со всеми (совместимыми) группами. Как и раньше, звукотип присоединяется к той из совместимых групп, в которой он имеет наибольший групповой вес. Такая постоянная проверка гарантирует оптимальность группировки, даже если бы в самом начале в качестве ядер была выбрана случайная пара звукотипов.

Все слова таблицы перекодируются в терминах нового алфавита: на пересечении строки (слова S_j) и столбца (группы H_e или звукотипа k_c) указываются порядковые номера следования этих элементов в данном слове.

Если идут два звукотипа одной группы подряд, то их естественно считать слившимися в один элемент, так что число элементов и порядок их следования в слове в результате перегруппировок могут меняться.

Аналогичная процедура производится со всеми остальными словами исходного словаря. Для каждого слова делается проверка - не сливается ли оно с каким-нибудь другим словом (проверка на достаточность алфавита). Если сливается - происходит перегруппировка звукотипов. С этой целью в каждой группе отмечаются звукотипы, удаление которых из данной группы устранило бы это слияние. Перегруппировка осуществляется за счет того из отмеченных звукотипов, групповой вес которого уменьшается при перемещении его в другую группу на минимальную величину.

Если ни один из вариантов перемещения в существующих группах не приносит желаемых результатов, то один из отмеченных звукотипов (обладающий минимальным групповым весом) становится ядром новой группы. Ясно, что такой способ введения новых групп обеспечивает минимизацию числа этих групп. Вместе с тем объединение элементов в группы по критерию максимума группового веса приводит к минимизации межгрупповых весов.

Действительно, пусть межгрупповой вес для групп H_n и H_e , содержащих соответственно K_n и K_e элементов, есть $D^{(n,e)}$,

а групповой вес элемента k_i с группой H_n больше, чем с группой H_e ($D_i^{(n)} > D_i^{(e)}$).

Присоединение k_i к H_e увеличило бы межгрупповой вес на $\Delta_e = D_i^{(n)} - \sum_{j=1}^{K_n} P_{ij}$, а присоединение k_i к H_n -

на $\Delta_n = D_i^{(e)} - \sum_{j=1}^{K_e} P_{ij}$. Так как по условию $D_i^{(n)} > D_i^{(e)}$,

то $\Delta_n < \Delta_e$ и $D^{(n)} + \Delta_e > D^{(n)} + \Delta_n$.

Следовательно, условие присоединения элемента к группе по критерию максимума группового веса одновременно минимизирует межгрупповые веса (т.е. среднюю вероятность ошибки распознавания групп) и описанный способ группировки удовлетворяет обязательным требованиям к алгоритмам таксономии [2] - максимальной связи между элементами группы и минимальной связи между группами.

Среднюю ошибку распознавания групп можно определить по следующей формуле:

$$P_{cp} = \sum_{e=1}^N q_e P_e,$$

где N - число групп;

P_e - ошибки I-го рода при распознавании элементов e -ой группы:

$$P_e = \sum_{j=1}^{K_0 - K_e} q_i P_{ij} + \sum_{i=1}^{K_e} q_i P_{ij}$$

Здесь P_{ij} - ошибки (в %) отнесения реализаций i -го звукотипа к j -тому, q_i - априорная вероятность появления i -го звукотипа в e -ой группе ($\sum_{i=1}^{K_e} q_i = 1$), а q_e - априорная вероятность любого звукотипа e -ой группы ($\sum_{e=1}^N q_e = 1$). Через K_0 обозначено общее число звукотипов, а через $K_0 - K_e$ число звукотипов не входящих в e -ую группу.

Блок-схемы описанного алгоритма группировки приводятся на рисунке 1. Здесь используются следующие обозначения и операторы:

- k - звукотип;
- H - группа звукотипов;
- S - вариант слова;
- i, j - индексы звукотипов в слове;
- e, n - индексы групп звукотипов;
- α - счетчик числа слов в словаре;

- β - счетчик числа звукотипов в слове;
- A_0 - оператор поиска несовместимых пар звукотипов;
- A_{zp} - оператор проверки на совместимость звукотипов со всеми группами;
- B - оператор формирования ядра новой группы;
- C - оператор проверки на ноль группового веса звукотипа со всеми группами;
- K - оператор нахождения пары (i и j) звукотипов, обладающих наибольшим значением $D_i \cdot D_j (1 - P_{ij})$;
- F - оператор поиска группы, с которой данный звукотип имеет максимальный групповой вес;
- E_D - оператор присоединения звукотипа к группе, в которой он обладает наибольшим групповым весом;
- R_1 - оператор, проверяющий не кончилось ли слово;
- R_2 - оператор, проверяющий не кончился ли список слов в таблице;
- G_1 - оператор перекодировки словаря (в терминах нового алфавита);
- G_2 - оператор перегруппировки, заключающейся в перенесении элемента k_i из H_e в H_n , если $D_i^e > D_i^n$;
- \mathcal{L} - оператор объединения совместных групп.

Проиллюстрируем применение описанного алгоритма на примере таблиц I и 2.

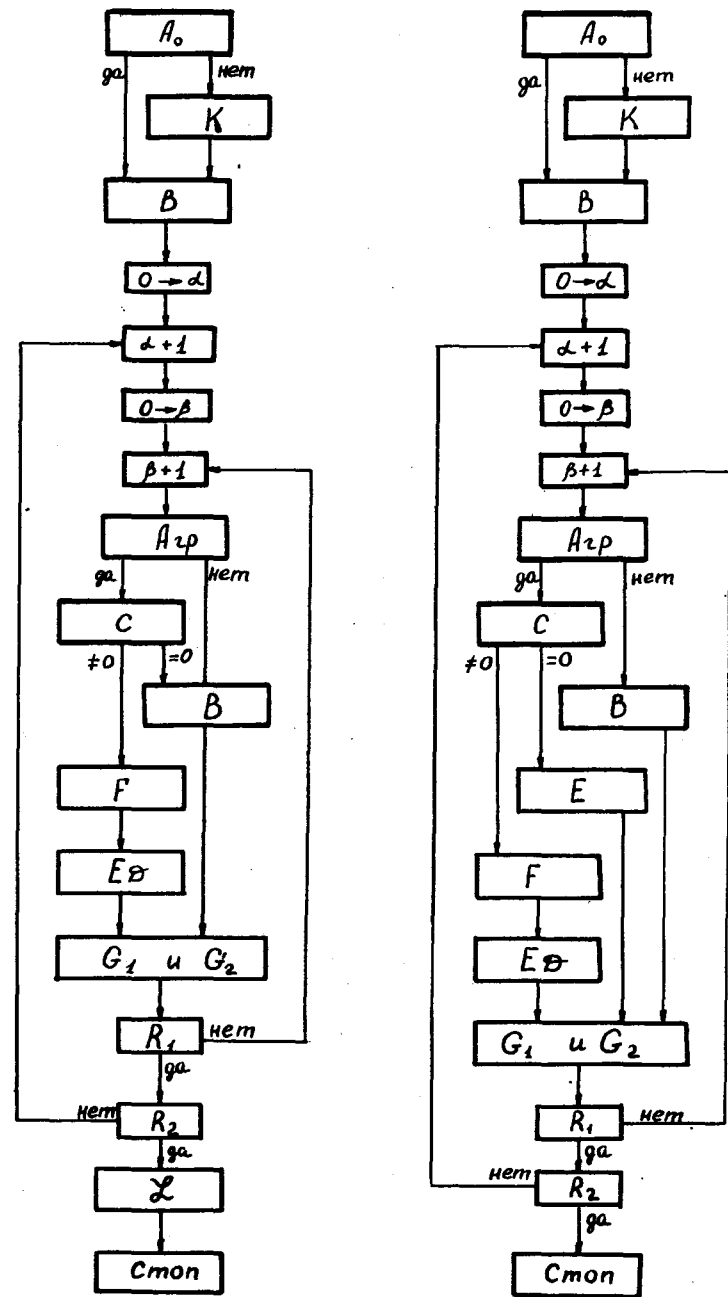
Для простоты будем считать вероятности появления всех звукотипов равными, а ошибки I и II рода одинаковыми.

Словарь состоит из слов "ноль", "семь", "сеть", "пять".

Т а б л и ц а I

	Н	О	Л'	С'	Э	М'	П'	А	Т'	D_i
1	Н	-	5			10				15
2	О	-			15			20		35
3	Л'	5	-			10				15
4	С'			-				10		10
5	Э		15		-			10		25
6	М'	10		10		-			5	25
7	П'						-	20		30
8	А		20			10		-		30
9	Т'				10		5	30	-	45

н о л'



Блок-схема 1.

Блок-схема 2.

Таблица 2

S \ H	Н	О	Л'	С'	Э	М'	Т'	П'	А
НОЛ'	1	2	3						
С'ЭМ'				1	2	3			
С'ЭТ'				1	2		3		
П'АТ'							3	1	2

Из таблицы 2 легко видеть, что исходный набор звуко типов удовлетворяет условию достаточности.

В качестве ядер следует выбрать несовместимую пару звуко типов М' и Т'. Элементы слов "НОЛ'" и "С'ЭМ'" присоединяются в соответствии с блок-схемой 2 к этим ядрам, как показано в таблице 3.

Таблица 3

S \ H	Н	М'	Т'	П'	А	
	НОЛ'	Н Л'	С'			
НОЛ'		1				
С'ЭМ'		2	1			
С'ЭТ'		2	1,3			
П'АТ'			3	1	2	

Окончательный результат минимизации числа групп показан в таблице 4.

Таблица 4

S \ H	Э	А	
	О Л' Н М'	П С' Т'	
НОЛ'	1		
С'ЭМ'	2	1	
С'ЭТ'	2	1,3	
П'АТ'		1	

Ясно, что меньшего количества групп, чем 2, сделать не удастся. Средняя вероятность межгрупповой ошибки будет (см. таблицу 1)

$$\begin{aligned}
 P_{op} &= q_1 P_1 + q_2 P_2 = q_1 \sum_i^5 \sum_j^4 q_i P_{ij} + \\
 &+ q_2 \sum_{i=1}^5 \sum_{j=1}^4 q_i P_{ji} = \\
 &= q_1 (q_M \sum_{j=1}^4 P_{Mj} + q_N \sum_{j=1}^4 P_{Nj} + q_{Л'} \sum_{j=1}^4 P_{Л'j} + q_{С'} \sum_{j=1}^4 P_{С'j} + q_{Э} \sum_{j=1}^4 P_{Эj}) + \\
 &+ q_2 (q_T \sum_{i=1}^5 P_{Ti} + q_C \sum_{i=1}^5 P_{Ci} + q_M \sum_{i=1}^5 P_{Mi} + q_A \sum_{i=1}^5 P_{Ai}) = \\
 &= \frac{5}{9} [\frac{1}{3}(0+0+0+20+10)] + \frac{4}{9} [\frac{1}{4}(0+0+0+30)] = 6,7\%.
 \end{aligned}$$

Для алгоритма по блок-схеме II промежуточный результат представлен таблицей 5, а окончательный - таблицей 6.

Таблица 5

S \ H	Н Л' М'	П' Т'	Э А О	С'
НОЛ'	1,3		2	
С'ЭМ'	3		2	1
С'ЭТ'		3	2	1
П'АТ'		1,3	2	

Таблица 6

S \ H	Н Л' М'	П' Т'	С' Э А О
НОЛ'	1,3		2
С'ЭМ'	2		1
С'ЭТ'		2	1
П'АТ'		1,3	2

Средняя вероятность ошибки при распознавании групп H_1, H_2 и H_3 будет $P_{op} \approx 2,2\%$.

Использование описанных алгоритмов оказалось особенно эффективным для обработки таблиц с большим количеством слов. Так, для конкретного словаря, состоящего из 40 слов, оказалось достаточным распознавать всего четыре группы звуко типов - 1) гласные, 2) шипящие, 3) взрывные, 4) остальные согласные. Вообще говоря, число и состав групп зависят от конкретного словаря, однако, можно ожидать, что число групп будет тем меньше, чем больше длинных слов будет входить в словарь.

Л и т е р а т у р а

Г. В.Н. Елкина, Н.Г. Загоруйко. Об алфавите объектов распознавания. - Вычислительные системы, г. Новосибирск, 1966, вып. 22.

Поступила в редакцию
21.VII.1967г.