Пастухова Д.А.

УПОРЯДОЧЕНИЕ ВАРИАНТОВ ТЕКСТОВ ПРИ ИХ СОСТАВЛЕНИИ И ПЕРЕВОДЕ

Литературный или технический текст состоит из размещений элементов. Под элементами понимаются абзацы, предложения, слова, морфемы, буквы. Автором предлагается методика упорядочения всех вариантов текстов с применением систем счисления, основания которых отличаются от десяти. Упорядочение облегчает составление программ и исключает возможность ошибок (повторений и пропусков вариантов).

Любой текст (литературный или технический) состоит из размещений элементов. Под элементами подразумеваются абзацы, предложения, слова, морфемы, буквы. После определения количества всех возможных (по правилам грамматики данного языка) размещений возникает проблема отыскания самих этих размещений.

Автором предлагается методика расположения всех вариантов текстов с применением систем счисления с основаниями, отличными от 10. Назовем их q-ичными системами счисления, где q – равно количеству элементов.

Теория систем счисления дана в работе С.В. Фомина /1/.

Расположение всех вариантов размещения по порядку значительно упрощает задачу составления программы для ЭВМ, помогает при обычном отыскании вариантов и исключает возможность ошибки.

Пример 1. Пусть требуется определить, сколько различных размещений по три можно составить из последовательности элементов «soft».

Если на расположение элементов не наложены никакие ограничения, то для определения количества размещений воспользуемся формулой, взятой из работы /2/.

$$A_q^n = q(q-1)(q-2)\ldots(q-n+1), \qquad (1)$$

где A_q^n – количество размещений из «q» элементов по «n».

Напомним, что правая часть формулы (1) содержит «n» сомножителей.

В нашем случае получаем:

$$A_4^3 = 4 \cdot 3 \cdot 2 = 24$$

варианта размещения.

Для отыскания этих вариантов каждому элементу присвоим цифру по порядку возрастания, начиная от нуля. Получим

$$\begin{matrix} s & o & f & t \\ 0 & 1 & 2 & 3 \end{matrix} \qquad (2)$$

Если в исходной последовательности имеются одинаковые элементы, то им будут соответствовать одинаковые цифры. Если число различных элементов превышает 10, то этим элементам можно присвоить специальные символы, означающие числа 10, 11, 12 и т. д. Чтобы не усложнять задачу, вместо специальных символов применим числа (10), (11), (12) и т. д., взятые в скобки.

Чтобы упорядочить все возможные перестановки, применим системы счисления с основаниями, отличными от десяти. Основание системы счисления должно равняться числу элементов в исходной последовательности. Тогда каждому размещению соответствует определенное число «D», записанное в системе счисления, основание которой в общем случае отлично от десяти.

После этого по известным формулам /1/ переводим это число «D» в понятную всем десятичную систему счисления. Теперь появится возможность расположить все варианты размещений по порядку возрастания числа «D». При этом ни один вариант размещения не будет повторен и ни один вариант не будет упущен.

В нашем примере 1 применяем четверичную систему счисления, так как в исходном варианте имеем четыре различных элемента.

Найдем упомянутые 24 варианта размещений и расположим их по порядку возрастания числа «D» (таблица 1).

Напомним, что в четверичной системе счисления применяются всего четыре символа: 0; 1; 2; 3. Следующее по порядку число 4, принадлежащее десятичной системе записывается как 10. Правила перевода любого числа из десятичной системы в q-ичную и обратно даны в работе /1/.

Пример, рассмотренный в таблице 1, был на упорядочение вариантов размещений без повторений.

Пример 2. Если в исходной последовательности имеется «n» различных элементов и из них нужно составить размещения, состоящие из «k» элементов при допущении, что каждый элемент может повторяться до «k» раз, то такие размещения называются размещениями с повторениями. Так пусть требуется определить число размещений по k=2 элемента из исходной последовательности элементов «short».

Таблица 1. Размещения без повторений

№ п.п.	В четверичной системе	Вариант размещения	Число «D» в десятичной системе счисления
1	2	3	4
1	012	sof	$2\cdot4^0+1\cdot4^1+0\cdot4^2=6$
2	013	sot	$3\cdot4^0+1\cdot4^1+0\cdot4^2=7$
3	021	sfo	$1\cdot4^0+2\cdot4^1+0\cdot4^2=9$
4	023	sft	$3\cdot4^0+2\cdot4^1+0\cdot4^2=11$
5	031	sto	$1\cdot4^0+3\cdot4^1+0\cdot4^2=13$
6	032	stf	$2\cdot4^0+3\cdot4^1+0\cdot4^2=14$
7	102	osf	$2\cdot4^0+0\cdot4^1+1\cdot4^2=18$
8	103	ost	$3\cdot4^0+0\cdot4^1+1\cdot4^2=19$
9	120	ofs	$0\cdot4^0+2\cdot4^1+1\cdot4^2=24$
10	123	oft	$3\cdot4^0+2\cdot4^1+1\cdot4^2=27$
11	130	ots	$0\cdot4^0+3\cdot4^1+1\cdot4^2=28$
12	132	otf	$2\cdot4^0+3\cdot4^1+1\cdot4^2=30$
13	201	fso	$1\cdot4^0+0\cdot4^1+2\cdot4^2=33$
14	203	fst	$3\cdot4^0+0\cdot4^1+2\cdot4^2=35$
15	210	fos	$0\cdot4^0+1\cdot4^1+2\cdot4^2=36$
16	213	fot	$3\cdot4^0+1\cdot4^1+2\cdot4^2=39$
17	230	fts	$0\cdot4^0+3\cdot4^1+2\cdot4^2=44$
18	231	fto	$1\cdot4^0+3\cdot4^1+2\cdot4^2=45$
19	301	tso	$1\cdot4^0+0\cdot4^1+3\cdot4^2=49$
20	302	tsf	$2\cdot4^0+0\cdot4^1+3\cdot4^2=50$
21	310	tos	$0\cdot4^0+1\cdot4^1+3\cdot4^2=52$
22	312	tof	$2\cdot4^0+1\cdot4^1+3\cdot4^2=54$
23	320	tfs	$0\cdot4^0+2\cdot4^1+3\cdot4^2=56$
24	321	tfo	$1\cdot4^0+2\cdot4^1+3\cdot4^2=57$

Таблица 2. Размещения с повторениями

№ п.п.	В пятеричной системе	Вариант размещения	Число «D» в десятичной системе счисления
1	2	3	4
1	00	ss	$0\cdot5^0+0\cdot5^1=0$
2	01	sh	$1\cdot5^0+0\cdot5^1=1$
3	02	so	$2\cdot5^0+0\cdot5^1=2$
4	03	sr	$3\cdot5^0+0\cdot5^1=3$
5	04	st	$4\cdot5^0+0\cdot5^1=4$
6	10	hs	$0\cdot5^0+1\cdot5^1=5$
7	11	hh	$1\cdot5^0+1\cdot5^1=6$
8	12	ho	$2\cdot5^0+1\cdot5^1=7$
9	13	hr	$3\cdot5^0+1\cdot5^1=8$
10	14	ht	$4\cdot5^0+1\cdot5^1=9$
11	20	os	$0\cdot5^0+2\cdot5^1=10$
12	21	oh	$1\cdot5^0+2\cdot5^1=11$
13	22	oo	$2\cdot5^0+2\cdot5^1=12$
14	23	or	$3\cdot5^0+2\cdot5^1=13$
15	24	ot	$4\cdot5^0+2\cdot5^1=14$
16	30	rs	$0\cdot5^0+3\cdot5^1=15$
17	31	rh	$1\cdot5^0+3\cdot5^1=16$
18	32	ro	$2\cdot5^0+3\cdot5^1=17$
19	33	rr	$3\cdot5^0+3\cdot5^1=18$
20	34	rt	$4\cdot5^0+3\cdot5^1=19$
21	40	ts	$0\cdot5^0+4\cdot5^1=20$
22	41	th	$1\cdot5^0+4\cdot5^1=21$
23	42	to	$2\cdot5^0+4\cdot5^1=22$
24	43	tr	$3\cdot5^0+4\cdot5^1=23$
25	44	tt	$4\cdot5^0+4\cdot5^1=24$

Для определения количества размещений с повторениями известна формула /2/

$$\overline{A}_n^k = n^k, \tag{3}$$

где черточка над «A» указывает на то, что размещения берутся с повторениями.

В указанной последовательности n=5 отличных друг от друга элементов. Тогда по формуле (3) получим

$$\overline{A}_5^2 = 5^2 = 25$$

вариантов размещений с допущением повторений.

Для нахождения указанных 25 вариантов размещений каждому элементу присвоим цифру

$$\begin{matrix} s & h & o & r & t \\ 0 & 1 & 2 & 3 & 4 \end{matrix} \tag{4}$$

В этом случае для отыскания всех вариантов размещений и расположения их в определенном порядке приходится применять пятеричную систему счисления (таблица 2).

Следует отметить, что в указных примерах под элементами множеств «soft» и «short» подразумеваются не только конкретные буквы латинского алфавита, но и другие элементы языка: морфемы, слова и т. д.

Из вышесказанного можно сделать следующие выводы:

– если применить системы счисления, основание которых отличается от десяти, то появляется возможность расположения всех вариантов размещений по вполне определенному порядку;

– упорядочение вариантов размещений элементов текста значительно упрощает составление программ для ЭВМ и исключает возможность ошибок (повторений или пропусков вариантов);

– при упорядочении для каждого варианта размещения записывается число «D» в q-ичной системе счисления и переводится в десятичную систему счисления. Упорядочение производится по порядку возрастания числа «D»;

– в случае упорядочения вариантов размещений с повторениями числа «D» представляют ряд натуральных чисел без пробелов, включая и «0» (см. таблицу 2). Если же мы имеем дело с размещениями без повторений или другими ограничениями, обусловленными грамматическими правилами данного языка, то числа «D» являются также натуральными числами, но между ними имеются пробелы (см. таблицу 1).

Список использованной литературы:

1. Фомин С.В. Системы счисления., «Наука», М., 1963.
2. Виленкин Н.Я. Комбинаторика., «Наука», М., 1963., 328 с.