辞書式配列の単語とその番号のアルゴリズム

札幌旭丘高校　中村文則

> Ex) 6 つの文字 A,B,C,D,E,F のすべてを用いて作られる文字列(単語)を ABCDEF からアルファベット順に並べる。次の問いに答えよ。
> (1) 555 番目の文字列は何か。
> (2) CEDFAB は最初から数えて何番目にあるか。

辞書式配列の代表的(頻出)問題であり、順列の場合の数の原理を理解する好材である。その求め方のオートメーション化は、「辞書式配列のちょっとした小手技」で示したが、それを樹形図の枝葉の広がりをイメージしながらシンプルなアルゴリズムに改良した。以下、その方法を述べる。

◯番号が表す文字列

自然数 p を

$$p=\sum_{k=1}^{n} a_k k! \qquad (0 \leqq a_k \leqq k\ , k \in Z\ )$$

の形に分解することを、階乗展開とよぶことにする。

$(n+1)$ 個の文字列の階乗展開は、 $p \leqq (n+1)!$ であるから、 p を $n!$ で割った余りを r_1 とし、次に r_2 を $(n-1)!$ で割り、以下同様にその余りを階乗で割ることで求められる。

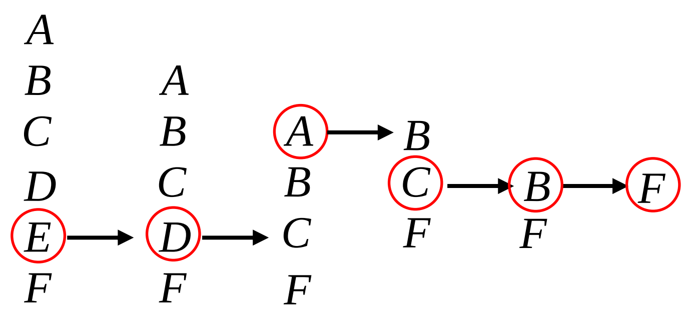
例えば、555 は、$555=4\times5!+75$ より、$r_1=75$。以下、$75=3\times4!+3$ 、$3=1\times2!+1$ となるから、次のように分解することができる。

$$\begin{array}{r|l} & 555 \\ \hline & 75 \\ \hline & 3 \end{array} \quad =4\times5!+3\times4!+0\times3!+1\times2!+1\times1!$$

この分解が表す文字列を次の手順により調べる。

① アルファベット順の 1 番目の文字列を縦に書く
② 階乗展開の $\alpha\times n!+\beta$ の部分を見て、文字列 $\alpha+1$ 番目にチェック(◯)を入れる。
③ チェックした文字以外の文字列を、右側に縦に書く。
④ ②③を $\beta\neq0$ の間続ける。
⑤ $\beta=0$ のとき、α 番目にチェックを入れ、文字列が残っている場合は、アルファベット順の最後の文字列を書く。

555 は次のように表される。



そして、チェック(◯)のついた文字を左から追っていくと求める文字列が得られる。

Ex) 階乗展開は、n 進法変換の要領で、 p を 1,2,3,4,5,⋯で割り、その商と余りを求めていくことで鮮やかに得ることができる。この方法は、札幌旭丘高校の菅原先生が考案したものである。

$$\begin{array}{r|rl} 2 & 555 & \\ \hline 3 & 277 & \cdots 1 \\ \hline 4 & 92 & \cdots 1 \\ \hline 5 & 23 & \cdots 0 \\ \hline & 4 & \cdots 3 \end{array}$$

$$\begin{aligned} 555 &= 2\times277+1 \\ &=2\times(3\times92+1)+1 \\ &=2\times3\times(4\times23+0)+2\times1+1 \\ &=2\times3\times4\times(5\times4+3)+2\times3\times0+2\times1+1 \\ &=4\times5!+3\times4!+0\times3!+1\times2!+1\times1! \end{aligned}$$

◯文字列が表す番号

前述の方法を用いて、与えられた文字列がアルファベット順に数えて何番目にあるかについても調べてみよう。

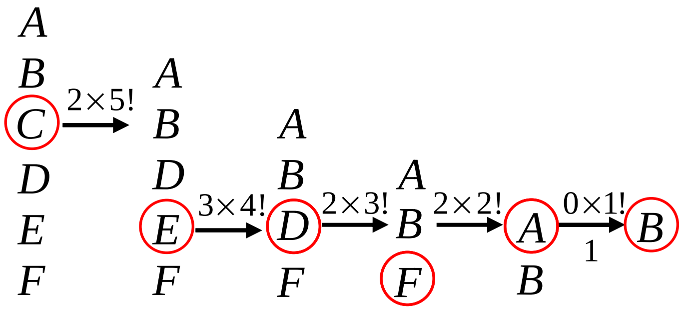
文字列(単語)CEDFAB を例にとる。

①アルファベット順の 1 番目の文字列を縦に書く

② 文字列の文字(左から順に)にチェック(◯)を入れる。

③ チェックを入れた文字列の上にある文字数をαとし、チェックを入れた文字以外の文字数をnとし、
$\alpha \times n!$を計算する。

④ ②③を繰り返す。文字列の最後の文字のときは、1 を加える。



$p = 2\times5! + 3\times4! + 2\times3! + 2\times2! + 0\times1! + 1 = 329$ 　より、CEDFAB は 329 番目にある。

◯アルゴリズム原理の理解

この方法では「◯と⇒」だけを用いて操作を繰り返すことで目的の番号や文字列に到達することができるが、そのアルゴリズムを理解しないで機械的な処理で終わってしまうことは避けたい。

辞書式配列の原理は単純であり、6 個の文字の配列であれば、$6! = 720$ を表す樹形図から求める文字列を含む辞典の枝葉の部分を拡大(Zoom Up)していくだけである。言い換えると、$6! = 6\times5!$ とみることで、元の辞典を索引が A から F までの同じ辞書数$5!$である 6 冊の辞典に分冊することになる。次にその 1 冊を$5! = 5\times4!$より、辞書数$4!$の辞典 5 冊に分冊する。この操作を続けていくことがアルゴリズムなのである。

このように、辞書数$(n+1)!$である辞典は辞書数$n!$である$(n+1)$冊に分冊され、文字列の番号が

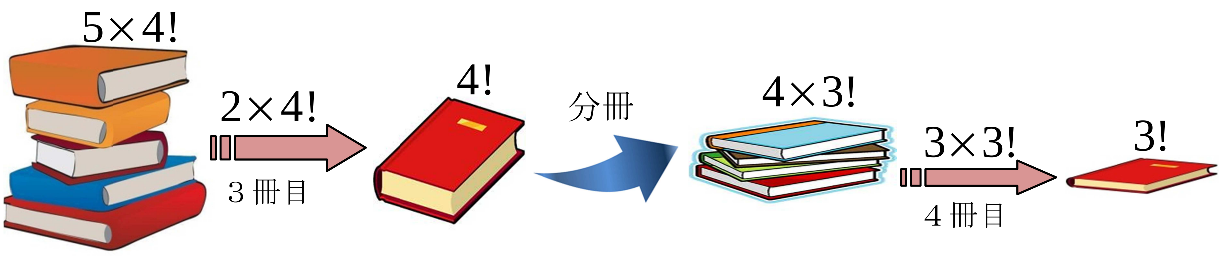
$$\alpha\times n! + \beta \ (0 \leqq \alpha \leqq n,\ 0 \leqq \beta < n!,\ \alpha, \beta \in Z)$$

で表されるならば、$(n+1)$冊の分冊に対して

$\beta \neq 0$ のとき、α 冊の辞典が数え終わり、$\alpha+1$ 冊目の辞典の β 番目の文字列

$\beta = 0$ のとき、α 冊目の辞典の最後の文字列

となる。そして、$\beta \neq 0$ のときは、文字列を含む辞典をまた分冊すればよい。例えれば、視点を、樹木の幹から枝、葉へと移していくのである。



この方法は、アルゴリズムの思考補助であり、Zoom Up した樹形図を書き並べて解答を終わらせるべきではない。樹形図から解答を起こすことが必要なのである。

番号から文字列を調べるには、 $p = 555$ のときは、 $p = 4\times5! + 3\times4! + 0\times3! + 2\times2! + 1\times1!$ より、

A□□□□□,　B□□□□□, C□□□□□, D□□□□□の文字列の総数は$4\times5!$(480)

EA□□□□, EB□□□□, EC□□□□の文字列の総数は$3\times4!$(480+72=552)

ここまで数えれば、アルゴリズムに委ねる必要もない。ED□□□□の最初から 3 番目が求めるものであるから、

EDABCF,　EDABFC,　EDACBF

逆に、文字列から番号を調べるには、CEDFAB のときは、

A□□□□□, B□□□□□の辞書数は$2\times5!$ (総数は 240)

CA□□□□, CB□□□□,CD□□□□の辞書数は$3\times4!$(総数は 240+72=312)

CEA□□□, CEB□□□の辞書数は$2\times3!$(総数は 312+12=324)

CED□□□の辞書数は$3!$であり、CEDFAB は最後の辞書の 1 つ前であるから、324+5=329

柔軟な思考はアルゴリズムに勝るものであり、これを養うことを目標とすべきなのである。

◯アナグラムを楽しむ

有意単語(意味のある単語)の文字の順番を入れ替えて、別の有意単語を生成することをアナグラム(anagram)という。

女優である「ともさかりえ」は歌手としては「さかともえり」として活動していた。小説では、ドラキュラ(Doralura)の使徒名はアルカード(Alucard)である。数学者チャールズ・ドジスンは巧妙なアナグラムを組み、ルイス・キャロルとして文壇に登場し名作を著した。

辞書式配列の問題はもともとの単語が有意単語であれば、辞書の中にアナグラムがある可能性はある。アナグラムを楽しむと考えればこの問題は生徒にはとても興味あるものになるのではないだろうか。

Lovers(恋人たち)を並べ替えると Solver(解法)である。「恋人たちはお互いに心を解き合うものである」として「辞書式配列の小手技」では、Solver の番号(538)および 163 番目の単語(Lovers)を求めている。

Listen(聞く)のアナグラムは Silent(沈黙)。「生徒に聞いても沈黙が返ってくるだけ」、あるいは「沈黙しているくらいなら聞け」とでもしてみる。(このアナグラムは本校の定期試験で出題した)。

日本語のアナグラムはさらに効果的である。

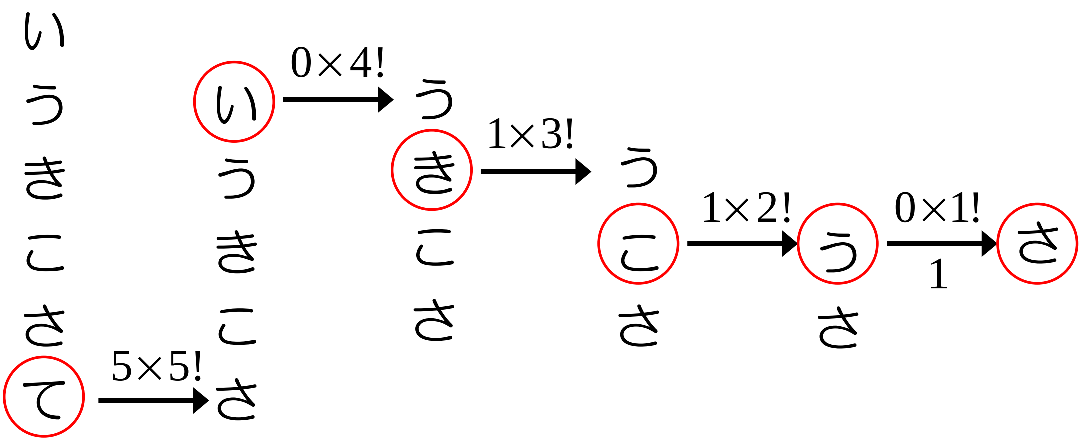
特に、人の名前はインパクトはあるが、「名前で遊ぶ」ことだから、対象を生徒にするとアナグラムがニックネーム(アダ名)になってしまうこともある。先生や有名人の名前で遊ぶのが無難であろう。

例えば、「まえだあつこ」(前田敦子)は知っているよね、と問いかける。では、その 566 番目はどんな単語になるだろう。生徒はアナグラムの何かの単語を期待するはずだ。正解は「だつこあまえ」(抱っこして甘え)である。
このように授業を展開していけば、生徒はきっと夢中になって調べるのではないだろうか。
ひとつ例を挙げて、このレポートのまとめとする(本校定期考査の問題候補として考えたものである)。

> Ex) 「ていきこうさ」の 6 つの文字をすべて使って五十音に並べる。次の問いに答えよ。
> (1)「ていきこうさ」は何番目の単語か。
> (2) 356 何番目の単語は何か。

解) ここでは樹形図のみを示す。1 番目の単語は「いうきこさて」である。

(1)

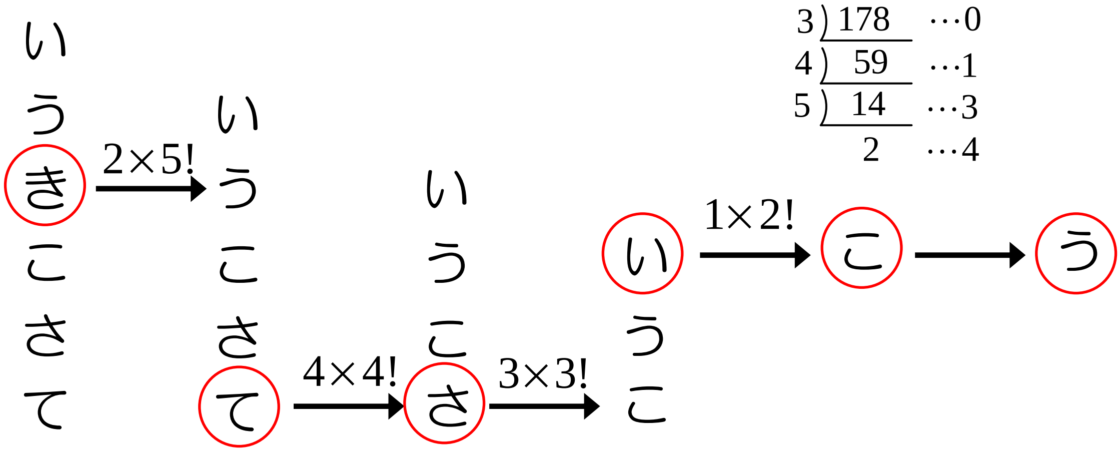


$5\times5!+0\times4!+1\times3!+1\times2!+0\times1!+1=609$

「ていきこうさ」は 609 番目の単語である。

(2) $356=2\times5!+4\times4!+3\times3!+1\times2!+0\times1!$ である。

2) 356
3) 178 …0
4) 59 …1
5) 14 …3
2 …4



「ていきこうさ」(定期考査) のアナグラムは「きてさいこう」(来て、最高)。努力が報われるといいのだが。