

# 挿入ソートの解析

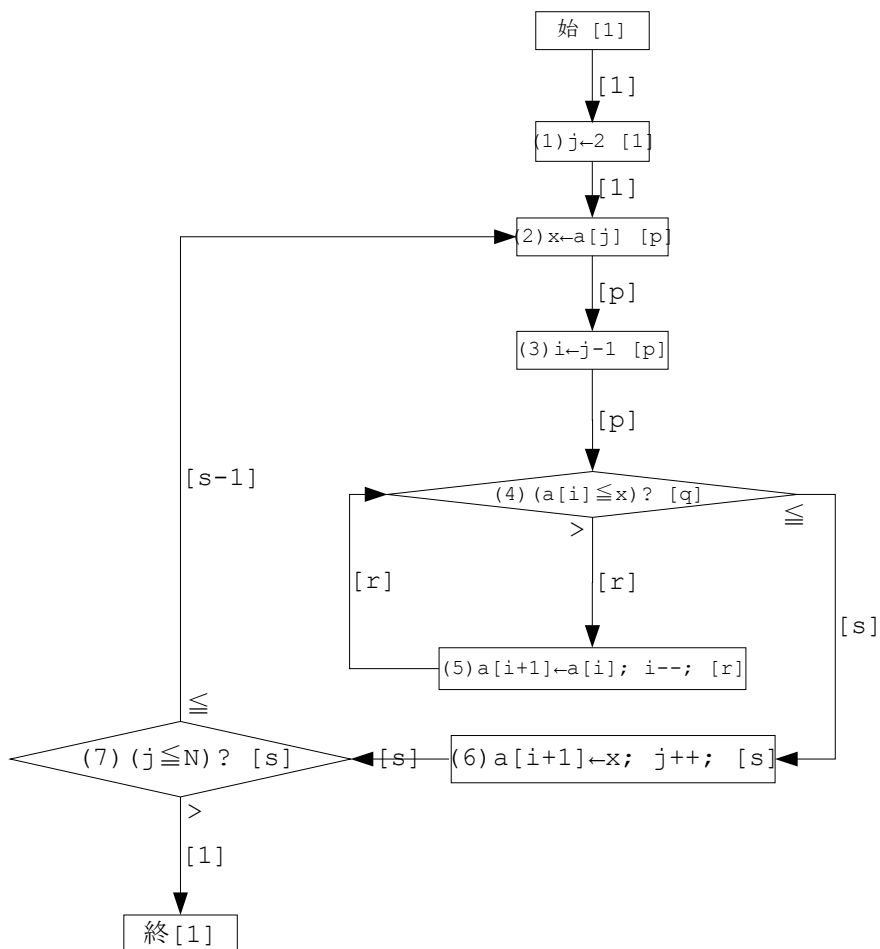
環境情報学部 4 年  
石原 和音 (Kazuto Ishihara)  
[t00069ki@sfc.keio.ac.jp](mailto:t00069ki@sfc.keio.ac.jp)  
学籍番号 70050607

## プログラムの内容

以下がプログラムである。左から順に、プログラムのステップ、その実行回数、ラベル、実行内容となっている。

- (1)<sub>[1]</sub>  $j \leftarrow 2;$
- (2)<sub>[p]</sub> **L1:**  $x \leftarrow a[j];$
- (3)<sub>[p]</sub>  $i \leftarrow j-1;$
- (4)<sub>[q]</sub> **L2:**  $\text{if } (a[i] \leq x) \text{ goto L3};$
- (5)<sub>[r]</sub>  $a[i+1] \leftarrow a[i]; i--;$  **goto L2;**
- (6)<sub>[s]</sub> **L3:**  $a[i+1] \leftarrow x; j++;$
- (7)<sub>[s]</sub>  $\text{if } (j \leq N) \text{ goto L1};$

## 流れ図



## パラメータ同士の関係

パラメータ  $p, q, r, s$  について考える。

1. まず、(2)と(7)に注目する。  $j$  と  $N$  を比較している(7)は、  $[s]$  回実行される。ここから、プログラムの終了へと向かう  $[1]$  回を除く、  $[s-1]$  回が(2)へ向かう。  
(2)は  $[p]$  回実行される。ここに流れ込むのは、先ほどの(7)からの  $[s-1]$  回に加えて、(1)から流れ込む  $[1]$  回である。よって  $(s-1)+1=p$  すなわち  $s=p$  である。
2. 次に、変数  $j$  に注目する。  $j$  には最初、  $2$  の値が入り、(6)によって  $[s]$  回インクリメントされ、終了時には  $j > N$  である。この場合、  $j$  は  $2$  から  $N+1$  まで変化する。それが  $[s]$  回インクリメントされた結果なのであるから、  $s=(N+1)-2$  つまり  $s=N-1$  である。
3. 次に、(4)に注目する。(4)は  $[q]$  回実行される。流れ込んでいるのは、(3)からの  $[p]$  回と(5)からの  $[r]$  回である。流れ出ているのは、(5)への  $[r]$  回と(6)への  $[s]$  回である。よって、  $q=p+r=r+s$  である。

以上より、パラメータ  $p, q, r, s$  は独立ではないこと、またそれぞれの関係が分かった。まとめると

$$p=N-1$$

$$q=r+N-1$$

$$r=q-N+1$$

$$s=N-1$$

となる。4つのパラメータは、  $q$  または  $r$  の1つのパラメータによって簡単に表すことが出来ることがわかる。

## パラメータの意味

### パラメータ $s=p$ の意味

パラメータ  $s, p$  は、ソート済みの領域に、新たな要素  $x$  を挿入する回数を表している。  $N$  の長さの配列において、まず、最初の  $1$  個の要素をソート済みと見なし、次からは、残りの  $N-1$  個の要素を  $s=p$  回挿入するのである。よって、  $s=p$  は、最小値・最大値・平均値ともに常に  $N-1$  である。

### パラメータ $r$ の意味

パラメータ  $r$  は、ソート済みの領域に  $x$  を挿入するとき、挿入のために(ソート済みの領域の一部の)要素をシフトする回数である。

$r$  の最大値は、  $x$  が常に、ソート済みの全ての要素より小さかった場合で、シフトはソート済みの領域の要素全てに対して実行される。よって、シフトの回数は  $\sum_{k=1}^{N-1} k = \frac{(1+(N-1))(N-1)}{2} = \frac{N^2-N}{2}$  で、これが  $r$  の最大値である。

$r$  の最小値は、  $x$  が常に、ソート済みの全ての要素より大きかった場合で、この場合、シフトは不要なので  $r=0$  が  $r$  の最小値である。

平均的なシフトの回数は、ソート済みの要素の約半分が  $x$  より大きく、残り

半分が  $x$  より小さいだろうと予想できるので、粗く見積もって、

$$\sum_{k=1}^{N-1} \frac{k}{2} = \frac{(1+(N-1))(N-1)}{4} = \frac{N^2-N}{4} \text{ が } r \text{ の平均値である。}$$

### パラメータ $q$ の意味

パラメータ  $q$  は、(4)の実行回数であり、これは、新たに挿入される要素  $x$  と、ソート済みの領域の要素との比較の回数である。 $q$  は  $r$  とほぼ同様であるが、挿入の度に、シフトの回数+1回の比較が必要となるので、 $q=r+\text{シフトの回数}=r+N-1$  である。これは、パラメータ同士の関係から  $q=r+N-1$  が言えたこととも合致する。

$q$  の最大値は  $\frac{N^2-N}{2}+(N-1)=\frac{N^2+N-2}{2}$ 、 $q$  の最小値は  $0+(N-1)=N-1$ 、 $q$  の平均値は、 $r$  と同様に粗く見積もって

$$\sum_{k=1}^{N-1} \left(\frac{k}{2}+1\right) = \frac{(1+(N-1))(N-1)}{4} + (N-1) = \frac{1}{4}(N^2+3N-4) \text{ である。}$$