アルゴリズムとデータ構造b
13 - 計算量

大見　嘉弘

今日の内容

- 計算量
  - 計算量とは
  - 計算量の例(基本交換法)
  - O記法
  - 処理時間の比較
  - 計算量の例(探索)

計算量とは

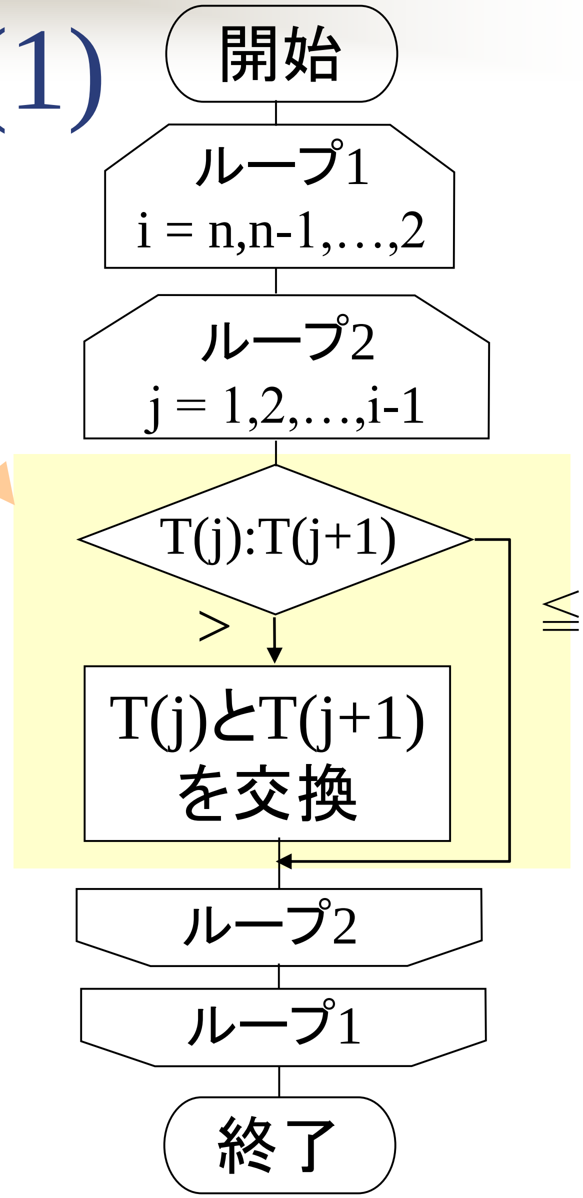
- アルゴリズムの良し悪しを客観的に測る指標

- 実行に必要な計算機資源(時間、メモリ等)を比較
  - 時間計算量　実行に要する時間を評価
  - 空間計算量　必要とするメモリ量を評価
- 時間計算量のみを評価することが多く、単に計算量というと時間計算量を指す場合もある

計算量の例(基本交換法)(1)

- 基本交換法の処理時間
  - ループ2を何回繰返すかで決まる
  - データ数をnとすると、
    i=nの時、n-1回
    i=n-1の時、n-2回
    …
    i=3の時、2回
    i=2の時、1回
  - 合計で $(n-1)+(n-2)+\ldots+2+1$

$$= \sum_{i=1}^{n-1} i = \frac{(n-1)n}{2} = \frac{1}{2}n^2 - \frac{1}{2}n$$



計算量の例(基本交換法)(2)

- ループ2の繰り返し回数

$$\boxed{\frac{1}{2}n^2 - \frac{1}{2}n} \rightarrow \boxed{\frac{1}{2}n^2} \rightarrow \boxed{n^2}$$

最も大きい(高次の)項のみ残す　　定数倍率を省略する

- 計算量を評価する場合は、このように最も高次の項のみで倍率を省略したものを比較する
- ここで、基本交換法はn^2のオーダーといい、$O(n^2)$と書く

O記法 (ビッグオー記法)

- 実行回数(時間)を大まかに表現する指標
- あるアルゴリズムを実行する時間(実行回数)が、an^3+bn^2+cn+d (a,b,c,dは定数)だとすると、

1. もっとも高次の項だけを選ぶ。つまり an^3
   (nが増えてくると、an^3以外の項の影響が減ってくる
   →無視できるようになる)
2. 定数倍率を省略する。つまり n^3 とする。
   そのアルゴリズムの時間計算量はO(n^3)である。
   (定数倍にはあまり意味がない。例えば、アルゴリズムAがBより常に2倍速いとしても、コンピュータが2倍速くなれば、Bは以前のAと同じ時間で実行できる)

O記法の例

速

- $O(1)$ ハッシュ法(衝突なしの場合)
- $O(\log n)$ 二分探索
- $O(n)$ 線型探索、ビンソート
- $O(n \log n)$ クイックソート、シェルソート
- $O(n^2)$ 基本交換法、基本選択法、基本挿入法
- $O(n^3)$
- $O(n^k)$

- $O(2^n)$,$O(n!)$ 巡回セールスマン問題

遅

O表記におけるlog

- 通常2を底とする対数とする
- つまり O(log n)はO($\log_2 n$)
- 結局のところ$\log_2 n = \log n / \log 2$
なので、O(log n) = O($\log_2 n$)である。
- 対数のおさらい
$n = 2^x$ の場合、$x = \log_2 n$

処理時間の比較(1)

$\log n$	n	$n \log n$	n^2	2^n
0	1	0	1	2
3.3	10	33.2	100	1024
4.3	20	86.4	400	1048576
5.6	50	282.2	2500	1.13×10^{15}
6.6	100	664	10000	1.27×10^{30}
9.9	1000	9966	1000000	1.07×10^{301}
13.3	10000	132877	100000000	2.00×10^{3010}

105秒

3570年

10^{20}年

8×10^{290}年

1秒とすると…

13秒

1秒

2時間47分

100秒

処理理時間の比較(2)

- n=100の時の処理時間を1とした時の倍率

n	$\log n$	n	$n \log n$	n^2	n^3	2^n
1	0	0.01	0	0.0001	0.000001	1.57×10^{-30}
10	0.5	0.1	0.05	0.01	0.001	8.06×10^{-28}
100	1	1	1	1	1	1
1000	1.5	10	15	100	1000	8.43×10^{270}
10000	2.0	100	200	10000	1000000	1.57×10^{2980}

計算量の例(探索)(1)

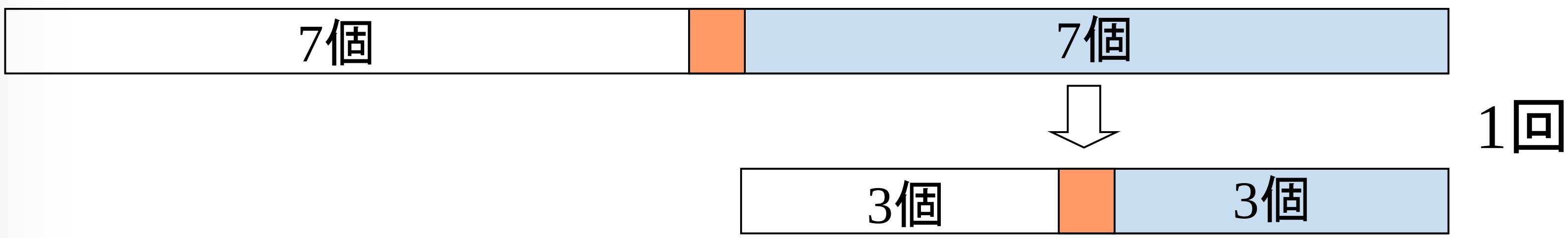
- 線型探索
  - 最悪の場合:n回
    (探索失敗(全てのデータと比較))
  - 平均: n/2回
  - したがって線型探索の計算量はO(n)

計算量の例(探索)(2)

- 二分探索
  - 1回の繰り返しで対象データ数が半分になる。　$n \to n/2 \to n/4 \to n/8 \to n/16 \to n/32 \to \ldots$
    (厳密には次のスライドのように半分の値の小数点以下を切り捨てた数となる)
  - x回繰返した時の対象データ数は、$n/2^x$
  - x回繰返したら対象データ数が1になったとすると、$n/2^x=1$ これからxを求める
  - $2^x = n$ であるから $x = \log_2 n$
  - つまり二分探索の計算量はO(log n)である。

計算量の例(探索)(3)

データ数が15個の場合の二分探索



- データ数8〜15まで最悪4回
(log$_2$8=3, log$_2$15=3.91)
- データ数16〜31まで最悪5回
(log$_2$16=4, log$_2$31=4.95)
- つまり繰り返し回数は、$[\log_2 n]+1$となる。
(ただし[x]はxの小数点以下を切り捨てた整数値)