アルゴリズムとデータ構造

http://www-ui.is.s.u-tokyo.ac.jp/~takeo

五十嵐 健夫

takeo@acm.org

目標

効率のよいプログラムを書くため
基本的な知識・技法を学ぶ

効率＝実行時間とメモリ使用量

コンピュータサイエンスの基礎。

具体的内容

基本的なアルゴリズムとデータ構造を学ぶ

ソート、グラフ、検索、など

計算量の解析について学ぶ

計算量の意味、その計算方法、など

進め方

教科書に沿う

情報処理シリーズ11 データ構造とアルゴリズム
A.V. エイホ, J.E. ホップクロフト 著, 大野義夫訳
培風館 (ISBN4-563-00791-9 C3355)

基本的に教科書が理解できればOK。

成績

毎回出席を取る。

期末テスト (1/25 予定)

ネット上に過去問あり

スケジュール（仮）

10/11 **アルゴリズムの設計と解析 擬似言語、実行時間**
10/18 **基本的な抽象データ型 リスト, スタック, 待ち行列**
10/25 (**休講**)
11/1 **ソート 前半 バブルソート, 内挿ソート, 選択ソート, クイックソート, その時間解析**
11/8 **ソート 後半 ヒープソート, ビンソート, 基数ソート, マージソート** (**シェルソート**)

11/15 **木 二分木, ハフマンコード**
11/22 **集合の基本操作 ハッシュ, ハッシュの効率, 優先度付き待ち行列**
11/29 **集合の高度な表現方法 2分探索木, その時間解析, トライ**, 2-3**木**, MFSET, LCS
12/6 **有向グラフ ダイクストラ, フロイド**, DAG, **強成分**
12/13 **無向グラフ プリム, クラスカル, 関節点, 最大マッチング**
12/20 **解析法 再帰方程式の解法** (**動的計画法**)
1/10 **設計法 欲張り、α－β枝刈り、分岐制約探索法、**
1/17 [**テスト**]

演習との関係

アルゴリズムとデータ構造演習
火曜日４限

Ｃ言語による講義の演習
＋
関数型言語の勉強

自己紹介

五十嵐　健夫　　情報科学科　助教授

専門：　ユーザインタフェース
　　インタラクティブＣＧ

電子ホワイトボード、電子カルテ
３次元モデリング、アニメーション
音声によるインタフェース　など

アルゴリズムとは

問題を解く手順

1) **問題を定式化（モデル化）する**
2) **解法をアルゴリズムとして記述する**
3) **アルゴリズムにしたがって問題を解く**

定義：

「明瞭な意味を持ち、有限時間内の有限な計算で実行できるような命令を有限個並べた形で記述される問題の解法」

アルゴリズムとは

モデル化の例（交差点）

- **グラフ構造の利用**
- **配色問題**
- **発見的アルゴリズム**

段階的詳細化の例（配色問題）

- **文章で書いたアルゴリズム**
- **擬似言語のプログラム**
- **プログラミング言語による記述**

抽象データ型

数学モデルと操作の組

（手続きをカプセル化する）

整数：　加減乗除
集合：　和・積・差集合
リスト：　空にする、取り出す、追加する

プログラムの実行時間

2つの目標

- **わかりやすい。構造がシンプルである。**
- **実行時間が早く、メモリを消費しない。**

実行時間を決める要素

- **入力データの性質・大きさ**
- **コンパイラの質**
- **ハードウェアの性質**
- **アルゴリズムの計算量**

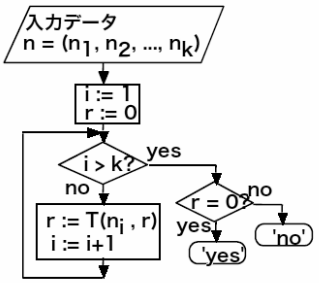
アルゴリズムによって計算時間が変わる例

線形探索と2分探索　n vs $\log_2 n$

3の倍数判定　　n vs $\log_{10} n$

3の倍数判定　　時間＝$\log_{10} n$ (桁数に比例)

フローチャート



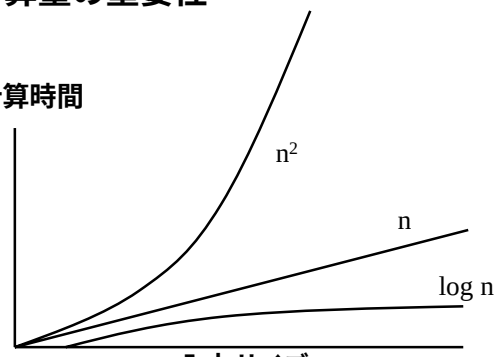
※除算を用いずに解く方法

剰余表 $T(n_i, r)$

n_i \ r	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1
3	0	1	2
4	1	2	0
5	2	0	1
6	0	1	2
7	1	2	0
8	2	0	1
9	0	1	2

計算量の重要性

計算時間



入力サイズ　n

計算量の重要性

n=100のときの時間を1[sec]として計算 ($O(2^n)$を除く)

計算量 \ n	10	50	100	500	1000	10000
O(1)	1	1	1	1	1	1
O(log n)	0.5	0.85	1	1.35	1.5	2
O(n)	0.1	0.5	1	5	10	100
O(n log n)	0.05	0.42	1	6.75	15	200
$O(n^2)$	0.01	025	1	25	100	2.78[h]
$O(2^n)$	0.001	34.8[y]	–	–	–	–

アルゴリズムの選択が重要（秒ー時間ー年）。
ハードウェア・コンパイラ・チューニングなどは小手先。

オーダー記法

「問題のサイズnに対して、予測される計算量の上界値を、定数部分を省略して表現する方法」

正確には、
「アルゴリズムの実行時間がO(f(n))である」とは
「正の定数c, n_0が存在して、n_0以上のnに対しては
$T(n) <= c\ f(n)$となる。」
ただし $T(n)$は大きさnの入力のプログラムの実行時間

（Ωは下界）

オーダー記法

例：　線形探索　$T(n) = an + b$　　…. O(n)
　　2分探索　$T(n) = a \log n + b$　　… O(log n)

　　3の倍数判定　割り算　$T(n) = an/3+b$ … O(n)
　　除算なし　$T(n) = a \log n + b$　　… O(log n)

オーダー記法

$O(1)<O(\log n)<O(n^a)<O(n \log n)<O(n^b)<O(\alpha^n)<O(n!)$

$0<a<1,\ 1<b,\ \alpha>0$

プログラムの実行時間

和と積の法則

$T_1(n) + T_2(n)$ … $O(\max(f_1(n), f_2(n))$

$T_1(n) \times T_2(n)$ … $O(f_1(n) \times f_2(n))$

解析は難しいことも多い。

いくつかの規則

- 一連の文は和の公式 ＝ 最も遅い部分に依存
- if 文は、長い方に依存
- ループは、ループの回数と最長の内部実行時間の積
- 再帰手続き＝再帰方程式を解く

アルゴリズムの選択の注意点

使用回数	多い場合にはオーダーに注意
入力サイズ	大きい場合にはオーダーに注意
保守	保守が必要なら読みやすさ優先
メモリ	外部記憶が使えるか
安定性、精度	数値アルゴリズムで重要

よいプログラミングの習慣

計画的に設計する。 段階的詳細化。

オーダーを意識する。

カプセル化・モジュール化する。

既存プログラムを活用する。

汎用性のある・応用の利くコードを書く。

まとめ

講義の進め方
アルゴリズムとは
モデル化と段階的詳細化
計算量の話 オーダー記法