

計算量

計算資源

- コンピュータ上の計算では次のものを資源（リソース）として利用する
 - CPU (Central Processing Unit)
 - プログラムを順に読み込んで解釈して実行する
 - メモリ（記憶装置）
 - プログラムや変数などを記憶する

➡ 「計算の規模」はその計算に必要な時間（CPUに関係）とメモリで表現する

計算時間

- 実時間
- CPU時間：プログラムが処理するためにCPUを占有した時間の累計
 - 1GHzのCPU ➡ 1秒間に $1 * 1000 * 1000 * 1000$ 回の振動
(K: キロ, M: メガ, G: ギガ, T: テラ, P: ペタ)
 - 整数の足し算が1クロック ➡ $1 \text{クロック} \times 1 / 1,000,000,000 = 1 \text{ ナノ秒}$
(m: ミリ, μ : マイクロ, n: ナノ, p: ピコ)
- ユーザCPU時間：アプリケーションが使うCPU時間
- システムCPU時間：システムが使うCPU時間

- CPU性能指標

- MIPS (million instructions per second): 1秒間に何百万回の命令が実行できるか
- FLOPS (floating point number operations per second) : 1秒間に何回の浮動小数点演算が実行できるか

時間計算量

- 実時間やCPU時間による評価はCPUなどの要因に依存
- 計算手順（アルゴリズム）そのものの評価をしたい
 - 時間計算量
 - 入力サイズ n が入ってきたときの処理にかかる手数（ステップ数）

- ある人の年齢を当てる問題
 - 質問できる内容： x 歳以上ですか？
 - 年齢の幅は 0 歳から 100 歳とする
- 何回の質問で当てることができるか？

- 素朴なやり方
 - 1 から 100 まで順番に聞いていく
 - 1 歳以上ですか？
 - 2 歳以上ですか？
 - ...
 - 99 歳以上ですか？
 - 100 歳以上ですか？
- 問題
 - 最小の質問回数は？
 - 最大の質問回数は？

- 少し賢い方法
 - 半分ずつ候補を絞るように聞いていく
 - 50歳以上ですか？
 - Yes: 50-100の間
 - No: 0-49の間
 - 50歳以上ですか？
 - 24歳以上ですか？
 -
- 問題
 - 最小の質問回数はいくつ？
 - 最大の質問回数はいくつ？

- 問題

- 鶴は千年，亀は万年生きると言われています
(注意：実際はどちらも長生きはしますがそこまで長く生きません)
- 鶴と亀の年齢を当てる問題で最悪の場合，素朴な方法だと人間の何倍？
- 鶴と亀の年齢を当てる問題で最悪の場合，少し賢い方法だと人間の何倍？

- 時間計算量

- 入力サイズが増えたときに計算ステップ（通常は最悪の場合の計算ステップ）が何倍くらいになるか？
- 年齢当て問題の場合
 - 入力: 何歳まで考えるか（候補となる年齢の数）
 - 計算ステップ：何回質問するか

- オーダー，ランダウの記号
 - ある値（計算の場合は無限大が多い）に近づけた場合に影響のある項

$$f(x) = O(g(x)), \quad \lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = \text{定数}$$

- 例

$$f(x) = 2x^2 + 3x + 1 = O(x^2) \quad \text{as } x \rightarrow \infty$$

- 定数倍や x^2 よりも次数の低い項を無視していることに注意！

- 問題

- 年齢当て問題で素朴な方法の時間計算量をランダウの記号（オーダー）を用いて表せ
- 年齢当て問題で少し賢い方法の時間計算量をランダウの記号（オーダー）を用いて表せ

なぜ計算量が必要？

- データが増えた時にかかる計算の見積もり
- n 個のデータがあるときに単回帰分析に必要な時間計算量: $O(n)$
 - X, Y の平均: $O(n)$
 - X, Y の分散, 共分散: $O(n)$
 - 100万個のデータに対する単回帰分析は100個のデータの単回帰分析の1万倍
 - $1 \text{ (ms)} \times 10000 = 10 \text{ (sec)}$
- n 個のデータに対するある分析: $O(n^2)$
 - 100万個のデータに対するある分析は100個のデータのある分析の1万 × 1万倍
 - $1 \text{ (ms)} \times 10000^2 = 100,000 \text{ (sec)} \approx 28 \text{ (hours)}$
- 巨大なデータの分析
 - ▶ 時間計算量および空間計算量のオーダーが低いアルゴリズム