

第三回課題

sample programs

trie.c

test.awk

レポート・成績について

- ほぼ毎回、プログラミング課題を出題する予定
 - 効率の良い計算機実験のためのツールを使ってみる
 - アルゴリズムの実装
 - ライブラリの利用…など
- 3回以上、レポートのファイルとプログラムのソースを添付し、メールで提出のこと
 - E-mail: algorithm2014@edu.jar.jp
 - サブジェクト 「アルゴリズムとプログラム実践講座・レポート」
 - 学生証番号と名前は、メールの本文にも書いてください。
 - メル:次の授業の直前の日曜日深夜（講評の都合上。メル後も受付ます）
 - プログラムやレポートは（見本として）公開することができます。適宜、作者名やコピーライトをいれておいてください。公開不可の場合は、プログラムの冒頭にその旨、コメントをいれておいてください。
 - 質問・作問提案も歓迎（作問については採用の場合は別途加点）
 - サンプルプログラムは「初心者向け」です。上級者は無視してください。

推奨環境など

- Linux, Mac, (Windows+Cygwin)
- 仮想マシン環境(VMware, VirtualBox, Parallels)
 - 余裕があれば、いろいろな組み合わせを試して比較してみると面白いと思います
- 言語
 - 自由。ただし、一般的でない言語については、上記いずれかのOS 上にインストール可能なものの

第三回 課題(その1)

Sample programs

trie.c test.awk

単語数を数える

たとえば「プロジェクト・グーテンベルグ」や
「青空文庫」から、書籍のデータを収集し、
必要があれば、茶筅等、適当な前処理を行い
出現した単語のヒストグラム等を作成する。

Hash (and/or) Trie を実装、連想配列など
言語が用意する物と性能比較をする。

作成したヒストグラム等から、なんとか面白い
結果をひねり出してみる。☺

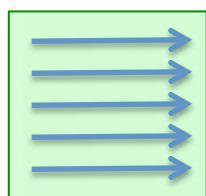
この課題の狙うところ

- Trie か hash か（両方でも♪）実装してみる
- 連想リスト等と性能を比べてみる
- 現実の大きめなデータを扱ってみる
(予想外に、面倒な処理が発生することもある)
- 大きなデータの中に、何か面白い性質が潜んでいないか探してみる

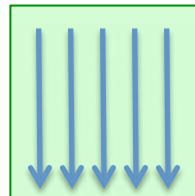
第三回 課題(その2)

先週の課題：行列のアクセスの時間の計測

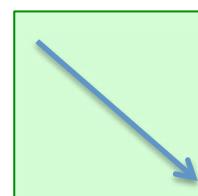
- 正方形に、(a)行順全要素、(b)列順全要素、(c)対角要素のみ、(d)1行のみ (e)1列のみ代入を行ったあと、代入した要素の読み出しを行うプログラムを書き、かかった時間を計測、比較考察せよ。
- 実験環境のうち、ハードウェア、OS、言語、コンパイラオプションなど実験環境のうち、いくつかを変更して、比較考察せよ。
- 実験環境について、CPU のバージョン・動作周波数・メモリサイズ・カーネルのバージョンなど、他人に再現可能なように記述すること。**



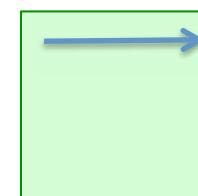
(a)



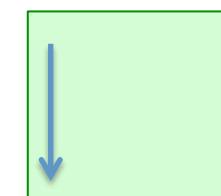
(b)



(c)



(d)



(e)

先週のレポート+α

出題者の本音

もうひといき ♪

どなたか、読んで面白いレポート
書いてくれないかなあ…

サンプルコードに関する注意

- 授業の際に提示するコードはヒントであって必ずしも良いコードでないことが多い。決してお手本というわけではない。
- たとえば、第二回の Java のサンプルは、出力にかかる時間が大きく何を計りたいかわからなくなりがち
- たとえば html への書き出しは、分散環境で複数人で使うことを真面目に考えるときは、工夫しないといけない。