

# アルゴリズムとデータ構造

グループワーク

# グループワークのねらい

- 計算機のリソースを意識した問題解決に取り組む.
- グループ内のメンバーと様々なアイデアを交換しながら問題に取り組む.
- 手法の設計の意図, その実現方法を的確に説明できる力を養う.

# スケジュール概要

- ガイダンス，議論 2023/12/22 15:05 – 16:45
  - 開発内容の説明，メンバー顔合わせ，議論
  - 対面により実施（63-202教室）
- 開発・発表準備期間 2023/12/22～2024/1/19
- 開発議論，発表準備 2024/1/12 15:05 – 16:45
  - 対面により実施（63-202教室）
- 発表会 2024/1/19 15:05 – 18:40
  - 対面により実施（63-202教室）
- 報告書等の提出〆切 2024/1/26

※ 2024年1月19日5限は，別の授業と重なっている人の参加は免除します。

# 評価方法

- グループ全体の評価
  - 成果物の性能
  - 成果発表の質
- 個別の評価
  - グループワーク中の各人の貢献度

# 評価方法

- グループ全体の評価

- 成果物の性能

- 教員が計測 → 成果発表会の日に公開

- 成果発表の質

- 皆さんが互いに評価, 教員の評価

- 個別の評価

- グループワーク中の各人の貢献度

- 皆さんが互いに評価
    - 報告書

# 評価方法

## ● グループ全体の評価

### ○ 成果物の性能

● 教員が計測 → 成果物の性能

### ○ 成果発表の質

● 皆さんが互いに評価, 教員の評価

#### 評価のポイント:

- 発表はわかりやすかったか？
- 提案手法は妥当だったか？
- 提案手法を正しく実装できていたか？
- 結果の分析は妥当か？

## ● 個別の評価

- 各人は, Moodleの“発表評価”から自分が所属していないグループの中で特に良かった発表に投票する. 自グループへの投票は無効.
- また, なぜ投票したのかコメントも書く.
- コメントは集計後, 皆さんにシェアします.

# 評価方法

- グループ全体の評価
  - 成果
    - 教
    - 成果
    - 皆
  - 各人は、Moodleの“作業評価”からメンバーの貢献度を評価します。
  - 貢献度が高いと感じた上位3人を順位をつけて選んでください。（6人グループの場合は4人）
  - 自分を選んでOKです。
- 個別の評価
  - グループワーク中の各人の貢献度
    - 皆さんが互いに評価
    - 報告書

# キックオフメモの提出（本日×切）

- 以下の内容を含む文書を作成して、グループの代表者がMoodle「キックオフメモ」に提出。
  - グループ番号
  - 代表者（提出係）
  - メンバー
  - 出欠報告
  - 議事録・作業計画
    - 例）本日の議論の内容，スケジュール表，作業項目の列挙等．
- PDF形式で作成．
- 分量：A4サイズで1～2枚程度でOK（必要に応じて増量して構わない．）



# 1月12日報告

- 以下の内容を含む文書を作成して、グループの代表者がMoodle に提出。
  - グループ番号
  - 代表者（提出係）
  - メンバー
  - 出欠報告
  - 進捗報告
  - 議事録・今後の作業計画
- PDF形式で作成.
- 分量：A4サイズで1～2枚程度でOK（必要に応じて増量して構わない。）

# 成果物提出のルール

- プログラムのソースを提出.
- 提出するプログラムはC言語で記述すること.
- 指定環境（後述）で動作確認すること.
- プログラムは指定の仕様（後述）に従うこと.
- 使用メモリの上限は20Gbyte. （それ以上の場合は実行を保証しない.）
- 提出：
  - プログラムはadsb\_グループID.cと命名して, Moodleの「成果物提出」にグループの代表者が提出.
    - 例：グループ番号がG1の場合は adsb\_G1.c を提出する.

# 性能評価

- 実行速度（CPU時間），ピークメモリ，精度（後述）を評価指標とする。
- 各グループの得点は以下により求める。
  - 各指標の順位の総和を加算。
  - 精度について，1～5位にはそれぞれ -30, -15, -10, -5, -2 を加算。
  - 実行速度・ピークメモリについて，1～5位にはそれぞれ，-15, -10, -5, -2, -1を加算。
  - また，精度に関しては，それぞれ最下位から数えて3番目までのグループにそれぞれ 30, 20, 10を加算。
  - 複数チームが同順位の場合，順位が $x$ ，チーム数が $y$ であった場合， $x, x+1, \dots, x+y-1$ までの順位に相当する得点の平均を付与します。
    - 例）速度順が，A班, B班 = C班, D班, ...であった場合，A班は  $1-15=-14$ ，B，C班は  $\{(2-10) + (3-5)\}/2 = -5$ ，Dは  $4-2=2$  となります。
  - 中間計測に参加するグループには，-1を加算。  
(不具合確認のためにも，参加をお勧めします。)

# 性能評価（つづき）

- 例えば，実行速度で1位，ピークメモリで5位，スコアで6位，中間計測に参加した場合の得点は，  
 $1 - 15 + 5 - 1 + 6 - 1 = -5$
- 得点は低いほど良い。

# 中間計測

- 参加の是非は自由.
- 2024/1/9, 23:59までに途中結果を提出したグループに関しては、本番と同様の方法で計測を行って、結果を公表します.
  - グループIDを知られたくない場合は、コードネームを使用可能. 提出時に要望してください.
- 提出方法：本番の時と同じフォーマットでMoodleの「中間計測用提出」からグループの代表者が提出.

# 成果発表

- 成果物に関する発表をする.
  - どのような方針で取り組んだのか？
  - 方針を実現するためにどのような方法論を用いたのか？その方法論を用いた根拠は？
  - 実際にそれはうまくいったのか？
  - うまくいった（若しくはうまくいかなかった）要因の分析など
- 登壇＆ポスターにて発表. グループメンバーの全員に発表の機会がある. （120～150分程度を予定. ）  
※ 対面授業に関する大学の方針に変更があった場合はオンライン
- 発表資料は2024/1/17までにMoodleの「成果発表会用資料提出」にグループの代表者が提出.
- 成果発表の詳細は後日案内します.

# 報告書の提出

- 以下の内容を含む文書（pdf）を作成し，2024/1/26までに各自がMoodleの「報告書提出」に提出。
  - 提案手法の説明
  - 提案手法の評価
  - 自分の貢献（どんな役割を果たしたかを具体的に説明。）
  - 発表会での質疑応答
    - 自分のチームの発表のみならず，他のチームの発表に参加した際の質疑応答についてもまとめる。
    - どんな質問をされて，どのように答えたか．どんな質問をして，どんな回答を得たか？それについてどう考察したか？
  - 考察
    - 作業を進める上で難しかったこと，またそれをどうやって解決したか。
    - 提案手法について，どのような改善が望めるか。 など。

# 提出〆切 & 作業スケジュール

- 2023/12/12, 23:59
  - 提出物： キックオフメモ
  - 提出先： Moodle 「キックオフメモ」
- 2024/1/12, 23:59
  - 提出物： 1月12日報告
  - 提出先： Moodle 「1月12日報告」
- 2024/1/9, 23:59 (オプション)
  - 提出物： 中間計測用のプログラム群 (cファイル1つ)
  - 提出先： Moodle 「中間計測用提出」
- 2023/1/15, 23:59
  - 提出物： 最終評価用のプログラム群 (cファイル1つ)
  - 提出先： Moodle 「成果物提出」
- 2023/1/17, 23:59
  - 提出物： 発表資料
  - 提出先： Moodle 「成果発表会用資料提出」



# 提出〆切 & 作業スケジュール（続き）

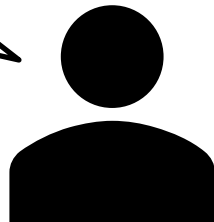
- 2023/1/19, 23:59
  - 作業： 発表評価（Moodle「発表評価」）  
※ ただし，グループの発表評価は授業時間中に行うこと。
- 2024/1/26, 23:59
  - 提出物： 報告書
  - 提出先： Moodle「報告書提出」
  - 作業： 作業評価（Moodle「作業評価」）

課題：気の合う人はどれくらい？

- Yes/Noで回答するとても長いアンケートを実施します.

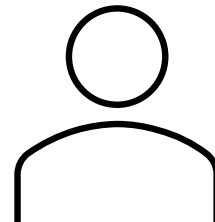
1. Yes
2. Yes
3. Yes
4. Yes
5. Yes
6. Yes
7. No

...










1. ビールが好き
2. ワインが好き
3. 日本酒が好き
4. カレーが好き
5. ラーメンが好き
6. カレーは激辛が好き
7. ラーメンは豚骨より味噌









...



- たくさんの人に回答してもらいました.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
	Y	Y	Y	Y	Y	Y	N	N	N	N
	Y	N	Y	N	Y	Y	N	Y	N	N
	Y	N	Y	Y	Y	Y	N	N	N	N
	Y	N	N	Y	Y	Y	N	N	N	N
	Y	N	Y	Y	N	Y	N	N	Y	N
	Y	N	Y	Y	Y	Y	N	Y	N	N
	Y	N	Y	Y	Y	Y	N	N	N	Y

- ある人物Aさんは、自分と同じ回答をした人がどれだけいるか興味を持ちました。

	Y	Y	Y	Y	Y	Y	N	N	Y	N
	Y	Y	Y	Y	Y	Y	N	N	N	N
	Y	N	Y	N	Y	Y	N	Y	N	N
	Y	N	Y	Y	Y	Y	N	N	N	N
	Y	N	N	Y	Y	Y	N	N	N	N
	Y	N	Y	Y	N	Y	N	N	Y	N
	Y	N	Y	Y	Y	Y	N	Y	N	N
	Y	N	Y	Y	Y	Y	N	N	N	Y

● Q3～Q6まで同じ回答したのは何人？

同類は4人！

	Y	Y	Y	Y	Y	Y	N	N	Y	N
	Y	Y	Y	Y	Y	Y	N	N	N	N
	Y	N	Y	N	Y	Y	N	Y	N	N
	Y	N	Y	Y	Y	Y	N	N	N	N
	Y	N	N	Y	Y	Y	N	N	N	N
	Y	N	Y	Y	N	Y	N	N	Y	N
	Y	N	Y	Y	Y	Y	N	Y	N	N
	Y	N	Y	Y	Y	Y	N	N	N	Y









● Q2～Q8まで同じ回答したのは何人？

同類は1人！

	Y	Y	Y	Y	Y	Y	N	N	Y	N
	Y	Y	Y	Y	Y	Y	N	N	N	N
	Y	N	Y	N	Y	Y	N	Y	N	N
	Y	N	Y	Y	Y	Y	N	N	N	N
	Y	N	N	Y	Y	Y	N	N	N	N
	Y	N	Y	Y	N	Y	N	N	Y	N
	Y	N	Y	Y	Y	Y	N	Y	N	N
	Y	N	Y	Y	Y	Y	N	N	N	Y









● Q5～Q10まで同じ回答したのは何人？

同類はいない...

	Y	Y	Y	Y	Y	Y	N	N	Y	N
	Y	Y	Y	Y	Y	Y	N	N	N	N
	Y	N	Y	N	Y	Y	N	Y	N	N
	Y	N	Y	Y	Y	Y	N	N	N	N
	Y	N	N	Y	Y	Y	N	N	N	N
	Y	N	Y	Y	N	Y	N	N	Y	N
	Y	N	Y	Y	Y	Y	N	Y	N	N
	Y	N	Y	Y	Y	Y	N	N	N	Y



- 任意の質問の範囲 $Q_i \sim Q_j$ に関して, Aさんと同一の回答をした人の数を計算せよ.

	Y	Y	Y	Y	Y	Y	N	N	Y	N
	Y	Y	Y	Y	Y	Y	N	N	N	N
	Y	N	Y	N	Y	Y	N	Y	N	N
	Y	N	Y	Y	Y	Y	N	N	N	N
	Y	N	N	Y	Y	Y	N	N	N	N
	Y	N	Y	Y	N	Y	N	N	Y	N
	Y	N	Y	Y	Y	Y	N	Y	N	N
	Y	N	Y	Y	Y	Y	N	N	N	Y

各データについて：

- $N$ 人の回答者がいる． 各人の回答
  - $x_1, x_2, \dots, x_M, (x_i \in \{0, 1\})$
- 回答の一致を調査する  $P$ 個の質問範囲
  - 開始位置： $l_1, l_2, \dots, l_P, (l_i \in \{0, \dots, M\})$
  - 範囲の大きさ： $s_1, s_2, \dots, s_p, (s_i \in \{1, \dots, K\})$
- MODEL1
  - $N : 10^3, M : 10^4, P : 3 \times 10^4, K : 8$
- MODEL2
  - $N : 10^3, M : 10^4, P : 5 \times 10^4, K : 300$

# 配布データについて

- Aさんを含む全員の回答

- ファイル名：data

- フォーマット

Aさんの回答（改行）

その他一人目（改行）

その他二人目（改行） ...

- 質問の範囲

- ファイル名：range

- フォーマット

モデル名 <= MODEL1もしくはMODEL2

$s_1, l_1$ （改行）

$s_2, l_2$ （改行） ...

- 正解

- ファイル名：answer

$s_1, l_1, d_1$ （改行）

$s_2, l_2, d_2$ （改行） ...

これらのデータをMODEL1とMODEL2それぞれについて、10セット配布します。（「グループワーク」⇒「資料」⇒ all.zip）

12/22配布のファイルには、モデル名が書かれていません。  
12/23以降に差し替えます。

dは正解（Aさんと回答の一致する人数）

# 例（配布物とはパラメータが異なる）

**data:**

01001010011001111010

01001010011011110010

01001110011001111010

01001010001000111110

01001010011001111011

01001010011101111010

**range:** （重複する範囲が含まれることもある. ）

**MODEL1**

11, 2

9, 11

9, 8

5, 15

# 提出に関して

- プログラムは一つのC言語ファイルとすること
- 以下のように実行できるプログラムであること
  - <プログラム名> <正解ファイル名> <質問範囲ファイル名> <出力ファイル名>
  - コマンドライン引数から、各ファイル名を取得できるプログラムであること。（固定ファイルを入力としない）
  - つまり、以下のように実行できるプログラムであること  
`./solve ./data0 ./range0 ./myans`
- 出力フォーマット
  - 正解ファイル（answer）と同じフォーマットでファイルに出力すること
    - 開始位置, 質問範囲の大きさ, 回答一致の人数（改行）
  - 質問範囲ファイル（range）に対して、各質問範囲の順序<sub>29</sub>が同じでなければいけない

# 性能評価

- 評価データ

- サンプルデータを生成したのと同じモデルを用いて作成。（つまり、サンプルデータと同質のデータを使って評価。）

- 実行速度

- 計測環境（予定）

- OS: Ubuntu 20.04.1 LTS, gcc: v9.4.0, CPU: AMD Ryzen Threadripper 3970X, build-essentialのみで開発できるプログラムを対象とする

- 30秒で打ち切り

- メモリ20Gを超えたプログラムの実行は保証しない

- 精度

$\sum |d_{p,t} - d_{p,u}|$  ただし、 $p$ 番目の正解と提出物の出力それぞれ  $d_{p,t}$ ,  $d_{p,u}$  とする。

- **MODEL2を主問題としますが、MODEL1とMODEL2は別々に評価して結果を公表します。**

# 提出時の動作確認

- 提出物は、以下の環境で動作することを確かめてから提出すること.
- OS : ubuntu-20.04.1-desktop-amd64  
開発 : build-essential
- Linux環境をお持ちでない方は、インストールして利用することを推奨しますが、インストールせずに利用することもできます. (次を参照)

1. <https://releases.ubuntu.com/20.04/ubuntu-20.04.1-desktop-amd64.iso> より入手
2. VirtualBoxをインストール
3. 仮想マシンの作成





## 4. isoイメージを設定

「光学ドライブ」をクリックして  
「ディスクファイルを選択」をクリックした後で  
ubuntu-20.04.1-desktop-amd64.isoを設定

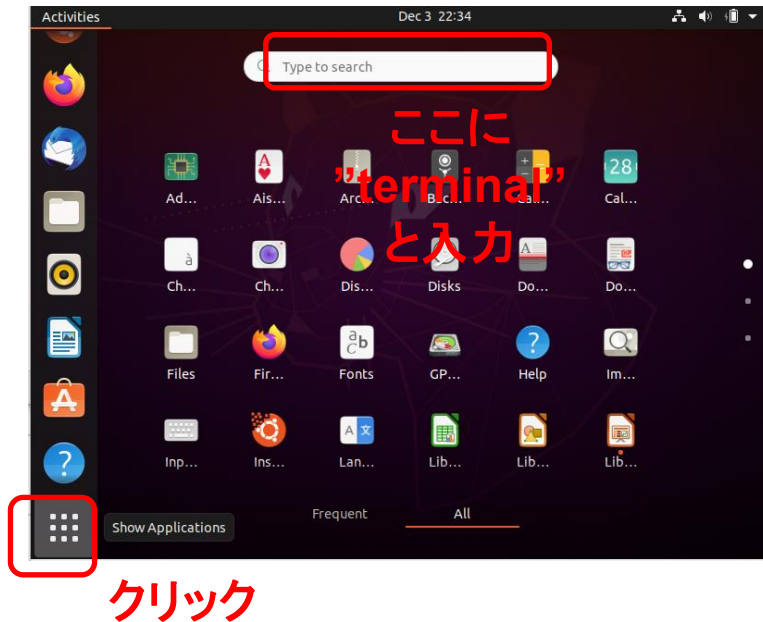


## 5. 仮想マシンの電源を入れる

クリック



## 6. Live CDを起動



## 7. terminalを立ち上げて以下を実行

```
sudo apt-get install build-essential
```

# どんな方法で解くか？

- 数多くの質問範囲を調べなければならない．どうすれば高速化できるか？
- MODEL1とMODEL2では問題の性質が少し異なる．同じ方法で解くのか，別の方法を用いるのか．

- ぜひ活発な議論を
- ソースの共有
  - github (<https://github.com/>)
  - Dropbox
  - Google Drive

# グループワーク進め方

- 進め方は自由ですが、過去に見られたケースをいくつか紹介します。

## その1

1. 議論により問題に対する理解を深め、様々なアイデアを出す。
2. 個人で取り組み、一番良い解法をグループの解法とする。

## その 2

1. 議論により問題に対する理解を深め、様々なアイデアを出す.
2. 個人,あるいはペアで解法を考え,それを元に方針を見出す.
3. 分担して実装し, テストデータでの性能を確認しながら, さらに良い手法がないか議論をする.
4. 発表資料を分担して準備する.

# グループディスカッション

- 本日不在のメンバーに後日連絡を取る必要がある場合は、グループフォーラムをご利用ください。