

主に数値計算系の

卒論の書き方

2021/12/09

慶應義塾大学理工学部物理情報工学科
渡辺

卒論を書く前に

Q: 卒論で一番大事なのは？

1. 結果の新規性
2. 適切な引用
3. 体裁
4. ストーリー



卒論を書く前に

Q: 卒論で一番大事なものは？

もちろん
大事ですが

1. 結果の新規性
2. 適切な引用
3. 体裁
4. ストーリー
5. バックアップ



期日までに提出することが最も大事

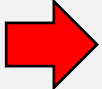
バックアップ方法

バックアップは**定期的**にとる

バックアップ頻度＝データが飛んだ時の手戻りの時間
最低でも毎日バックアップすること

バックアップは**リモート**にとる

自分のPCの別フォルダにコピーするのはダメ
USBへのコピーも信用できない

 **GitHubを使うことを強く推奨**

※ BoxでもDropboxでもなんでもよいが、どうせバージョン管理するので

GitHubによる卒論執筆管理

卒論執筆用リポジトリ

TeXファイル
bibファイル
図のファイル
図を作るためのデータ
図を作るスクリプト等

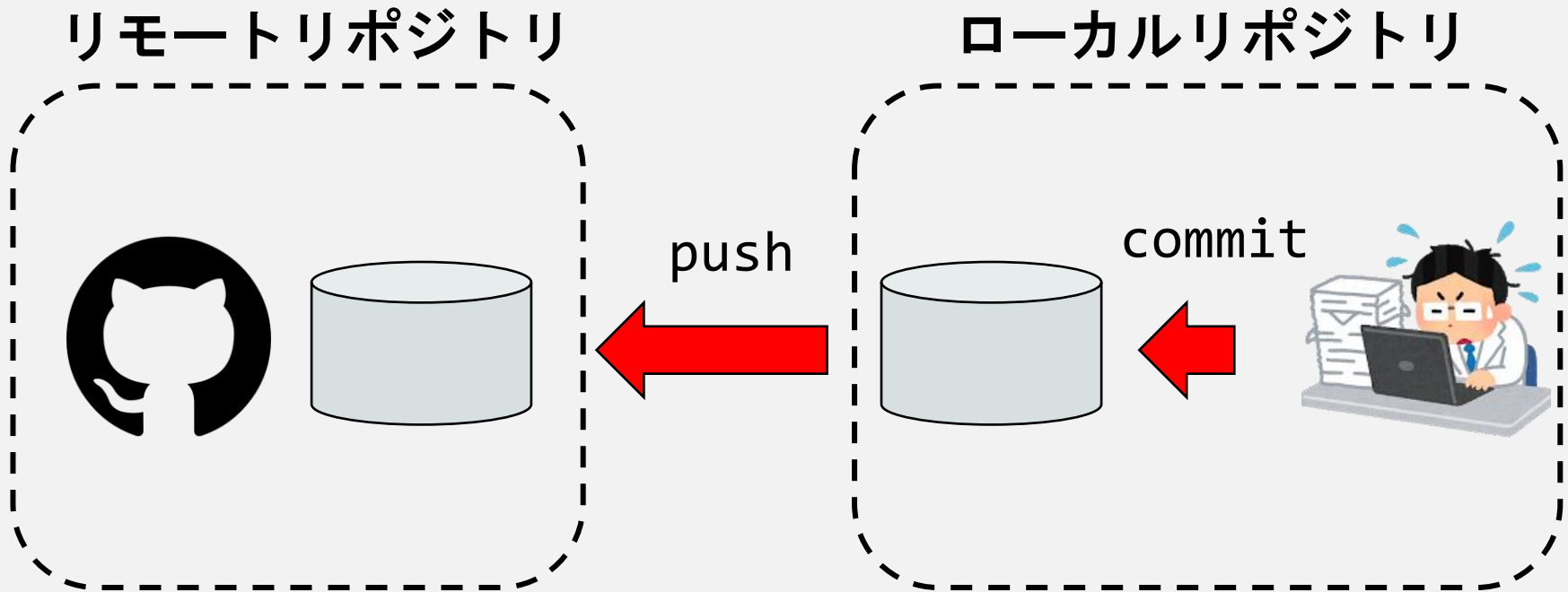
プログラム開発用リポジトリ

プログラムソース
インプットファイル
ジョブスクリプト等

- 卒論執筆用とプログラム開発のリポジトリを分ける
- 出力データはプログラム開発リポジトリにいない
- 図を作るためのデータは卒論執筆用リポジトリに入れる
- 大きなデータ(ダンプなど)はリポジトリに入れない

GitHubによる卒論執筆管理

Git/GitHubを使う場合のバックアップ＝push



コミットとプッシュの頻度



15分～1時間程度に一回

「ひとかたまり」の仕事ごとにコミット&プッシュ

いちいちパスワードを入力しないで済むようにssh-agentを使うこと

卒論を書く前に

Q: **卒論の中身**で一番大事なのは？

1. 手法の独自性
2. 結果の新規性
3. 十分なサーベイ
4. 適切な引用



卒論を書く前に

Q: **卒論の中身**で一番大事なのは？

もちろん
大事ですが

1. 手法の独自性
2. 結果の新規性
3. 十分なサーベイ
4. 適切な引用
5. **ストーリー**



「一つの論文」として、一貫したストーリーを持つことが大事

卒業論文の構成

背景・目的

なぜこの研究を行うか(イントロダクション)

手法

この研究をどのように行うか

結果

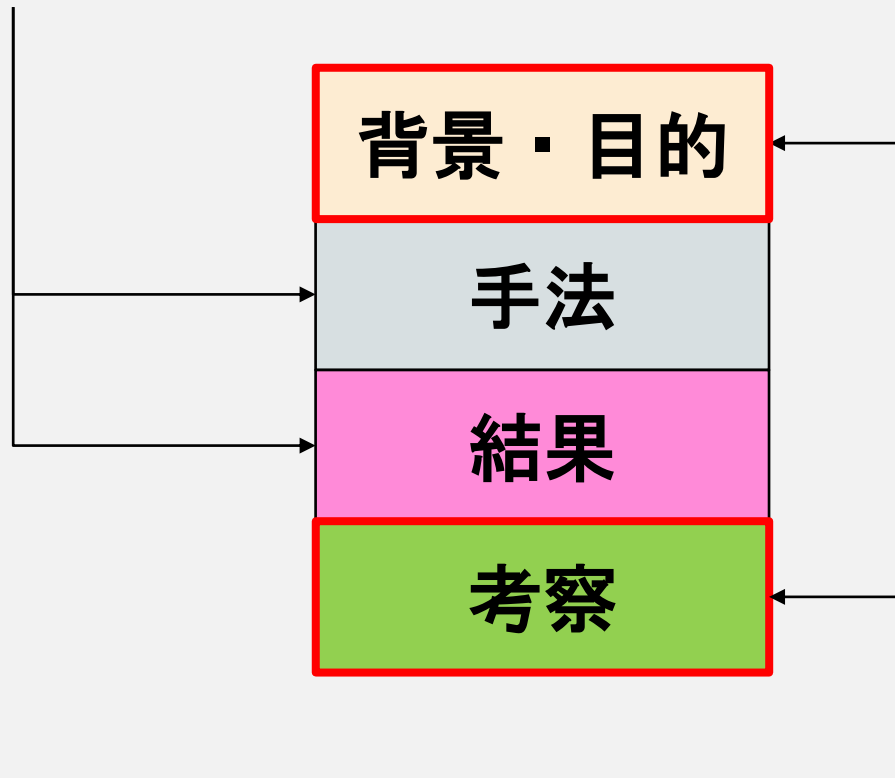
どのような結果が出たか

考察

この研究にはどんな意味があるか

卒業論文の構成

自分のやったことをがんばって書きがちだが



卒論はこっちを書く練習

卒論のストーリー

ありがちなパターン

背景

いまAという分野が注目されており、分野AではXという手法が主流である

目的

手法Xの精度を改善する

手法

手法Xを修正した手法X'を考案した

結果

X'はXに比べて20%精度が改善した

考察

手法Xの精度改善という目的を達成できた

対応している？

一見、目的と考察が対応しているように見える

卒論のストーリー

イントロダクション(背景+目的)の役割：
読者に「この研究は必要だ」と納得させる

背景

いまAという分野が注目されており、
分野AではXという手法が主流である

分野Aが注目されているのはなぜか？
手法Xが主流なのはなぜか？

目的 手法Xの精度を改善する

なぜ手法Xの精度を改善する必要があるのか？
どのくらい改善したいのか？

卒論のストーリー

背景

いまAという分野が注目されており、分野AではXという手法が主流である



背景

近年、Aという分野は発展を続けており、様々な応用例が提案されている[1-5]。Aという分野では、X、Y、Zという多くの手法が提案されたが、計算量と精度のバランスから、現在は手法Xが主流である[6,7]。

分野Aが注目されているのはなぜか？

→ 多くの応用成功例があるから

手法Xが主流なのはなぜか？

→ バランスが良いから

適切に引用しながらストーリーに説得力をもたせる

卒論のストーリー

目的 手法Xの精度を改善する



目的 手法Xを分野Bにも適用したいが、分野Bで実用的に使うには最低でも正解率〇〇%は必要となる。そのためには精度を40%改善したい。

なぜ手法Xの精度を改善する必要があるのか？

→ 分野Bにも適用したいが、現状では精度が足りないからどのくらい改善したいのか？

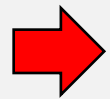
→ 40%は改善したい

読者が納得するような「ストーリー」を作る

卒論のストーリー

考察では「イントロで提示した問題」に答える

手法Xを分野Bに適用したいから精度を改善したい



「考案した手法X'は分野Bに適用できるのか？」
に答えなくてはならない

対応しているか？

本研究により20%という精度改善を得られた。しかし、分野Bに利用するためにはさらなる改善が不可欠である。より精度を改善するには・・・

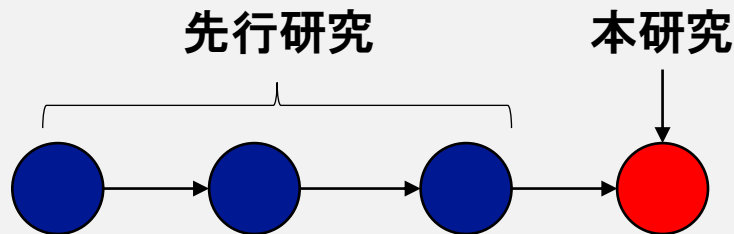
卒論のストーリー

研究の「ストーリー」は、概ねこんな形となる

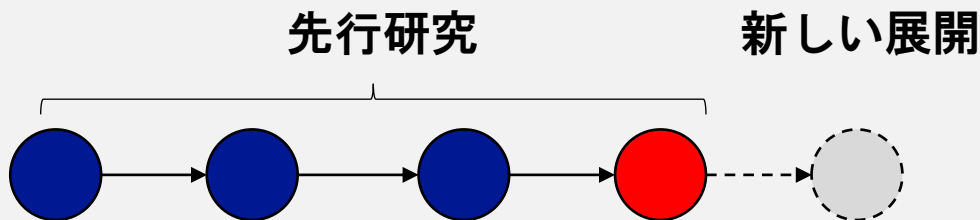
- [大きな背景] なにか大きな目的Aを達成したい
- [小さな背景] その目的Aのサブタスクaを達成したい
- [先行研究紹介] そのためにこんな研究が行われてきた
- [問題提起] でもまだこんな不満がある
- [目的の説明] その不満をこんな形で改善したい
- [手法の説明] 具体的な不満の改善方法
- [結果の説明] どれくらい不満は解決できたのか？
- [まとめと考察] この研究はどんな意味を持つか？

研究の流れ

研究には「流れ」があり、新しい研究は、その一部となる



いずれこの研究も「先行研究」の一部となる



考察には

- この研究は「研究の流れ」においてどんな意味を持つか
 - この研究の先にどんな展開があるか
- を書く

まとめ

卒業論文で重要なのは 「バックアップ」と「ストーリー」

「がんばったこと」ではなく「読者が知りたいこと」を書く
「大きな研究の流れ」の中の「本研究の位置づけ」を明確に

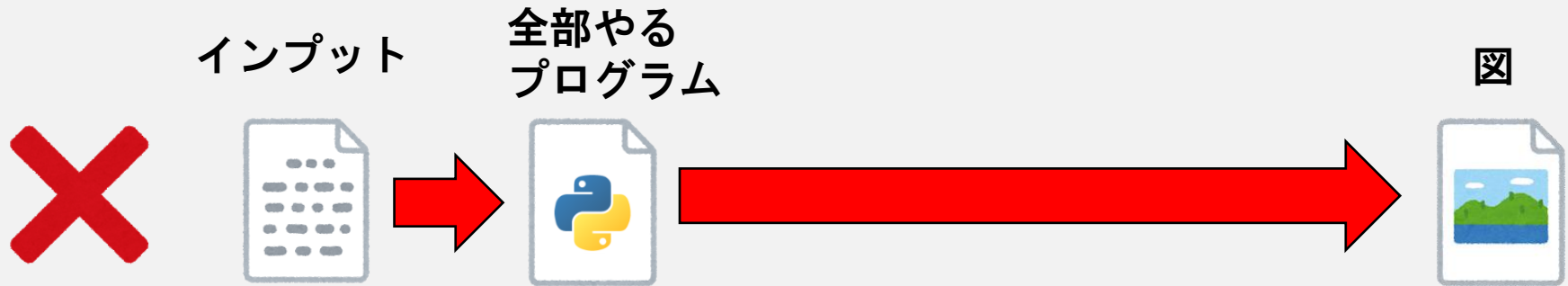


上記を実施するには論文をたくさん読む必要があります。がんばりましょう・・・

以下、テクニカルな注意点

図とデータ

計算と図の作成を一度にやらない



- 線の太さや軸のフォントなど、図はなんども作り直すから
- 結果ファイルと図の作成プログラムを論文リポジトリに入れたたい

図とデータ

図はデータからコマンド一発で作る

手順の記録

Excel等で作ると、後でどうやってその図を作ったか忘れてしまう。最初は面倒でも、スクリプトにしておけば、後から何をしたかがわかる

修正が容易

「図のフォントを全部大きくして」と言われても、スクリプトで作っていればsed一発で済む

図とデータ

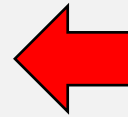
図とスクリプトのファイル名は揃える

pressure.plt

```
set term pdf
set out "pressure.pdf"
set xlabel "t"
set ylabel "P"
p "pressure.dat" pt 6 t "Data"
```

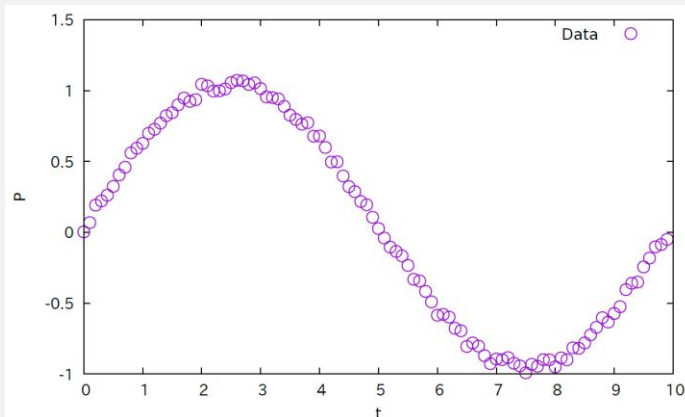
pressure.dat

```
0.000000 0.041263
0.100000 0.143985
0.200000 0.188462
0.300000 0.241296
0.400000 0.264365
0.500000 0.309186
0.600000 0.444013
....
```



\$ gnuplot pressure.plt

pressure.pdf



hoge.pdfを作りたいければhoge.pltや
hoge.pyを探せばよい

図とデータ

コマンド一発で図が全てできるのが望ましい

pressure.datとtemperature.datから
pressure.pltとtemperature.pltを使って
pressure.pdfとtemperature.pdfを作る

} Makefile

```
all: pressure.pdf temperature.pdf

%.pdf: %.plt %.dat
    gnuplot $<
```

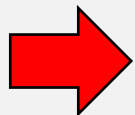
「そのディレクトリに入ってmakeしたら必要なものが揃う」
という状況を作る

※ Pythonやシェルスクリプト等でも良い。とにかくコマンド一発で。

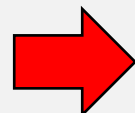
図の作り方

図の役割は主に**比較**

理論値と実験値が**合っている/合っていない**
提案手法が既存手法よりも性能が**向上した/していない**
パラメータを変化させると計算量が**増えた/減った**
etc.

 比較したいものを一つの図にまとめる

図一つにつき、メッセージ一つ

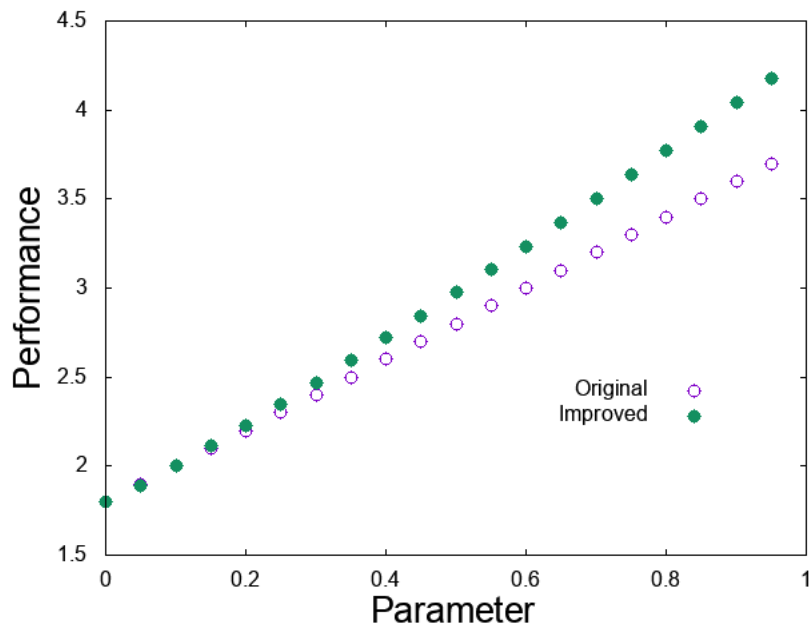
 複数の情報を一つの図に詰め込まない

図の作り方

主張

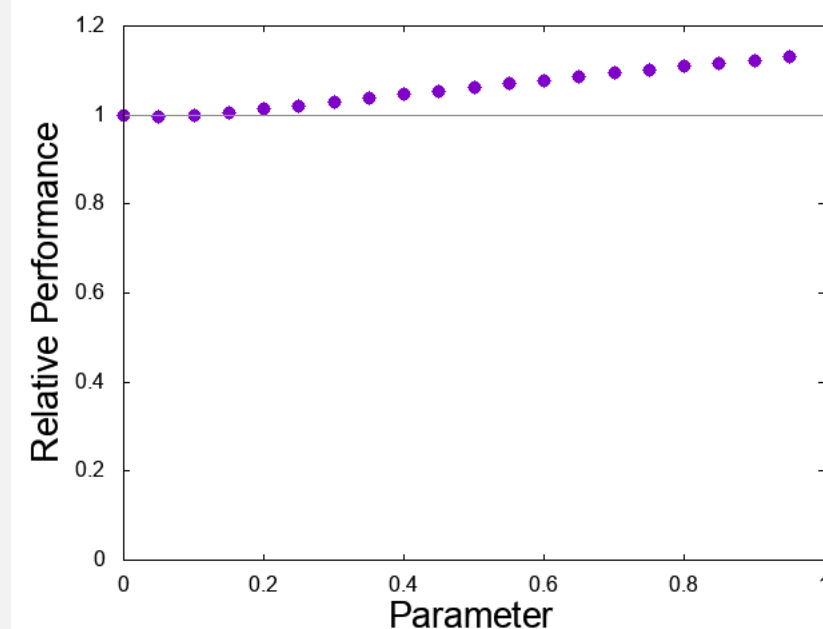
パラメタが大きいところで既存手法Xに比べて提案手法X'の方が性能が良い

それぞれの性能の絶対値をプロット



どれくらいよくなったのか
わかりづらい

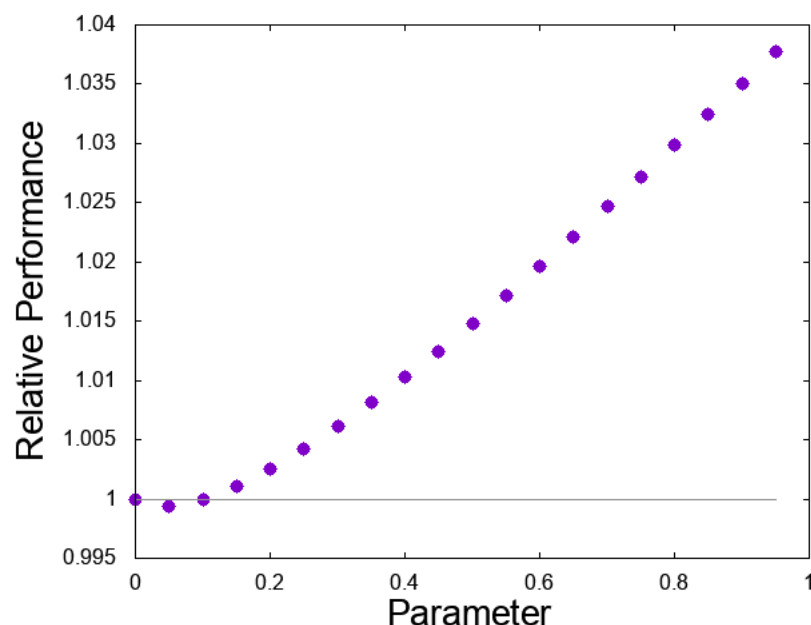
X'とXの性能比をプロット



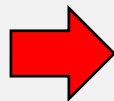
性能向上が10%以上20%未満で
あることがすぐにわかる

図の軸の「範囲」は適切に

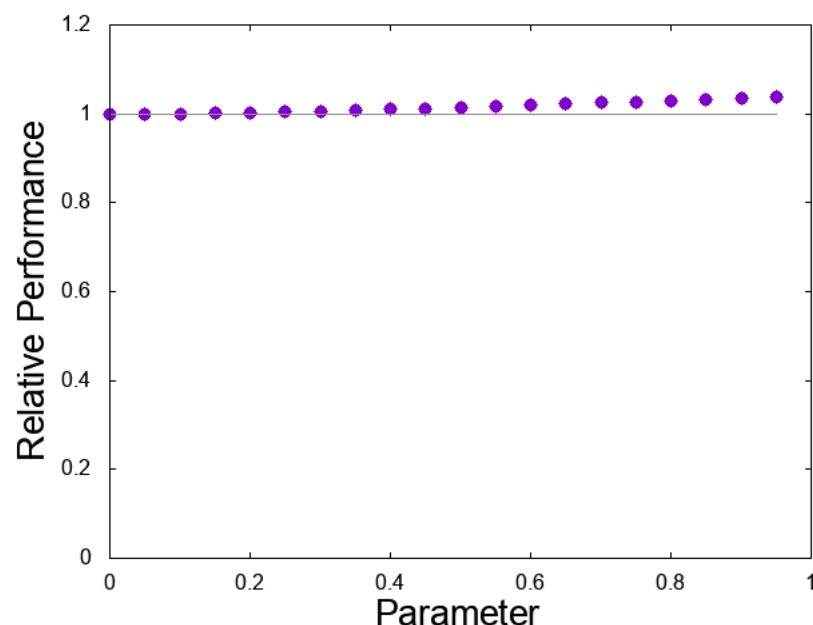
X'/Xの性能比 (不適切版)



一見、大幅に性能向上しているように見えるが . . .



X'/Xの性能比 (適切版)



y軸の範囲を0からとると性能向上はさほどでもないことがわかる

比較したいものを同じ図に

主張

手法Xに比べ、手法Yは**精度**を上げやすいが**計算量**も増える

図1(a) 手法Xの精度と計算時間

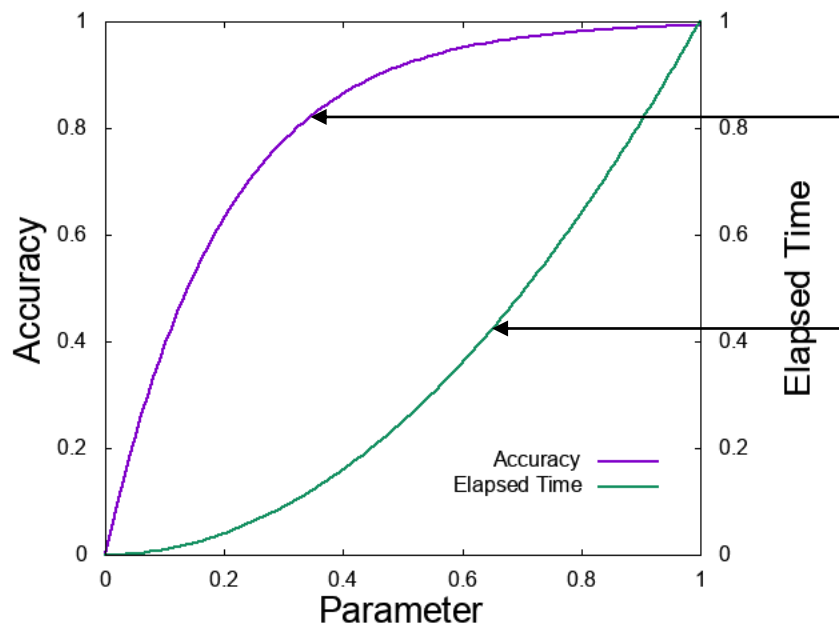
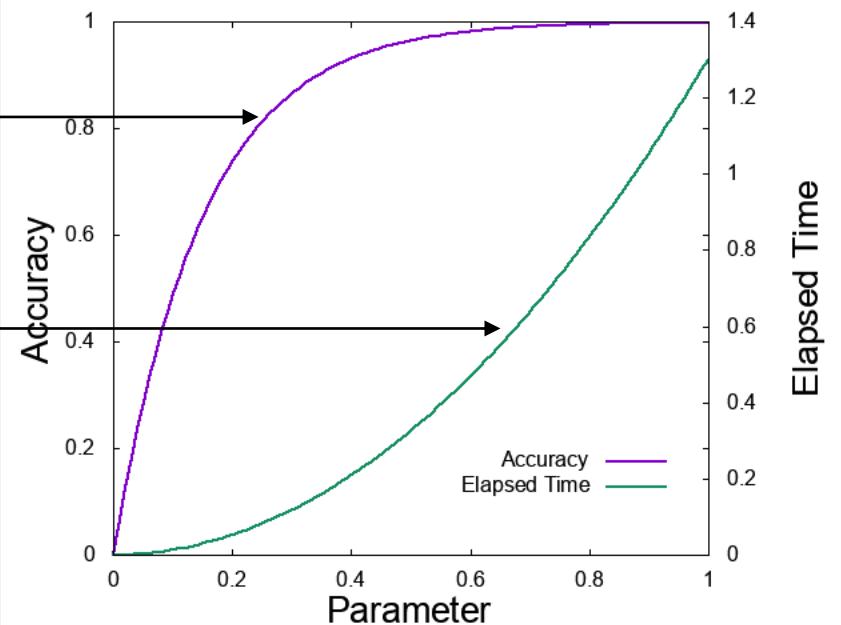


図1(B) 手法Yの精度と計算時間



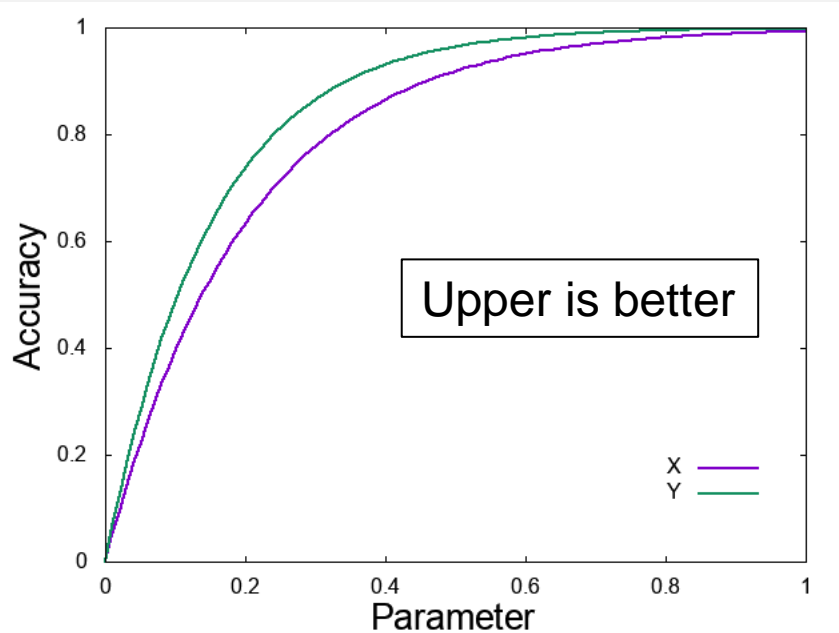
- 比較したいものが別の図に分かれていて分かりづらい
- 高いと良いもの(精度)と低いと良いもの(計算時間)が混在している

比較したいものを同じ図に

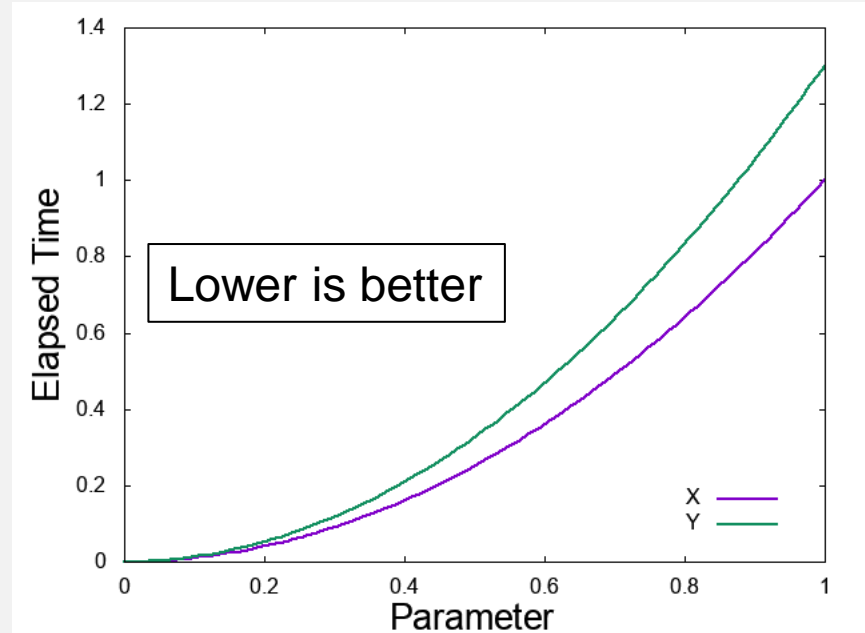
主張

手法Xに比べ、手法Yは**精度**を上げやすいが**計算量**も増える

手法X,Yの精度



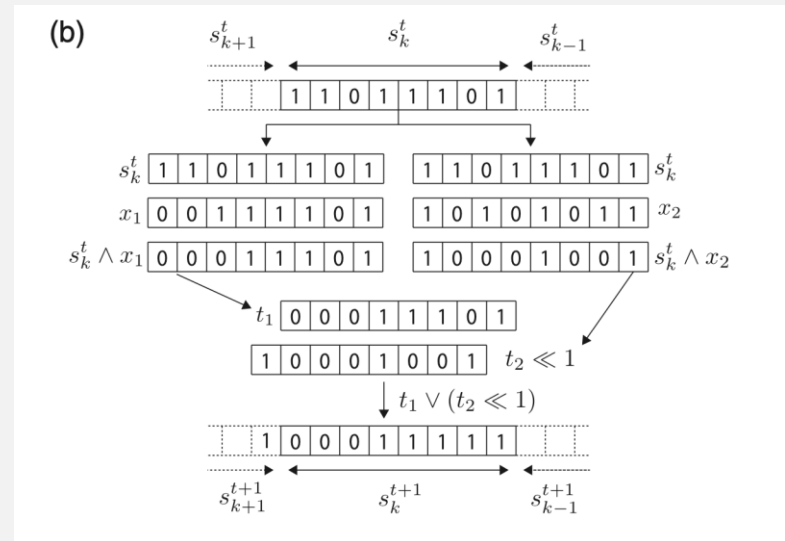
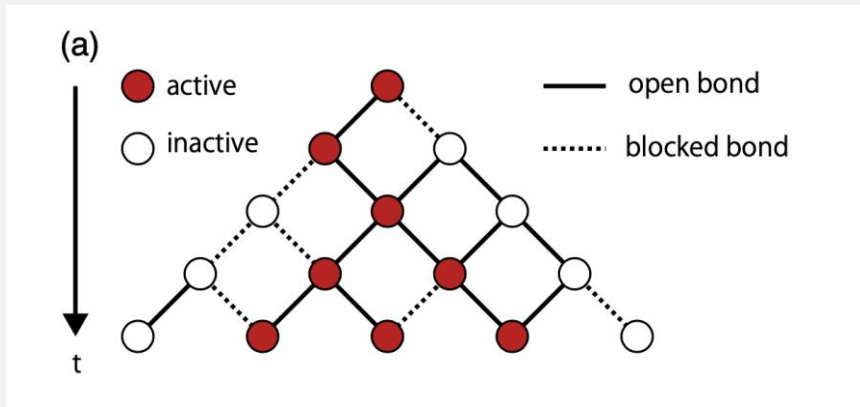
手法X,Yの計算時間



主張したいことが明確になるように図を作る

ポンチ絵の作り方

ポンチ絵：読者の理解を助けるための概念図



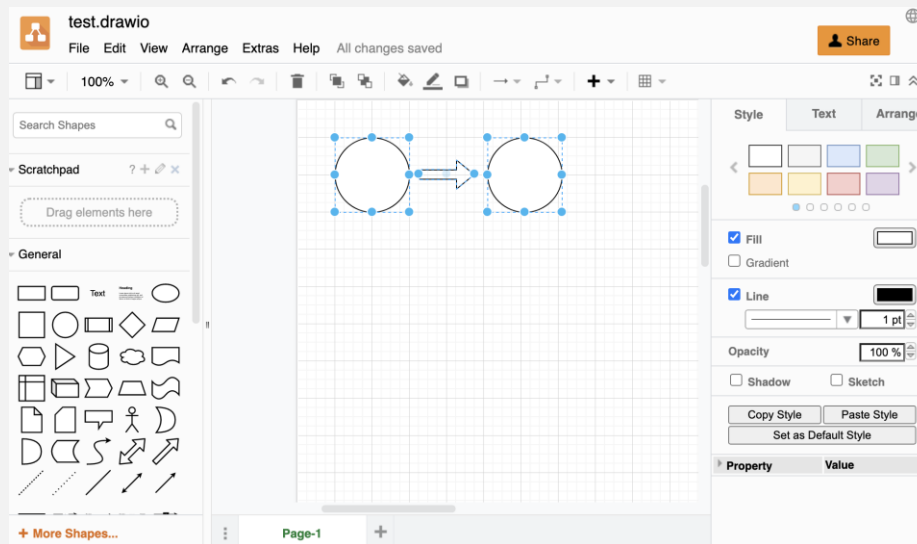
HW, S. Morita, S. Todo, N. Kawashima, J. Phys. Soc. Jpn. **88**, 024004 (2019)

複雑な手順や、実験のセットアップ等、文章だけでは
わかりづらいものを図示する

ポンチ絵の作り方

ポンチ絵を作る手段

- PowerPointで作ってPNGで「図として保存」
- Adobe Illustratorで作成してPDFで保存
- draw.ioで作成してExport asでPDFで保存



好きな方法で作成して良いが「元のファイル」を一緒にバージョン管理すること (PowerPointならpptx、イラレならai等)

※ (僕がいうのもなんだが)卒論の図では「いらすとや」の多用などは避ける

引用の仕方

引用の無い文章は「著者のオリジナル」とみなされる

Aという手法にはBという問題がある。

→ 著者がそう思っている。

Aという手法にはBという問題がある[4]。

→ 文献[4]の著者がそう指摘している。

Aという手法にはBという問題がある[4-8]。

→ 多くの人が問題だと認識している。

引用の仕方

引用するのは原則として書籍か査読論文

 できるだけウェブサイトの引用は避ける

- ・ 「まっとうなサイト」なら、文献が引用されているはず
- ・ 必ず「原典」にあたって、内容を確認する

 arXivを引用する際は出版されていないか確認する

- ・ 機械学習の論文はarXivを引用することが多いが、有名な論文はカンファレンスペーパーになっていることが多い
- ・ タイトル等で検索してみる

体裁等についてはBibTeXを使えば自動的に満たされるはずなので略

引用の仕方

コピペはダメ！ゼッタイ！



引用の仕方

たとえ引用をしていても「コピペ」はしてはいけない

- 引用は、その文献を読んで**自分の言葉で**理解したことを書く
- 文章をそのまま載せたい場合は、カギカッコで示すか、quotation 環境を使うなど「引用である」ことがわかるようにする

※ 理系で参考文献の文章をそのまま引用することは少ないはず

目立つのが、語句説明のWikipediaからのコピペ

※ たまに投稿論文で見つけてびっくりする

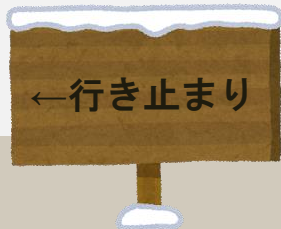
- Wikipediaに頼りたくなるのは**文献の読み込み不足**
- 教科書や論文をちゃんと読み、**自分の言葉で書く**こと

結果について

「当初の見込みに沿った結果が出るのが成功」 **ではない**
「やったけれどできなかった」は**立派な成果**

- 研究の目的は「人類の知に資する」こと
- 「ネガティブな結果」も、後進の試行錯誤を減らす

今回の研究結果



変に取り繕ったりせず
結果は誠実に書くこと