

# コンピュータ概論 A/B

-- 情報と検索エンジン --

数学科 栗野 俊一 (TA: 北野拓也 [院生 2 年])

2016/05/10 コンピュータ概

# 伝言

---

## 私語は慎むように !!

### □ 担任からの連絡

○ 学生証での出席は済ませましたか？

▶ 入口の脇の出席装置に学生証を翳す

○ Web 履修登録変更

▶ 履修登録に変更がある場合は、木曜日(05/12)迄

○ 数学科オリエンテーション関係

▶ 木曜日に水野先生に問い合せてください

### □ 席は自由です：できるだけ前に詰めよう

### □ 色々な、お知らせについて

○ 栗野の Web Page に注意する事

<http://edu-gw2.math.cst.nihon-u.ac.jp/~kurino>

○ 水野先生の「数学科 1 年生の連絡ページ(2016 年度版)」

<https://sites.google.com/a/nihon-u.ac.jp/cst-math2016/home>

### □ 補習

○ 本日(2016/05/10) はありません / 次回(2016/05/17)は予定しています。

# 前回(2016/04/26)の内容

---

## □ 前回(2016/04/26)の内容

### ○ フリーウェアのインストール/セットアップ

- ▶ [comp-20160510] 中の [soft] に、インストールパッケージが入っている
- ▶ 積極的にインストールしておく(特に firefox)

### ○ Skype のセットアップ

- ▶ Skype のインストール/Skype ID の取得/コンタクトの提出
- ▶ 「kurino-2016-math-cst-nihon-u」をコンタクトに登録しましたか？

### ○ IT 資産管理

### ○ 「コンピュータ概論 A/B」ガイダンス

- ▶ 成績は演習と試験で / 毎回コンピュータとネットワークを利用する
- ▶ 人に迷惑をかける行為には厳しく対応 / 講義中はしゃべるな
- ▶ 友達と相談しながら学習しよう / 習うより慣れろの精神で

# 本日(2016/05/10)の予定

---

## □ 本日(2016/05/10)の予定

- 情報と検索エンジン

## □ 本日(2016/05/10)の目標

- 「情報」という言葉の一次理解
- 検索エンジンの利用
- CST Portal を利用した課題提出

## □ 演習

- CST Portal を利用した課題提出

- ▷ 2015/04/12 の課題 / 今回 (2016/05/10) の課題

- ソフトのインストール

- ▷ 話を聞きながら、各自、静かに進める

# 本日の課題 (2016/05/10)

---

## □ 前回の課題

- Note-PC の setup のためなし

## □ 今回 (2016/05/10) の課題

- 1. 2015/04/12 の課題を行う
- 2. CST Portal に以下の課題の結果を提出しなさい
  - ▶ ファイル名 : 20160510-QQQQ.txt (QQQQ は学生番号)
  - ▶ 表題 : URL の紹介
  - ▶ 内容 : 最近、自分が気になっている単語 ( 10 個 ) に関する URL を三つずつ紹介する
  - ▶ サンプル(20160510-9999.txt)を参照のこと
- Wikipedia を紹介する場合は、一次情報も併記する事

# 情報とは

---

## □ 情報とは

○ 複数の可能性ある現象に対して、その内の一つあるいは一部を選んだもの

▷ (天気) 今日の午後は雨だ(晴の可能性があった)

▷ (数学) 二等辺三角形の内角和は180度だ (180度と異なる可能性があった)

○ 「正しい」情報と「正しくない」情報

▷ 正しい情報: 事実を記述したもの (じゃあ、「事実って何?」って話さない)

## □ 情報の量(情報が多いか少いか)

○ 情報そのものから計る事ができる情報量(客観的/情報理論)

▷ その情報が表現する現象が「起きる可能性が低け」れば、それを得た時の「情報量が多い」

▷ (ギャンブル) 「穴馬が勝つ」と言う情報は、「本命が勝つ」と言う情報より情報量が多い

▷ (サイコロ) 「ゾロ目が出た」と言う情報は、「そうでない」と言う情報より情報量が多い

▷ 「学」としては、こちらを利用する

▷ bit : 共に 1/2 の確率の二つの事象 (Yes/No, 裏/表, 偶数/奇数) の内の一つが示された時の情報量(情報の最小単位)

○ 受け手の知識量から計られる情報「価値(主観的/一般的に利用される)」

▷ 受け手が知らない情報かどうか (報された時に知っていた情報は価値が低い)

▷ 受け手に価値のある情報かどうか (興味がない情報の価値は低い)

▷ 「現実の世界」ではこちらが利用される事が多い (主観が入る)

# 情報の量と質(メタ情報)

---

## □ 情報「量」の定義

- 事象  $E$  が確率  $P(E)$  で起きるとする
- 「事象  $E$  が起きた」という「情報」の情報量( $I(P)$ )は次の式で表現される

$$I(E) = \log_2 \frac{1}{P(E)} = -\log_2 (P(E))$$

## □ 情報の「質」

- 「情報」そのものに関する情報 (メタ情報)
  - ▶ その情報は「正しいかどうか (正しくなければ価値がない)」を判定する材料
- 例 1 : 誰から聞いた? (情報ソース)
  - ▶ 彼奴(あいつ)の情報じゃ、信頼できない
- 例 2 : 何語で書かれているか? (情報の形式)
  - ▶ 楔形文字で書かれているんじゃ、読みとれない
- 例 3 : どうやって入手した(情報の入手方法)
  - ▶ 不法な手段で入手した証拠は、裁判では取り上げられない
- 例 4 : それは何時の話? (情報の有効な範囲)
  - ▶ バーゲンは、昨日でお仕舞いだった..

## □ 「情報(の内容)」だけでなく、その「量」と「質」にも気にする

# 「データ」と「(狭義な)情報」

---

- 「データ」と「(狭義な)情報」: 相対的な違いである事に注意
  - データ: 客観的な観測結果を記述した(形をしている)物
    - ▶ (芸能) 川谷絵音が離婚
    - ▶ (天気) 2016/05/08 の千葉県の気温最高気温: 26.6 度
    - ▶ (数学) 地面に三角形 ABC を描画、角度の測定結果は、約 30, 60, 90 度
  - (狭義な)情報: 一定の価値感(主観を含む)に基き、データを「解釈」した結果
    - ▶ (芸能) ベッキーとよりを戻すんじゃないか?
    - ▶ (天気) 5 月初旬で、暑くなりそうだ
    - ▶ (数学) 三つの内角 A, B, C の和は約 180 度
- 一次情報/二次情報
  - 一次情報: そのデータ(情報)が最初に作られた形での情報
  - 二次情報: 一次情報や他の二次情報に基いて作られた情報
  - 「Wikipedia からの引用」が「ダメ」な理由
    - ▶ Wikipedia は二次情報だから => 一次情報に当ろう (cf. 論文)



# 「データ」から「(狭義な)情報」へ

---

## □「データ」から「(狭義な)情報」へ

○データを「処理(分析)する」事によって「(狭義な)情報」が得られる

▶データは「取る」もので、情報は「作る」もの

○「処理する」には、「判断基準」が必要となる

▶(芸能) そもそも、離婚の原因は不倫騒動だったんだから..

▶(天気) 去年の千葉市の5月の平均最高気温は、20度程度なので、それに比較すると..

▶(数学) 平面幾何学の公理と論理を利用して証明できるので..

○「基準」が異れば、「同じデータから異なる情報」が得られる

▶(芸能) いや、「お友達」だったんだよね..

▶(天気) 沖縄出身なので、これでも寒く感じる？

▶(数学) 球面幾何学では、180度より大きくなってしまう

以下、「データ」と「(狭義な)情報」をまとめて単に「(広義な)情報」と呼ぶ

# 情報の形

---

## □ 情報の記述を構成する要素

### ○ 5W1H

- ▶ Who : 誰が
- ▶ What/Whom : 何を
- ▶ When : 何時
- ▶ Where : 何処で
- ▶ Why : なぜ(どんな目的で)
- ▶ How To : どうやって

### ○ 5W2H : 5W1H に 次を追加

- ▶ How Much ( How Many ) : いくらで (どのくらいで..)

## □ 知識の形式

### ○ What : それは何であるか ?

- ▶ (数学) 定義/公理

### ○ How to : それはどうやってやるのか ?

- ▶ (数学) 証明/計算

# 知識を身に付ける

---

## □ 情報の種類

- 仮定/決め(定義)/事実：知らないと判らない情報(一般と特殊がある)
  - ▷ 「調べて」入手するしかない ( 検索する !! ) / 考えてはならない(無駄)
- 結論/導かれる物(定理)/予想：考えれば解る情報(複雑)
  - ▷ 「考えて」入手する事が望ましい / 調べると時間がかかるかもしれない
  - ▷ 調べても入手できるが、コストや確実性に問題がある

## □ 知らないと判らない情報の入手方法

- 特定な対象と結びついており、その対象毎に異なる情報
  - ▷ cf. 個人情報 ( 年齢など.. )
  - ▷ その情報の対象に確認する ( 典型的な質問の一つ )
- 多くの人 ( 場合によっては全ての人 ) と関係があり、共通となる情報
  - ▷ cf. 年の数え方
  - ▷ Web を検索する

## □ 考えれば解る情報の入手方法

- 考えるのが一番 ( 「数学」そのもの )
- 調べても良いが、身に付かない ( cf. レポートの結果をコピーする )
  - ▷ 正しいかどうかの判定が必要(結局、考える事が必要に..)

# 検索エンジンの使い方

---

## □まずやること

- 「グーグルの上手な使い方」で検索する

- ▶「自分に合った」方法を身に付ける

## □「検索」とは

- 欲しい情報に関連するキーワードを入力し、ページを表示させる

- ページの内容を「精査」して、欲しい情報が記載されたページを篩(ふる)う

## □良いキーワードとは

- 欲しい情報と関連が深い

- ▶欲しい情報が記載されているページには高い頻度で現れる

- ▶欲しい情報関係ないページに現れる可能性は低い

- 情報量が多い

- ▶専門用語は良いキーワードになる ( cf. 活断層 )

- ▶良く利用される言葉はキーワードとして不適切 ( cf. 離婚 )

- まずは、良いキーワードを探せ ( サンプルング )

- ▶思い付くキーワードで表示されるページから良いキーワードを探す

# 検索結果の利用方法

---

- 検索結果は「正しい」とは限らない
  - 内容が正しいかどうかを \*確認\* してから利用する
- 結果の「正しさ」のチェック
  - 複数の結果を比較しろ (より便利なページもある)
    - ▶ 検索リストのトップから幾つか(最低 3 できれば 10 位)を見て確認
  - 内容を確認し、「矛盾」がないかを調べろ
    - ▶ ページ内のキーワードが適切に利用されているか
    - ▶ 論理的な説明になっているか
  - 検索の前に「答を予想」して、それと比較しろ
    - ▶ 予想とかけ離れているなら変かもしれない
    - ▶ 予想と逆の結果も検索してみる (それが間違っている事を確認する)
  - 一次情報を参照しているか (一次情報元でなければ..)
    - ▶ 根拠のない推論をしている可能性はないか

# 「理解」とは

---

## □「理解」とは

○その「知識」が他の「知識」と沢山の「関連付け」が起きている状態

▶ Web が何故「知識の表現」に適切かという、関連付け (link) があるから

▶ 新しい事(3割)を理解するには、ある程度(7割)、それに関連して知っている必要がある(3割法則)

▶ それ(新しい事)は、既に知っている何 (三つ考える) と関係があるか？

## □学習の為の基本 (無知に気付いた時がチャンス)

○「知らない事」を放置するな

▶ それは何か「調べて」みよう

▶ (数学)定義や公理は覚えよう：何度も調べている内に「覚えてしまう」のが理想的

○「解らない事」を放置するな

▶ 答が得られるまで「考え」よう

▶ (数学) 証明は、try しよう：常に「何故そうするのか？」という意識を持とう

○「出来ない事」を放置するな

▶ 手順を覚えて、繰り返し「練習」しよう

▶ (数学) 色々な計算を行ってみよう：計算結果から判る事実というものもある

# 「コピー・ペースト」の良し悪し

---

## □ 情報の特質

- 「情報」の最大の利点は「コピーできる」事

  - ▶ cf. 「活版印刷」は、「ルネッサンスの三大発明」

- 「情報」の効率的な利用法は「コピーする」事

  - ▶ cf. 「学習」とは、「過去の成果」を「頭にコピー」する事 / 「応用」とは、「頭の中の情報」を「現実にコピー」する事

## □ 「コピペ」の問題点 (なぜ、「STAP 細胞」騒ぎが起きたのか ?)

- 「情報」と「情報の表現(コピペの対象)」は異なる

  - ▶ 「情報のコピー」と「表現のコピー(コピペ)」は異なる

- 「情報の表現」は単独では意味を持たない(他の表現と関連して意味を持つ)

  - ▶ 文脈(他の情報)を無視した「表現のコピー」は、「元の情報」を表現しない

  - ▶ cf. 友達のやった課題を提出(ついでに、提出者の名前も友達の名前をコピペ) / 友達が結婚した、自分も結婚したいので友達の妻にプロポーズ

- 正しい「コピペ」をするには「文脈の理解」が不可欠

  - ▶ 正しく「情報をコピペ」するには、「表現の変更」が必要な場合がある(「表現のコピペ」でだめ)

  - ▶ cf. 「I like it」, 「Me too」は良いが、「I love you」, 「Me too」は駄目(「I love you too」が正しい)

## □ 正しく内容が理解できているなら「コピペ」は強力な道具

- 「コピペ」するなら、「内容を理解して」から

  - ▶ 必要に応じて(コピペの後に..)「表現の変更」をする

# [演習 1] google を利用してみよう

---

## □ google のページ

- url : <http://www.google.co.jp>

- ▷ お気に入り(ブックマーク)に入れておこう

## □ グーグルの使い方

- 「グーグルの使い方」というキーワードで検索してみる

- 候補は複数に当たってみる : 「タブ」を有効利用しよう

## □ Web Page の扱い方

- まずは、「ブラウジング(拾い読み)」をする

- ▷ いきなり頭から読むのは非効率 ( 当り外れがある [対] 教科書 )

## □ 本当にそのページは、読む価値があるのか？

- ページの内容は、「求めている物事」に関係するか？

- 内容は適切？ (著作権に注意)

- ▷ 量(網羅性はあるか)、表現(理解り易いか)、便利性(転用可能か)

- 内容は \*正しい\* か？

- ▷ 論理的？ / 根拠はあるか？ / 他の情報と矛盾していないか？

- 最低、三つのページをブラウズし、比較する必要がある

- ▷ [悪例] 「I'm Feeling Lucky」を見て「ない」と言う...



# [演習 2] 課題レポートを作成しよう

---

## □ 今週 (2016/05/10) の課題

- 最近、自分が気になっている単語 ( 10 個 ) に関する URL を三つずつ紹介する

## □ 解題を解く手順

- 検索キーワードを 10 個考える
- Google 等の検索エンジンを利用してキーワード検索する
  - ▶ 表示されたページを幾つかブラウズする
  - ▶ キーワードの説明として適切だと思われるページの URL をコピーする
- 「メモ帳」に、URL をコピーペーストする
- 形式は、サンプル(20160510-9999.txt)を真似して作成する
- ファイル名を「20160510-QQQQ.txt (QQQQ は学生番号)」にする
- CST Portal に提出する

# CST Portal を利用したレポート提出

---

## □ 課題の提出

- 主に、CST Portal を利用する
- 毎回、「講義連絡」で「レポート提出」を要求する
  - ▶ 課題を作成して CST Portal 経由で提出する事

## □ URL

- <https://newportal.cst.nihon-u.ac.jp/cst/top.do>
  - ▶ 「コンピュータ概論」の Page のリンクからたどる
  - ▶ CST で検索して理工学部の Web Page からリンクをたどる

## □ 出席

- 出席は成績に影響しないが、折角なので、やる
  - ▶ 「講義」を開くと、「出席ボタン」があるので押す
  - ▶ 「出席パスワード」を聞かれるので、「20160510」と入力する

# 補習について

---

## □ 補習の時間

- 内容：数学科 1 年の講義のサポートを行う
  - ▶ 基本は「質問の時間」とするが、講義をする事も有り得る
- 日時：前期は、基本、火曜日 5 限 (コンピュータ概論 A の後)
  - ▶ 不定期：毎週とは限らない ( Web を参照の事 )
- 場所：基本は 1211 GW (コンピュータ概論 A の教室)
  - ▶ 未確定：毎回 1211 とは限らない ( Web を参照の事 )
- 対象：基本 数学科 1 年生だが、それとは限定しない
  - ▶ 少なくともコンピュータ概論の受講者は対象となる

# 「説明」

---

## □ 説明とは

○「新しい知識」を「(聞き手の)既存の知識」と結びつける「内容」を示す事

▷ 事例(相手が知っている事)から始める

▷ その後に、それと \*関連付け\* て、「定義(What)=新しい知識」を述べる

▷ ※「定義を述べ」て初めて「説明が終」る(「例示」だけでは説明と言わない)

○「相手の知っている(筈..)の事」を知らないと「説明」できない

○ 説明の例

▷ 「素数とは、2,3,5,7 (例示)の様に、『約数が自分自身と 1 の二つだけの自然数(定義)』の事」

▷ 「素数は無限にある」、「任意の自然数は素数の積として一意に表現できる」

# 「理解」

---

## □ 理解とは

### ○ 「自分の知識」と「新しい知識」を関連付ける作業

- ▶ この新しい知識は、自分のどの知識と関係があるか？
- ▶ 新しい知識に当てはまる、「自分の知っている事例」は何か？
- ▶ ある程度、知識がないと、新しい知識を「理解する事」ができない
- ▶ ※ 知識が有るば有る程、新しい知識を理解する事が容易になる

### ○ 理解の例

- ▶ 「13 も自分自身(13)と 1 の二つの約数しかないから素数(正例が言える)」
- ▶ 「14 は、14 と 1 の他に 7, 2 も約数なので素数でない(反例が言える)」
- ▶ 「素数の定義(新知識)」と、「13, 14 の約数(旧知識)」の「関係付け」ができた

### ○ 応用(理解の先)

- ▶ 従来知識と新しい知識を組合せて、与えられていない知識を作り出す事
- ▶ cf. 新しい「演算法」を学ぶと、「式の値を求める」事ができるようになる

# 「解答」

---

## □ 解答とは

- 基本は、「定義(What)」を述べる事

- ▷ 最初に「定義」を述べて、(必要なら..)次に「例」や「性質」を述べる

- 解答の例(「素数とは何か?」という質問に対する答え)

- ▷ 「素数とは『約数が自分自身と1の二つしかない自然数』の事(で、例は13や17)」

## □ 「説明」と「解答」の違い：「述べる順序(表現:定義が後か先か)」が異なる

- 内容は同じ：共に、「定義」を述べる事

- 状況(メタ情報)が異なる

- ▷ 説明→相手が定義を知らない

- ▷ 解答→相手が定義を知っている

- 解答と説明を混同してはいけない(「説明(の鸚鵡返し)」は「解答」ではない)

- ▷ 「説明→解答」は、「コピペ」してはいけない、典型的な例(表現を変える必要がある)

- ▷ ましてや、「例示だけ(無様なコピペ)」は「説明でも解答でもない」