

# ファイルの拡張子とサイズ

情報科学の世界 2

2023 年度前期

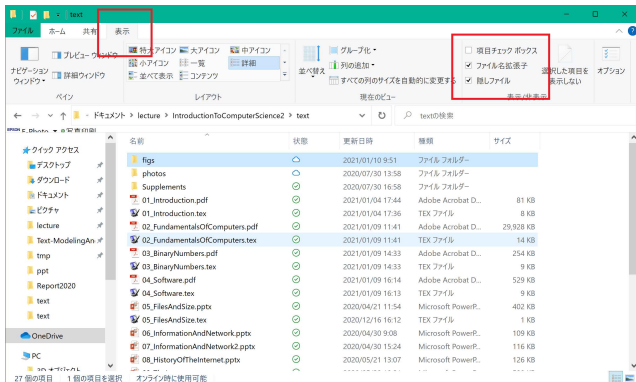
佐賀大学工学部 只木進一

- ① 拡張子とファイルの種類
- ② アプリケーションとファイル
- ③ ファイルサイズ

# 拡張子の表示

- ファイルの拡張子 (file extensions) は、**ファイルの種類**を示す重要な情報
  - アプリケーションとの関係付け
- **拡張子を表示するように設定**を変更しよう
  - explorer で設定を変更
  - 「表示」で「ファイル名拡張子」にチェック
- 拡張子に対応した特別な内部構造データ
- ウィルスはアイコンを偽るため、アイコンを信用してはいけない

# 拡張子の表示



- ファイルエクスプローラーの「表示」メニュー
- ファイル拡張子、隠しファイルをチェック

# 主要なファイル拡張子

- テキスト  
txt、csv、html、pdf
- 画像  
jpg、gif、png
- 動画  
mpg、wmv
- Office 関連 docx、xlsx、pptx

# 主要なファイル拡張子

- 実行形式

exe

メールに添付されている場合には、開かないように

- アーカイブ

zip

フォルダや複数ファイルをまとめて一つに

# Office ファイルには大きく 2 種類

- 3 文字の拡張子: 古いタイプ
  - doc, xls, ppt
  - バイナリファイル
- 4 文字の拡張子
  - zip ファイル
  - 中身はテキスト
  - サイズは小さめで、かつ安全性が高い

# アプリケーションとファイル

- アプリケーションによって扱えるファイルに制限がある
  - 対応関係の設定がある
- ファイルをメールで送る場合の注意
  - 何をしてほしい？見るだけ？編集して欲しい？
  - 相手はアプリケーションを持っている？



# 例：テキストファイル

- 文字の表現
  - 数字やアルファベット (半角)→1 バイト
  - かなや漢字 (全角)→2 バイト
- 日本語はコードに注意
  - JIS、SJIS、EUC、UTF

# 例：テキストファイル

- txt ファイルは素直に文字数の大きさになる（見えていない文字もある）
- メールは、ヘッダと本文から構成される
  - ヘッダに文字コード指定
- HTML も、ヘッダと本文
  - ヘッダに文字コード指定
- 改行コード

OS	改行コード名	改行コード	
Windows	CRLF	\r\n	0x0d0a
Mac, Linux	LF	\n	0x0a

# ファイルサイズ

- 半角英数文字 → 1Byte=8bit
- 全角文字 (JIS や SJIS) → 2Byte=16bit
- 1kByte = 1024Byte
- 1MByte = 1024kByte
- 1GByte = 1024MByte
- 1TByte = 1024GByte

# データ圧縮

- 可逆圧縮
  - 圧縮 (compress) と展開 (decompress) が可逆的
  - データ中のパターンなどを利用
  - 対象によって手法を選択
- 非可逆圧縮
  - 圧縮と伸長 (expand) は非可逆
  - 伸長時の影響を最小化

# 画像ファイル：XGA ディスプレイの場合

- 1024 × 768 の点
- 各点に何色表示できるか
  - 8bit (256 色)、16bit (65336 色)、24bit(1677 万色)
- 24bit (3Byte) だと
  - $24 \times 1024 \times 768 \text{bit} = 18.87 \text{Mb} = 2.36 \text{MB}$
- jpg だと 0.2MB に圧縮される

# 色深度

- 8ビット
  - 赤 3ビット、緑 3ビット、青 2ビット
  - ヒトは青への感度が赤と緑に比して鈍感
- 16ビット
  - 赤 5ビット、緑 6ビット、青 5ビット
- 24ビット
  - RGB 各 8ビット

# 演習

各自の PC の、画面解像度、色深度を確認しなさい。

# 24 ビットカラーの例

r	g	b	
255	128	0	Orange
0	128	64	Green
128	128	128	Grey
64	0	128	Purple



# 「1000 万画素」のデジタルカメラ

- 3648 × 2736 画素
- 24 ビット画像は 29.94MB
- 実際の jpeg ファイルは 2.86MB

# 音声データ

- アナログ信号のデジタル化
  - サンプルング: 決まった時間間隔でサンプルング
  - 整数へ変換
  - 符号化
- CD の場合
  - 44.1kHz でサンプルング
  - 16bit で符号化
  - 2ch
  - 1 時間で約 605.6MB

# 課題

- 自身の PC で拡張子を表示する設定に変更する
- Explorer で「詳細」を表示し、ファイルサイズを確かめる
  - Office ファイル
  - 画像など