最終更新　2020.7.11

# ファイルのしくみ

奈良教育大学　薮哲郎

　コンピューターの仕組みを理解する上で一番重要なのが「ファイル」という概念です。スマホ（Android, iOS）ではユーザーがファイルを直接操作することはできません。アプリを介して操作します。それに対して、Windowsはエクスプローラーというアプリでユーザーが直接ファイルを操作することを許しています。Mac OSはWindowsとスマホの中間に位置します。Mac OSではファインダーというアプリを使ってファイルを操作できますが、Windowsに比べるとブラックボックス化の度合いが大きいです。

　本実習ではWindowsについて学びます。Windowsで学んだことは、他のOSにも応用できます。

（授業を受ける前に以下のことを行っておいて下さい）

(1)フリーのバイナリエディタFavBinEditをインストールしてください。「FavBinEdit」で検索してベクターのサイトへ行き、FavBinEdit-1-2-4.exeをダウンロードし、ダブルクリックしてインストールして下さい。C:\Program Files (x86)\FavBinEdit以下にインストールされます。

　自分のパソコンに余計なアプリをインストールしたくない人は、代わりにインストール不要のバイナリエディタStirlingを使って下さい。ただしStirlingを使うにはlzhファイルを解凍する必要があります。Windows 10ではデフォルトでlzhファイルを解凍できないので、7-zipなどのアプリのインストールが必要です。結局、何かをインストールする必要があります。Stirlingを使う人は本文中のFavBinEditをStirlingに置き換えて下さい。

(2) 以下のファイルを「右クリック」→「対象をファイルに保存」で、ローカルのパソコンにダウンロードしておいて下さい。

　「sjis.txt」「utf8.txt」「a.bin」「拡張子に結びつけられたexeファイルを調べる（shell\_exec.exe）」「システムフォルダなどの場所を表示する (kasou.exe)」「Wordファイル」「Excelファイル」「PowerPointファイル」「テキストファイル」「htmlファイル」「bitmap形式画像ファイル」「jpeg形式画像ファイル」「PDFファイル」

## ファイルの分類

　コンピューターはアプリ、データなど全てを「ファイル」という単位で管理します。全てのファイルはバイト（0～255までの数値を表すことが出来る）の羅列です。ファイルのイメージを図2.1に示します。



図2.1　ファイルのイメージ

課題1

(1) sjis.txtを「右クリック」→「名前を付けて保存」し、FavBinEditで開きなさい。sjis.txtの内容をバイナリ（16進数の数値）で表示します。「半角文字の"a"」「改行文字」「全角文字の"あ"」がどのような数値で表されるか、それぞれ書きなさい。

(2) utf8.txtを「右クリック」→「名前を付けて保存」し、FavBinEditで開きなさい。(1) と同じ質問に答えなさい。

(3) 「file2020.docx（本ファイル）」, 「FavBinEdit.exe」を開き、ファイルの先頭から16×3バイトの内容をコピペしなさい。FavBinEdit.exeはC:\Program Files (x86)\FavBinEditの中にあります。

(4) a.binは数値 -50を4バイト実数で表したときのメモリイメージです。IEEE754での表現形式はどうなりますか？　a.binをFavBinEditで開き、確認しなさい。

　ファイルには名前が付いています。ファイルの名前のうち、最後のピリオドから後の部分を拡張子と呼び、ファイルの種類を表します。ファイルを大きく二つに分けると次のように分けることが出来ます。

1. 実行型ファイル（アプリ本体）

2. データファイル

　実行型ファイルは拡張子がexeやdllとなっているファイルです。「アプリ = 実行型ファイル」です。FavBinEdit.exeは実行型ファイルです。ダブルクリックすると起動します（実行します）。

　データファイルはdocxを拡張子に持つWordファイル、xlsxを拡張子に持つExcelファイル、pptxを拡張子に持つPowerPointファイル、txtを拡張子に持つテキストファイルがその代表です。

　データファイルをダブルクリックすると、そのデータファイルを読み書きするためのアプリが起動し、データファイルを読み込みます。

## データファイルの種類

　データファイルは次のように分類できます。

◆ 仕様が公開されており、多くのソフトで読み書き可能なファイル

　　txt（テキスト）, pdf（文書）, jpg（画像）, bmp（画像）, wav（音声）, mp3（音声）など

◆ 仕様は公開されているが、ソフト間の互換性は完全ではないファイル

　　docx（Word）, xlsx（Excel）, pptx（PowerPoint）など

◆ アプリケーション固有の形式のファイル

　　psd（photoshop）, ai（illustrator）など

　仕様が公開されている形式の中でもテキストファイルは特に重要です。テキストファイルは文字だけが入ったファイルです。「文字」と「改行記号」だけを含むので、シンプルであり、多くのアプリケーションが「読み込み」に対応しています。拡張子はtxtです。テキストファイルを読み書きするプログラムはエディタと呼ばれます。「メモ帳」はエディタの一つです。

　Webページを表すhtml（あるいはhtm）ファイルもテキストファイルの一種です。Webサイトを表示し、ブラウザの「ソースを表示」のメニューで内容を見ることができます。

確認

　docxファイルをメモ帳で開きます。「ファイル」→「開く」のときに「テキスト文書 (\*.txt)」ではなく「全てのファイル (\*.\*)」を選ぶと、docxファイルも選べます。訳が分からない文字列が表示されます。

　docxはShift\_JIS あるいはutf-8で書かれたファイルではありませんが、メモ帳はそれをShift\_JISあるいはutf-8で書かれたテキストファイルであるとみなして、表示します。しかし、docxファイルはそうではありません。結果として、無意味な文字列が表示されます。

　テキストファイル以外のファイルをバイナリファイルと呼びます。バイナリファイルのうち、jpg, bmp, wav, mp3などは仕様が公開されているので多数のソフトウェアで読み書きができます。例えば、bmpはペイント、Word, IEなど多数のアプリケーションソフトがサポートしています。一方、docx, xlsx, pptxなどのファイルはOffice Open XMLという仕様に基づいていますが、docxはWord以外のアプリケーションで読み込むと、レイアウトが崩れたりします。

課題2

　docx, xlsx, pptxはzipファイルであり、多数のファイルを1つにまとめたものです。拡張子をzipに変更してから展開すると、中のテキスト、画像などを取り出すことができます。

　本ファイルは「ファイル」→「オプション」→「詳細設定」で「イメージのサイズと画質」について「ファイル内のイメージを圧縮しない」にチェックを入れています。ゆえに、本ファイル中に高解像度の画像が内包されています。展開し、その中から、図2.2の高解像度のたぬきの画像を取り出し、縦横のサイズを答えなさい。

（補足）

　デフォルトでは「ファイル内のイメージを圧縮しない」にチェックが入っていません。その場合の解像度は96 dpi, 150 dpi, 220 dpi, 330 dpi, 高品質, のどれかを選択することになります。docxが内包する画像は解像度が落ちたものとなり、オリジナルの画像は存在しません。

　また、docxファイルに含まれている画像が高解像度でも、クリップボードにコピーすると、解像度が劇的に落ちます（私の環境では344×258になります）。



図2.2　たぬきの画像

## 拡張子と登録されているプログラムの関連づけ

　データファイルはその拡張子が登録されている場合、ダブルクリックすると、登録されているアプリケーションが起動し、そのデータファイルを開きます。例えば、○○.docxというファイルをダブルクリックするとWordが起動し、そのファイルを開きます。

　登録されている拡張子の確認や変更は以下のようにできます。1つめの方法は「スタート」→「設定」→「アプリ」→「既定のアプリ」から操作します。2つめの方法は「その拡張子を持ったファイルを右クリック」→「プロパティ」→「全般」で「変更」をクリックします。

　ダブルクリックしたときに起動されるプログラムは図2.3のshell\_exec.exeで確認することができます。shell\_exec.exeを実行して下さい。

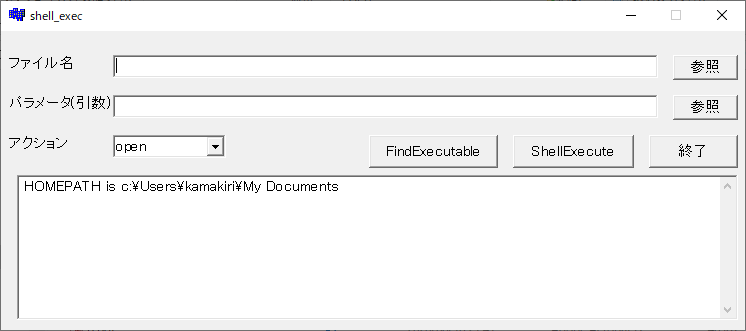


図2.3　shell\_exec.exeの外観

　ファイル名としてデータファイルを指定します。データファイルをドラッグ＆ドロップすることも出来ます。「FindExecutable」のボタンを押すと、そのデータファイルに結びつけられているアプリケーションがあれば、表示します。「ShellExecute」のボタンを押すとそのアプリケーションを起動し、ファイルを開きます。

課題3

　拡張子docx, txt, xlsx, pptx, jpg, html, pdf, bmpに結びつけられているアプリケーションの名前をフルパスで述べなさい（アプリケーションの名前をC:\Program Files (x86)\.....\○○.exeのように述べる）。ただし、ストアアプリで開かれるファイルについては、"Find Executable is fail." と表示されます。その場合は「不明」と答えなさい。

　サンプルファイルは学習用Webサイトに用意してあります。

ファイルの拡張子

　ファイルの拡張子は通常は1～4文字であり、そのファイルがどういう種類のファイルであるかを表しています。代表的なものは以下の通りです。

　exe 実行型ファイル

　dll プログラムを実行するときに必要なライブラリ

　txt テキストファイル

　htm htmlタグを含むテキストファイル

　docx Wordファイル

　xlsx Excelファイル

　pptx PowerPointファイル

　bmp Bitmap形式の画像ファイル

　jpg Jpeg形式の画像ファイル

　png PNG形式の画像ファイル

　wav 無圧縮の音声ファイル

　mp3 mp3形式の音声ファイル

　m4a AAC形式の音声ファイル　iTunesのデフォルト

　mp4 mpeg4形式の動画ファイル

## 仮想と実体

　Windowsはコンピューターを扱いやすくするために、仮想的な世界を作り、それを使用者に見せています。例えば、デスクトップは仮想的なものです。

　一方で、コンピューターは「ファイル」で情報を管理しており、階層化されたフォルダにファイルを格納します。

　エクスプローラーを起動すると、PCの下に「ダウンロード」「デスクトップ」「ドキュメント」「ピクチャ」「ビデオ」「ミュージック」「ローカルディスク(C:)」などのアイコンがあります。このうち「ローカルディスク(C:)」のように、末尾に括弧でくくったローマ字1字があるのは実体を持ちます。すなわち、物理的に存在するドライブに対応します。Windowsでは「ハードディスク」「DVDマルチドライブ」「USBメモリ」などにドライブレターと呼ばれるローマ字1文字を割り当てます。

　それ以外の「デスクトップ」などはWindowsが仮想的に作っているものです。「ドキュメント」「ピクチャ」などは1個のフォルダに対応します。その実体はどこにあるのでしょうか？　ファイルをバックアップするときは「実体がどこにあるか」を知っておく必要があります。

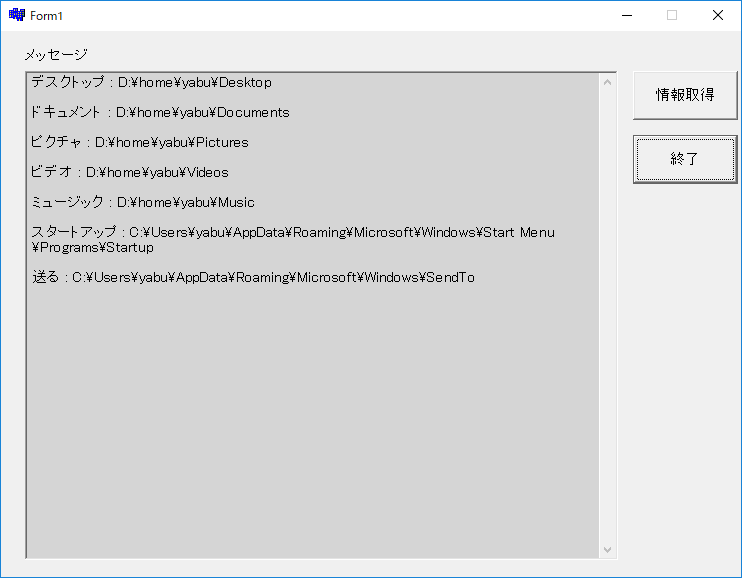


図2.4　kasou.exeの外観

　図2.4に示すkasou.exeを起動し、「情報取得」ボタンを押して下さい。デスクトップ、ドキュメントなどの実体がどこにあるかを表示します。

「送る」メニューにメモ帳を追加してパソコンを便利にする

　ファイルを右クリックして出現するメニューの「送る」にメモ帳を入れておくと、便利です。メモ帳を「送る」のメニューに追加するには、以下のように操作します。

1. ショートカットの作成 …… メモ帳の実体はC:\Windows\System32\notepad.exeです。右クリックしてデスクトップにショートカットを作成してください。

2. 1. で作成したショートカットを「送る」に対応するフォルダにコピーしてください。そのフォルダはkasou.exeで求めることができます。

　実は「送る」に対応するフォルダは次の方法でも開くことができます。こちらの方が簡単です。

　　　　スタートメニューを右クリック → ファイル名を指定して実行

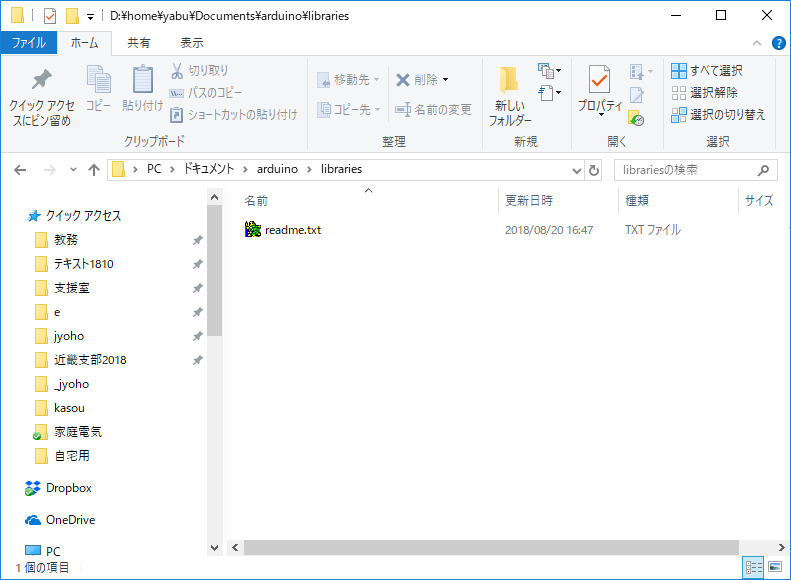
　　　　→ shell:sendto と入れる

（追加の知識）

　shell:startupと入れると、スタートアップフォルダ（パソコン起動時に実行したいアプリがある場合、そのアプリへのショートカットを入れておくと、自動起動する）。

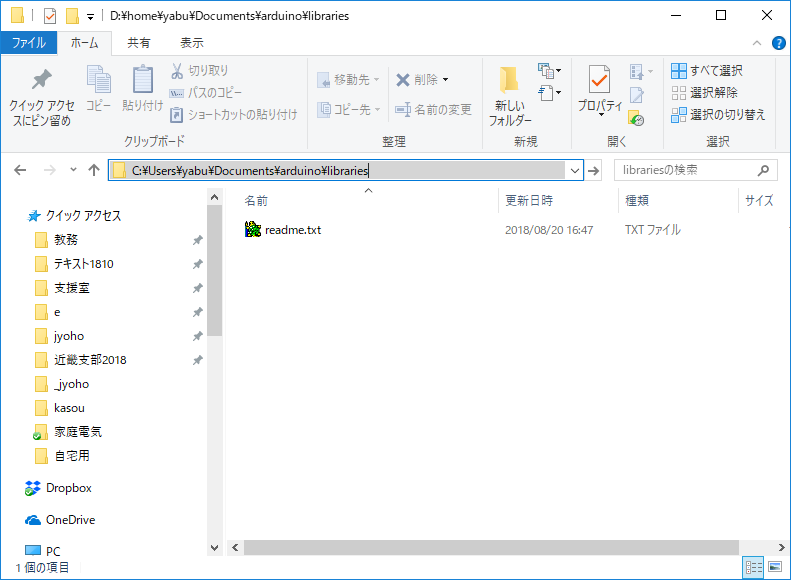
「ドキュメント」フォルダがどこにあるかを知る

　エクスプローラーでドキュメントの直下にあるフォルダを何か開いてください。その状態でエクスプローラーのファイルのパス名を表示しているエディットボックスの右側の空白部分をクリックしてください。その場所の実体がある場所が表示されます。この様子を図2.5に示します。



ここをクリック

(a) クリックする場所



(b) 実体の場所が表示される

図2.5　実体のある場所を表示する