

情報の電子化 (4) 画像、音声

かつらだ まさし
桂田 祐史

2002 年 6 月 20 日

うっかりして、情報の電子化の話題を一つ忘れていました...レポート課題 X¹ の落とし前をつける必要があるのです。これを課題 7 とします (以下説明)。なお、6 月末~~々~~切りという課題が多くなったので、少し延長して、7 月 3 日とします。

1 レポート課題 7 のための説明の補足

1.1 最近の外部記憶媒体の容量

- (最近に使われなくなってきたが) フロッピーディスクは、2DD が 720KB, 2HD が 1.44MB
- CD-ROM (CD-R, CD-RW など同様) は 650 ~ 700MB²
- DVD-ROM/R/RW は CD-ROM の約 7 倍 (片面シングル・レイヤーの場合) 4.7GB
- パソコン用ハードディスクは数十 GB 程度で、100GB を超えるものも珍しくない。

1.2 テキスト・ファイルのサイズ

通常、日本語は 1 文字 2 バイトで表現されている。だから日本語のテキスト・ファイルのサイズは大ざっぱに言って、文字数 × 2 バイトと考えれば良い (実際には空白や改行、また英数字など 1 バイトで表現される文字などもあるので、あくまでも概算である)。

文書の文字数を概算するのは簡単だから (一行 文字、一ページ 行、 ページ)、文書の文字情報をテキスト・ファイルにした際のサイズの概算ができることになる。

2 レポート課題 7

以下のことを調べよ。~~々~~切は 7 月 3 日とする。(ファイルのサイズについての感覚を身につけてもらうの主旨であって、自分で計算すること。)

¹第 5 回 <http://www.math.meiji.ac.jp/~mk/syori2-2005/jouhousyori2-2002-05/node6.html>

²音楽用 CD では、当初記録時間が 74 分までだったが、最近では 80 分のもも出て来たことに対応している。

(0) フロッピー・ディスクを使ったことがあるか?(教えて下さい。アンケートのつもりです。)

(1) 自分が触れるコンピューター (情報科学センターの Windows 環境, Solaris 環境, 自宅のパソコン) にあるファイル³のサイズについて調べよ。バイト数以外に、フロッピーや CD-R にどれくらい入るかを記せ。

(a) 文書ファイル

レポート、メール、C プログラム、 \TeX のソース (.tex) など。

(単にサイズだけ書いてもあまり意味がない。長いものもあれば短いものもあるのだから。例えば「印刷して何ページくらいの文書が何バイトになる」等の情報を添えること。)

ワープロソフトの文書ファイルなどを調べてみるのも良い。

(b) 実行可能プログラム

- 自分が普段使っているプログラムをいくつか選び、そのプログラム・ファイルのサイズを調べよ (複数のファイルからなるプログラムもあるので、正しく調べるのは難しいかもしれないが)。

- C で書いたプログラムを持っている場合、コンパイル前 (ソースプログラム) と後 (実行可能プログラム、あるいは機械語プログラム) でどうサイズが変わるか。

(c) 現在、自分が持っているファイルの総容量。

```
isc-xas06% du -k ~
```

とすると、ホームディレクトリ (~) 下のファイルの容量を KB 単位で教えてくれる。ただし /home/ユーザー名/.snapshot の分は差し引くこと。)

(2) 自分が持っている本を一冊選び、その文字情報を記憶するファイルを作った場合、サイズはどれくらいになるか計算せよ。フロッピー・ディスクに記憶する場合、何枚必要か? また CD-R には何冊分記憶できるか。

(3) (もし出来れば) 画像ファイル、音声ファイルなど。

(記録の形式、画像の場合は図の大きさ&色数、音声の場合は時間等も分かる範囲で調べる。)

3 マルチメディア情報の電子化

文書はテキスト・ファイルとして電子化できることを詳しく見てきたが (なお、付録 A にも目を通して欲しい)、画像や音声についてはどうだろうか?

³自分以外の持ち物でも構わない。例えば、私はホームディレクトリ (~re00018) を解放している (読み出し許可にしてある) ので、そこにある Gutenberg テキストや、解析概論 I の講義ノートの \LaTeX ファイルなどを調べることも出来る。

3.1 画像の電子化

コンピュータのディスプレイは、色 (明暗込み) をつけた点 (ピクセルあるいはドット) の集まりとして構成されている (カラー・ディスプレイ上の 1 ピクセルは R, G, B (red, green, blue) に光る 3 つの小さな点で構成されることが多い)。各ピクセルの色情報を数値化して、それを集めることによって画像情報をデジタル・データとして表現できることになる。

例えば、実習に用いているパソコンのディスプレイ画面は、縦横約 1000 ピクセルずつ⁴、つまりおよそ 100 万ピクセルからなっていて、各ピクセルは約 1600 万色の色がつけられる。ピクセル数は最近のパソコン、ワークステーションでは大体この程度である (やや高級なもので 1600 × 1200 程度)。色数は、1 ピクセルの色を 8 ビットで表わす場合 256 色で、16 ビットの場合は 65536 色、24 ビットの場合は約 1600 万色となる⁵。

この素朴な方法で 1 画面分の画像を記録したファイルのサイズはいくらになるか計算してみよう。すぐに分かることは

画像を記録したファイルはサイズが大きい。

そこで、

画像の記録にはデータの圧縮が不可欠

になる。デジタル・データの圧縮法は次のように大きく二つに分類できる。

デジタル・データの圧縮法

- (1) 可逆な圧縮 オリジナルのデータが完全に復元できる圧縮法
- (2) 不可逆な圧縮 オリジナルのデータが完全には復元できない圧縮法— うまくやると圧縮後のデータのサイズを非常に小さくすることができる

文書データやプログラム・データなどは可逆な圧縮をすることになる。gzip, compress, zip, LHa など可逆な圧縮をする。

画像の場合は、オリジナルと多少違っていても、十分役立つことが多いため、不可逆な圧縮が採用されることが多い。色々な圧縮法があり、画像データのフォーマットも多くの種類がある。

画像データは不可逆な圧縮をすることでサイズをかなり小さくできる。

画像圧縮技術は特許などがからんで「取り扱い注意」の部分がある。

例 (GIF の場合)

WWW ページでも良く使われている GIF 画像は、UNISYS という企業が特許を持っている圧縮法 (LZW) を利用することが多い。UNISYS と契約した会社の作成したソフトで GIF 画像を作成する場合は問題ないが、そうでない場合は特許料を支払う必要がある。大した問題

⁴UNIX 環境では、xdpyinfo コマンドを実行すると調べられる。例えば、この文章を書いているパソコンの画面は、横 1024, 縦 768 ピクセル。

⁵ $2^8 = 256$, $2^{16} = 65536 \approx 6$ 万 5 千, $2^{24} = 16777216 \approx 1600$ 万。

ではないと思われるかもしれないが、例えばフリーソフトは閉め出されることになる。今のところ、データを再生するプレーヤー（画像の場合はビューアー？）の開発を制限しているケースはないようだが、万プレーヤーも制限されるようになると、「作成したデータは本当に作成者のものなのか？」心配になってくる（筆者の杞憂であれば良いのだが）。

X や OpenWindows が動いている場合に試せる実験

```
isc-xas06% xdpinfo  
isc-xas06% xwd > image.xwd
```

(結果は省略)

この後マウスの入力待ちになる。ウィンドウを選択して左ボタンをクリックする。

```
isc-xas06% ls -l image.xwd  
(省略)
```

サイズをしてみる。

ルートウィンドウを選択した場合

全画面が記録されて約 3 MB

試しに画面に表示してみよう。

```
isc-xas06% xwud -in image.xwd
```

```
isc-xas06% cat image.xwd | xwdtopnm | pnmtopng > image.png
```

PNG というフォーマットに変換

```
isc-xas06% ls -l image.png
```

サイズをしてみる。

```
-rw-r--r-- 1 re00018 re00018 21723 Jun 20 09:27 image.png
```

かなり小さくなった！

```
isc-xas06% display image.xwd
```

表示してみる。

```
isc-xas06% rm image.xwd
```

終了は、マウス右ボタン, Quit,

ディスクの無駄使いなので削除

(情報科学センターは pnmtopng を用意していない。

```
set path=($path ~re00018/syori2/bin)
```

としてから試してみよう。あるいは GIF 形式に変換する ppmtogif を試してみよう。)

これは動画では特に顕著である。テレビなどでは、1 秒間に約 30 の画像を次々に映すことで動画を表現している。

ワークステーションの画像をテレビのように 1 秒間に約 30 回描き換えるとして、1 時間にどれだけの量のデータが必要となるか？

3.2 音声の電子化

音声は適当な時間間隔で圧力変化を記録することで電子化されるのが基本である。

音楽用 CD (CD-DA) では、44.1kHz のサンプリング・レートで⁶、圧力を 16 ビット (2 バイト) の数値で (普通は 2 チャンネルのステレオで) 記録する。特にデータの圧縮はされないため⁷、1 分間、ステレオで録音されたデータの容量は、

$$44.1 \times 10^3 \times 2 \times 2 \times 60B = 10584000B \approx 10.1MB$$

もの巨大なサイズになる。

音声データに関しても、画像データと同様のことが言える。

音声データも不可逆な圧縮をすることでサイズをかなり小さくできる。

⁶要するに 1 秒間に 44.1×1000 回測定する。人間は通常 20kHz までの音を聴くことができると言われていて、20kHz の音を記録するには、最低でも倍の 40kHz のサンプリング・レートが必要なことから、この程度の値を採用した (らしい)。

⁷何と言っても、CD の誕生は NEC PC-9801 が出たか出ないかの、1980 年代初頭ですから。http://www.sony.co.jp/Fun/SH/2-9/h1.html, http://www.oneoffcd.com/info/historycd.cfm など。

音声データを圧縮する方法も特許がからむことが多い。

最近では著作権保護機能なども組み込まれるようになっている。

MD の場合

MD (MiniDisc) では、^{アトミック}ATRAC (Adaptive TRansform Acoustic Cording) という圧縮技術を採用している。CD-DA の場合の約 1/5 程度のサイズになるということで、容量約 140MB の MD メディアに CD-DA とほぼ同じ長さの音声を記録できる。なお ATRAC3 (MD の LP モード (MDLP)) では、CD-DA の約 1/10 程度のサイズになる。

MP3 の場合

mp3 (MPEG 1 Audio Layer-3) という圧縮技術では、44.1kHz, 16bit ステレオで、約 1/10 程度のサイズ (つまり 1 分 1 MB 程度) に圧縮できる (例えば 3 分 44.39 秒の曲が 3,590,559 バイト)。圧縮・復元の方法は公開されているが、実際に圧縮するプログラム (エンコーダー) は特許料を払わずに作成することはできない (?)。

A データ・フォーマットの選択についての注意 — 文書を例として

文書を電子化するための形式 (データ・フォーマット) として、テキスト・ファイルを紹介したが、それは非常に単純なもので、文書の中にどういう文字がどういう順番で現れるかだけしか表現できない。

一方で世の中で実際に使われている (印刷された、あるいは手で書かれた) 文書には、レイアウトや、文字の大きさ、色なども重要な要素である場合がある。さらには図、写真などが含まれる文書も多い。これらの文書を電子化するには、テキスト・ファイルでは不十分であることは明らかであろう。

そのため、このような文書を扱うソフト (ワードプロセッサ、DTP ソフト、WWW ブラウザー、 $\text{T}_\text{E}_\text{X}$ のような文書整形システム, etc.) では、独自のデータ・フォーマットが採用されている。

独自形式のデータ・フォーマットで電子化された文書は、それを読むためにもそれに対応したソフトを必要とする。そのため次のような問題が生じうる。

- データを送った相手に対応したソフトを持っていないかもしれない。
- そもそもコンピューター環境によってはソフトが存在しないこともある。
- 現在はソフトがあっても、将来はそうでないかもしれない⁸ — 保存の問題。

⁸例えば、私の最初の二つの論文はいずれもコンピューターを用いて書いたが、現在はそれをコンピューターの画面で読んだり、印刷したりするのは難しい。部屋の片隅でほりをかぶっているパソコンを動かせば読めるかもしれないが、そのソフトが対応しているプリンターで、今でもちゃんと動くものは果たして、このキャンパスの中に存在するだろうか？

それほどではなくても、ソフトの細かいバージョンの違いによるトラブルは日常茶飯事である。

- 辛うじて読めるにしても、柔軟な利用が困難なことがある。
(例えば文書の中の単語の検索などが簡単に出来るか？よそのソフトのデータに簡単に取り込めるか？)

自分が入手・作成するデータのフォーマットとして何を選択するか、慎重に考える必要がときどき生じる。

ソフトを作るのは企業である場合が多く、企業の利益を守るために、データのフォーマットを非公開にしたりするなど、必ずしもユーザーの利益優先では考えてくれないことを肝に命じておこう。

きれいに見栄え良く表示・印刷できることは気持の良いことで、時には最も重要なことであるが(少し考えれば例はいくらでも見つかる)、ものによっては、多くの人が手軽に利用できること、また長期にわたって保存できることの方が重要な場合もある。そういうわけでテキスト・ファイルもかなり役立つことがある⁹。

B データの圧縮用コマンド

B.1 UNIX の場合

UNIX では、可逆な圧縮をするための一般的なコマンドとして、compress, gzip, bzip2 などがある。compress は以前はよく使われたが、やはり特許の問題であまり使われなくなってきた。最近は gzip が良く使われている。

<pre>isc-xas06% cd filter isc-xas06% cp alice29.txt alice29.doc isc-xas06% gzip alice29.doc isc-xas06% ls -l alice*</pre>	圧縮する alice29.doc の代わりに alice29.doc.gz という 小さなファイルが出来る ファイルのサイズを調べる。
<pre>isc-xas06% wc alice* isc-xas06% gzip -d alice29.doc.gz isc-xas06% ls -l alice* isc-xas06% diff alice29.doc alice29.txt</pre>	復元する。 alice29.doc.gz が無くなり、alice29.doc が現れる。 変化がないかチェックする (変化はないはず)。

英文テキスト・ファイルなどでは 1/3 程度までにサイズが小さくなる。また gzip -d で decompress (圧縮の解除) すると、完全に元に戻る。

B.2 Windows の場合

Windows では、複数のファイルを一つにまとめてから圧縮をするコマンド (アーカイバー (archiver) と呼ばれる) が普及している (Lha, WinZip, Cab, ...) ¹⁰。私の最近のお気に入り、

⁹実際、私の最初の論文は、テキスト・ファイルとしても保存しておいたので、数式以外の文章の部分は今でも読むことができる。その手の「抜け道」については配慮することが重要である。

¹⁰UNIX では、複数のファイルを一つにまとめるコマンドとして、tar が良く使われている。tar と例えば gzip を組み合わせることで Windows のアーカイバーと同様のことができる。このあたりはプログラムの設計思想の

Lhaca デラックス版 (<http://www1.sphere.ne.jp/app/Lhaca/> から入手可能) である。情報科学センターの Windows 2000 環境には、圧縮機能を除いた (解凍機能のみを備えた) Lhasa がインストールされているが、この選択はちょっと残念¹¹。

C 先週言い忘れたこと

実は私は漠然と夢見ていることがある。数学版 Gutenberg あるいは青空文庫が作れないだろうか?ということである。特に日本人にとって、数学の古典に触れる機会はかなり限られているが (よほど歴史とお金のある大学でないと、19 世紀以前の文献はほとんど持っていない)、それだけに実現できたらどんなに素敵だろうと思う。

権利の問題をどう解決するか、入力をどうするか、種となるテキストをどこで調達するか、とても片手間に出来ることではないのだが...

必要となる技術 (例えば T_EX はその一つ) は揃っているような気がするだけに何とかならないかな、と徒然に考える。

違いであろう (単一の機能を持ったプログラムを複数組み合わせることで複雑な仕事をしようという UNIX と、一つのプログラムで色々なことに対応しようとする Windows)。

¹¹率直に言うと...見識を疑う。