

8章「アーカイブ/圧縮/エンコード」

中島康彦

§8. 0 今日の作業ディレクトリを作る

1. % `cd` ⇒ ホームディレクトリへ移動
2. % `mkdir chap08` ⇒ ディレクトリchap08を作成
3. % `cd chap08` ⇒ ディレクトリchap08へ移動
4. % `pwd` ⇒ 作業ディレクトリの確認

ネットワーク経由でデータを送受信する際に必要です。

§8. 1 アーカイブ

- ▶ **tar** オプション アーカイブ ファイル等
アーカイブに対する操作を行う
オプション

<code>c</code>	⇒	アーカイブの作成
<code>t</code>	⇒	アーカイブ内のファイル一覧の表示
<code>x</code>	⇒	アーカイブからファイルの取り出し
<code>v</code>	⇒	詳細な情報を表示
<code>f</code>	⇒	アーカイブファイルの指定
<code>% (cd ~; tar cvf - Bin) > Bin.tar</code>	⇒	一時的にホームに移動しBin.tarを作成
<code>% ls -l</code>	⇒	アーカイブBin.tarができた
<code>% tar tvf Bin.tar</code>	⇒	Bin.tar内のファイル一覧の表示
<code>% tar xvf Bin.tar</code>	⇒	Bin.tarからファイルの取り出し
<code>% ls</code>	⇒	ディレクトリBinができている
<code>% ls Bin</code>	⇒	Binの中にファイルもできている
<code>% rm -rf Bin</code>	⇒	Binを削除する

§8. 2 圧縮/復元

- ▶ **compress** ファイル(省略時は標準入力)
ファイルを圧縮し, ファイル.zに置き換える.
標準入力の場合, 標準出力へ圧縮する.

<code>% ls -l</code>	⇒	Bin.tarのサイズを確認する
<code>% compress Bin.tar</code>	⇒	Bin.tar.zができる
<code>% ls -l</code>	⇒	サイズが小さくなった

- ▶ **uncompress** ファイル.z(省略時は標準入力)
ファイル.zを復元し, ファイルに置き換える.
標準入力の場合, 標準出力へ復元する.

<code>% uncompress Bin.tar.Z</code>	⇒	Bin.tarができる
<code>% ls -l</code>	⇒	元に戻った

§8. 2 圧縮/復元(続き)

- ▶ **gzip** ファイル(省略時は標準入力)
ファイルを圧縮し、ファイル.gzに置き換える。
標準入力の場合、標準出力へ圧縮する。
 - % **ls -l** ⇒ Bin.tarのサイズを確認する
 - % **gzip Bin.tar** ⇒ Bin.tar.gzができる
 - % **ls -l** ⇒ compressよりもサイズが小さい

- ▶ **gunzip** ファイル.gz(省略時は標準入力)
ファイル.gzを復元し、ファイルに置き換える。
標準入力の場合、標準出力へ復元する。
 - % **gunzip Bin.tar.gz** ⇒ Bin.tarができる
 - % **ls -l** ⇒ 元に戻った

§8. 2 圧縮/復元(続き)

- ▶ **lha** オプション アーカイブ ファイル等
UNIXよりもDOSで一般的なコマンド
オプション
 - c** ⇒ アーカイブの作成
 - l** ⇒ アーカイブ内のファイル一覧の表示
 - x** ⇒ アーカイブからファイルの取り出し

 - % **(cd ~; lha c - Bin) > Bin.lha** ⇒ 一時的にホームに移動しBin.lhaを作成
 - % **ls -l** ⇒ Bin.lhaができた。gzipと同程度に小さい
 - % **lha l Bin.lha** ⇒ Bin.lha内のファイル一覧の表示
 - % **lha x Bin.lha** ⇒ Bin.lhaからファイルの取り出し
 - % **ls** ⇒ ディレクトリBinができている
 - % **ls Bin** ⇒ Binの中にファイルもできている
 - % **rm -rf Bin** ⇒ Binを削除する
-

§8.3 エンコード/デコード

- ▶ メールによる直接送信が可能なのはテキストのみ。
- ▶ 画像や音声などの様々な添付ファイルは、一旦、テキストに変換されている。
- ▶ `uuencode`は、以下の太枠内の文字のみからなるテキストに変換する。

(下位4ビット)

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	NUL	SOH	STX	ETX	EOT	ENQ	ACK	BEL	BS	HT	NL	VT	NP	CR	SO	SI
1	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
2	SP	!	"	#	\$	%	&	'	()	*	+	,		.	/
3	O	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
6	'	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7	p	q	r	s	t	u	v	w	x	y	z	{		}		DEL

§8.3 エンコード/デコード(続き)

- ▶ `uuencode` ファイル(省略時は標準入力) 内部名
制御文字を含まないテキストに変換する。
 - ※ `gzip Bin.tar` ⇒ `Bin.tar.gz`ができる
 - ※ `ls -l` ⇒ `Bin.tar.gz`のサイズを確認する
 - ※ `uuencode Bin.tar.gz Bin2.tar.gz > Bin.uu`
⇒ `Bin.tar.gz`から`Bin.uu`を作る
 - ※ `ls -l` ⇒ `Bin.uu`のサイズを確認する
- ▶ `uudecode` ファイル(省略時は標準入力)
内部名のファイルへ復元する。
-pを付けると標準出力へ復元するものもある。
 - ※ `uudecode Bin.uu` ⇒ `Bin2.tar.gz`を復元する
 - ※ `gunzip Bin2.tar.gz` ⇒ `Bin2.tar`を復元する
 - ※ `tar xvf Bin2.tar` ⇒ `Bin`を復元する
 - ※ `ls -l` ⇒ `Bin`が元に戻った

§8.4 応用 ... ディレクトリ全体をメール送信する

以上の方法を組み合わせることにより、ディレクトリ全体をメール送信することができる。

▶ `tar cf - Bin|gzip|uuencode Bin3.tar.gz|mail ...`

tar: Binからアーカイブを作成し標準出力へ
gzip: 標準入力を圧縮し標準出力へ
uuencode: 制御文字を含まない形式に変換
mail: メール送信

▶ `tar cf - Bin|gzip|uuencode Bin3.tar.gz>Bin.uu`
▶ `webmail`

いきなり送信せずにBin.uuに保存し、メール送信時に読み込む。

§8.4 応用 ... 受信メールからデータを復元する

受信メールからのデータの復元

▶ Bin.uuに保存したとする
▶ `uudecode -o - Bin.uu|gunzip|tar xvf -`

uudecode: Bin.uuをデコードしgzip形式を標準出力へ
gunzip: gzip形式を復元し標準出力へ
tar: アーカイブ形式からファイルを取り出す

§8.5 暗号化

▶ crypt 暗号キー

標準入力を暗号化して標準出力へ

標準入力がある暗号の場合は解読して標準出力へ

	0	1	2	3	4	5	6	7	8	9	A	B	C	(下位4ビット)		
0	NUL	SOH	STX	ETX	EOT	ENO	ACK	BEL	BS	HT	NL	VT	NP	CR	SO	SI
1	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
2	SP	!	"	#	\$	%	&	'	()	*	+	,		.	/
3	O	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
6	'	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7	p	q	r	s	t	u	v	w	x	y	z	{		}		DEL

```
% echo ABCDEFG | xd -c
```

```
0: 41 42 43 44 45 46 47 0A ⇒ 16進ダンプ
```

```
% echo ABCDEFG | crypt kyoto | xd -c
```

```
0: F4 44 41 1B B1 E3 FB BE ⇒ 暗号化後
```

§8.6 暗号化の応用 ... 暗号化してメール送信する

§8.4の方法では誰でも読める。暗号化すれば、暗号を知っている人だけが読める。

```
▶ tar cf - Bin|gzip|crypt kyoto|uencode Bin3.tar.gz|mail ...
```

tar: Binからアーカイブを作成し標準出力へ

gzip: 標準入力を圧縮し標準出力へ

crypt: kyotoを暗号キーにして暗号化

uencode: 制御文字を含まない形式に変換

mail: メール送信

```
▶ tar cf - Bin|gzip|crypt kyoto|uencode Bin3.tar.gz>Bin.uu
```

```
▶ webmail
```

いきなり送信せずにBin.uuに保存し、メール送信時に読み込む。

§8. 6 暗号化の応用 ... 受信メールからデータを復元する

受信メールからのデータの復元

- ▶ `Bin.uu`に保存したとする
- ▶ `uudecode -o - Bin.uu|crypt kyoto|gunzip|tar xvf -`

`uudecode`: `Bin.uu`をデコードしgzip形式を標準出力へ
`crypt`: `kyoto`を解読キーにして解読
`gunzip`: gzip形式を復元し標準出力へ
`tar`: アーカイブ形式からファイルを取り出す

§8. 7 例題

`data08`は、何らかのファイルである。今までに説明したコマンドを駆使して、意味のあるテキスト(一行)を取り出せ。

- ▶ `file`コマンドにより種類の見当が付く。
- ▶ 解読キーには「`kyoto`」を用いること。
- ▶ 端末の種類によっては日本語コードの不一致により読めない場合がある。 `nkf`コマンドを試すこと。
- ▶ `nkf`を使っても必要なコマンド数は4を越えない。

§8. 8 今日の課題

data08は何かのアーカイブである. 今までに説明したコマンドを駆使してテキストファイルを取り出し, 内容を示せ. なお使用したコマンド列も示すこと.

▶ 解読キーには「kyoto」を用いること.

宛先: nakashim@econ.kyoto-u.ac.jp

件名: unix1-学生番号

名前(忘れずに)

今日はここまで