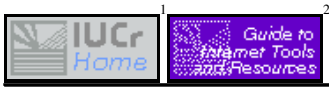

Packing, compressing and encoding files

It is very convenient and efficient for the storage and transfer of files via ftp to pack, compress and encode them. The reasons for so doing are:

- **Packing**: Application software very often covers a whole directory tree of files. If this whole tree structure is packed together into a single file, the task of transfering the software is made that much easier. Often the last modification date of the file is stored within the packed form.
- **Compression** can result in a 30 to 70% reduction in the size of a file depending on the tree structure and the algorithm used. This will save storage space on disk and floppy disks, and reduce the bandwidth necessary for the transfer of the file.
- **Encoding** is frequently used in the transfer of files by e–mail. E–mail systems are in many cases limited to 7–bit ASCII transfer making them *per se* unsuitable for binary programs. To get around this problem, the binary file is first encoded into a text format and then sent by e–mail. Moreover as the binary files tend to be rather long, the encoding algorithm also splits the encoded text file into several sections which are sent as separate e–mail messages.

Available Software

The most popular and widely–used programs for packing, compressing and encoding files are indicated in the list below.

File Extension	Program Name	Pack	Compress	Encode	File Type	Operating System
.arc	ARC	X	X		binary	MS-DOS
.bin	MacBinary			X	binary	Macintosh
.cpt	Compactor Pro	X	X		binary	Macintosh
.gz	GNU zip		X		binary	UNIX
.hqx	BinHex4	X		X	text	Macintosh
.lha,.lzh	Lempel-Ziv-Huffman	X	X		binary	Atari
.msa	Magic Shadow Archiver	X	X		binary	Atari
.sea	Self Extracting Archive	X	X		binary	Macintosh
.shar	Shell Archive	X			text	UNIX
.sit	Stuffit	X	X		binary	Macintosh
.tar	tar	X			binary	UNIX
.uue	uuencode			X	text	UNIX
.Z	compress		X		binary	UNIX
.zip	PKZIP	X	X		binary	MS-DOS
.zoo	Zoo	X	X		binary	MS-DOS

In most cases one identifies the compression or packing algorithm from the file extension. A file may be packed by one program and then compressed by another. For example, `.tar.Z` means that file structure was first packed by *tar* and then compressed with *compress*. In some cases one see only one extension is appended, *e.g.* only `.uue`. But when such a file is decoded (with *uuencode*), a new file is produced with another extension, *e.g.* `.Z`.

The above software can be obtained from many anonymous ftp servers spread around the world. Locations can be identified by using an Archie server.[3] On the anonymous ftp servers one frequently finds a FAQ document giving much fuller information on packing, compressing and encoding.

Image, video and audio

Other compression techniques exist for image, video and audio files. These can be both lossless and lossy algorithms. GIF, JPEG and MPEG are standards in this area.

[1] http://www.iucr.ac.uk/iucr–top/welcome.html
[2] int.index.html
[3] int.archie.html

[Index][4] – 20th June 1996 – © Howard Flack[5] – **Not to be copied or reproduced without permission**

[4] int.index.html
[5] ../ahdf/Howard.Flack.html