

How to Make Packages of Files on UNIX

Norman Matloff

October 21, 2002

©1992-2002, N.S. Matloff

Contents

1 Introduction

One often needs to put a group of files into one single package file. This is typically for the purpose of copying the files from one machine to another.

2 Tar Files

On UNIX systems, the most commonly used pack/unpack program is **tar**. One uses the **c** (“create”) command to pack, and the **x** (“extract”) command to unpack.

For example, suppose I have files A, B and C, of which I wish to put copies into a file g.tar. I could type

```
tar cf g.tar A B C
```

The **c** option means I am creating a tar-file, while the **f** says that the next item on my command line will be the tar-file name. So, I am creating a tar-file named g.tar from the files A, B and C.

Important note: Make sure to include the command-line argument which specifies your tar-file name, in this example g.tar. In the example above, if we were to mistakenly type

```
tar cf A B C
```

then tar would think we wish A to be the tar-file. That would destroy the contents of the original A!

I can now copy the single file g.tar to another machine. At that machine, I can type

```
tar xf g.tar
```

to extract all the files from g.tar.

By the way, if say, B had been a subdirectory, then the g.tar file would also include all the files within B, and the B subdirectory and its files would then be created by the unpacking at Machine 2.

When you unpack a tar-file, in some cases you should use the **p** option, e.g. as

```
tar xpf g.tar
```

This preserves the original file permissions.

3 File Compression

In many cases, a tar-file will also be compressed, both to reduce network transfer time and to save space on the disk. On UNIX machines, this compression is usually done with **gzip** or **bzip2**. (The latter is newer and generally more efficient at compression.)

For instance, if in the example above I type

```
gzip g.tar
```

Similarly, if I use **gzip**, the original file `g.tar` will be replaced by compressed file with a `.gz` suffix added to its name. I can uncompress by using the **-d** option. In other words

```
gzip -d g.tar.gz
```

would replace `g.tar.gz` by the original uncompressed file `g.tar`.

An acceptable suffix is `.tgz`, which means a file that has first been created as `.tar` and then compressed by **gzip**. So, to unpack the file `y.tgz` one could type

```
gzip -d y.tgz  
tar xf y.tar
```

The syntax of using the **bzip2** program is like that of **gzip**.

4 Unpacking Windows Packages

Programs for Windows are generally bundled and compressed using the **zip** program (not **gzip**). You can spot them by their `.zip` prefix.

Most UNIX machines include a program **unzip** which can unpack `.zip` files. You may wish to use this program with the **-a** option, which changes Windows text files to UNIX format.¹

¹The only difference is in end-of-line and end-of-file indicators.