

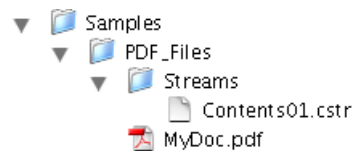
Å. FILE SPECIFICATION

- A content stream's data may be associated with an external file.
- The path to the external file is specified by the *F* entry in the stream dictionary.
 - Any bytes between the *stream* and *endstream* keywords are ignored.
 - The *Length* entry is still required, representing the number of bytes between *stream* and *endstream*, as usual.
 - This will likely be zero.

ABSOLUTE & RELATIVE FILE SPECS

- PDF supports both relative and absolute file specifications.
 - An *absolute* file spec must supply the entire path to the file, starting at the root of the file system.
 - A *relative* file spec supplies a path that starts at the directory in which the PDF file resides.

- For example, if the file *MyDoc.pdf*, at right, contains a reference to the file *Contents01.cstr*, the reference could take one of the two following forms:



- *Absolute:* (/Samples/PDF_Files/Streams/Contents01.cstr)
- *Relative:* (Streams/Contents01.cstr)

FILE SPEC STRINGS AND DICTIONARIES

- A file specification within a stream dictionary may be either a string or a dictionary.
- A string specifies a *simple file specification*, which describes the pathname of the file in an environment-independent format.

This is a chapter from the student notes for the Acumen Training course, "PDF File Content and Structure 2."

You can get more information on this class from the Acumen Training website.

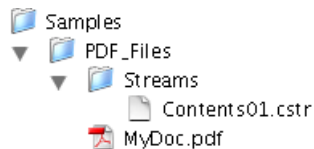
- The viewing software must convert this to a platform-specific file spec at viewing time.
- A dictionary may contain a simple or a *full file specification*.
 - A full file spec provides platform-specific file name and location.
 - It also provides support for embedded files.

FILE SPEC STRINGS

- A string file spec provides either a relative or absolute pathname.
- Subdirectory descents are indicated by slashes within the pathname.
 - These must be converted at viewing time to a character appropriate to the viewer's specific file system.
- Slashes that are part of the pathname must be preceded in the string by a backslash.
 - This means that the slash must be preceded by *two* backslashes in the PDF file.

Absolute Filespecs

- An absolute filespec supplies the complete path from the root directory to the target file.
- Absolute pathnames must start with a slash character:

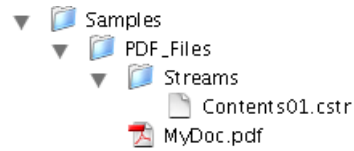


```
(/Samples/PDF_Files/Streams/Contents01.cstr)
```

Relative Filespecs

- A relative filespec takes the PDF file's position within the file system as its starting point.

- In *myDoc.pdf*, in the illustration at right, *Contents01.cstr* could be identified by the relative pathname



(Streams/Contents01.cstr)

- The component “../” at the start of a pathname indicates an ascent in the file hierarchy.

- Thus, in *myDoc.pdf*, at right,

../target.pdf

would be the equivalent of

/Samples/target.pdf

- Instances of “../” may be repeated to move increasingly higher in the file structure hierarchy.

FILE SPEC DICTIONARIES

- A file spec dictionary offers two main advantages over a file spec string:
 - You may include platform-specific file specifications for Mac, Windows, and UNIX environments.
 - You may specify information describing embedded files.
 - These are files that, though logically they are separate files, actually reside within the PDF file.

File Spec Dictionary Contents

Type

/Filespec

- Must be “/Filespec”.

F

(str)

- The platform-independent pathname, as described earlier.
- This is required if the *DOS*, *Mac*, and *Unix* entries are all missing.

Table a-1 File Spec Dictionary Contents

<i>Key</i>	<i>Value</i>	<i>Meaning</i>
Type	/Filespec*	
F	(str) [°]	Platform-ind. pathname
DOS	(str)*	DOS pathname
Mac	(str)*	Mac pathname
Unix	(str)*	UNIX pathname
ID	[(s ₁) (s ₂)]*	File ID strings
V	bool	True if volatile
EF	dict	Embedded files
*Optional		°Conditionally required

DOS (str)

Mac (str)

Unix (str)

- These three entries supply the platform-specific pathnames for the corresponding environment.

ID [(s₁) (s₂)]

- An array of two strings generated from the contents of the target file.
- These are used as an identifier to ensure that the file being referenced has not been changed.
- We talked about ID arrays in the earlier PDF class. (Yes, we did!)

V Boolean

- If *true*, indicates that the target file is volatile and subject to change at arbitrary times.
- In particular, a PDF viewer should not cache the contents of this file.

EMBEDDED FILES

- A file referenced in the *F* entry of a stream may itself be embedded as a stream in the PDF file.

- The *EF* entry in a Filespec dictionary indicates associates a stream with each of the paths specified with the *F*, *Mac*, *DOS*, and *Unix* keywords.
 - You must still provide these paths, but the file contents will actually be embedded in the PDF file.
- The extent to which a PDF viewer pays attention to these files is somewhat indeterminate.

EF Dictionary Contents

- The *EF* dictionary contains up to four entries, each an indirect reference to the stream representing an embedded file.

Table 1-1 EF Dictionary Contents

<i>Key</i>	<i>Value</i>	<i>Meaning</i>
F	stream*	Embedded file stream
DOS	(str)*	DOS embedded file stream
Mac	(str)*	Mac embedded file stream
Unix	(str)*	UNIX embedded file stream