

## IEEE Standards Interpretations for IEEE Std 1003.2™-1992 IEEE Standard for Information Technology--Portable Operating System Interfaces (POSIX®)-- Part 2: Shell and Utilities

Copyright © 1996 by the Institute of Electrical and Electronics Engineers, Inc. 3 Park Avenue New York, New York 10016-5997 USA All Rights Reserved.

Interpretations are issued to explain and clarify the intent of a standard and **do not** constitute an alteration to the original standard. In addition, interpretations are not intended to supply consulting information. Permission is hereby granted to download and print one copy of this document. Individuals seeking permission to reproduce and/or distribute this document in its entirety or portions of this document must contact the IEEE Standards Department for the appropriate license. Use of the information contained in this document is at your own risk.

IEEE Standards Department, Copyrights and Permissions, 445 Hoes Lane, Piscataway, New Jersey 08855-1331, USA

### Interpretation Request #71

**Topic:** diff -c **Relevant Clauses:** 4.17,6.1.4

The format specified for diff -c output by IEEE Std 1003.2-1992 subclause 4.17.6.1.4) is not compatible with historical diff -c output. This difference both breaks existing applications that process context diff output and makes it impossible for a conforming patch utility to process diff -c output in either form.

Historically a line consisting of 15 asterisks ("\*\*\*\*\*") was output at the start of the output of each set of differences. The standard requires instead that the line consisting of 15 asterisks be output only once immediately after the file name identification. For example historical context diff output (with one line of context) looks like this: `*** t1 Thu May 5 11:15:51 1994 --- t2 Thu May 5 11:15:51 1994  
*****  
*** 4,6 **** abacus - Abaddon abaff --- 4,5 ---- *****  
*** 8,10 **** abandon ! abandonned abase --- 7,9 ---- abandon ! abandoned abase  
*****  
*** 13 **** --- 12,14 ---- abatis + abattoir + abaxial` but the standard requires that it looks like this: `*** t1 Thu May 5 11:15:51 1994 --- t2 Thu May 5 11:15:51 1994  
*****  
*** 4,6 **** abacus - Abaddon abaff --- 4,5 ----  
*** 8,10 **** abandon ! abandonned abase --- 7,9 ---- abandon ! abandoned abase  
*** 13 **** --- 12,14 ---- abatis + abattoir + abaxial`

This also affects the patch utility because its input is defined in terms of diff output. It makes context diff output, in either the historical format or the format the standard requires, impossible to process by patch. The line with 15 asterisks is key to disambiguating whether a line in the format of "\*\*\* %d,%d \*\*\*\*\n" is part of a "hunk" or is the part of the filename identification, matching expression "\*\*\* filename timestamp" (see 5.22.7.1). Traditional patch considered a line in the format of "\*\*\* %d,%d \*\*\*\*\n" file

identification unless it was preceded by a line of 15 asterisks, in which case it considered it the start of a new hunk. Note, patch does not interpret filename identification in terms of diff output, only hunks are.

It has its own rules that require that two expressions be recognized as filename identification, neither of which include a line consisting of 15 asterisks. Also, the timestamp can not be used to disambiguate between the two lines because it is required that the timestamp not affect the processing of a patchfile (see PASC Interpretation reference 1003-92.2 #14). Would it be possible to change 4.17.6.1.4 by: deleting lines 3352-3353. deleting line 3362 and inserting in its place: First, a string of 15 asterisks shall be output: "\*\*\*\*\*\n" Next, the range of lines in file1 shall be written in following format:

### **Interpretation Response**

The standard states behavior for the diff -c option and conforming implementations must conform to this. However, concerns have been raised about this which are being referred to the sponsor.

### **Rationale for Interpretation**

None.