

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 #88

Topic: BREs **Relevant Clauses:** 2.8.3.3

In the description of backreferences in Basic Regular Expressions (BREs) in POSIX.2, subclause 2.8.3.3 states (in part): The backreference expression `\n` shall match the same (possibly empty) string of characters as was matched by a subexpression enclosed between `\(` and `\)` preceding the `n`. My question revolves around the use of the phrase “(possibly empty)”. In particular, consider the following BRE: `a\b)*c\1` Does this BRE match the string `ÓacÔ`? The BRE `Óa\b)*cÔ` clearly matches `ÓacÔ`. In this case, `\(b\)` fails to match, but `\(b\)*` matches zero or more instances of `ØbØ` (and therefore matches anything). Adding the backreference requires that the backreference match what the subexpression matched (which was nothing). One can read the standard as meaning that this is an empty match, and that therefore the backreference will (in this case) match an empty string. Alternatively, one can read the standard as saying that the backreference fails to match because the subexpression it references failed to match. Which of these interpretations is correct?

Interpretation Response

This request is substantially identical to interpretation #43 part 15, and the resolution to that interpretation applies in this case.

Rationale for Interpretation

None.