Draft addition to ISO/IEC 24772-2

26 October 2024

6.66 Code representation differs between compiler view and reader view

6.66.1 Description of application vulnerability

<<< This needs an Ada Uniode expert to answer!!>>

The vulnerability as described in ISO/IEC 24772-1:2024 6.66 applies to Ada. Ada allows the direction-changing Unicode characters in string literals <<<???>>> or comments and hence is susceptible to the ensuing vulnerability described in ISO/IEC 24772-1:2024 6.66.

<<<I have no idea what Ada says about these characters in identifiers or in white space. Without diving deep into the characterizations of Unicode, I cannot figure out from the Ada manual which Ada or Unicode Normalization rule would prevent their occurrence. Maybe they are deemed separators, so that they cannot be inside lexemes. That still leaves them in white space and comments. Whether they are allowed in string literals and comments, I merely guessed. >>>

Equally, the issue of <CR> without <LF> applies to Ada programs. A small mitigation is that these characters are not allowed in string literals and cannot be part of a comment, as they indicate singly or in combination the end of line and hence the comment. However, they can appear in white space of the program text in arbitrary order.

6.66.2 Avoidance mechanisms for language users

To avoid the vulnerability or mitigate its ill effects Ada software developers can apply the avoidance mechanisms of ISO/IEC 24772-1:2024 6.66.5.