Fortran preprocessor requirements

INCITS/Fortran JoR

[2023-10-24 Tue 11:27]

1 Summary

We summarize the potential requirements for the Fortran 202y preprocessor.

Each heading is of the form `[shorthand-name] One-line description`. There may be an additional tag (category) on the line (such as `:output:`). The section body contains information on

- Where the requirement came from in Fortran discussions and posts
- Current status (TBD, accepted by JoR, accepted by WG5, rejected by JoR, rejected by WG5, etc.)
- Where the reference appears in the C 2018 standard *ISO/IEC 9899:2018 Information technology – Programming languages – C*. (We will update for 2024 when ratified.) We write these references with
  - an opening square bracket `[` and the letter `C`, followed by
  - the section marker `§` and the section number, followed by
  - (optionally) a `¶` followed by the paragraph number, followed by
  - (optionally) a bullet symbol · and the bullet number or letter, followed by
  - (optionally) the section title for redundant context, followed by
  - a closing square bracket.

For example, `[C§5.1.1.2¶1·2 Translation phases]` references section 5.1.1.2 of the C 2018 standard, paragraph 1, second bullet point. This is in the "Translation phases" section.

1.1 Translation phases

Similar to the C standard, we define phases of text processing related to preprocessing. The preprocessor performs the following phases on the input source program before the processor transforms the program for use.

This requires implementations to behave as if these separate preprocessing phases occur, even though they are typically folded together in practice. Source input need not necessarily be stored as files, nor need there be any one-to-one correspondence between these entities and any external representation. The description is conceptual only, and does not specify any particular implementation.

1. Continuation lines in the input source are interpreted, producing a sequence of logical lines (introducing new-line characters for end-of-line indicators). [C§5.1.1.2¶1·1-2]

2. The preprocessor decomposes the logical lines into preprocessing tokens and sequences of white-space characters and comments. This input shall not end in a partial preprocessing token. New-line characters are retained. For the proper handling of compiler directives, comments are retained. Whether each nonempty sequence of white-space characters other than new-line is retained or replaced by one space character is implementation-defined. [C§5.1.1.2¶1·3]

3. Preprocessing directives are executed and macro invocations are expanded. A #include preprocessing directive causes the named header or source file to be processed from phase 1 through phase 3, recursively. All preprocessing directives are then deleted. [C§5.1.1.2¶1·4]

4. Each preprocessing token is converted into a token for subsequent handling by the processor.

2 Phase 1 Continuation handling

2.1 [c-cont] C-style line continuations in directives

- Source: Fla1
- Status: TBD
- C reference: §5.1.1.2¶1·2 Translation phases

In fixed-form and free-from source code, delete a backslash \ immediately followed by a newline character.

From The C standard:

> Each instance of a backslash character (\) immediately followed by a new-line character is deleted, splicing physical source lines to form logical source lines. Only the last backslash on any physical source line shall be eligible for being part of such a splice. A source file that is not empty shall end in a new-line character, which shall not be immediately preceded by a backslash character before any such splicing takes place.

2.2 [fortran-cont-fixed] Process Fortran line continuation in directives CONTINUATION

- Source: Che1
- Status: TBD
- C reference: N/A

In fixed-form input, a character in column 6 that is not blank or the digit zero continues the line with the previous line, even if the previous line is a directive line, or the continuation of a directive line.

2.3 [fortran-cont-free] Fortran line continuation in directives in free form

- Source: Che1
- Status: TBD
- C reference: §6.10.

In free-form input, an `&` character as the last character on a directive line indicates the directive continues on the next line.

2.4 [fortran-cont-free-trim] Fortran line continuation in directives in free form CONTINUATION

- Source: Che1
- Status: TBD

- C reference: §6.10.

In free-form input, an `&` character as the last character on a directive line indicates the directive continues on the next line. When the first non-blank character on the next line is also an `&`, the characters between the ampersands are deleted.

2.5 [c-comment-strip] Strip C-style `/* ... */` comments

- Source: Fla1
- Status: TBD
- C reference: §6.10.

2.6 [comment-definition-cont] Comment lines in definitions with continuation lines

- Source: Fla1
- Status: TBD
- C reference: §6.10.

2.7 [comment-bang] Recognize the comment character '!' in directives

- Source: Che1
- Status: TBD
- C reference: §6.10.

3 Phase 2 Tokenization

3.1 [tokens-case-insensitive] Case insensitive tokens

- Source: Fla1
- Status: TBD
- C reference: §6.10.

3.2 [spaces-end-token] Spaces significant in determining tokens

- Source: Fla1
- Status: TBD
- C reference: §6.10.

3.3 [replace-trigraph] Trigraph sequences replaced

- Source: Che1
- Status: TBD
- C reference: §6.10.

4 Phase 3 Directive processing

4.1 [non-directive] # non-directive

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.2 [#if] # if

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.3 [#ifdef] # ifdef

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.4 [#ifndef] # ifndef

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.5 [#elif] # elif

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.6 [#else] # else

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.7 [#endif] # endif

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.8 [#include] # include

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.9 [#include-computed] # include (computed)

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.10 [#define-id] # define id replacement-list

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.11 [#define-id-function] # define id (id-list) replacement-list

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.12 [#define-id-0-varargs] # define id (...) replacement-list

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.13 [#define-id-n-varargs] # define id (id-list , ...) replacement-list

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.14 [#undef] # undef

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.15 [#line] # line

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.16 [#error] # error

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.17 [#pragma] # pragma

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.18 [#newline] # new-line

- Source: cpp
- Status: TBD
- C reference: §6.10.

4.19 [#show] # show

- Source: Lio3
- Status: TBD
- C reference: §6.10.

4.20 [#import] # import VARNAME

- Source: Lio3
- Status: TBD
- C reference: §6.10.

4.21 [#output] # output filename [–append]

- Source: Lio3
- Status: TBD
- C reference: §6.10.

5 Expressions

5.1 [#-operator] `#`

- Source: cpp
- Status: TBD
- C reference: §6.10.

5.2 [##-operator] `##`

- Source: cpp
- Status: TBD
- C reference: §6.10.

5.3 [defined-operator] `defined`

- Source: cpp
- Status: TBD
- C reference: §6.10.

5.4 [bang-operator] `!`

- Source: cpp
- Status: TBD
- C reference: §6.10.

5.5 [c-expressions] C-style expressions

- Source :::
- Status: TBD
- C reference: §6.10.

5.6 [fortran-expressions] Fortran-style expressions

- Source :::
- Status: TBD
- C reference: §6.10.

6 Predefined macros

6.1 [file-process-date] `__DATE__`

- Source: cpp
- Status: TBD
- C reference: §6.10.

6.2 [file-name-context] `__FILE__`

- Source: cpp
- Status: TBD
- C reference: §6.10.

6.3 [line-number-context] `__LINE__`

- Source: cpp
- Status: TBD
- C reference: §6.10.

6.4 [fortran-conform] `__STDFORTRAN__`

- Source: cpp-ish
- Status: TBD
- C reference: §6.10.

6.5 [hosted-implementation] `__STDFORTRAN_HOSTED__`

- Source: cpp-ish
- Status: Not accepted
- C reference: §6.10.

6.6 [fortran-version] `__STDFORTRAN_VERSION__`

- Source: cpp-ish
- Status: TBD
- C reference: §6.10.

6.7 [file-process-time] `__TIME__`

- Source :::
- Status: TBD
- C reference: §6.10.

6.8 [stringify-macro] `STRINGIFY`

- Source: Clu1
- Status: Not accepted
- C reference: §6.10.

6.9 [scope-macro] `__SCOPE__`

- Source: Clu1, Lio1
- Status: Not accepted
- C reference: §6.10.

6.10 [vendor-macro] `__VENDOR__`

- Source: Clu1
- Status: Not accepted
- C reference: §6.10.

6.11 [no-undecorated-std-definitions] undecorated names (no _) defined by preprocessor

- Source: Lio2
- Status: TBD
- C reference: §6.10.

7 Expansion

7.1 [fixed-no-expand-col-1] No expansion of `C` (or `D`) in column 1

- Source: Ble1
- Status: TBD
- C reference: §6.10.

7.2 [fixed-expand-col-1] Expansion of `C` (or `D`) in column 1

- Source: Ble1, Fla1
- Status: TBD
- C reference: §6.10.

7.3 [fixed-no-expand-col-6] No expansion of column 6

- Source: Kli1
- Status: TBD
- C reference: §6.10.

7.4 [fixed-strip-col-1-comments] Strip Column 1 `C` comments from expanded text

- Source: Fla1
- Status: TBD
- C reference: §6.10.

7.5 [no-expand-string] No expansion in strings

- Source: Ble1, Fla1
- Status: TBD
- C reference: §6.10.

7.6 [no-expand-hollerith] No expansion in Hollerith

- Source: Ble1
- Status: TBD
- C reference: §6.10.

7.7 [no-expand-implicit-char-list] No expansion in `IMPLICIT` single-character specifiers

- Source: Ble1
- Status: TBD
- C reference: §6.10.

7.8 [no-expand-format] No expansion of `FORMAT` specifiers

- Source: Ble1, Fla1
- Status: TBD
- C reference: §6.10.

7.9 [expand-comments] Expansion in comments

- Source: Ble1
- Status: TBD
- C reference: §6.10.

7.10 [expand-directives] Expansion in directives (e.g., OpenMP)

- Source: Ble1
- Status: TBD
- C reference: §6.10.

7.11 [preprocess-fortran-include] Expand `INCLUDE` lines as if `#include`

- Source: Fla1, Jor1
- Status: TBD
- C reference: §6.10.

8 Output form

8.1 [fixed-clip-input] Right margin clipping at column 72

- Source: Fla1
- Status: TBD
- C reference: §6.10.

8.2 [fixed-no-directive-clip] No right margin clipping on directive lines

- Source: Fla1
- Status: TBD
- C reference: §6.10.

8.3 [fixed-output-conform] Expanded text reflects fixed-format rules

- Source: Fla1
- Status: TBD
- C reference: None

9 Sources

- cpp: *cpp* if in the C standard (2018), *cpp-ish* if in C standard, but "Fortranized".
- Ble1: JoR Email threads from Rich Bleikamp re: tutorial *[2022-08-08 Mon 21:34]*.
- Che1: Email from Daniel Chen to JoR *[2022-07-29 Fri 11:08]*.
- Clu1: Email from Tom Clune *[2022-08-01 Mon 10:48]*.
- Fla1: LLVM Flang Preprocessing.md [`https://github.com/llvm/llvm-project/blob/main/flang/docs/Preprocessing.md`]
- Jor1: JoR meeting on preprocessors *[2022-08-22 Mon 10:00]*.
- Jor2: JoR meeting on preprocessors *[2022-09-20 Tue 13:00]*.
- Kli1: Private communication in his head.
- Lio1: Email from Steve Lionel *[2022-08-01 Mon 13:52]*.
- Lio2: JoR discussion forum `https://j3-fortran.org/forum/viewtopic.php?p=561`
- Lio3: JoR discussion forum `https://j3-fortran.org/forum/viewtopic.php?p=562`

10 References

- Jor email re: cpp tutorial for October meeting?
- INCITS+ISO+IEC+9899+2018+(2019)
- LLVM Flang Preprocessing.md[`https://github.com/llvm/llvm-project/blob/main/flang/docs/Preprocessing.md`