Proposal for C2X  
WG14 N2489

Title: Min-max functions  
Author, affiliation: C FP group  
Date: 2020-02-21  
Proposal category: Technical  
Reference: N2400, N2454, N2488

Problem:

The 2019 revision of IEEE 754 replaces the minNum, maxNum, minNumMag, and maxNumMag operations from IEEE 754-2008, which are supported by the fmin, fmax, fminmag, and fmaxmag functions in draft C2X, with two sets of new min-max operations:

minimum, maximum, minimumMag, maximumMag

and

minimumNum, maximumNum, minimumMagNum, maximumMagNum

These two sets differ primarily in their treatment of NaN operands. None of the new operations is equivalent to a min-max operations from IEEE 754-2008.

The IEEE 754 committee took this unusual step because it believed the min-max operations from IEEE 754-2008 were seriously flawed and problematic for their intended use. The new min-max operations are optional in IEEE 754-2019, but only because the 2019 revision was restricted from adding requirements. The next revision is expected to make the new operations mandatory.

The suggested changes below update draft C2X to support the new min-max operations in IEEE 754-2019 as new `<math.h>` functions. Also, given the 754 thinking about the min-max operations from IEEE 754-2008, this proposal removes the fminmag and fmaxmag functions, which were just recently introduced into draft C2X via TS 18661-1. This proposal does not remove the fmin and fmax functions, for compatibility reasons.

See also N2488 regarding C support for IEEE 754-2019.

Suggested changes

In the table “Preferred quantum exponents” in 5.2.4.2.3#7, replace:

fmin, fmax, fminmag, fmaxmag
with

\texttt{fmin, fmax, fminimum, fmaximum, fminimum\_mag, fmaximum\_mag, fminimum\_num, fmaximum\_num, fminimum\_mag\_num, fmaximum\_mag\_num}

In the table “Operation binding” in F.3, replace:

| minNum      | fmin       | 7.12.12.3, F.10.9.3 |
| maxNum      | fmax       | 7.12.12.2, F.10.9.2 |
| minNumMag   | fminmag    | 7.12.12.5, F.10.9.5 |
| maxNumMag   | fmaxmag    | 7.12.12.4, F.10.9.4 |

with:

| fmax        | 7.12.12.2, F.10.9.2 |
| fmin        | 7.12.12.3, F.10.9.3 |
| maximum     | fmaximum     | 7.12.12.4, F.10.9.4 |
| minimum     | fminimum     | 7.12.12.5, F.10.9.4 |
| maximumMagnitude | fmaximum\_mag | 7.12.12.6, F.10.9.4 |
| minimumMagnitude | fminimum\_mag | 7.12.12.7, F.10.9.4 |
| maximumNumber | fmaximum\_num | 7.12.12.8, F.10.9.5 |
| minimumNumber | fminimum\_num | 7.12.12.9, F.10.9.5 |
| maximumMagnitudeNumber | fmaximum\_mag\_num | 7.12.12.10, F.10.9.5 |
| minimumMagnitudeNumber | fminimum\_mag\_num | 7.12.12.11, F.10.9.5 |

In F.3, after #2, insert:

[2a] The \texttt{fmin} and \texttt{fmax} functions provide the minNum and maxNum operations specified in IEC 60559:2011.

Remove 7.12.12.4 and 7.12.12.5 and all references to \texttt{fmaxmag} and \texttt{fminmag} functions.

After 7.12.12.3, add:

7.12.12.4 The \texttt{fmaximum} functions

Synopsis

[1] \#include <math.h>
   double fmaximum(double x, double y);
   float fmaximumf(float x, float y);
   long double fmaximuml(long double x, long double y);
   \#ifdef __STDC_IEC_60559_DFP__
   _Decimal32 fmaximumd32(_Decimal32 x, _Decimal32 y);
   _Decimal64 fmaximumd64(_Decimal64 x, _Decimal64 y);
   _Decimal128 fmaximumd128(_Decimal128 x, _Decimal128 y);
   \#endif
   \#endif
Description

[2] The \texttt{fmaximum} functions determine the maximum value of their arguments. For these functions, \(+0\) is considered greater than \(-0\). These functions differ from the \texttt{fmaximum_num} functions only in their treatment of NaN arguments (see F.10.9.4, F.10.9.5).

Returns

[5] The \texttt{fmaximum} functions return the maximum value of their arguments.

7.12.12.5 The \texttt{fminimum} functions

Synopsis

[1] \#include <math.h>
   double fminimum(double x, double y);
   float fminimumf(float x, float y);
   long double fminimuml(long double x, long double y);
   \#ifdef \_STDC\_IEC\_60559\_DFP\_
      \_Decimal32 fminimumd32(\_Decimal32 x, \_Decimal32 y);
      \_Decimal64 fminimumd64(\_Decimal64 x, \_Decimal64 y);
      \_Decimal128 fminimumd128(\_Decimal128 x, \_Decimal128 y);
   \#endif

Description

[2] The \texttt{fminimum} functions determine the minimum value of their arguments. For these functions, \(-0\) is considered less than \(+0\). These functions differ from the \texttt{fminimum_num} functions only in their treatment of NaN arguments (see F.10.9.4, F.10.9.5).

Returns

[5] The \texttt{fminimum} functions return the minimum value of their arguments.

7.12.12.6 The \texttt{fmaximum\_mag} functions

Synopsis

[1] \#include <math.h>
   double fmaximum_mag(double x, double y);
   float fmaximum_magf(float x, float y);
   long double fmaximum_magl(long double x, long double y);
   \#ifdef \_STDC\_IEC\_60559\_DFP\_
      \_Decimal32 fmaximum\_magd32(\_Decimal32 x, \_Decimal32 y);
      \_Decimal64 fmaximum\_magd64(\_Decimal64 x, \_Decimal64 y);
      \_Decimal128 fmaximum\_magd128(\_Decimal128 x, \_Decimal128 y);
   \#endif
Description

[2] The \texttt{fmaximum\_mag} functions determine the value of the argument of maximum magnitude: \(x\) if \(|x| > |y|\), \(y\) if \(|y| > |x|\), and \texttt{fmaximum}(x, y) otherwise. These functions differ from the \texttt{fmaximum\_mag\_num} functions only in their treatment of NaN arguments (see F.10.9.4, F.10.9.5).

Returns

[5] The \texttt{fmaximum\_mag} functions return the value of the argument of maximum magnitude.

7.12.12.7 The \texttt{fminimum\_mag} functions

Synopsis

[1] \#include <math.h>
    double fminimum\_mag(double x, double y);
    float fminimum\_magf(float x, float y);
    long double fminimum\_magl(long double x, long double y);
    #ifdef __STDC_IEC_60559_DFP__
        _Decimal32 fminimum\_magd32(_Decimal32 x, _Decimal32 y);
        _Decimal64 fminimum\_magd64(_Decimal64 x, _Decimal64 y);
        _Decimal128 fminimum\_magd128(_Decimal128 x, _Decimal128 y);
    #endif

Description

[2] The \texttt{fminimum\_mag} functions determine the value of the argument of minimum magnitude: \(x\) if \(|x| < |y|\), \(y\) if \(|y| < |x|\), and \texttt{fminimum}(x, y) otherwise. These functions differ from the \texttt{fminimum\_mag\_num} functions only in their treatment of NaN arguments (see F.10.9.4, F.10.9.5).

Returns

[5] The \texttt{fminimum\_mag} functions return the value of the argument of minimum magnitude.

7.12.12.8 The \texttt{fmaximum\_num} functions

Synopsis

[1] \#include <math.h>
    double fmaximum\_num(double x, double y);
    float fmaximum\_numf(float x, float y);
    long double fmaximum\_numl(long double x, long double y);
#ifdef __STDC_IEC_60559_DFP__
  _Decimal32 fmaximum_numd32(_Decimal32 x, _Decimal32 y);
  _Decimal64 fmaximum_numd64(_Decimal64 x, _Decimal64 y);
  _Decimal128 fmaximum_numd128(_Decimal128 x, _Decimal128 y);
#endif

Description

[2] The \texttt{fmaximum\_num} functions determine the maximum value of their numeric arguments. They determine the number if one argument is a number and the other is a NaN. These functions differ from the \texttt{fmaximum} functions only in their treatment of NaN arguments (see F.10.9.4, F.10.9.5).

Returns

[5] The \texttt{fmaximum\_num} functions return the maximum value of their numeric arguments.

7.12.12.9 The \texttt{fminimum\_num} functions

Synopsis

[1] #include <math.h>
  double fminimum_num(double x, double y);
  float fminimum_numf(float x, float y);
  long double fminimum_numl(long double x, long double y);
  #ifdef __STDC_IEC_60559_DFP__
  _Decimal32 fminimum_numd32(_Decimal32 x, _Decimal32 y);
  _Decimal64 fminimum_numd64(_Decimal64 x, _Decimal64 y);
  _Decimal128 fminimum_numd128(_Decimal128 x, _Decimal128 y);
  #endif

Description

[2] The \texttt{fminimum\_num} functions determine the minimum value of their numeric arguments. They determine the number if one argument is a number and the other is a NaN. These functions differ from the \texttt{fminimum} functions only in their treatment of NaN arguments (see F.10.9.4, F.10.9.5).

Returns

[5] The \texttt{fminimum\_num} functions return the minimum value of their numeric arguments.
7.12.12.10 The fmaximum_mag_num functions

Synopsis

[1] #include <math.h>
    double fmaximum_mag_num(double x, double y);
    float fmaximum_mag_numf(float x, float y);
    long double fmaximum_mag_numl(long double x,
                                long double y);
    #ifdef __STDC_IEC_60559_DFP__
    __Decimal32 fmaximum_mag_numd32(__Decimal32 x,
                               __Decimal32 y);
    __Decimal64 fmaximum_mag_numd64(__Decimal64 x,
                               __Decimal64 y);
    __Decimal128 fmaximum_mag_numd128(__Decimal128 x,
                               __Decimal128 y);
    #endif

Description

[2] The fmaximum_mag_num functions determine the value of a numeric argument of maximum magnitude. They determine the number if one argument is a number and the other is a NaN. These functions differ from the fmaximum_mag functions only in their treatment of NaN arguments (see F.10.9.4, F.10.9.5).

Returns

[5] The fmaximum_mag_num functions return the value of a numeric argument of maximum magnitude.

7.12.12.11 The fminimum_mag_num functions

Synopsis

[1] #include <math.h>
    double fminimum_mag_num(double x, double y);
    float fminimum_mag_numf(float x, float y);
    long double fminimum_mag_numl(long double x,
                                long double y);
    #ifdef __STDC_IEC_60559_DFP__
    __Decimal32 fminimum_mag_numd32(__Decimal32 x,
                               __Decimal32 y);
    __Decimal64 fminimum_mag_numd64(__Decimal64 x,
                               __Decimal64 y);
    __Decimal128 fminimum_mag_numd128(__Decimal128 x,
                               __Decimal128 y);
    #endif
Description

[2] The \texttt{fminimum\_mag\_num} functions determine the value of a numeric argument of minimum magnitude. They determine the number if one argument is a number and the other is a NaN. These functions differ from the \texttt{fminimum\_mag} functions only in their treatment of NaN arguments (see F.10.9.4, F.10.9.5).

Returns

[5] The \texttt{fminimum\_mag\_num} functions return the value of a numeric argument of minimum magnitude.

NOTE The \texttt{fmax} and \texttt{fmin} functions are similar to the \texttt{fmaximum\_num} and \texttt{fminimum\_num} functions, though may differ in which signed zero is returned when the arguments are differently signed zeros and in their treatment of signaling NaNs (see F.10.9.5).

Remove F.10.9.4 and F.10.9.5.

After F.10.9.3 add:

F.10.9.4 The \texttt{fmaximum}, \texttt{fminimum}, \texttt{fmaximum\_mag}, and \texttt{fminimum\_mag} functions

These functions treat NaNs like other functions in <\texttt{math.h}> (see F.10). They differ from the corresponding \texttt{fmaximum\_num}, \texttt{fminimum\_num}, \texttt{fmaximum\_mag\_num}, and \texttt{fminimum\_mag\_num} functions only in their treatment of NaNs.

F.10.9.5 The \texttt{fmaximum\_num}, \texttt{fminimum\_num}, \texttt{fmaximum\_mag\_num}, and \texttt{fminimum\_mag\_num} functions

These functions return the number if one argument is a number and the other is a quiet or signaling NaN. If both arguments are NaNs, a quiet NaN is returned. If an argument is a signaling NaN, the "invalid" floating-point exception is raised (even though the function returns the number when the other argument is a number).

In F.2.1, change:

[4] Any operator or \texttt{<math.h>} function that raises an "invalid" floating-point exception, if delivering a floating type result, shall return a quiet NaN.

to:

[4] Any operator or \texttt{<math.h>} function that raises an "invalid" floating-point exception, if delivering a floating type result, shall return a quiet NaN, unless explicitly specified otherwise.

In 7.25#9, in the list of type-generic macros, remove: \texttt{fminmag} and \texttt{fmaxmag}. 
In 7.25#9, in the list of type-generic macros, add: \texttt{fmaximum, fminimum, fmaximum\_num, fminimum\_num, fmaximum\_mag, fminimum\_mag, fmaximum\_mag\_num, and fminimum\_mag\_num}.