



IEEE Standard for Floating-Point Arithmetic

IEEE Computer Society

Sponsored by the
Microprocessor Standards Committee

IEEE
3 Park Avenue
New York, NY 10016-5997, USA
29 August 2008

IEEE Std 754™-2008
(Revision of
IEEE Std 754-1985)

Currently in preview, click buy full version

IEEE Std 754™-2008

(Revision of
IEEE Std 754-1985)

IEEE Standard for Floating-Point Arithmetic

Sponsor

**Microprocessor Standards Committee
of the
IEEE Computer Society**

Approved 12 June 2008

IEEE-SA Standards Board

Currently in preview, click buy full version

Abstract: This standard specifies interchange and arithmetic formats and methods for binary and decimal floating-point arithmetic in computer programming environments. This standard specifies exception conditions and their default handling. An implementation of a floating-point system conforming to this standard may be realized entirely in software, entirely in hardware, or in any combination of software and hardware. For operations specified in the normative part of this standard, numerical results and exceptions are uniquely determined by the values of the input data, sequence of operations, and destination formats, all under user control.

Keywords: arithmetic, binary, computer, decimal, exponent, floating-point, format, interchange, NaN, number, rounding, significand, subnormal

The Institute of Electrical and Electronics Engineers, Inc.
3 Park Avenue, New York, NY 10016-5997, USA

Copyright © 2008 by the Institute of Electrical and Electronics Engineers, Inc.
All rights reserved. Published 29 August 2008. Printed in the United States of America.

IEEE is a registered trademark in the U.S. Patent & Trademark Office, owned by the Institute of Electrical and Electronics Engineers, Incorporated.

PDF: ISBN 978-0-7381-5752-8 STD95802
Print: ISBN 978-0-7381-5753-5 STDPD95802

No part of this publication may be reproduced in any form, in an electronic retrieval system or otherwise, without the prior written permission of the publisher.

IEEE Standards documents are developed within the IEEE Societies and the Standards Coordinating Committees of the IEEE Standards Association (IEEE-SA) Standards Board. The IEEE develops its standards through a consensus development process, approved by the American National Standards Institute, which brings together volunteers representing varied viewpoints and interests to achieve the final product. Volunteers are not necessarily members of the Institute and serve without compensation. While the IEEE administers the process and establishes rules to promote fairness in the consensus development process, the IEEE does not independently evaluate, test, or verify the accuracy of any of the information or the soundness of any judgments contained in its standards.

Use of an IEEE Standard is wholly voluntary. The IEEE disclaims liability for any personal injury, property or other damage, of any nature whatsoever, whether special, indirect, consequential, or compensatory, directly or indirectly resulting from the publication, use of, or reliance upon this, or any other IEEE Standard document.

The IEEE does not warrant or represent the accuracy or content of the material contained herein, and expressly disclaims any express or implied warranty, including any implied warranty of merchantability or fitness for a specific purpose, or that the use of the material contained herein is free from patent infringement. IEEE Standards documents are supplied "AS IS".

The existence of an IEEE Standard does not imply that there are no other ways to produce, test, measure, purchase, market, or provide other goods and services related to the scope of the IEEE Standard. Furthermore, the viewpoint expressed at the time a standard is approved and issued is subject to change brought about through developments in the state of the art and comments received from users of the standard. Every IEEE Standard is subjected to review at least every five years for revision or reaffirmation. When a document is more than five years old and has not been reaffirmed, it is reasonable to conclude that its contents, although still of some value, do not wholly reflect the present state of the art. Users are cautioned to check to determine that they have the latest edition of any IEEE Standard.

In publishing and making this document available, the IEEE is not suggesting or rendering professional or other services for, or on behalf of, any person or entity. Nor is the IEEE undertaking to perform any duty owed by any other person or entity to another. Any person utilizing this, and any other IEEE Standards document, should rely upon his or her independent judgment in the exercise of reasonable care in any given circumstances or, as appropriate, seek the advice of a competent professional in determining the appropriateness of a given IEEE standard.

Interpretations: Occasionally questions may arise regarding the meaning of portions of standards as they relate to specific applications. When the need for interpretations is brought to the attention of IEEE, the Institute will initiate action to prepare appropriate responses. Since IEEE Standards represent a consensus of concerned interests, it is important to ensure that any interpretation has also received the concurrence of a balance of interests. For this reason, IEEE and the members of its societies and Standards Coordinating Committees are not able to provide an instant response to interpretation requests except in those cases where the matter has previously received formal consideration. A statement, written or oral, that is not processed in accordance with the IEEE-SA Standards Board Operations Manual shall not be considered the official position of IEEE or any of its committees and shall not be considered to be, nor be relied upon as, a formal interpretation of the IEEE. At lectures, symposia, seminars, or educational courses, an individual presenting information on IEEE standards shall make it clear that his or her views should be considered the personal views of that individual rather than the formal position, explanation, or interpretation of the IEEE.

Comments for revision of IEEE Standards are welcome from any interested party, regardless of membership affiliation with IEEE. Suggestions for changes in documents should be in the form of a proposed change of text, together with appropriate supporting comments. Comments on standards and requests for interpretations should be submitted to the following address:

Secretary, IEEE-SA Standards Board
445 Hoes Lane
Piscataway, NJ 08854
USA

Permission to photocopy portions of any individual standard for internal or personal use is granted by The Institute of Electrical and Electronics Engineers, Inc., provided that the appropriate fee is paid to Copyright Clearance Center. To arrange for payment of licensing fee, please contact Copyright Clearance Center, Customer Service, 222 Rosewood Drive, Danvers, MA 01923 USA; +1 978 750 8400. Permission to photocopy portions of any individual standard for educational classroom use can also be obtained through the Copyright Clearance Center.

Introduction

This introduction is not part of IEEE Std 754-2008, IEEE Standard for Floating-Point Arithmetic.

This standard is a product of the Floating-Point Working Group of, and sponsored by, the Microprocessor Standards Committee of the IEEE Computer Society.

This standard provides a discipline for performing floating-point computation that yields results independent of whether the processing is done in hardware, software, or a combination of the two. For operations specified in the normative part of this standard, numerical results and exceptions are uniquely determined by the values of the input data, the operation, and the destination, all under user control.

This standard defines a family of commercially feasible ways for systems to perform binary and decimal floating-point arithmetic. Among the desiderata that guided the formulation of this standard were:

- a) Facilitate movement of existing programs from diverse computers to those that adhere to this standard as well as among those that adhere to this standard.
- b) Enhance the capabilities and safety available to users and programmers who, although not expert in numerical methods, might well be attempting to produce numerically sophisticated programs.
- c) Encourage experts to develop and distribute robust and efficient numerical programs that are portable, by way of minor editing and recompilation, onto any computer that conforms to this standard and possesses adequate capacity. Together with language controls, it should be possible to write programs that produce identical results on all conforming systems.
- d) Provide direct support for
 - execution-time diagnosis of anomalies
 - smoother handling of exceptions
 - interval arithmetic at a reasonable cost.
- e) Provide for development of
 - standard elementary functions such as *exp* and *cos*
 - high precision (multiword) arithmetic
 - coupled numerical and symbolic algebraic computation.
- f) Enable rather than preclude further refinements and extensions.

In programming environments, this standard is also intended to form the basis for a dialog between the numerical community and programming language designers. It is hoped that language-defined methods for the control of expression evaluation and exceptions might be defined in coming years, so that it will be possible to write programs that produce identical results on all conforming systems. However, it is recognized that utility and safety in languages are sometimes antagonists, as are efficiency and portability.

Therefore, it is hoped that language designers will look on the full set of operation, precision, and exception controls described here as a guide to providing the programmer with the ability to portably control expressions and exceptions. It is also hoped that designers will be guided by this standard to provide extensions in a completely portable way.

Notice to users

Laws and regulations

Users of these documents should consult all applicable laws and regulations. Compliance with the provisions of this standard does not imply compliance to any applicable regulatory requirements. Implementers of the standard are responsible for observing or referring to the applicable regulatory requirements. IEEE does not, by the publication of its standards, intend to urge action that is not in compliance with applicable laws, and these documents may not be construed as doing so.

Copyrights

This document is copyrighted by the IEEE. It is made available for a wide variety of both public and private uses. These include both use, by reference, in laws and regulations, and use in private self-regulation, standardization, and the promotion of engineering practices and methods. By making this document available for use and adoption by public authorities and private users, the IEEE does not waive any rights in copyright to this document.

Updating of IEEE documents

Users of IEEE standards should be aware that these documents may be superseded at any time by the issuance of new editions or may be amended from time to time through the issuance of amendments, corrigenda, or errata. An official IEEE document at any point in time consists of the current edition of the document together with any amendments, corrigenda, or errata then in effect. In order to determine whether a given document is the current edition and whether it has been amended through the issuance of amendments, corrigenda, or errata, please visit the IEEE Standards Association Web site at <http://ieeexplore.ieee.org/xpl/standards.jsp>, or contact the IEEE at the address listed previously.

For more information about the IEEE Standards Association or the IEEE standards development process, visit the IEEE-SA Web site at <http://standards.ieee.org>.

Errata

Errata, if any, for this and all other standards can be accessed at the following URL: <http://standards.ieee.org/reading/ieee/updates/errata/index.html>. Users are encouraged to check that URL for errata periodically.

Interpretations

Current interpretations can be accessed at the following URL: <http://standards.ieee.org/reading/ieee/interp/index.html>.

Patents

Attention is called to the possibility that implementation of this standard may require use of subject matter covered by patent rights. By publication of this standard, no position is taken with respect to the existence or validity of any patent rights in connection therewith. A patent holder or patent applicant has filed a statement of assurance that it will grant licenses under these rights without compensation or under reasonable rates, with reasonable terms and conditions that are demonstrably free of any unfair discrimination to applicants desiring to obtain such licenses. Other Essential Patent Claims may exist for which a statement of assurance has not been received. The IEEE is not responsible for identifying Essential Patent Claims for which a license may be required, for conducting inquiries into the legal validity or scope of Patent Claims, or for determining whether any licensing terms or conditions provided in connection with submission of a Letter of Assurance, if any, or in any licensing agreements are reasonable or non-discriminatory. Users of this standard are expressly advised that determination of the validity of any patent rights, and the risk of infringement of such rights, is entirely their own responsibility. Further information may be obtained from the IEEE Standards Association.

Participants

The following participants in the Floating-Point Working Group contributed to the development of this standard:

Dan Zuras, *Chair*
Mike Cowlshaw, *Editor*

Alex Aiken	David Gustafson	Ian Ollmann
Matthew Applegate	Michel Hack	Michael Parks
David Bailey	John R. Harrison	Tom Pittman
Steve Bass	John Hauser	Eric Postpischil
Dileep Bhandarkar	Yozo Hida	Jason Riedy
Mahesh Bhat	Chris N. Hinds	Eric M. Schwarz
David Bindel	Graydon Hoare	David Scott
Sylvie Boldo	David G. Mough †	Don Senzig
Stephen Canon	Jerry Hook	Ilya Sharapov
Steven R. Carlough	Jim F. Johnson	Jim Shearer
Marius Cornea	Michael Ingrassia	Michael Siu
Mike Cowlshaw †	David W. James	Ron Smith
John H. Crawford	Dick James	Chuck Stevens
Joseph D. Darcy	William Kahan †	Peter Tang
Debjit Das Sarma	John Kapernick	Pamela J. Taylor
Marc Daumas	Richard Karpinski	James W. Thomas
Bob Davis †	Jeff Kidder †	Brandon Thompson
Mark Davis	Plamen Koev	Wendy Thrash
Dick Delp	Ren-Cang Li	Neil Toda
Jim Demmel	Zhishun Alex Liu	Son Dao Trong
Mark A. Felt	Raymond Mak	Leonard Tsai †
Hossam A. G. Faraway	Peter Markstein †	Charles Tsen
J.P. Farnoy	David Matula	Fred Tydeman
Richard Fateman	Guillaume Melquiond	Liang Kai Wang
Eric Feng	Nobuyoshi Mori	Scott Westbrook
William E. Ferguson	Ricardo Morin	Steve Winkler
Alex Fit-Florea	Ned Nedialkov	Anthony Wood
Laurent Fournier	Craig Nelson	Umit Yalcinalp
Chip Freitag	Stuart Oberman	Fred Zemke
Ivan Godard	Jon Okada	Paul Zimmermann
Roger A. Golliver †		Dan Zuras †

† identifies those who also participated in the Ballot Review Committee.

The following individual members of the balloting committee voted on this standard. Balloters might have voted for approval, disapproval, or abstention.

Ali Al Awazi
Tomo Bogataj
Mark Brown
Steven R. Carlough
Juan C. Carreon
Danila Chenetsov
Srinivas Chennupaty
Danila Chernetsov
Naveen Cherukuri
Keith Chow
Glenn Colon-Bonet
Tommy P. Cooper
Robert Corbett
Marius Cornea
Mike Cowlshaw
John H. Crawford
Bob Davis
Florent de Dinechin
Kenneth Dockser
Ulrich Drepper
Sourav K. Dutta
Carol T. Eidt
Bo Einarsson
Mark A. Erle
Hossam A. H. Fahmy
John W. Fendrich
Warren E. Ferguson
Andrew Fieldsend
Rabiz N. Foda
Roger A. Golliver
Sergiu R. Goma
Randall C. Groves
Scott A. Gudgel
Michel Hack

John R. Harrison
Peter L. Harrod
Barry E. Hedquist
Rutger Heunks
Chris N. Hinds
Werner Hoelzl
Dennis Horwitz
David G. Hough
Jerome Huck
David V. James
Muzaffer Kal
Piotr Karocki
Mark J. Knight
Theodore E. Kubaska
Susan Land
Christoph Lauter
Shawn M. Leard
David J. Leciston
Solomon Lee
Vincent Lefevre
Vincent J. Lipsio
Zhishun A. Liu
William Lumpkins
Peter Markstein
Edward M. McCall
George J. Miao
Gary L. Michel
James Moore
Jean-Michel Muller
Bruce A. Nitz
Michael Newman
Charles K. Ngethe
Gregory D. Peterson

Urich Pohl
Subburajan Ponnuswamy
Jose Puthenkulam
Marko Radmilac
Gary S. Robinson
Robert A. Robinson
Michael D. Rush
M. S. Sachdev
Sridhar Samudrala
Randy Saunders
Bartien Sayogo
Thomas Schesig
Michael J. Schmitz
Stephen C. Schwartz
Eric M. Schwarz
Matthew Smith
Mitchell W. Smith
Thomas E. Starai
Walter Struppler
Ying T. Tang
Pamela J. Taylor
James W. Thomas
Wendy Thrash
Leonard Tsai
Charles Tsen
S. Tulasidas
Srinivasa R. Vemuru
Steven Wallach
Paul R. Work
Forrest D. Wright
Oren Yuen
Janusz Zalewski
Alexandru Zamfirescu
Dan Zuras

When the IEEE-SA Standards Board approved this standard on 12 June 2008, it had the following membership:

Robert M. Grow, *Chair*
Thomas Prevost, *Vice Chair*
Steve M. Mills, *Past Chair*
Judith Gorman, *Secretary*

Victor Berman
Richard DeBlasio
Andy Drozd
Mark Epstein
Alexander Gelman
William Goldbach
Arnie Greenspan
Ken Hanus

Jim Hughes
Richard Hulett
Young Kyun Kim
Joseph L. Koepfinger*
John Kulick
David J. Law
Glenn Parsons

Ron Petersen
Chuck Powers
Narayanan Ramachandran
Jon Walter Rosdahl
Anne-Marie Sahazizian
Malcolm Thaden
Howard Wolfman
Don Wright

*Member Emeritus

Also included are the following non-voting IEEE-SA Standards Board liaisons:

Satish K. Aggarwal, *NRC Representative*
Michael H. Kelley, *NIST Representative*

Lisa Perry
IEEE Standards Project Editor

Malia Zaman
IEEE Standards Program Manager, Technical Program Development

Contents

1. Overview	1
1.1 Scope	1
1.2 Purpose	1
1.3 Inclusions	1
1.4 Exclusions	2
1.5 Programming environment considerations	2
1.6 Word usage	2
2. Definitions, abbreviations, and acronyms	3
2.1 Definitions	3
2.2 Abbreviations and acronyms	5
3. Floating-point formats	6
3.1 Overview	6
3.2 Specification levels	7
3.3 Sets of floating-point data	7
3.4 Binary interchange format encodings	9
3.5 Decimal interchange format encodings	10
3.6 Interchange format parameters	13
3.7 Extended and extendable precisions	14
4. Attributes and rounding	15
4.1 Attribute specification	15
4.2 Dynamic modes for attributes	15
4.3 Rounding-direction attributes	16
5. Operations	17
5.1 Overview	17
5.2 Decimal exponent calculation	18
5.3 Homogeneous general-computational operations	19
5.4 formatOf general-computational operations	21
5.5 Quiet-computational operations	23
5.6 Signaling-computational operations	24
5.7 Non-computational operations	24
5.8 Details of conversions from floating-point to integer formats	26
5.9 Details of operations to round a floating-point datum to integral value	27
5.10 Details of totalOrder predicate	28
5.11 Details of comparison predicates	29
5.12 Details of conversion between floating-point data and external character sequences	30
6. Infinity, NaNs, and sign bit	34
6.1 Infinity arithmetic	34
6.2 Operations with NaNs	34
6.3 The sign bit	35
7. Default exception handling	36
7.1 Overview: exceptions and flags	36
7.2 Invalid operation	37
7.3 Division by zero	37
7.4 Overflow	37
7.5 Underflow	38
7.6 Inexact	38
8. Alternate exception handling attributes	39
8.1 Overview	39
8.2 Resuming alternate exception handling attributes	39
8.3 Immediate and delayed alternate exception handling attributes	40

9. Recommended operations	41
9.1 Conforming language- and implementation-defined functions	41
9.2 Recommended correctly rounded functions	42
9.3 Operations on dynamic modes for attributes	46
9.4 Reduction operations	46
10. Expression evaluation	48
10.1 Expression evaluation rules	48
10.2 Assignments, parameters, and function values	48
10.3 preferredWidth attributes for expression evaluation	49
10.4 Literal meaning and value-changing optimizations	50
11. Reproducible floating-point results	51
Annex A (informative) Bibliography	53
Annex B (informative) Program debugging support	55
Index of operations	57

IEEE Standard for Floating-Point Arithmetic

IMPORTANT NOTICE: This standard is not intended to assure safety, security, health, or environmental protection in all circumstances. Implementers of the standard are responsible for determining appropriate safety, security, environmental, and health practices or regulatory requirements.

This IEEE document is made available for use subject to important notices and legal disclaimers. These notices and disclaimers appear in all publications containing this document and may be found under the heading “Important Notice” or “Important Notices and Disclaimers Concerning IEEE Documents”. They can also be obtained on request from IEEE or viewed at <http://standards.ieee.org/IPR/disclaimers.html>.

1. Overview

1.1 Scope

This standard specifies formats and methods for floating-point arithmetic in computer systems—standard and extended functions with single, double, extended, and extendable precision—and recommends formats for data interchange. Exception conditions are defined and standard handling of these conditions is specified.

1.2 Purpose

This standard provides a method for computation with floating-point numbers that will yield the same result whether the processing is done in hardware, software, or a combination of the two. The results of the computation will be identical, independent of implementation, given the same input data. Errors, and error conditions, in the mathematical processing will be reported in a consistent manner regardless of implementation.

1.3 Inclusions

This standard specifies:

- Formats for binary and decimal floating-point data, for computation and data interchange.
- Addition, subtraction, multiplication, division, fused multiply add, square root, compare, and other operations.
- Conversions between integer and floating-point formats.
- Conversions between different floating-point formats.
- Conversions between floating-point formats and external representations as character sequences.
- Floating-point exceptions and their handling, including data that are not numbers (NaNs).