

Information Technology — Programming Languages — Guidance to Avoiding Vulnerabilities in Programming Languages through Language Selection and Use

Élément introductif — Élément principal — Partie n: Titre de la partie

Warning

This document is not an ISO International Standard. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an International Standard.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

Document type: International standard
Document subtype: if applicable
Document stage: (40) approval stage
Document language: E

Copyright notice

This ISO document is a working draft or committee draft and is copyright-protected by ISO. While the reproduction of working drafts or committee drafts in any form for use by participants in the ISO standards development process is permitted without prior permission from ISO, neither this document nor any extract from it may be reproduced, stored or transmitted in any form for any other purpose without prior written permission from ISO.

Requests for permission to reproduce this document for the purpose of selling it should be addressed as shown below or to ISO's member body in the country of the requester:

*ISO copyright office
Case postale 56, CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org*

Reproduction for sales purposes may be subject to royalty payments or a licensing agreement.

Violators may be prosecuted.

Contents

Page

Foreword	viii
Introduction	ix
1 Scope	1
1.1 In Scope	1
1.2 Not in Scope	1
1.3 Approach	1
1.4 Intended Audience	1
1.5 How to Use This Document	2
2 Normative references	4
3 Terms and definitions	5
4 Symbols (and abbreviated terms)	7
5 Vulnerability issues	8
5.1 Issues arising from incomplete or evolving language specifications	8
5.2 Issues arising from human cognitive limitations	10
5.3 Issues arising from a lack of predictable execution	11
5.4 Issues arising from the lack of portability and interoperability	11
5.5 Issues arising from inadequate language intrinsic support	12
5.6 Issues arising from language features prone to erroneous use	12
6. Programming Language Vulnerabilities	13
6.1 Obscure Language Features [BRS]	13
6.2 Unspecified Behaviour [BQF]	14
6.3 Undefined Behaviour [EWF]	15
6.4 Implementation-defined Behaviour [FAB]	17
6.5 Deprecated Language Features [MEM]	18
6.6 Pre-processor Directives [NMP]	19
6.7 Choice of Clear Names [NAI]	21
6.8 Choice of Filenames and other External Identifiers [AJN]	22
6.9 Unused Variable [XYR]	24
6.10 Identifier Name Reuse [YOW]	25
6.11 Type System [IHN]	27
6.12 Bit Representations [STR]	29
6.13 Floating-point Arithmetic [PLF]	30
6.14 Enumerator Issues [CCB]	32
6.15 Numeric Conversion Errors [FLC]	34
6.16 String Termination [CJM]	36
6.17 Boundary Beginning Violation [XYX]	37
6.18 Unchecked Array Indexing [XYZ]	39
6.19 Unchecked Array Copying [XYW]	40
6.20 Buffer Overflow [XZB]	41
6.21 Pointer Casting and Pointer Type Changes [HFC]	42
6.22 Pointer Arithmetic [RVG]	43
6.23 Null Pointer Dereference [XYH]	44
6.24 Dangling Reference to Heap [XYK]	45
6.25 Templates and Generics [SYM]	47
6.26 Inheritance [RIP]	49
6.27 Initialization of Variables [LAV]	50
6.28 Wrap-around Error [XYY]	52
6.29 Sign Extension Error [XZI]	53

6.30	Operator Precedence/Order of Evaluation [JCW].....	54
6.31	Side-effects and Order of Evaluation [SAM]	55
6.32	Likely Incorrect Expression [KOA]	57
6.33	Dead and Deactivated Code [XYQ]	58
6.34	Switch Statements and Static Analysis [CLL].....	60
6.35	Demarcation of Control Flow [EOJ].....	61
6.36	Loop Control Variables [TEX].....	62
6.37	Off-by-one Error [XZH].....	63
6.38	Structured Programming [EWD]	65
6.39	Passing Parameters and Return Values [CSJ].....	66
6.40	Dangling References to Stack Frames [DCM].....	68
6.41	Subprogram Signature Mismatch [OTR]	70
6.42	Recursion [GDL].....	71
6.43	Returning Error Status [NZN].....	72
6.44	Termination Strategy [REU]	74
6.45	Extra Intrinsic [LRM]	75
6.46	Type-breaking Reinterpretation of Data [AMV]	76
6.47	Memory Leak [XYL].....	78
6.48	Argument Passing to Library Functions [TRJ]	79
6.49	Dynamically-linked Code and Self-modifying Code [NYY]	80
6.50	Library Signature [NSQ]	82
6.51	Unanticipated Exceptions from Library Routines [HJW]	83
7.	Application Vulnerabilities	85
7.1	Adherence to Least Privilege [XYN]	85
7.2	Privilege Sandbox Issues [XYO]	85
7.3	Executing or Loading Untrusted Code [XYS].....	87
7.4	Unspecified Functionality [BVQ].....	87
7.5	Distinguished Values in Data Types [KLLK]	88
7.6	Memory Locking [XZX]	90
7.7	Resource Exhaustion [XZP]	90
7.8	Injection [RST].....	92
7.9	Cross-site Scripting [XYT].....	94
7.10	Unquoted Search Path or Element [XZQ].....	96
7.11	Improperly Verified Signature [XZR].....	97
7.12	Discrepancy Information Leak [XZL]	97
7.13	Sensitive Information Uncleared Before Release [XZK].....	98
7.14	Path Traversal [EWR].....	99
7.15	Missing Required Cryptographic Step [XZS]	101
7.16	Insufficiently Protected Credentials [XYM]	101
7.17	Missing or Inconsistent Access Control [XZN].....	102
7.18	Authentication Logic Error [XZO]	103
7.19	Hard-coded Password [XYP].....	104
A.	Guideline Recommendation Factors	106
A.1	Factors that need to be covered in a proposed guideline recommendation	106
A.2	Language definition	106
A.3	Measurements of language usage.....	106
A.4	Level of expertise.....	106
A.5	Intended purpose of guidelines	106
A.6	Constructs whose behaviour can vary	106
A.7	Example guideline proposal template	107
B.	Guideline Selection Process	108
B.1	Cost/Benefit Analysis	108
B.2	Documenting of the selection process.....	108
C.	Skeleton template for use in proposing programming language vulnerabilities	109
C.1	6.<x> <short title> [<unique immutable identifier>].....	109
D.	Skeleton template for use in proposing application vulnerabilities.....	111

D.1	7.<x> <short title> [<unique immutable identifier>].....	111
E.	Vulnerability Outline	112
F.	Skeleton template for use in proposing language specific information for vulnerabilities	115
F.1	Identification of standards	115
F.2	General Terminology	115
F.3	<language annex>.3.<x> <short title><unique immutable identifier>	115
	Bibliography.....	117

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO/IEC TR 24772, which is a Technical Report of type 3, was prepared by Joint Technical Committee ISO/IEC JTC 1, Subcommittee SC 22, Programming Languages.

Introduction

All programming languages contain constructs that are incompletely specified, exhibit undefined behaviour, are implementation-dependent, or are difficult to use correctly. The use of those constructs may therefore give rise to *vulnerabilities*, as a result of which, software programs can execute differently than intended by the writer. In some cases, these vulnerabilities can compromise the safety of a system or be exploited by attackers to compromise the security or privacy of a system.

This Technical Report is intended to provide guidance spanning multiple programming languages, so that application developers will be better able to avoid the programming constructs that lead to vulnerabilities in software written in their chosen language and their attendant consequences. This guidance can also be used by developers to select source code evaluation tools that can discover and eliminate some constructs that could lead to vulnerabilities in their software or to select a programming language that avoids anticipated problems.

It should be noted that this Technical Report is inherently incomplete. It is not possible to provide a complete list of programming language vulnerabilities because new weaknesses are discovered continually. Any such report can only describe those that have been found, characterized, and determined to have sufficient probability and consequence.

Furthermore, in order to focus its limited resources, the working group developing this report decided to defer treatment of several subject areas until future editions of the report. These subject areas include:

- Object-oriented language features
- Concurrency
- Numerical analysis (although some simple items regarding the use of floating point are described in PLP)
- Scripting languages
- Some issues related to inter-language operability

It should also be noted that the first edition of this Technical Report contains few annexes describing language-specific application of its guidance because it was important to create a consensually validated baseline of the report before asking the language-specific committees to invest effort in the creation of language-specific guidance.

1 Information Technology — Programming Languages — Guidance to Avoiding Vulnerabilities in Programming
2 Languages through Language Selection and Use

3 **1 Scope**

4 **1.1 In Scope**

5 This Technical Report specifies software programming language vulnerabilities that should be avoided in the
6 development of systems where assured behaviour is required for security, safety, mission critical and business
7 critical software. In general, this guidance is applicable to the software developed, reviewed, or maintained for any
8 application.

9 **1.2 Not in Scope**

10 This Technical Report does not address software engineering and management issues such as how to design and
11 implement programs, use configuration management tools, use managerial processes, and perform process
12 improvement. Furthermore, the specification of properties to be assured is not treated.

13 Vulnerabilities are described in a generic manner that is applicable to a broad range of programming languages.
14 Annexes apply the guidance to some particular programming languages.

15 The specification of an application is *not* within the scope.

16 **1.3 Approach**

17 The impact of the guidelines in this Technical Report is likely to be highly leveraged in that they are likely to affect
18 many times more people than the number that worked on them. Therefore guidelines such as these have the
19 potential to make large savings, for a small cost, or to generate large unnecessary costs, for little benefit. For these
20 reasons, the writers of this Technical Report have taken a cautious approach to creating guideline
21 recommendations. New guideline recommendations can be added over time, as experience and evidence is
22 accumulated.

23 **1.4 Intended Audience**

24 The intended audience for this Technical Report is those who are concerned with assuring the existence of a
25 critical property in the software of their system; that is, those who are developing, qualifying, or maintaining a
26 software system and need to avoid language constructs that could cause the software to execute in a manner other
27 than intended.

28 As described in the following paragraphs, developers of applications that have clear safety, security or mission
29 criticality are expected to be aware of the risks associated with their code and could use this document to ensure
30 that their development practices address the issues presented by the chosen programming languages, for example
31 by subsetting or providing coding guidelines.

32 That should not be taken to mean that other developers can ignore this document. A weakness in an application
33 that of itself has no direct criticality may provide the route by which an attacker gains control of a system or may
34 otherwise disrupt co-hosted applications that are safety, security or mission critical.

35 It is hoped that such developers would use this document to ensure that common vulnerabilities are removed or at
36 least minimized from all applications.

37 Several specific audiences for this International Technical Report have been identified and are described below.

1 **1.4.1 Safety-Critical Applications**

2 Users who may benefit from this document include those developing, qualifying, or maintaining a system where it is
3 critical to prevent behaviour that might lead to:

- 4 • loss of human life or human injury, or
- 5 • damage to the environment.

6 **1.4.2 Security-Critical Applications**

7 Users who may benefit from this document include those developing, qualifying, or maintaining a system where it is
8 critical to exhibit security properties of:

- 9 • confidentiality,
- 10 • integrity, and
- 11 • availability.

12
13 **1.4.3 Mission-Critical Applications**

14 Users who may benefit from this document include those developing, qualifying, or maintaining a system where it is
15 critical to prevent behaviour that might lead to:

- 16 • property loss or damage, or
- 17 • economic loss or damage.

18
19 **1.4.4 Business-Critical Systems**

20 Users who may benefit from this document include those developing, qualifying, or maintaining a system whose
21 correctness and integrity are essential for successful operation of a business or enterprise.

22 **1.4.5 Modeling and Simulation Applications**

23 People who may benefit from this document include those who are primarily experts in areas other than
24 programming but need to use computation as part of their work. Such people include scientists, engineers,
25 economists, and statisticians. They require high confidence in the applications they write and use because of the
26 increasing complexity of the calculations made (and the consequent use of teams of programmers each
27 contributing expertise in a portion of the calculation), or to the costs of invalid results, or to the expense of individual
28 calculations implied by a very large number of processors used and/or very long execution times needed to
29 complete the calculations. These circumstances give a consequent need for high reliability and motivate the need
30 felt by these programmers for the guidance offered in this document.

31 **1.5 How to Use This Document**

32 This Technical Report gathers language-independent descriptions of programming language vulnerabilities, as well
33 as selected application vulnerabilities, which have occurred in the past and are likely to occur again. Because new
34 vulnerabilities are always being discovered, it is anticipated that the document will be revised and new descriptions
35 added. For that reason, a scheme that is distinct from document sub-clause numbering has been adopted to
36 identify the vulnerability descriptions. Each description has been assigned an arbitrarily generated, unique three-
37 letter code. These codes should be used in preference to sub-clause numbers when referencing descriptions.

38 The main part of the document contains descriptions that are intended to be language-independent to the greatest
39 possible extent. Future editions will include annexes that apply the generic guidance to particular programming
40 languages.

41 The document has been written with several possible usages in mind:

- 1 • Programmers familiar with the vulnerabilities of a specific language can reference the guide for more
2 generic descriptions and their manifestations in less familiar languages.
- 3 • Tool vendors can use the three-letter codes as a succinct way to “profile” the selection of vulnerabilities
4 considered by their tools.
- 5 • Individual organizations may wish to write their own coding standards intended to reduce the number of
6 vulnerabilities in their software products. The guide can assist in the selection of vulnerabilities to be
7 addressed in those standards and the selection of coding guidelines to be enforced.
- 8 • Organizations or individuals selecting a language for use in a project may want to consider the
9 vulnerabilities inherent in various candidate languages.

10 The following clauses include suggestions for ways of avoiding the vulnerabilities. It should be noted that these
11 include techniques that can be applied during development, and those that must be implemented as run-time
12 checks. The former are likely to be appropriate to all applications. For some applications, it is relatively more
13 important to ensure that potential run-time errors are eliminated during development because there may be
14 insufficient opportunity to recover from them. Source code checking tools can be used to automatically enforce
15 some coding rules and standards.

1 **2 Normative references**

2 The following referenced documents are indispensable for the application of this document. For dated references,
3 only the edition cited applies. For undated references, the latest edition of the referenced document (including any
4 amendments) applies.

1 3 Terms and definitions

2 For the purposes of this document, the following terms and definitions apply. Other terms are defined where they
3 appear in *italic* type. Programming language keywords or syntactic tokens appear in `courier` font.

4 3.1 Language Vulnerability

5 A *property* (of a programming language) that can contribute to, or that is strongly correlated with, application
6 vulnerabilities in programs written in that language.

7 **Note:** The term "property" can mean the presence or the absence of a specific feature, used singly or in
8 combination. As an example of the absence of a feature, encapsulation (control of where names may be
9 referenced from) is generally considered beneficial since it narrows the interface between modules and can
10 help prevent data corruption. The absence of encapsulation from a programming language can thus be
11 regarded as a vulnerability. Note that a property together with its complement may both be considered
12 language vulnerabilities. For example, automatic storage reclamation (garbage collection) is a vulnerability
13 since it can interfere with time predictability and result in a safety hazard. On the other hand, the absence of
14 automatic storage reclamation is also a vulnerability since programmers can mistakenly free storage
15 prematurely, resulting in dangling references. On the other hand, the absence of automatic storage
16 reclamation is also a vulnerability because the program can run out of storage, or programmers can mistakenly
17 free storage prematurely, resulting in dangling references.

18 3.2 Application Vulnerability

19 A security vulnerability or safety hazard, or defect.

20 3.3 Security Vulnerability

21 A weakness in an information system, system security procedures, internal controls, or implementation that could
22 be exploited or triggered by a threat.

23 3.4 Safety Hazard

24 IEC61508 part 4: defines a "Hazard" as a "potential source of harm", where "harm" is "physical injury or damage to
25 the health of people either directly or indirectly as a result of damage to property or to the environment".

26 IEC61508 cites ISO/IEC Guide 51 as the source for the definition.

27 **Note:** IEC61508 is titled "Functional safety of electrical/electronic/ programmable electronic safety-related
28 systems", with part 4 being "Definitions and abbreviations". Hence within IEC61508 the "safety" context of
29 "safety hazard" is assumed.

30 **Note:** Some derived standards, such as UK Defence Standard 00-56, broaden the definition of "harm" to
31 include material and environmental damage (not just harm to people caused by property and environmental
32 damage).

33 3.5 Safety-critical software

34 Software for applications where failure can cause very serious consequences such as human injury or death.
35 IEC61508 part 4: defines "Safety-related software" as "software that is used to implement safety functions in a
36 safety-related system.

37 **Note:** Notwithstanding that in some domains a distinction is made between safety-related (may lead to any
38 harm) and safety-critical (life threatening), this Technical Report uses the term *safety-critical* for all
39 vulnerabilities that may result in safety hazards.

1 **3.6 Software quality**

2 The degree to which software implements the requirements described by its specification and the degree to which
3 the characteristics of a software product fulfils its requirements.

4 **3.7 Predictable Execution**

5 The property of the program such that all possible executions have results that can be predicted from the source
6 code, the relevant language-defined implementation characteristics and knowledge of the universe of execution.

7 **Note:** In some environments, this would raise issues regarding numerical stability, exceptional processing, and
8 concurrent execution.

9 **Note:** Predictable execution is an ideal that must be approached keeping in mind the limits of human
10 capability, knowledge, availability of tools, etc. Neither this nor any standard ensures predictable execution.
11 Rather this standard provides advice on improving predictability. The purpose of this document is to assist a
12 reasonably competent programmer to approach the ideal of predictable execution.

13 **Note:** The following terms are used in relation to “Predictable execution”

- 14 • **Unspecified behaviour:** A situation where the implementation of a language will have to make some
15 choice from a finite set of alternatives, but that choice is not in general predictable by the programmer, for
16 example, the order in which sub-expressions are evaluated in an expression in many languages.
- 17 • **Implementation-defined behaviour:** A situation where the implementation of a language will have to
18 make some choice, and it is required that this choice be documented and available to the programmer.
- 19 • **Undefined behaviour:** A situation where the definition of a language can give no indication of what
20 behaviour to expect from a program – it may be some form of catastrophic failure (a ‘crash’) or continued
21 execution with some arbitrary data.

22 **Note:** This document includes a section on **Unspecified functionality**. This notion is related to neither
23 unspecified behaviour, which is a characteristic of an application, nor the language used to develop the
24 application.

25

1 **4 Symbols (and abbreviated terms)**

2 None.

1 5 Vulnerability issues

2 Software vulnerabilities are unwanted characteristics of software that may allow software to behave in ways that
3 are unexpected by a reasonably sophisticated user of the software. The expectations of a reasonably
4 sophisticated user of software may be set by the software's documentation or by experience with similar software.
5 Programmers introduce vulnerabilities into software by failing to understand the expected behaviour (the software
6 requirements), or by failing to correctly translate the expected behaviour into the actual behaviour of the software.

7 This document does not discuss a programmer's understanding of software requirements or the completeness of
8 the requirements. This document does not discuss software engineering issues per se. This document does not
9 discuss configuration management, build environments, code-checking tools, nor software testing. This document
10 does not discuss the classification of software vulnerabilities according to safety or security concerns. This
11 document does not discuss the costs of software vulnerabilities, or the costs of preventing them.

12 This document does discuss a reasonably competent programmer's failure to translate the understood
13 requirements into correctly functioning software. This document does discuss programming language features
14 known to contribute to software vulnerabilities. That is, this document discusses issues arising from those features
15 of programming languages found to increase the frequency of occurrence of software vulnerabilities. The intention
16 is to provide guidance to those who wish to specify coding guidelines for their own particular use.

17 A programmer writes source code in a programming language to translate the understood requirements into
18 working software. The programmer selects and codes constructs specified by a programming language with the
19 intention of achieving a written expression of the desired behaviour.

20 A program's expected behaviour might be stated in a complex technical document, which can result in a complex
21 sequence of features of the programming language. Software vulnerabilities occur when a reasonably competent
22 programmer fails to understand the totality of the effects of the language features combined to construct the
23 software. The overall software may be a very complex technical document itself (written in a programming
24 language whose definition is also a complex technical document).

25 The recommendations contained in this Technical Report might also be considered to be code quality issues. Both
26 kinds of issues might be addressed through the use of a systematic development process, use of
27 development/analysis tools and thorough testing.
28

29 5.1 Issues arising from incomplete or evolving language specifications

30 While there are many millions of programmers in the world, there are only several hundreds of authors engaged in
31 designing and specifying those programming languages defined by international standards. The design and
32 specification of a programming language is very different from programming. Programming involves selecting and
33 sequentially combining features from the programming language to (locally) implement specific steps of the
34 software's design. In contrast, the design and specification of a programming language involves (global)
35 consideration of all aspects of the programming language. This must include how all the features will interact with
36 each other, and what effects each will have, separately and in any combination, under all foreseeable
37 circumstances. Thus, language design has global elements that are not generally present in any local
38 programming task.

39 The creation of the abstractions which become programming language standards therefore involve consideration of
40 issues unneeded in many cases of actual programming. Therefore perhaps these issues are not routinely
41 considered when programming in the resulting language. These global issues may motivate the definition of subtle
42 distinctions or changes of state not apparent in the usual case wherein a particular language feature is used.
43 Authors of programming languages may also desire to maintain compatibility with older versions of their language
44 while adding more modern features to their language and so add what appears to be an inconsistency to the
45 language.

46 For example, some languages may allow a subprogram to be invoked without specifying the correct signature of
47 the subprogram. This may be allowed in order to keep compatibility with earlier versions of the language where
48 such usage was permitted, and despite the knowledge that modern practice demands the signature be specified.

1 Specifically, the programming language C does not require a function prototype be within scope¹. The
 2 programming language Fortran does not require an explicit interface. Thus, language usage is improved by coding
 3 standards specifying that the signature be present.

4 A reasonably competent programmer therefore may not consider the full meaning of every language feature used,
 5 as only the desired (local or subset) meaning may correspond to the programmer's immediate intention. In
 6 consequence, a subset meaning of any feature may be prominent in the programmer's overall experience.

7 Further, the combination of features indicated by a complex programming goal can raise the combinations of
 8 effects, making a complex aggregation within which some of the effects are not intended.

9 **5.1.1 Compiler Selection**

10 Compiler selection is important to ensure a system operates safely and securely. Compilers are important as they
 11 are the intermediary between the human readable source code and the machine readable binary code. This crucial
 12 step is often overlooked and compilers, unless coming from a trusted source and developed according to agreed
 13 standards, should be treated as any other commercial off the shelf software that has an unknown pedigree.

14 Often, developers analyze the source code to detect any code that can negatively impact security or safety. This
 15 aims to solve one part of the problem. After the source has been compiled, we need to be sure that the compiler
 16 did not insert any logic (maliciously or inadvertently) into the binary that compromises the systems security or
 17 safety. This is especially important because this type of vulnerability will be inserted into every piece of software
 18 that the compiler processes.

19 To combat against this, developers of security or safety critical systems should only use compilers from a trusted
 20 source. The trusted source should also provide evidence that the compiler is free from anomalous behaviour;
 21 similar to the way RTCA's DO-178B defines qualifiable tools. In addition, developers of critical software can
 22 perform source to binary traceability to ensure the compiler has not inserted any undesired logic into the binary
 23 code.

24 If a compiler has many options, then developers should consider if the options to be used in the project are the
 25 same options used when the compiler was validated.

26 **5.1.2 Issues arising from unspecified behaviour**

27 While every language standard attempts to specify how software written in the language will behave in all
 28 circumstances, there will always be some behaviour that is not specified completely. In any circumstance, of
 29 course, a particular compiler will produce a program with some specific behaviour (or fail to compile the program at
 30 all). Where a programming language construct is insufficiently defined, different compilers may generate different
 31 behaviours from the same source code. The authors of language standards often have an interpretations or
 32 defects process in place to treat these situations once they become known, and, eventually, to specify one
 33 particular behaviour. However, the time needed by the process to produce corrections to the language standard is
 34 often long, as careful consideration of the issues involved is needed.

35 When programs are compiled with only one compiler, the programmer may not be aware when behaviour not
 36 specified by the standard has been produced. Programs relying upon behaviour not specified by the language
 37 standard may behave differently when they are compiled with different compilers. An experienced programmer
 38 may choose to use more than one compiler, even in one environment, in order to obtain diagnostics from more
 39 than one source. In this usage, any particular compiler must be considered to be a different compiler if it is used
 40 with different options (which can give it different behaviour), or is a different release of the same compiler (which
 41 may have different default options or may generate different code), or is on different hardware (which may have a
 42 different instruction set). In this usage, a different computer may be the same hardware with a different operating
 43 system, with different compilers installed, with different software libraries available, with a different release of the
 44 same operating system, or with a different operating system configuration.

¹ This feature has been deprecated in the 1999 version of the ISO C Standard.

1 **5.1.3 Issues arising from implementation-defined behaviour**

2 In some situations, a programming language standard may specifically allow compilers to support a range of
 3 behaviours to a given language feature or combination of features. This may enable a more efficient execution on
 4 a wider range of hardware, or enable use of the programming language in a wider variety of circumstances.

5 In order to allow use on a wide range of hardware, for example, many languages do not specify the amount of
 6 storage reserved for language-defined entities such as variables. The degree to which a diligent programmer may
 7 obtain information on the amount of storage reserved for entities varies among languages.

8 The authors of language standards are encouraged to provide lists of all allowed variations of behaviour (as many
 9 already do). Such a summary will benefit applications programmers, those who define applications coding
 10 standards, and those who make code analysis tools.

11 **5.1.4 Issues arising from undefined behaviour**

12 In some situations, a programming language standard may specify that program behaviour is undefined. While the
 13 authors of language standards naturally try to minimize these situations, they may be inevitable when attempting to
 14 define software recovery from errors, or other situations recognized as being incapable of precise definition.

15 An example of undefined behaviour, in many languages, is the use of the value of a variable to which there has not
 16 yet been an assignment.

17 **5.2 Issues arising from human cognitive limitations**

18 The goal in the creation of a programming language is to provide a tool that may be used by the writers of software
 19 to manipulate data and produce a desired result. Some programming languages are general purpose languages,
 20 while others are targeted to specific tasks or needs. Even general purpose languages may be created with a
 21 specific user in mind, for example C was created by engineers for programmers implementing system software,
 22 while COBOL was created for business analysts and business programmers.

23 All humans are different. Constructs that may be easily understood by mathematicians may be confusing to a
 24 business analyst. Also, similar constructs in different languages (the use of '=' in COBOL for comparison, and in C
 25 for assignment), and similar constructs within the same language (the use of parentheses for both function
 26 parameter lists, and for array subscripts) can add to this confusion. In addition, many languages provide multiple
 27 syntaxes to accomplish the same task, and coders will chose those syntaxes that make the most sense to them,
 28 again adding additional confusion in program creation and maintenance.

29 Humans are also fallible, and can only comprehend a limited number of interactions within a process before that
 30 process must be subdivided into smaller segments. In addition, stress, whether internal pressures and deadlines
 31 or external influences totally unrelated to the task at hand, can interfere with the process of software creation.
 32 These factors combined with language constructs that may be confusing can lead to failures due to human
 33 fallibility. These failures can include:

- 34 • Cognitive shortcomings, which are failures of design and implementation. These may be
 - 35 ○ faulty reasoning and
 - 36 ○ incomplete solutions due to lack of time and effort in comprehending the task.
- 37 • Knowledge shortcomings, which are failures of training and environment. These can include
 - 38 ○ incomplete and/or incorrect knowledge of appropriate language semantics,
 - 39 ○ incomplete and/or incorrect knowledge of how the chosen syntax will be executed by a particular
 40 implementation, and
 - 41 ○ incomplete and/or incorrect knowledge of the internal and external interactions of the various
 42 software components involved.
- 43 • Judgment shortcomings, which are failures of choice and will, and may include
 - 44 ○ selection of simpler but vulnerable constructs in place of more robust but more time consuming
 45 solutions and

- 1 ○ selection of terse constructs, which may be less understandable, in place of verbose, maintainable
2 solutions.

3 This technical report identifies issues related to the use of programming languages that can increase the likelihood
4 of errors due to cognitive limitations, and recommends ways that can be used to mitigate or eliminate these errors.
5 Some of the mechanisms recommended in this technical report include reducing the cognitive effort necessary in
6 reading existing source code, reducing the amount of knowledge needed by readers of existing source code, and
7 reducing the probability that incorrect developer knowledge will result in unpredictable code execution.

8 **5.3 Issues arising from a lack of predictable execution**

9 If a reasonably competent programmer has a good understanding of the intent of a program after reading source
10 code as far as a particular line of code, the programmer ought to have a good understanding of the state of the
11 program after reading the next line of code. However, some features, or, more likely, some combinations of
12 features, of programming languages are associated with relatively decreased rates of the programmer's
13 maintaining their understanding as they read through a program. It is these features and combinations of features
14 that are indicated in this document, along with ways to increase the programmer's understanding as code is read.

15 Here, the term understanding means the programmer's recognition of all effects, including subtle or unintended
16 changes of state, of any language feature or combination of features appearing in the program. This view does not
17 imply that programmers only read code from beginning to end. It is simply a statement that a line of code changes
18 the state of a program, and that a reasonably competent programmer ought to understand the state of the program
19 both before and after reading any line of code. It is only to a first approximation that code is read and understood
20 line by line.

21 **5.4 Issues arising from the lack of portability and interoperability**

22 The representation of characters, the representation of true/false values, the set of valid addresses, the properties
23 and limitations of any (fixed-point or floating-point) numerical quantities, and the representation of programmer-
24 defined types and classes may vary among hardware, among languages (affecting inter-language software
25 development), and among compilers of a given language. These variations may be the result of hardware
26 differences, operating system differences, library differences, compiler differences, or different configurations of the
27 same compiler (as may be set by environment variables or configuration files). In each of these circumstances,
28 there is an additional burden on the programmer because part of the program's behaviour is indicated by a factor
29 that is not a part of the source code. That is, the program's behaviour may be indicated by a factor that is invisible
30 when reading the source code. Compilation control schemes (IDE projects, make, and scripts) further complicate
31 this situation by abstracting and manipulating the relevant variables (target platform, compiler options, libraries, and
32 so forth).

33 Many compilers of standard-defined languages also support language features that are not specified by the
34 language standard. These non-standard features are called extensions. For portability, the programmer must be
35 aware of the language standard, and use only constructs with standard-defined semantics. The motivation to use
36 extensions may include the desire for increased functionality within a particular environment, or increased efficiency
37 on particular hardware. There are well-known software engineering techniques for minimizing the ill effects of
38 extensions; these techniques should be a part of any coding standard where they are needed, and they should be
39 employed whenever extensions are used. These issues are software engineering issues and are not further
40 discussed in this document.

41 Some language standards define libraries that are available as a part of the language definition. Such libraries are
42 an intrinsic part of the respective language and are called intrinsic libraries. There are also libraries defined by
43 other sources and these are called non-intrinsic libraries.

44 The use of non-intrinsic libraries to broaden the software primitives available in a given development environment is
45 a useful technique, allowing the use of trusted functionality directly in the program. Libraries may also allow the
46 program to bind to capabilities provided by an environment. However, these advantages are potentially offset by
47 any lack of skill on the part of the designer of the library (who may have designed subtle or undocumented changes
48 of state into the library's behaviour), and implementer of the library (who may not have implemented the library
49 identically on every platform), and even by the availability of the library on a new platform. The quality of the

1 documentation of a third-party library is another factor that may decrease the reliability of software using a library in
2 a particular situation by failing to describe clearly the library's full behaviour. If a library is missing on a new
3 platform, its functionality must be recreated in order to port any software depending upon the missing library. The
4 re-creation may be burdensome if the reason the library is missing is because the underlying capability for a
5 particular environment is missing.

6 Using a non-intrinsic library usually requires that options be set during compilation and linking phases, which
7 constitute a software behaviour specification beyond the source code. Again, these issues are software
8 engineering issues and are not further discussed in this document.

9 **5.5 Issues arising from inadequate language intrinsic support**

10 Many languages are created to facilitate programming within an application domain. Some languages are
11 specifically designed for programming of business applications, numerical computation or systems programming.
12 Problems can arise when, for example, a language being used to implement a real-time, multi-threaded system
13 lacks key features that are needed such as a way of enforcing mutual exclusion. Such features can be provided by
14 the programming environment in the form of libraries, but the definition of such libraries may be proprietary and
15 inclined to change in later releases. A vendor may even decide to withdraw support entirely for such a library. Also,
16 such a library may not be verified and validated to the same standard as the compiler and the application being
17 developed.

18 The use of *OOP* (Object-Oriented Programming) language features may well be highly appropriate for
19 implementing a *GUI* (Graphical User Interface) but at the same time, other features such as dynamic memory
20 management, heap utilization, inefficient data representation, and dynamic polymorphism are very unsuitable for
21 implementing a solid real-time safety-critical system. If a language is so intrinsically bound to OOP, it is sensible to
22 seek an alternative language for implementing some applications. Conversely, if the problem domain is GUI
23 development, it will often (although not always) be advisable to use a language that has OOP features.

24 Some potential problems may be preventable through the use of stronger types, or the use of controls such as
25 array bounds checking or integer checking to avoid overflows. These stronger restraints on a language have a cost
26 both in performance and in the flexibility to perform certain operations. Language designers must strike a balance
27 between restraints in the language, performance and flexibility causing some languages to lean heavily toward one
28 or more extremes in pursuit of some language attributes. The intrinsic support provided by a language can help
29 considerably in avoiding vulnerabilities, but such support can cause the utility of programming within a particular
30 application domain to diminish.

31 **5.6 Issues arising from language features prone to erroneous use**

32 Certain language constructs are relatively simple and straightforward to use. Other ones are complex to use or
33 easily misused in a legal, but unintended way. Programmers may use floating point variables and pointers without
34 fully understanding the nuances of the data representation. Rarely needed constructs or constructs that can be
35 substituted for a series of simpler constructs can be used without a complete understanding of the full effects of the
36 constructs.

37 Syntactic language features that are not intolerant of common typo errors can produce some problems that are
38 notoriously difficult to find. One common example of this is that C permits an unintentional assignment to be
39 performed in a Boolean expression by the accidental use of a single "=" (assignment) instead of the intended "=="
40 test for equality. It then allows the resulting value to be treated as a Boolean.

1 **6. Programming Language Vulnerabilities**

2 The standard for a programming language provides definitions for that language's constructs. This Technical
 3 Report will in general use the terminology that is most natural to the description for each individual vulnerability,
 4 relying upon the individual standards for terminology details. In general, the reader should be aware that "method",
 5 "function", and "procedure" could denote similar constructs in different languages, as can "pointer" and "reference".
 6 Situations described as "undefined behaviour" in some languages are known as "unbounded behaviour" in others.

7 **6.1 Obscure Language Features [BRS]**

8 **6.1.1 Description of application vulnerability**

9 Every programming language has features that are obscure, difficult to understand or difficult to use correctly. The
 10 problem is compounded if a software design must be reviewed by people who may not be language experts, such
 11 as, hardware engineers, human-factors engineers, or safety officers. Even if the design and code are initially
 12 correct, maintainers of the software may not fully understand the intent. The consequences of the problem are
 13 more severe if the software is to be used in trusted applications, such as safety or mission critical ones.

14 Misunderstood language features or misunderstood code sequences can lead to application vulnerabilities in
 15 development or in maintenance.

16 **6.1.2 Cross reference**

17 JSF AV Rules: 84, 86, 88, and 97
 18 MISRA C 2004: 3.2, 10.2, 13.1, 20.6-20.12, 12.10, and 17.5
 19 MISRA C++ 2008: 0-2-1, 2-3-1, and 12-1-1
 20 CERT/CC guidelines: FIO03-C, MSC05-C, MSC30-C, and MSC31-C.
 21 ISO/IEC TR 15942:2000: 5.4.2, 5.6.2 and 5.9.3

22 **6.1.3 Mechanism of failure**

23 The use of obscure language features can lead to an application vulnerability in several ways:

- 24 • The original programmer may misunderstand the correct usage of the feature and could utilize it incorrectly
 25 in the design or code it incorrectly.
- 26 • Reviewers of the design and code may misunderstand the intent or the usage and overlook problems.
- 27 • Maintainers of the code cannot fully understand the intent or the usage and could introduce problems
 28 during maintenance.

29 **6.1.4 Applicable language characteristics**

30 This vulnerability description is intended to be applicable to any language.

31 **6.1.5 Avoiding the vulnerability or mitigating its effects**

32 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 33 • Individual programmers should avoid the use of language features that are obscure or difficult to use,
 34 especially in combination with other difficult language features. Organizations should adopt coding
 35 standards that discourage use of such features or show how to use them correctly.
- 36 • Organizations developing software with critically important requirements should adopt a mechanism to
 37 monitor which language features are correlated with failures during the development process and during
 38 deployment.
- 39 • Organizations should adopt or develop stereotypical idioms for the use of difficult language features, codify
 40 them in organizational standards, and enforce them via review processes.

- 1 • Avoid the use of complicated features of a language.
- 2 • Avoid the use of rarely used constructs that could be difficult for entry-level maintenance personnel to
- 3 understand.
- 4 • Static analysis can be used to find incorrect usage of some language features.

5 It should be noted that consistency in coding is desirable for each of review and maintenance. Therefore, the
6 desirability of the particular alternatives chosen for inclusion in a coding standard does not need to be empirically
7 proven.

8 **6.1.6 Implications for standardization**

9 In future standardization activities, the following items should be considered:

- 10 • Language designers should consider removing or deprecating obscure, difficult to understand, or difficult to
- 11 use features.
- 12 • Language designers should provide language directives that optionally disable obscure language features.

13 **6.1.7 Bibliography**

14 Hatton 17: Use of obscure language features

15 **6.2 Unspecified Behaviour [BQF]**

16 **6.2.1 Description of application vulnerability**

17 The external behaviour of a program whose source code contains one or more instances of constructs having
18 unspecified behaviour may not be fully predictable when the source code is (re)compiled or (re)linked.

19 **6.2.2 Cross reference**

20 JSF AV Rules: 17-25

21 MISRA C 2004: 1.3, 1.5, 3.1 3.3, 3.4, 17.3, 1.2, 5.1, 18.2, 19.2, and 19.14

22 MISRA C++ 2008: 5-0-1, 5-2-6, 7-2-1, and 16-3-1

23 CERT/CC guidelines: MSC15-C

24 Also see guideline recommendations: Undefined Behaviour [EWF] and Implementation-defined Behaviour [FAB].

25 **6.2.3 Mechanism of failure**

26 Language specifications do not always uniquely define the behaviour of a construct. When an instance of a
27 construct that is not uniquely defined is encountered (this might be at any of compile, link, or run time)
28 implementations are permitted to choose from the set of behaviours allowed by the language specification. The
29 term 'unspecified behaviour' is sometimes applied to such behaviours, (language specific guidelines need to
30 analyze and document the terms used by their respective language).

31 A developer may use a construct in a way that depends on a subset of the possible behaviours occurring. The
32 behaviour of a program containing such a usage is dependent on the translator used to build it always selecting the
33 'expected' behaviour.

34 Many language constructs may have unspecified behaviour and unconditionally recommending against any use of
35 these constructs may be impractical. For instance, in many languages the order of evaluation of the operands
36 appearing on the left- and right-hand side of an assignment is unspecified, but in most cases the set of possible
37 behaviours always produce the same result.

38 The appearance of unspecified behaviour in a language specification is a recognition by the language designers
39 that in some cases flexibility is needed by software developers and provides a worthwhile benefit for language
40 translators; this usage is not a defect in the language.

1 The important characteristic is not the internal behaviour exhibited by a construct (e.g., the sequence of machine
 2 code generated by a translator) but its external behaviour (i.e., the one visible to a user of a program). If the set of
 3 possible unspecified behaviours permitted for a specific use of a construct all produce the same external effect
 4 when the program containing them is executed, then rebuilding the program cannot result in a change of behaviour
 5 for that specific usage of the construct.

6 For instance, while the following assignment statement contains unspecified behaviour in many languages (i.e., it is
 7 possible to evaluate either the `A` or `B` operand first, followed by the other operand):

8 `A = B;`

9 in most cases the order in which `A` and `B` are evaluated does not effect the external behaviour of a program
 10 containing this statement.

11 **6.2.4 Applicable language characteristics**

12 This vulnerability is intended to be applicable to languages with the following characteristics:

- 13 • Languages whose specification allows a finite set of more than one behaviour for how a translator handles
 14 some construct, where two or more of the behaviours can result in differences in external program
 15 behaviour.

16 **6.2.5 Avoiding the vulnerability or mitigating its effects**

17 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 18 • Use language constructs that have specified behaviour.
- 19 • Ensure that a specific use of a construct having unspecified behaviour produces a result that is the same
 20 for all of the possible behaviours permitted by the language specification.
- 21 • When developing coding guidelines for a specific language all constructs that have unspecified behaviour
 22 should be documented and for each construct the situations where the set of possible behaviours can vary
 23 shall be enumerated.

24 **6.2.6 Implications for standardization**

25 In future standardization activities, the following items should be considered:

- 26 • Languages should minimize the amount of unspecified behaviours, minimize the number of possible
 27 behaviours for any given "unspecified" choice, and document what might be the difference in external
 28 effect associated with different choices.

29 **6.2.7 Bibliography**

30 [None]

31 **6.3 Undefined Behaviour [EWF]**

32 **6.3.1 Description of application vulnerability**

33 The external behaviour of a program containing an instance of a construct having undefined behaviour, as defined
 34 by the language specification, is not predictable.

35 **6.3.2 Cross reference**

36 JSF AV Rules: 17-25

- 1 MISRA C 2004: 1.3, 1.5, 3.1 3.3, 3.4, 17.3, 1.2, 5.1, 18.2, 19.2, and 19.14
- 2 MISRA C++ 2008: 2-13-1, 5-2-2, 16-2-4, and 16-2-5
- 3 CERT/CC guidelines: MSC15-C
- 4 See guideline recommendations: Unspecified Behaviour [BQF] and Implementation-defined Behaviour [FAB].

5 **6.3.3 Mechanism of failure**

6 Language specifications may categorize the behaviour of a language construct as undefined rather than as a
7 semantic violation (i.e., an erroneous use of the language) because of the potentially high implementation cost of
8 detecting and diagnosing all occurrences of it. In this case no specific behaviour is required and the translator or
9 runtime system is at liberty to do anything it pleases (which may include issuing a diagnostic).

10 The behaviour of a program built from successfully translated source code containing a construct having undefined
11 behaviour is not predictable. For example, in some languages the value of a variable is undefined before it is
12 initialized.

13 **6.3.4 Applicable language characteristics**

14 This vulnerability is intended to be applicable to languages with the following characteristics:

- 15 • Languages that do not fully define the extent to which the use of a particular construct is a violation of the
16 language specification.
- 17 • Languages that do not fully define the behaviour of constructs during compile, link and program execution.

18 **6.3.5 Avoiding the vulnerability or mitigating its effects**

19 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 20 • Ensuring that undefined language constructs are not used.
- 21 • Ensuring that a use of a construct having undefined behaviour does not operate within the domain in which
22 the behaviour is undefined. When it is not possible to completely verify the domain of operation during
23 translation a runtime check may need to be performed.
- 24 • When developing coding guidelines for a specific language all constructs that have undefined behaviour
25 should be documented. The items on this list might be classified by the extent to which the behaviour is
26 likely to have some critical impact on the external behaviour of a program (the criticality may vary between
27 different implementations, e.g., whether conversion between object and function pointers has well defined
28 behaviour).

29 **6.3.6 Implications for standardization**

30 In future standardization activities, the following items should be considered:

- 31 • Language designers should minimize the amount of undefined behaviour to the extent possible and
32 practical.
- 33 • Language designers should enumerate all the cases of undefined behaviour.
- 34 • Language designers should provide language directives that optionally disable obscure language features.

35 **6.3.7 Bibliography**

36 [None]

1 **6.4 Implementation-defined Behaviour [FAB]**

2 **6.4.1 Description of application vulnerability**

3 Some constructs in programming languages are not fully defined (see Unspecified Behaviour [BQF]) and thus
 4 leave compiler implementations to decide how the construct will operate. The behaviour of a program whose
 5 source code contains one or more instances of constructs having implementation-defined behaviour, can change
 6 when the source code is recompiled or relinked.

7 **6.4.2 Cross reference**

8 JSF AV Rules: 17-25

9 MISRA C 2004: 1.3, 1.5, 3.1 3.3, 3.4, 17.3, 1.2, 5.1, 18.2, 19.2, and 19.14

10 MISRA C++ 2008: 5-2-9, 5-3-3, 7-3-2, and 9-5-1

11 CERT/CC guidelines: MSC15-C

12 ISO/IEC TR 15942:2000: 5.9

13 Also see guideline recommendations: Unspecified Behaviour [BQF] and Undefined Behaviour [EWF].

14 **6.4.3 Mechanism of failure**

15 Language specifications do not always uniquely define the behaviour of a construct. When an instance of a
 16 construct that is not uniquely defined is encountered (this might be at any of translation, link-time, or program
 17 execution) implementations are permitted to choose from a set of behaviours. The only difference from
 18 unspecified behaviour is that implementations are required to document how they behave.

19 A developer may use a construct in a way that depends on a particular implementation-defined behaviour
 20 occurring. The behaviour of a program containing such a usage is dependent on the translator used to build it
 21 always selecting the 'expected' behaviour.

22 Some implementations provide a mechanism for changing an implementation's implementation-defined behaviour
 23 (e.g., use of `pragmas` in source code). Use of such a change mechanism creates the potential for additional
 24 human error in that a developer may be unaware that a change of behaviour was requested earlier in the source
 25 code and may write code that depends on the implementation-defined behaviour that occurred prior to that explicit
 26 change of behaviour.

27 Many language constructs may have implementation-defined behaviour and unconditionally recommending against
 28 any use of these constructs may be completely impractical. For instance, in many languages the number of
 29 significant characters in an identifier is implementation-defined. Developers need to choose a minimum number of
 30 characters and require that only translators supporting at least that number, *N*, of characters be used.

31 The appearance of implementation-defined behaviour in a language specification is recognition by the language
 32 designers that in some cases implementation flexibility provides a worthwhile benefit for language translators; this
 33 usage is not a defect in the language.

34 **6.4.4 Applicable language characteristics**

35 This vulnerability is intended to be applicable to languages with the following characteristics:

- 36 • Languages whose specification allows some variation in how a translator handles some construct, where
 37 reliance on one form of this variation can result in differences in external program behaviour.
- 38 • Language implementations may not be required to provide a mechanism for controlling implementation-
 39 defined behaviour.

40 **6.4.5 Avoiding the vulnerability or mitigating its effects**

41 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 1 • Document the set of implementation-defined features an application depends upon, so that upon a change
2 of translator, development tools, or target configuration it can be ensured that those dependencies are still
3 met
- 4 • Ensure that a specific use of a construct having implementation-defined behaviour produces an external
5 behaviour that is the same for all of the possible behaviours permitted by the language specification.
- 6 • Only use a language implementation whose implementation-defined behaviours are within a known subset
7 of implementation-defined behaviours. The known subset should be chosen so that the 'same external
8 behaviour' condition described above is met.
- 9 • Create highly visible documentation (e.g., at the start of a source file) that the default implementation-
10 defined behaviour is changed within the current file.
- 11 • When developing coding guidelines for a specific language all constructs that have implementation-defined
12 behaviour shall be documented and for each construct, the situations where the set of possible behaviours
13 can vary shall be enumerated.
- 14 • When applying this guideline on a project the functionality provided by and for changing its implementation-
15 defined behaviour shall be documented.
- 16 • Verify code behaviour using two different compilers with two different technologies.

17 **6.4.6 Implications for standardization**

18 In future standardization activities, the following items should be considered:

- 19 • Portability guidelines for a specific language should provide a list of common implementation-defined
20 behaviours.
- 21 • Language specifiers should enumerate all the cases of implementation-defined behaviour.
- 22 • Language designers should provide language directives that optionally disable obscure language features.

23 **6.4.7 Bibliography**

24 [None]

25 **6.5 Deprecated Language Features [MEM]**

26 **6.5.1 Description of application vulnerability**

27 All code should conform to the current standard for the respective language. In reality though, a language standard
28 may change during the creation of a software system or suitable compilers and development environments may not
29 be available for the new standard for some period of time after the standard is published. In order to smooth the
30 process of evolution, features that are no longer needed or which serve as the root cause of or contributing factor
31 for safety or security problems are often deprecated to temporarily allow their continued use but to indicate that
32 those features may be removed in the future. The deprecation of a feature is a strong indication that it should not
33 be used. Other features, although not formally deprecated, are rarely used and there exist other more common
34 ways of expressing the same function. Use of these rarely used features can lead to problems when others are
35 assigned the task of debugging or modifying the code containing those features.

36 **6.5.2 Cross reference**

37 JSF AV Rules: 8 and 11
38 MISRA C 2004: 1.1, 4.2, and 20.10
39 MISRA C++ 2008: 1-0-1, 2-3-1, 2-5-1, 2-7-1, 5-2-4, and 18-0-2

40 **6.5.3 Mechanism of failure**

41 Most languages evolve over time. Sometimes new features are added making other features extraneous.
42 Languages may have features that are frequently the basis for security or safety problems. The deprecation of
43 these features indicates that there is a better way of accomplishing the desired functionality. However, there is
44 always a time lag between the acknowledgement that a particular feature is the source of safety or security

1 problems, the decision to remove or replace the feature and the generation of warnings or error messages by
 2 compilers that the feature shouldn't be used. Given that software systems can take many years to develop, it is
 3 possible and even likely that a language standard will change causing some of the features used to be suddenly
 4 deprecated. Modifying the software can be costly and time consuming to remove the deprecated features.
 5 However, if the schedule and resources permit, this would be prudent as future vulnerabilities may result from
 6 leaving the deprecated features in the code. Ultimately the deprecated features will likely need to be removed
 7 when the features are removed.

8 **6.5.4 Applicable language characteristics**

9 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 10 • All languages that have standards, though some only have defacto standards.
- 11 • All languages that evolve over time and as such could potentially have deprecated features at some point.

12 **6.5.5 Avoiding the vulnerability or mitigating its effects**

13 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 14 • Adhere to the latest published standard for which a suitable compiler and development environment is
 15 available.
- 16 • Avoid the use of deprecated features of a language.
- 17 • Stay abreast of language discussions in language user groups and standards groups on the Internet.
 18 Discussions and meeting notes will give an indication of problem prone features that should not be used or
 19 should be used with caution.

20 **6.5.6 Implications for standardization**

21 In future standardization activities, the following items should be considered:

- 22 • Obscure language features for which there are commonly used alternatives should be considered for
 23 removal from the language standard.
- 24 • Obscure language features that have routinely been found to be the root cause of safety or security
 25 vulnerabilities, or that are routinely disallowed in software guidance documents should be considered for
 26 removal from the language standard.
- 27 • Language designers should provide language directives that optionally disable obscure language features.

28 **6.5.7 Bibliography**

29 [None]

30 **6.6 Pre-processor Directives [NMP]**

31 **6.6.1 Description of application vulnerability**

32 Pre-processor replacements happen before any source code syntax check, therefore there is no type checking –
 33 this is especially important in function-like macro parameters.

34 If great care is not taken in the writing of macros, the expanded macro can have an unexpected meaning. In many
 35 cases if explicit delimiters are not added around the macro text and around all macro arguments within the macro
 36 text, unexpected expansion is the result.

37 Source code that relies heavily on complicated pre-processor directives may result in obscure and hard to maintain
 38 code since the syntax they expect may be different from the expressions programmers regularly expect in a given
 39 programming language.

1 **6.6.2 Cross reference**

- 2 Holtzmann-8
- 3 JSF SV Rules: 26, 27, 28, 29, 30, 31, and 32
- 4 MISRA C 2004: 19.6, 19.7, 19.8, and 19.9
- 5 MISRA C++ 2008: 16-0-3, 16-0-4, and 16-0-5
- 6 CERT/CC guidelines: PRE01-C, PRE02-C, PRE10-C, and PRE31-C

7 **6.6.3 Mechanism of failure**

8 Readability and maintainability may be greatly increased if the language features available in the programming
9 language are used instead of a pre-processor directive.

10 While static analysis can identify many problems early; heavy use of the pre-processor can limit the effectiveness
11 of many static analysis tools, which typically work on the pre-processed source code.

12 In many cases where complicated macros are used, the program does not do what is intended. For example:

13 define a macro as follows,

```
#define CD(x, y) (x + y - 1) / y
```

14 whose purpose is to divide. Then suppose it is used as follows

```
a = CD (b & c, sizeof (int));
```

15 which expands into

```
a = (b & c + sizeof (int) - 1) / sizeof (int);
```

16 which most times will not do what is intended. Defining the macro as

```
#define CD(x, y) ((x) + (y) - 1) / (y)
```

17 will provide the desired result.

18 **6.6.4 Applicable language characteristics**

19 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 20 • Languages that have a lexical-level pre-processor.
- 21 • Languages that allow unintended groupings of arithmetic statements.
- 22 • Languages that allow cascading macros.
- 23 • Languages that allow duplication of side effects.
- 24 • Languages that allow macros that reference themselves.
- 25 • Languages that allow nested macro calls.
- 26 • Languages that allow complicated macros.

27 **6.6.5 Avoiding the vulnerability or mitigating its effects**

28 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 29 • Where it is possible to achieve the desired functionality without the use of pre-processor directives, this
30 should be done in preference to the use of pre-processor directives.
31

32 **6.6.6 Implications for standardization**

33 In future standardization activities, the following items should be considered:

- 1 • Standards should reduce or eliminate dependence on lexical-level pre-processors for essential functionality
 2 (such as conditional compilation).
 3 • Standards should consider providing capabilities to inline functions and procedure calls, to reduce the need
 4 for pre-processor macros.

5 **6.6.7 Bibliography**

6 [None]

7 **6.7 Choice of Clear Names [NAI]**

8 **6.7.1 Description of application vulnerability**

9 Humans sometimes choose similar or identical names for objects, types, aggregates of types, subprograms and
 10 modules. They tend to use characteristics that are specific to the native language of the software developer to aid
 11 in this effort, such as use of mixed-casing, underscores and periods, or use of plural and singular forms to support
 12 the separation of items with similar names. Similarly, development conventions sometimes use casing for
 13 differentiation (e.g., all uppercase for constants).

14 Human cognitive problems occur when different (but similar) objects, subprograms, types, or constants differ in
 15 name so little that human reviewers are unlikely to distinguish between them, or when the system maps such
 16 entities to a single entity.

17 Conventions such as the use of capitalization, and singular/plural distinctions may work in small and medium
 18 projects, but there are a number of significant issues to be considered:

- 19 • Large projects often have mixed languages and such conventions are often language-specific.
 20 • Many implementations support identifiers that contain international character sets and some language
 21 character sets have different notions of casing and plurality.
 22 • Different word-forms tend to be language and dialect specific, such as a pidgin, and may be meaningless
 23 to humans that speak other dialects.

24
 25 An important general issue is the choice of names that differ from each other negligibly (in human terms), for
 26 example by differing by only underscores, (none, "_ " "__"), plurals ("s"), visually identical letters (such as "l" and "1",
 27 "O" and "0"), or underscores/dashes ("-","_"). [There is also an issue where identifiers appear distinct to a human
 28 but identical to the computer, e.g., FOO, Foo, and foo in some computer languages.] Character sets extended with
 29 diacritical marks and non-Latin characters may offer additional problems. Some languages or their implementations
 30 may pay attention to only the first n characters of an identifier.

31 The problems described above are different from overloading or overriding where the same name is used
 32 intentionally (and documented) to access closely linked sets of subprograms. This is also different than using
 33 reserved names which can lead to a conflict with the reserved use and the use of which may or may not be
 34 detected at compile time.

35 Name confusion can lead to the application executing different code or accessing different objects than the writer
 36 intended, or than the reviewers understood. This can lead to outright errors, or leave in place code that may
 37 execute some time in the future with unacceptable consequences.

38 Although most such mistakes are unintentional, it is plausible that such usages can be intentional, if masking
 39 surreptitious behaviour is a goal.

40 **6.7.2 Cross Reference**

41 JSF AV Rules: 48-56
 42 MISRA C 2004: 1.4
 43 CERT/CC guidelines: DCL02-C

1 **6.7.3 Mechanism of Failure**

2 Calls to the wrong subprogram or references to the wrong data element (that was missed by human review) can
3 result in unintended behaviour. Language processors will not make a mistake in name translation, but human
4 cognition limitations may cause humans to misunderstand, and therefore may be easily missed in human reviews.

5 **6.7.4 Applicable language characteristics**

6 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 7 • Languages with relatively flat name spaces will be more susceptible. Systems with modules, classes,
8 packages can use qualification to disambiguate names that originate from different parents.
- 9 • Languages that provide preconditions, postconditions, invariances and assertions or redundant coding of
10 subprogram signatures help to ensure that the subprograms in the module will behave as expected, but do
11 nothing if different subprograms are called.
- 12 • Languages that treat letter case as significant. Some languages do not differentiate between names with
13 differing case, while others do.

14 **6.7.5 Avoiding the vulnerability or mitigating its effects**

15 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 16 • Implementers can create coding standards that provide meaningful guidance on name selection and use.
17 Good language specific guidelines could eliminate most problems.
- 18 • Use static analysis tools to show the target of calls and accesses and to produce alphabetical lists of
19 names. Human review can then often spot the names that are sorted at an unexpected location or which
20 look almost identical to an adjacent name in the list.
- 21 • Use static tools (often the compiler) to detect declarations that are unused.
- 22 • Use languages with a requirement to declare names before use or use available tool or compiler options to
23 enforce such a requirement.

24 **6.7.6 Implications for standardization**

25 In future standardization activities, the following items should be considered:

- 26 • Languages that do not require declarations of names should consider providing an option that does impose
27 that requirement.

28 **6.7.7 Bibliography**

29 Jones, Derek, "Some proposed language vulnerability guidelines" Submitted to the December 2006 Washington,
30 D.C. meeting of the ISO/IEC SC22 OWGV

31 Jones, Derek M., "The New C Standard (Identifiers)" www.coding-guidelines.com/cbook/sent792.pdf

32 **6.8 Choice of Filenames and other External Identifiers [AJN]**

33 **6.8.1 Description of application vulnerability**

34 Interfacing with the directory structure or other external identifiers on a system on which software executes is very
35 common. Differences in the conventions used by operating systems can result in significant changes in behaviour
36 when the same program is executed under different operating systems. For instance, the directory structure,
37 permissible characters, case sensitivity, and so forth can vary among operating systems and even among
38 variations of the same operating system. For example, Microsoft XP prohibits "/?:&*"<>|#%"; but UNIX, Linux, and
39 OS X operating systems allow any character except for the reserved character '/' to be used in a filename.

1 Some operating systems are case sensitive while others are not. On non-case sensitive operating systems,
 2 depending on the software being used, the same filename could be displayed, as “filename”, “Filename” or
 3 “FILENAME” and all would refer to the same file.

4 Some operating systems, particularly older ones, only rely on the significance of the first *n* characters of the file
 5 name. *n* can be unexpectedly small, such as the first 8 characters in the case of Win16 architectures which would
 6 cause “filename1”, “filename2” and “filename3” to all map to the same file.

7 Variations in the filename, named resource or external identifier being referenced can be the basis for various kinds
 8 of problems. Such mistakes or ambiguity can be unintentional, or intentional, and in either case they can be
 9 potentially exploited, if surreptitious behaviour is a goal.

10 **6.8.2 Cross Reference**

11 JSF AV Rules: 46, 51, 53, 54, 55, and 56
 12 MISRA C 2004: 1.4 and 5.1
 13 CERT/CC guidelines: MSC09-C and MSC10-C

14 **6.8.3 Mechanism of Failure**

15 The wrong named resource, such as a file, may be used within a program in a form that provides access to a
 16 resource that was not intended to be accessed. Attackers could exploit this situation to intentionally misdirect
 17 access of a named resource to another named resource.

18 **6.8.4 Applicable language characteristics**

19 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 20 • Any language providing for use of an *API* (Application Programming Interface) for external access of
 21 resources with varied naming conventions. In practice, this means all languages.
- 22 • A particular language interface to a system should be consistent in its processing of filenames or external
 23 identifiers. Consistency is only the first consideration. Even though it is consistent, it may consistently do
 24 something that is unexpected by the developer of the software interfacing with the system.

25 **6.8.5 Avoiding the vulnerability or mitigating its effects**

26 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 27 • Where possible, use an API that provides a known common set of conventions for naming and accessing
 28 external resources, such as POSIX, ISO/IEC 9945:2003 (IEEE Std 1003.1-2001).
- 29 • Analyze the range of intended target systems, develop a suitable API for dealing with them, and document
 30 the analysis.
- 31 • Ensure that programs adapt their behaviour to the platform on which they are executing, so that only the
 32 intended resources are accessed. This means that information on such characteristics as the directory
 33 separator string and methods of accessing parent directories need to be parameterized and not exist as
 34 fixed strings within a program.
- 35 • Avoid creating resource names that are longer than the guaranteed unique length of all potential target
 36 platforms.
- 37 • Avoid creating resources, which are differentiated only by the case in their names.

39 **6.8.6 Implications for standardization**

40 In future standardization activities, the following items should be considered:

- 41 • Language APIs for interfacing with external identifiers should be compliant with ISO/IEC 9945:2003 (IEEE
 42 Std 1003.1-2001).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

6.8.7 Bibliography

Jones, Derek, "Some proposed language vulnerability guidelines" Submitted to the December 2006 Washington, D.C. meeting of the ISO/IEC SC22 OWGV

6.9 Unused Variable [XYR]

6.9.1 Description of application vulnerability

A variable's value is assigned but never used, making it a dead store. As a variant, a variable is declared but neither read nor written to in the program, making it an unused variable. This type of error suggests that the design has been incompletely or inaccurately implemented.

Unused variables by themselves are innocuous, but can be combined with other vulnerabilities such as index bounds errors and buffer overflows and may mask errors or provide hidden channels.

6.9.2 Cross reference

CWE:
563. Unused Variable
MISRA C++ 2008: 0-1-4 and 0-1-6
CERT/CC guidelines: MSC13-C

6.9.3 Mechanism of failure

A variable is declared, but never used. It is likely that the variable is simply vestigial, but it is also possible that the unused variable points out a bug. This is likely to suggest that the design has been incompletely or inaccurately implemented.

A variable is assigned a value but this value is never used thereafter. The assignment is then generally referred to as a dead store. Note that this may be acceptable if the variable is a volatile variable, for which the assignment of a value triggers some external event.

A dead store is indicative of careless programming or of a design or coding error; either the use of the value was forgotten (almost certainly an error) or the assignment was performed even though it was not needed (unless there is a justification for it).

An unused variable or a dead store is very unlikely to be the cause of a vulnerability. However, since compilers diagnose unused variables routinely and dead stores occasionally, their presence is often an indication that compiler warnings are either suppressed or are being ignored by programmers. This observation does not hold for automatically generated code, where it is commonplace to find unused variables and dead stores, introduced to keep the generation process simple and uniform.

6.9.4 Applicable language characteristics

This vulnerability description is intended to be applicable to languages with the following characteristics:

- Dead stores are possible in any programming language that provides assignment. (Pure functional languages do not have this issue.)
- Unused variables (in the technical sense above) are possible only in languages that provide variable declarations.

6.9.5 Avoiding the vulnerability or mitigating its effects

Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 1 • Enable detection of unused variables and dead stores in the compiler. The default setting may be to
2 suppress these warnings.

3 **6.9.6 Implications for standardization**

4 In future standardization activates, the following items should be considered:

- 5 • Languages should consider requiring mandatory diagnostics for unused variables.

6 **6.9.7 Bibliography**

7 [None]

8 **6.10 Identifier Name Reuse [YOW]**

9 **6.10.1 Description of application vulnerability**

10 When distinct entities are defined in nested scopes using the same name it is possible that program logic will
11 operate on an entity other than the one intended.

12 When it is not clear which identifier is used, the program could behave in ways that were not predicted by reading
13 the source code. This can be found by testing, but circumstances can arise (such as the values of the same-named
14 objects being mostly the same) where harmful consequences occur. This weakness can also lead to vulnerabilities
15 such as hidden channels where humans believe that important objects are being rewritten or overwritten when in
16 fact other objects are being manipulated.

17 For example, the innermost definition is deleted from the source, the program will continue to compile without a
18 diagnostic being issued (but execution can produce unexpected results).

19 **6.10.2 Cross reference**

20 JSF AV Rules: 120 and 1359

21 MISRA C 2004: 5.2, 5.5, 5.6, 5.7, 20.1, 20.2

22 MISRA C++ 2008: 2-10-2, 2-10-3, 2-10-4, 2-10-5, 2-10-6, 17-0-1, 17-0-2, and 17-0-3

23 CERT/CC guidelines: DCL01-C and DCL32-C

24 **6.10.3 Mechanism of failure**

25 Many languages support the concept of scope. One of the ideas behind the concept of scope is to provide a
26 mechanism for the independent definition of identifiers that may share the same name.

27 For instance, in the following code fragment:

```
28       int some_var;
29       
30       {
31           {
32               int t_var;
33               int some_var; /* definition in nested scope */
34               
35               t_var=3;
36               some_var=2;
37           }
38       }
```

39 an identifier called `some_var` has been defined in different scopes.

1 If either the definition of `some_var` or `t_var` that occurs in the nested scope is deleted (e.g., when the source is
 2 modified) it is necessary to delete all other references to the identifier's scope. If a developer deletes the definition
 3 of `t_var` but fails to delete the statement that references it, then most languages require a diagnostic to be issued
 4 (e.g., reference to undefined variable). However, if the nested definition of `some_var` is deleted but the reference
 5 to it in the nested scope is not deleted, then no diagnostic will be issued (because the reference resolves to the
 6 definition in the outer scope).

7 An example of how interpretations of a programming language can differ, in the following code fragment:

```
8     int j = 100;
9     {
10        for (int j = 0; j < 10; j++) ;
11        std::cout << j << std::endl; // What is the value of j
12    }
```

13 According to ISO 14882:2003 (C++) standard the value printed for `j` should be 100, but in some implementations
 14 that do not conform to the current version of the standard it will be 10, as the loop counter `j` remains in-scope after
 15 the end of the loop statement.

16 In some cases non-unique identifiers in the same scope can also be introduced through the use of identifiers
 17 whose common substring exceeds the length of characters the implementation considers to be distinct. For
 18 example, in the following code fragment:

```
19     extern int global_symbol_definition_lookup_table_a[100];
20     extern int global_symbol_definition_lookup_table_b[100];
```

21 the external identifiers are not unique on implementations where only the first 31 characters are significant. This
 22 situation only occurs in languages that allow multiple declarations of the same identifier (other languages require a
 23 diagnostic message to be issued). (See, Choice of Filenames and other External Identifiers [AJN].)

24 A related problem exists in languages that allow overloading or overriding of keywords or standard library function
 25 identifiers. Such overloading can lead to confusion about which entity is intended to be referenced.

26 Definitions for new identifiers should not use a name that is already visible within the scope containing the new
 27 definition. Alternately, utilize language-specific facilities that check for and prevent inadvertent overloading of
 28 names should be used.

29 **6.10.4 Applicable language characteristics**

30 This vulnerability is intended to be applicable to languages with the following characteristics:

- 31 • Languages that allow the same name to be used for identifiers defined in nested scopes.
- 32 • Languages where unique names can be transformed into non-unique names as part of the normal tool
 33 chain.

34 **6.10.5 Avoiding the vulnerability or mitigating its effects**

35 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 36 • Ensure that a definition of an entity does not occur in a scope where a different entity with the same name
 37 is accessible and can be used in the same context. A language-specific project coding convention can be
 38 used to ensure that such errors are detectable with static analysis.
- 39 • Ensure that a definition of an entity does not occur in a scope where a different entity with the same name
 40 is accessible and has a type that permits it to occur in at least one context where the first entity can occur.
- 41 • Use language features, if any, which explicitly mark definitions of entities that are intended to hide other
 42 definitions.

- 1 • Develop or use tools that identify name collisions or reuse when truncated versions of names cause
- 2 conflicts.
- 3 • Ensure that all identifiers differ within the number of characters considered to be significant by the
- 4 implementations that are likely to be used, and document all assumptions.

6 6.10.6 Implications for standardization

7 In future standardization activities, the following items should be considered:

- 8 • Languages should require mandatory diagnostics for variables with the same name in nested scopes.
- 9 • Languages should require mandatory diagnostics for variable names that exceed the length that the
- 10 implementation considers unique.
- 11 • Languages should consider requiring mandatory diagnostics for overloading or overriding of keywords or
- 12 standard library function identifiers.

13 6.10.7 Bibliography

14 [None]

15 6.11 Type System [IHN]

16 6.11.1 Description of application vulnerability

17 When data values are converted from one data type to another, even when done intentionally, unexpected results
18 can occur.

19 6.11.2 Cross reference

20 JSF AV Rule: 148 and 183
21 MISRA C 2004: 6.1, 6.2, 6.3, 10.1, and 10.5
22 MISRA C++ 2008: 3-9-2, 5-0-3 to 5-0-14
23 CERT/CC guidelines: DCL07-C, DCL11-C, DCL35-C, EXP05-C and EXP32-C

24 6.11.3 Mechanism of failure

25 The *type* of a data object informs the compiler how values should be represented and which operations may be
26 applied. The *type system* of a language is the set of rules used by the language to structure and organize its
27 collection of types. Any attempt to manipulate data objects with inappropriate operations is a *type error*. A program
28 is said to be *type safe* (or *type secure*) if it can be demonstrated that it has no type errors [2].

29 Every programming language has some sort of type system. A language is *statically typed* if the type of every
30 expression is known at compile time. The type system is said to be *strong* if it guarantees type safety and *weak* if it
31 does not. There are strongly typed languages that are not statically typed because they enforce type safety with run
32 time checks [2].

33 In practical terms, nearly every language falls short of being strongly typed (in an ideal sense) because of the
34 inclusion of mechanisms to bypass type safety in particular circumstances. For that reason and because every
35 language has a different type system, this description will focus on taking advantage of whatever features for type
36 safety may be available in the chosen language.

37 Sometimes it is appropriate for a data value to be converted from one type to another *compatible* one. For
38 example, consider the following program fragment, written in no specific language:

```
39     float a;
40     integer i;
41     a := a + i;
```

1 The variable "i" is of integer type. It must be converted to the float type before it can be added to the data value.
 2 An implicit conversion, as shown, is called coercion. If, on the other hand, the conversion must be explicit, e.g., "a
 3 := a + float(i)", then the conversion is called a cast.

4 Type *equivalence* is the strictest form of type compatibility; two types are equivalent if they are compatible without
 5 using coercion or casting. Type equivalence is usually characterized in terms of *name type equivalence*—two
 6 variables have the same type if they are declared in the same declaration or declarations that use the same type
 7 name—or *structure type equivalence*—two variables have the same type if they have identical structures. There
 8 are variations of these approaches and most languages use different combinations of them [1]. Therefore, a
 9 programmer skilled in one language may very well code inadvertent type errors when using a different language.

10 It is desirable for a program to be type safe because the application of operations to operands of an inappropriate
 11 type may produce unexpected results. In addition, the presence of type errors can reduce the effectiveness of
 12 static analysis for other problems. Searching for type errors is a valuable exercise because their presence often
 13 reveals design errors as well as coding errors. Many languages check for type errors—some at compile-time,
 14 others at run-time. Obviously, compile-time checking is more valuable because it can catch errors that are not
 15 executed by a particular set of test cases.

16 Making the most use of the type system of a language is useful in two ways. First, data conversions always bear
 17 the risk of changing the value. For example, a conversion from integer to float risks the loss of significant digits
 18 while the inverse conversion risks the loss of any fractional value. Conversion of an integer value from a type with a
 19 longer representation to a type with a shorter representation risks the loss of significant digits. This can produce
 20 particularly puzzling results if the value is used to index an array. Conversion of a floating-point value from a type
 21 with a longer representation to a type with a shorter representation risks the loss of precision. This can be
 22 particularly severe in computations where the number of calculations increase as a power of the problem size. (It
 23 should be noted that similar surprises can occur when an application is retargeted to a machine with different
 24 representations of numeric values.)

25 Second, a programmer can use the type system to increase the probability of catching design errors or coding
 26 blunders. For example, the following Ada fragment declares two distinct floating-point types:

```
27     type Celsius is new Float;  
28     type Fahrenheit is new Float;
```

29 The declaration makes it impossible to add a value of type Celsius to a value of type Fahrenheit without explicit
 30 conversion.

31 **6.11.4 Applicable language characteristics**

32 This vulnerability is intended to be applicable to languages with the following characteristics:

- 33 • Languages that support multiple types and allow conversions between types.

34 **6.11.5 Avoiding the vulnerability or mitigating its effects**

35 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 36 • Take advantage of any facility offered by the programming language to declare distinct types and use any
 37 mechanism provided by the language processor and related tools to check for or enforce type compatibility.
- 38 • Use available language and tools facilities to preclude or detect the occurrence of coercion. If it is not
 39 possible, use human review to assist in searching for coercions.
- 40 • Avoid casting data values except when there is no alternative. Document such occurrences so that the
 41 justification is made available to maintainers.
- 42 • Use the most restricted data type that suffices to accomplish the job. For example, use an enumeration
 43 type to select from a limited set of choices (e.g., a switch statement or the discriminant of a union type)
 44 rather than a more general type, such as integer. This will make it possible for tooling to check if all
 45 possible choices have been covered.

- 1 • Treat every compiler, tool, or run-time diagnostic concerning type compatibility as a serious issue. Do not
2 resolve the problem by modifying the code by inserting an explicit cast, without further analysis; instead
3 examine the underlying design to determine if the type error is a symptom of a deeper problem.
4 • Never ignore instances of coercion; if the conversion is necessary, convert it to a cast and document the
5 rationale for use by maintainers.
6 • Analyze the problem to be solved to learn the magnitudes and/or the precisions of the quantities needed as
7 auxiliary variables, partial results and final results.

8 **6.11.6 Implications for standardization**

9 In future standardization activities, the following items should be considered:

- 10 • Language specifiers should standardize on a common, uniform terminology to describe their type systems
11 so that programmers experienced in other languages can reliably learn the type system of a language that
12 is new to them.
13 • Provide a mechanism for selecting data types with sufficient capability for the problem at hand.
14 • Provide a way for the computation to determine the limits of the data types actually selected.
15 • Language implementers should consider providing compiler switches or other tools to provide the highest
16 possible degree of checking for type errors.

17 **6.11.7 Bibliography**

18 [1] Robert W. Sebesta, Concepts of Programming Languages, 8th edition, ISBN-13: 978-0-321-49362-0, ISBN-10:
19 0-321-49362-1, Pearson Education, Boston, MA, 2008

20 [2] Carlo Ghezzi and Mehdi Jazayeri, Programming Language Concepts, 3rd edition, ISBN-0-471-10426-4, John
21 Wiley & Sons, 1998

22 **6.12 Bit Representations[STR]**

23 **6.12.1 Description of application vulnerability**

24 Interfacing with hardware, other systems and protocols often requires access to one or more bits in a single
25 computer word, or access to bit fields that may cross computer words for the machine in question. Mistakes can be
26 made as to what bits are to be accessed because of the “endianness” of the processor (see below) or because of
27 miscalculations. Access to those specific bits may affect surrounding bits in ways that compromise their integrity.
28 This can result in the wrong information being read from hardware, incorrect data or commands being given, or
29 information being mangled, which can result in arbitrary effects on components attached to the system

30 **6.12.2 Cross reference**

31 JSF AV Rules 147, 154 and 155

32 MISRA C 2004: 3.5, 6.4, 6.5, and 12.7

33 MISRA C++ 2008: 5-0-21, 5-2-4 to 5-2-9, and 9-5-1

34 CERT/CC guidelines: EXP38-C, INT00-C, INT07-C, INT12-C, INT13-C, and INT14-C

35 **6.12.3 Mechanism of failure**

36 Computer languages frequently provide a variety of sizes for integer variables. Languages may support short,
37 integer, long, and even big integers. Interfacing with protocols, device drivers, embedded systems, low level
38 graphics or other external constructs may require each bit or set of bits to have a particular meaning. Those bit
39 sets may or may not coincide with the sizes supported by a particular language implementation. When they do not,
40 it is common practice to pack all of the bits into one word. Masking and shifting of the word using powers of two to
41 pick out individual bits or using sums of powers of 2 to pick out subsets of bits (e.g., using $28=2^2+2^3+2^4$ to create
42 the mask 11100 and then shifting 2 bits) provides a way of extracting those bits. Knowledge of the underlying bit
43 storage is usually not necessary to accomplish simple extractions such as these. Problems can arise when
44 programmers mix their techniques to reference the bits or output the bits. Problems can arise when programmers

1 mix arithmetic and logical operations to reference the bits or output the bits. The storage ordering of the bits may
2 not be what the programmer expects.

3 Packing of bits in an integer is not inherently problematic. However, an understanding of the intricacies of bit level
4 programming must be known. Some computers or other devices store the bits left to right while others store them
5 right to left. The type of storage can cause problems when interfacing with external devices that expect the bits in
6 the opposite order. One problem arises when assumptions are made when interfacing with external constructs and
7 the ordering of the bits or words are not the same as the receiving entity. Programmers may inadvertently use the
8 sign bit in a bit field and then may not be aware that an arithmetic shift (sign extension) is being performed when
9 right shifting causing the sign bit to be extended into other fields. Alternatively, a left shift can cause the sign bit to
10 be one. Bit manipulations can also be problematic when the manipulations are done on binary encoded records
11 that span multiple words. The storage and ordering of the bits must be considered when doing bitwise operations
12 across multiple words as bytes may be stored in big endian or little endian format.

13 **6.12.4 Applicable language characteristics**

14 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 15 • Languages that allow bit manipulations

16 **6.12.5 Avoiding the vulnerability or mitigating its effects**

17 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 18 • Any assumption about bit ordering should be explicitly documented.
- 19 • The way bit ordering is done on the host system and on the systems with which the bit manipulations will
20 be interfaced should be understood.
- 21 • Bit fields should be used in languages that support them.
- 22 • Bit operators should not be used on signed operands.
- 23 • Localize and document the code associated with explicit manipulation of bits and bit fields.

24 **6.12.6 Implications for standardization**

- 25 • For languages that are commonly used for bit manipulations, an API for bit manipulations that is
26 independent of word size and machine instruction set should be defined and standardized.

27 **6.12.7 Bibliography**

28 [1] Hogaboom, Richard, *A Generic API Bit Manipulation in C*, Embedded Systems Programming, Vol 12, No 7, July
29 1999 <http://www.embedded.com/1999/9907/9907feat2.htm>

30 **6.13 Floating-point Arithmetic [PLF]**

31 **6.13.1 Description of application vulnerability**

32 Most real numbers cannot be represented exactly in a computer. To represent real numbers, most computers use
33 ANSI/IEEE Std 754. The bit representation for a floating-point number can vary from compiler to compiler and on
34 different platforms. Relying on a particular representation can cause problems when a different compiler is used or
35 the code is reused on another platform. Regardless of the representation, many real numbers can only be
36 approximated since representing the real number using a binary representation would require an endlessly
37 repeating string of bits or more binary digits than are available for representation. Therefore it should be assumed
38 that a floating-point number is only an approximation, even though it may be an extremely good one. Floating-point
39 representation of a real number or a conversion to floating-point can cause surprising results and unexpected
40 consequences to those unaccustomed to the idiosyncrasies of floating-point arithmetic.

1 that two values can be considered equal. Note that if the two values are very large, the “small enough”
2 difference can be a very large number.

- 3 • Use library functions with known numerical characteristics whenever possible.
- 4 • Unless the use of floating-point is simple an expert in numerical analysis should check the stability and
5 accuracy of the algorithm employed.
- 6 • Avoid the use of a floating-point variable as a loop counter. If necessary to use a floating-point value as a
7 loop control, use inequality to determine the loop control (i.e. <, <=, > or >=).
- 8 • Understand the floating-point format used to represent the floating-point numbers. This will provide some
9 understanding of the underlying idiosyncrasies of floating-point arithmetic.
- 10 • Manipulating the bit representation of a floating-point number should not be done except with built-in
11 language operators and functions that are designed to extract the mantissa and exponent.
- 12 • Do not use floating-point for exact values such as monetary amounts. Use floating-point only when
13 necessary such as for fundamentally inexact values such as measurements.
- 14 • Consider the use of decimal floating-point facilities when available.

15 6.13.6 Implications for standardization

16 In future standardization activates, the following items should be considered:

- 17 • Languages that do not already adhere to or only adhere to a subset of ANSI/IEEE 754 should consider
18 adhering completely to the standard. Examples of standardization that should be considered:
 - 19 ○ C, which predates ANSI/IEEE 745 and currently has is as optional in the 9899:1999 version,
20 should consider requiring ANSI/IEEE 754 for floating-point arithmetic.
 - 21 ○ Java should consider fully adhering to ANSI/IEEE 754 instead of a subset.
- 22 • Languages should consider providing a means to generate diagnostics for code that attempts to test
23 equality of two floating point values.
- 24 • Languages should consider standardizing their data type to ISO/IEC 10967-1:1994 and ISO/IEC 10967-
25 2:2001.

26 6.13.7 Bibliography

- 27 [1] Goldberg, David, *What Every Computer Scientist Should Know About Floating-Point Arithmetic*, ACM
28 Computing Surveys, vol 23, issue 1 (March 1991), ISSN 0360-0300, pp 5-48.
- 29 [2] IEEE Standards Committee 754. IEEE Standard for Binary Floating-Point Arithmetic, ANSI/IEEE Standard 754-
30 1985. Institute of Electrical and Electronics Engineers, New York, 1985.
- 31 [3] Bo Einarsson, ed. Accuracy and Reliability in Scientific Computing, SIAM, July 2005
32 <http://www.nsc.liu.se/wg25/book>
- 33 [4] GAO Report, *Patriot Missile Defense: Software Problem Led to System Failure at Dhahran, Saudi Arabia*, B-
34 247094, Feb. 4, 1992, <http://archive.gao.gov/t2pbat6/145960.pdf>
- 35 [5] Robert Skeel, *Roundoff Error Cripples Patriot Missile*, SIAM News, Volume 25, Number 4, July 1992, page 11,
36 <http://www.siam.org/siamnews/general/patriot.htm>
- 37 [6] *ARIANE 5: Flight 501 Failure*, Report by the Inquiry Board, July 19, 1996 <http://esamultimedia.esa.int/docs/esa-x-1819eng.pdf> (Press release is at: http://www.esa.int/esaCP/Pr_33_1996_p_EN.html and there is a link to the
38 report at the bottom of the press release)
- 39 [7] ISO/IEC 10967-3:2006. ISO/IEC Information technology – Language independent arithmetic – Part 3: Complex
40 integer and floating-point arithmetic and complex elementary numerical functions, ISO/IEC Standard 10967-3:2006,
41 International Organization for Standardization/International Electrotechnical Commission, May 2006
42 http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=37994
43

44 6.14 Enumerator Issues [CCB]

45 6.14.1 Description of application vulnerability

46 Enumerations are a finite list of named entities that contain a fixed mapping from a set of names to a set of integral
47 values (called the representation) and an order between the members of the set. In some languages there are no
48 other operations available except order, equality, first, last, previous, and next; in others the full underlying
49 representation operators are available, such as integer “+” and “-” and bit-wise operations.

1 Most languages that provide enumeration types also provide mechanisms to set non-default representations. If
 2 these mechanisms do not enforce whole-type operations and check for conflicts then some members of the set
 3 may not be properly specified or may have the wrong mappings. If the value-setting mechanisms are positional
 4 only, then there is a risk that improper counts or changes in relative order will result in an incorrect mapping.

5 For arrays indexed by enumerations with non-default representations, there is a risk of structures with holes, and if
 6 those indexes can be manipulated numerically, there is a risk of out-of-bound accesses of these arrays.

7 Most of these errors can be readily detected by static analysis tools with appropriate coding standards, restrictions
 8 and annotations. Similarly mismatches in enumeration value specification can be detected statically. Without such
 9 rules, errors in the use of enumeration types are computationally hard to detect statically as well as being difficult to
 10 detect by human review.

11 **6.14.2 Cross reference**

12 JSF AV Rule: 145
 13 MISRA C 2004: 9.3, 9.2, and 9.3
 14 MISRA C++ 2008: 8-5-3
 15 CERT/CC guidelines: INT09-C
 16 Holtzmann rule 6.

17 **6.14.3 Mechanism of failure**

18 As a program is developed and maintained the list of items in an enumeration often changes in three basic ways:
 19 new elements are added to the list; order between the members of the set often changes; and representation (the
 20 map of values of the items) change. Expressions that depend on the full set or specific relationships between
 21 elements of the set can create value errors that could result in wrong results or in unbounded behaviours if used as
 22 array indices.

23 Improperly mapped representations can result in some enumeration values being unreachable, or may create
 24 "holes" in the representation where undefinable values can be propagated.

25 If arrays are indexed by enumerations containing nondefault representations, some implementations may leave
 26 space for values that are unreachable using the enumeration, with a possibility of unnecessarily large memory
 27 allocations or a way to pass information undetected (hidden channel).

28 When enumerators are set and initialized explicitly and the language permits incomplete initializers, then changes
 29 to the order of enumerators or the addition or deletion of enumerators can result in the wrong values being
 30 assigned or default values being assigned improperly. Subsequent indexing or switch/case statements can result in
 31 illegal accesses and possibly unbounded behaviours.

32 **6.14.4 Applicable language Characteristics**

33 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 34 • Languages that do not require full coverage of an enumeration in a switch/case statement.
- 35 • Languages that provide a default choice in a switch/case statement.
- 36 • Languages that permit incomplete mappings between enumerator specification and value assignment, or
 37 that provide a positional-only mapping require additional static analysis tools and annotations to help
 38 identify the complete mapping of every literal to its value.
- 39 • Languages that provide a trivial mapping to a type such as integer require additional static analysis tools to
 40 prevent mixed type errors. They also cannot prevent illegal values from being placed into variables of such
 41 enumerator types. For example:

```
42     enum Directions {back, forward, stop};
43     enum Directions a = forward, b = stop, c = a+b;
```

- 1 In this example, `c` may have a value not defined by the enumeration, and any further use as that
 2 enumeration will lead to erroneous results.
- 3 • Some languages provide no enumeration capability, leaving it to the programmer to define named
 4 constants to represent the values and ranges.

5 **6.14.5 Avoiding the vulnerability or mitigating its effects**

6 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 7 • Use static analysis tools that will detect inappropriate use of enumerators, such as using them as integers
 8 or bit maps, and that detect enumeration definition expressions that are incomplete or incorrect. For
 9 languages with a complete enumeration abstraction this is the compiler.
- 10 • When a language requires full coverage of an enumeration in a switch/case statement, a default choice
 11 should not be provided. For languages that do not require full coverage, then a default choice with a
 12 possible error generation should be provided to ensure that there is full coverage.
- 13 • When named syntax is available for representation setting, coverage analysis can eliminate the order
 14 issues and the incomplete coverage issues as long as no default choice is given.

15 **6.14.6 Implications for standardization**

16 In future standardization activates, the following items should be considered:

- 17 • Languages that currently permit arithmetic and logical operations on enumeration types could provide a
 18 mechanism to ban such operations program-wide.
- 19 • Languages that provide automatic defaults or that do not enforce static matching between enumerator
 20 definitions and initialization expressions could provide a mechanism to enforce such matching.

21 **6.14.7 Bibliography**

22 [None]

23 **6.15 Numeric Conversion Errors [FLC]**

24 **6.15.1 Description of application vulnerability**

25 Certain contexts in various languages may require exact matches with respect to types [7]:

```
26     aVar := anExpression
27     value1 + value2
28     foo(arg1, arg2, arg3, ... , argN)
```

29 Type conversion seeks to follow these exact match rules while allowing programmers some flexibility in using
 30 values such as: structurally-equivalent types in a name-equivalent language, types whose value ranges may be
 31 distinct but intersect (for example, subranges), and distinct types with sensible/meaningful corresponding values
 32 (for example, integers and floats). Explicit conversions are called *type casts*. An implicit type conversion between
 33 compatible but not necessarily equivalent types is called *type coercion*.

34 Numeric conversions can lead to a loss of data, if the target representation is not capable of representing the
 35 original value. For example, converting from an integer type to a smaller integer type can result in truncation if the
 36 original value cannot be represented in the smaller size and converting a floating point to an integer can result in a
 37 loss of precision or an out-of-range value.

38 Type conversion errors can lead to erroneous data being generated, algorithms that fail to terminate, array bounds
 39 errors, and arbitrary program execution.

1 6.15.2 Cross reference

2 CWE:

3 192. Integer Coercion Error

4 MISRA C 2004: 10.1-10.6, 11.3-11.5, and 12.9

5 MISRA C++ 2008: 2-13-3, 5-0-3, 5-0-4, 5-0-5, 5-0-6, 5-0-7, 5-0-8, 5-0-9, 5-0-10, 5-2-5, 5-2-9, and 5-3-2

6 CERT/CC guidelines: FLP34-C, INT02-C, INT08-C, INT31-C, and INT35-C

7 6.15.3 Mechanism of failure

8 Numeric conversion errors results in data integrity issues, but they may also result in a number of safety and
9 security vulnerabilities.

10 Vulnerabilities typically occur when appropriate range checking is not performed, and unanticipated values are
11 encountered. These can result in safety issues, for example, then the Ariane 5 launcher failure occurred due to an
12 improperly handled conversion error resulting in the processor being shutdown [3].

13 Conversion errors can also result in security issues. An attacker may input a particular numeric value to exploit a
14 flaw in the program logic. The resulting erroneous value may then be used as an array index, a loop iterator, a
15 length, a size, state data, or in some other security critical manner. For example, a truncated integer value may be
16 used to allocate memory, while the actual length is used to copy information to the newly allocated memory,
17 resulting in a buffer overflow [6].

18 Numeric type conversion errors often lead to undefined states of execution resulting in infinite loops or crashes. In
19 some cases, integer type conversion errors can lead to exploitable buffer overflow conditions, resulting in the
20 execution of arbitrary code. Integer type conversion errors result in an incorrect value being stored for the variable
21 in question.

22 6.15.4 Applicable language characteristics

23 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 24 • Languages that perform implicit type conversion (coercion).
- 25 • Weakly typed languages that do not strictly enforce type rules.
- 26 • Languages that support logical, arithmetic, or circular shifts on integer values.
- 27 • Languages that do not generate exceptions on problematic conversions.

28 6.15.5 Avoiding the vulnerability or mitigating its effects

29 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 30 • The first line of defense against integer vulnerabilities should be range checking, either explicitly or through
31 strong typing. All integer values originating from a source that is not trusted should be validated for
32 correctness. However, it is difficult to guarantee that multiple input variables cannot be manipulated to
33 cause an error to occur in some operation somewhere in a program [6].
- 34 • An alternative or ancillary approach is to protect each operation. However, because of the large number of
35 integer operations that are susceptible to these problems and the number of checks required to prevent or
36 detect exceptional conditions, this approach can be prohibitively labor intensive and expensive to
37 implement.
- 38 • A language that generates exceptions on erroneous data conversions might be chosen. Design objects
39 and program flow such that multiple or complex casts are unnecessary. Ensure that any data type casting
40 that you must use is entirely understood to reduce the plausibility of error in use.
- 41 • The use of static analysis can often identify whether or not unacceptable numeric conversions will occur.

42 Verifiably in-range operations are often preferable to treating out of range values as an error condition because the
43 handling of these errors has been repeatedly shown to cause denial-of-service problems in actual applications.
44 Faced with a numeric conversion error, the underlying computer system may do one of two things: (a) signal some
45 sort of error condition, or (b) produce a numeric value that is within the range of representable values on that

1 system. The latter semantics may be preferable in some situations in that it allows the computation to proceed, thus
2 avoiding a denial-of-service attack. However, it raises the question of what numeric result to return to the user.

3 A recent innovation from ISO/IEC TR 24731-1 [8] is the definition of the `rsize_t` type for the C programming
4 language. Extremely large object sizes are frequently a sign that an object's size was calculated incorrectly. For
5 example, negative numbers appear as very large positive numbers when converted to an unsigned type like
6 `size_t`. Also, some implementations do not support objects as large as the maximum value that can be
7 represented by type `size_t`. For these reasons, it is sometimes beneficial to restrict the range of object sizes to
8 detect programming errors. For implementations targeting machines with large address spaces, it is recommended
9 that `RSIZE_MAX` be defined as the smaller of the size of the largest object supported or $(SIZE_MAX \gg 1)$, even
10 if this limit is smaller than the size of some legitimate, but very large, objects. Implementations targeting machines
11 with small address spaces may wish to define `RSIZE_MAX` as `SIZE_MAX`, which means that there is no object size
12 that is considered a runtime-constraint violation.

13 6.15.6 Implications for standardization

14 In future standardization activities, the following items should be considered:

- 15 • Languages should consider providing means similar to the ISO/IEC TR 24731-1 definition of `rsize_t` type
- 16 for C in order to restrict object sizes so as to expose programming errors.
- 17 • Languages should consider making all type conversions explicit or at least generating warnings for implicit
- 18 conversions where loss of data might occur.

19 6.15.7 Bibliography

- 20 [1] CERT. *CERT C Secure Coding Standard*. <https://www.securecoding.cert.org/confluence/x/HQE> (2007).
- 21 [2] CERT. *CERT C++ Secure Coding Standard*. <https://www.securecoding.cert.org/confluence/x/fQI> (2007).
- 22 [3] Lions, J. L. *ARIANE 5 Flight 501 Failure Report*². Paris, France: European Space Agency (ESA) & National
- 23 Center for Space Study (CNES) Inquiry Board, July 1996.
- 24 [4] Hatton 2003
- 25 [5] MISRA Limited. "*MISRA C*²: 2004 Guidelines for the Use of the C Language in Critical Systems." Warwickshire,
- 26 UK: MIRA Limited, October 2004 (ISBN 095241564X).
- 27 [6] Seacord, R. *Secure Coding in C and C++*. Boston, MA: Addison-Wesley, 2005. See
- 28 <http://www.cert.org/books/secure-coding>² for news and errata.
- 29 [7] John David N. Dionisio. Type Checking. <http://myweb.lmu.edu/dondi/share/pl/type-checking-v02.pdf>
- 30 [8] ISO/IEC TR 24731-1. *Extensions to the C Library, — Part 1: Bounds-checking interfaces*. Geneva, Switzerland:
- 31 International Organization for Standardization, April 2006.

32 6.16 String Termination [CJM]

33 6.16.1 Description of application vulnerability

34 Some programming languages use a termination character to indicate the end of a string. Relying on the
35 occurrence of the string termination character without verification can lead to either exploitation or unexpected
36 behaviour.

37 6.16.2 Cross reference

38 CWE:
39 170. Improper Null Termination
40 CERT/CC guidelines: STR03-C, STR31-C, STR32-C, and STR36-C

41 6.16.3 Mechanism of failure

42 String termination errors occur when the termination character is solely relied upon to stop processing on the string
43 and the termination character is not present. Continued processing on the string can cause an error or potentially

1 be exploited as a buffer overflow. This may occur as a result of a programmer making an assumption that a string
2 that is passed as input or generated by a library contains a string termination character when it does not.

3 Programmers may forget to allocate space for the string termination character and expect to be able to store an *n*
4 length character string in an array that is *n* characters long. Doing so may work in some instances depending on
5 what is stored after the array in memory, but it may fail or be exploited at some point.

6 **6.16.4 Applicable language characteristics**

7 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 8 • Languages that use a termination character to indicate the end of a string.
- 9 • Languages that do not do bounds checking when accessing a string or array.

10

11 **6.16.5 Avoiding the vulnerability or mitigating its effects**

12 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 13 • Do not rely solely on the string termination character.
- 14 • Use library calls that do not rely on string termination characters such as `strncpy` instead of `strcpy` in
15 the standard C library.

16 **6.16.6 Implications for standardization**

17 In future standardization activities, the following items should be considered:

- 18 • Eliminating library calls that make assumptions about string termination characters.
- 19 • Checking bounds when an array or string is accessed.
- 20 • Specifying a string construct that does not need a string termination character.

21 **6.16.7 Bibliography**

22 [None]

23 **6.17 Boundary Beginning Violation [XYX]**

24 **6.17.1 Description of application vulnerability**

25 A buffer underwrite condition occurs when an array is indexed outside its lower bounds, or pointer arithmetic results
26 in an access to storage that occurs before the beginning of the intended object.

27 **6.17.2 Cross reference**

28 CWE:

29 124. Boundary Beginning Violation ('Buffer Underwrite')

30 129. Unchecked Array Indexing

31 JSF AV Rule: 25

32 MISRA C 2004: 21.1

33 MISRA C++ 2008: 5-0-15 to 5-0-18

34 CERT/CC guidelines: ARR30-C, ARR32-C, and ARR38-C

35 **6.17.3 Mechanism of failure**

36 There are several kinds of failures (in all cases an exception may be raised if the accessed location is outside of
37 some permitted range):

- 1 • A read access will return a value that has no relationship to the intended value, e.g., the value of another
2 variable or uninitialized storage.
- 3 • An out-of-bounds read access may be used to obtain information that is intended to be confidential.
- 4 • A write access will not result in the intended value being updated and may result in the value of an
5 unrelated object (that happens to exist at the given storage location) being modified.
- 6 • When an array has been allocated storage on the stack an out-of-bounds write access may modify internal
7 runtime housekeeping information (e.g., a functions return address) which might change a programs
8 control flow.

9 **6.17.4 Applicable language characteristics**

10 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 11 • Languages that do not detect and prevent an array being accessed outside of its declared bounds (either
12 by means of an index or by pointer arithmetic).
- 13 • Languages that do not automatically allocate storage when accessing an array element for which storage
14 has not already been allocated.

15 **6.17.5 Avoiding the vulnerability or mitigating its effects**

16 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:.

- 17 • Use of implementation provided functionality to automatically check array element accesses and prevent
18 out-of-bounds accesses.
- 19 • Use of static analysis to verify that all array accesses are within the permitted bounds. Such analysis may
20 require that source code contain certain kinds of information, e.g., that the bounds of all declared arrays be
21 explicitly specified, or that pre- and post-conditions be specified.
- 22 • Sanity checks should be performed on all calculated expressions used as an array index or for pointer
23 arithmetic.

24 Some guideline documents recommend only using variables having an unsigned type when indexing an array, on
25 the basis that an unsigned type can never be negative. This recommendation simply converts an indexing
26 underflow to an indexing overflow because the value of the variable will wrap to a large positive value rather than a
27 negative one. Also some language support arrays whose lower bound is greater than zero, so an index can be
28 positive and be less than the lower bound.

29 In the past the implementation of array bound checking has sometimes incurred what has been considered to be a
30 high runtime overhead (often because unnecessary checks were performed). It is now practical for translators to
31 perform sophisticated analysis that significantly reduces the runtime overhead (because runtime checks are only
32 made when it cannot be shown statically that no bound violations can occur).

33 **6.17.6 Implications for standardization**

34 In future standardization activates, the following items should be considered:

- 35 • Languages that use pointer types should consider specifying a standardized feature for a pointer type that
36 would enable array bounds checking, if such a pointer type is not already in the standard.

37 **6.17.7 Bibliography**

38 [None]

1 **6.18 Unchecked Array Indexing [XYZ]**

2 **6.18.1 Description of application vulnerability**

3 Unchecked array indexing occurs when a value is used as an index into an array without checking that it falls within
4 the acceptable index range.

5 **6.18.2 Cross reference**

6 CWE:

7 129. Unchecked Array Indexing

8 JSF AV Rules: 164 and 15

9 MISRA C 2004: 21.1

10 MISRA C++ 2008: 5-0-15 to 5-0-18

11 CERT/CC guidelines: ARR30-C, ARR32-C, ARR33-C, and ARR38-C

12 **6.18.3 Mechanism of failure**

13 A single fault could allow both an overflow and underflow of the array index. An index overflow exploit might use
14 buffer overflow techniques, but this can often be exploited without having to provide "large inputs." Array index
15 overflows can also trigger out-of-bounds read operations, or operations on the wrong objects; i.e., "buffer
16 overflows" are not always the result. Unchecked array indexing, depending on its instantiation, can be responsible
17 for any number of related issues. Most prominent of these possible flaws is the buffer overflow condition. Due to
18 this fact, consequences range from denial of service, and data corruption, to arbitrary code execution. The most
19 common condition situation leading to unchecked array indexing is the use of loop index variables as buffer
20 indexes. If the end condition for the loop is subject to a flaw, the index can grow or shrink unbounded, therefore
21 causing a buffer overflow or underflow. Another common situation leading to this condition is the use of a function's
22 return value, or the resulting value of a calculation directly as an index in to a buffer. Unchecked array indexing can
23 result in the corruption of relevant memory and perhaps instructions, lead to the program halting, if the values are
24 outside of the valid memory area. If the memory corrupted is data, rather than instructions, the system might
25 continue to function with improper values. If the corrupted memory can be effectively controlled, it may be possible
26 to execute arbitrary code, as with a standard buffer overflow.

27 Language implementations might or might not statically detect out of bound access and generate a compile-time
28 diagnostic. At runtime the implementation might or might not detect the out-of-bounds access and provide a
29 notification at runtime. The notification might be treatable by the program or it might not be. Accesses might violate
30 the bounds of the entire array or violate the bounds of a particular extent. It is possible that the former is checked
31 and detected by the implementation while the latter is not. The information needed to detect the violation might or
32 might not be available depending on the context of use. (For example, passing an array to a subroutine via a
33 pointer might deprive the subroutine of information regarding the size of the array.)

34 Aside from bounds checking, some languages have ways of protecting against out-of-bounds accesses. Some
35 languages automatically extend the bounds of an array to accommodate accesses that might otherwise have been
36 beyond the bounds. However, this may or may not match the programmer's intent and can mask errors. Some
37 languages provide for whole array operations that may obviate the need to access individual elements thus
38 preventing unchecked array accesses.

39 **6.18.4 Applicable language characteristics**

40 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 41 • Languages that do not automatically bounds check array accesses.
- 42 • Languages that do not automatically extend the bounds of an array to accommodate array accesses.

43 **6.18.5 Avoiding the vulnerability or mitigating its effects**

44 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 1 • Include sanity checks to ensure the validity of any values used as index variables.
- 2 • The choice could be made to use a language that is not susceptible to these issues.
- 3 • When available, use whole array operations whenever possible.

4 **6.18.6 Implications for standardization**

5 In future standardization activities, the following items should be considered:

- 6 • Languages should consider providing compiler switches or other tools to check the size and bounds of
7 arrays and their extents that are statically determinable.
- 8 • Languages should consider providing whole array operations that may obviate the need to access
9 individual elements.
- 10 • Languages should consider the capability to generate exceptions or automatically extend the bounds of an
11 array to accommodate accesses that might otherwise have been beyond the bounds.

12 **6.18.7 Bibliography**

13 [None]

14 **6.19 Unchecked Array Copying [XYW]**

15 **6.19.1 Description of application vulnerability**

16 A buffer overflow occurs when some number of bytes (or other units of storage) is copied from one buffer to
17 another and the amount being copied is greater than is allocated for the destination buffer.

18 **6.19.2 Cross reference**

19 CWE:
20 121. Stack-based Buffer Overflow
21 JSF AV Rule: 15
22 MISRA C 2004: 21.1
23 MISRA C++ 2008: 5-0-15 to 5-0-18
24 CERT/CC guidelines: ARR33-C and STR31-C

25 **6.19.3 Mechanism of failure**

26 Many languages and some third party libraries provide functions that efficiently copy the contents of one area of
27 storage to another area of storage. Most of these libraries do not perform any checks to ensure that the copied
28 from/to storage area is large enough to accommodate the amount of data being copied.

29 The arguments to these library functions include the addresses of the contents of the two storage areas and the
30 number of bytes (or some other measure) to copy. Passing the appropriate combination of incorrect start
31 addresses or number of bytes to copy makes it possible to read or write outside of the storage allocated to the
32 source/destination area. When passed incorrect parameters the library function performs one or more unchecked
33 array index accesses, as described in [XYZ] Unchecked Array Indexing.

34 **6.19.4 Applicable language characteristics**

35 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 36 • Languages that contain Standard library functions for performing bulk copying of storage areas.
- 37 • The same range of languages having the characteristics listed in [XYZ] Unchecked Array Indexing.

1 **6.19.5 Avoiding the vulnerability or mitigating its effects**

2 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 3 • Only use library functions that perform checks on the arguments to ensure no buffer overrun can occur
- 4 (perhaps by writing a wrapper for the Standard provided functions). Perform checks on the argument
- 5 expressions prior to calling the Standard library function to ensure that no buffer overrun will occur.
- 6 • Use static analysis to verify that the appropriate library functions are only called with arguments that do not
- 7 result in a buffer overrun. Such analysis may require that source code contain certain kinds of information,
- 8 e.g., that the bounds of all declared arrays be explicitly specified, or that pre- and post-conditions be
- 9 specified as annotations or language constructs.

10 **6.19.6 Implications for standardization**

11 In future standardization activities, the following items should be considered:

- 12 • Languages should consider only providing libraries that perform checks on the parameters to ensure that
- 13 no buffer overrun can occur.
- 14 • Languages should consider providing full array assignment.

15 **6.19.7 Bibliography**

16 [None]

17 **6.20 Buffer Overflow [XZB]**

18 **6.20.1 Description of application vulnerability**

19 A buffer overflow arises when, due to unchecked array indexing or unchecked array copying, storage outside the
 20 buffer is accessed. Usually overflows describe the situation where such storage is then written. Depending on
 21 where the buffer is located, logically unrelated portions of the stack or the heap could be modified maliciously or
 22 unintentionally. Usually, buffer overflows describe accesses to contiguous memory beyond the end of the buffer
 23 data, as may arise when arrays are copied without length checks. However, accessing before the beginning of the
 24 buffer data is equally possible, dangerous and maliciously exploitable.

25 **6.20.2 Cross reference**

26 CWE:
 27 122. Heap-based Buffer Overflow
 28 JSF AV Rule: 15
 29 MISRA C 2004: 21.1
 30 MISRA C++ 2008: 5-0-15 to 5-0-18
 31 CERT/CC guidelines: ARR33-C, STR31-C and MEM35-C

32 **6.20.3 Mechanism of failure**

33 Overwriting adjacent data (or data at arbitrarily computed locations) outside the area allocated for an array leads to
 34 value failures of the application. The program statements causing the buffer overflow are often difficult to find. But
 35 not only data storage can be corrupted. Buffer overflow may also inadvertently or even maliciously overwrite
 36 function pointers that may be in memory, pointing them to the attacker's code. Even in applications that do not
 37 explicitly use function pointers, the run-time will usually store function pointers in memory. For example, object
 38 methods in object-oriented languages are generally implemented using function pointers in data structures that are
 39 kept in memory. Since the consequence of a buffer overflow can be targeted to cause arbitrary code execution, this
 40 vulnerability can be used to subvert any security service.

1 **6.20.4 Applicable language characteristics**

2 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 3 • Copying of arrays can be done without an automatic length check ensuring that source and target locations
4 are of the same size.
- 5 • Indexing of array elements can be done without an automatic check that the indexing is within the bounds
6 of the array.
- 7 • Accesses might violate the physical bounds of the entire array or violate the logical bounds of a particular
8 extent. The vulnerability is somewhat mitigated, if the former violation is checked for and detected by the
9 implementation although the latter is not.
- 10 • Languages that provide bounds checking but permit the check to be suppressed.
- 11 • The bounds of an array are not automatically extended to accommodate accesses that might otherwise
12 have been beyond the bounds. (This may or may not match the programmer's intent.)

13 **6.20.5 Avoiding the vulnerability or mitigating its effects**

14 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 15 • Use a language or compiler that performs automatic bounds checking on elements accesses and
16 automatic length checking on copying entire arrays.
- 17 • Use an abstraction library to add checks on top of library functions that copy arrays without length checks.
- 18 • Checks that prevent overflows can be disabled in some languages to increase performance. This option
19 should be used rarely.
- 20 • Implementation-defined checks that prevent overflows can be enabled in some languages that do not
21 require such checks. This option should be used whenever feasible.

22 **6.20.6 Implications for standardization**

23 In future standardization activities, the following items should be considered:

- 24 • Languages should provide safe copying of arrays as built-in operation.
- 25 • Languages should consider only providing array copy routines in libraries that perform checks on the
26 parameters to ensure that no buffer overrun can occur.
- 27 • Languages should perform automatic bounds checking on accesses to array elements. This capability may
28 need to be disabled at times for performance reasons.

29 **6.20.7 Bibliography**

30 [None]

31 **6.21 Pointer Casting and Pointer Type Changes [HFC]**

32 **6.21.1 Description of application vulnerability**

33 The code produced for access via a data or function pointer requires that the type of the pointer is appropriate for
34 the data or function being accessed. Otherwise undefined behaviour can occur. Specifically, "access via a data
35 pointer" is defined to be "fetch or store indirectly through that pointer" and "access via a function pointer" is defined
36 to be "invocation indirectly through that pointer." The detailed requirements for what is meant by the "appropriate"
37 type may vary among languages.

38 Even if the type of the pointer is appropriate for the access, erroneous pointer operations can still cause a fault.

39 **6.21.2 Cross reference**

40 CWE

1 136. Type Errors
 2 188. Reliance on Data/Memory Layout
 3 JSF AV Rules: 182 and 183
 4 MISRA C 2004: 11.1, 11.2, 11.3, 11.4, and 11.5
 5 MISRA C++ 2008: 5-2-2 to 5-2-9
 6 CERT/CC guidelines: INT11-C and EXP36-A
 7 Hatton 13: Pointer casts
 8

9 **6.21.3 Mechanism of failure**

10 If a pointer's type is not appropriate for the data or function being accessed, data can be corrupted or privacy can
 11 be broken by inappropriate read or write operation using the indirection provided by the pointer value. With a
 12 suitable type definition, large portions of memory can be maliciously or accidentally modified or read. Such
 13 modification of data objects will generally lead to value faults of the application. Modification of code elements such
 14 as function pointers or internal data structures for the support of object-orientation can affect control flow. This can
 15 make the code susceptible to targeted attacks by causing invocation via a pointer-to-function that has been
 16 manipulated to point to an attacker's payload.

17 **6.21.4 Applicable language characteristics**

18 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 19 • Pointers (and/or references) can be converted to different pointer types.
- 20 • Pointers to functions can be converted to pointers to data.

21 **6.21.5 Avoiding the vulnerability or mitigating its effects**

22 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 23 • Treat the compiler's pointer-conversion warnings as serious errors.
- 24 • Adopt programming guidelines (preferably augmented by static analysis) that restrict pointer conversions.
 25 For example, consider the rules itemized above from JSF AV, CERT/CC, Hatton, or MISRA C.
- 26 • Other means of assurance might include proofs of correctness, analysis with tools, verification techniques,
 27 etc.

28 **6.21.6 Implications for standardization**

29 In future standardization activities, the following items should be considered:

- 30 • Languages should consider creating a mode that provides a runtime check of the validity of all accessed
 31 objects before the object is read, written or executed.

32 **6.21.7 Bibliography**

33 Hatton 13: Pointer casts

34 **6.22 Pointer Arithmetic [RVG]**

35 **6.22.1 Description of application vulnerability**

36 Using pointer arithmetic incorrectly can result in addressing arbitrary locations, which in turn can cause a program
 37 to behave in unexpected ways.

1 **6.22.2 Cross reference**

- 2 JSF AV Rule: 215
- 3 MISRA C 2004: 17.1, 17.2, 17.3, and 17.4
- 4 MISRA C++ 2008: 5-0-15 to 5-0-18
- 5 CERT/CC guidelines: EXP08-C

6 **6.22.3 Mechanism of failure**

7 Pointer arithmetic used incorrectly can produce:

- 8 • Addressing arbitrary memory locations, including buffer underflow and overflow.
- 9 • Arbitrary code execution.
- 10 • Addressing memory outside the range of the program.

11 **6.22.4 Applicable language characteristics**

12 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 13 • Languages that allow pointer arithmetic.

14 **6.22.5 Avoiding the vulnerability or mitigating its effects**

15 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 16 • Use pointer arithmetic only for indexing objects defined as arrays.
- 17 • Prefer indexing for accessing array elements rather than using pointer arithmetic.
- 18 • Limit pointer arithmetic calculations to the addition and subtraction of integers.

19 **6.22.6 Implications for standardization**

20 [None]

21 **6.22.7 Bibliography**

22 [None]

23 **6.23 Null Pointer Dereference [XYH]**

24 **6.23.1 Description of application vulnerability**

25 A null-pointer dereference takes place when a pointer with a value of `NULL` is used as though it pointed to a valid
26 memory location. This is a special case of accessing storage via an invalid pointer.

27 **6.23.2 Cross reference**

- 28 CWE:
- 29 476. NULL Pointer Dereference
- 30 JSF AV Rule 174
- 31 CERT/CC guidelines: EXP34-C

1 **6.23.3 Mechanism of failure**

2 When a pointer with a value of `NULL` is used as though it pointed to a valid memory location, then a null-pointer
3 dereference is said to take place. This can result in a segmentation fault, unhandled exception, or accessing
4 unanticipated memory locations.

5 **6.23.4 Applicable language characteristics**

6 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 7 • Languages that permit the use of pointers and that do not check the validity of the location being accessed
8 prior to the access.
- 9 • Languages that allow the use of a `NULL` pointer.

10 **6.23.5 Avoiding the vulnerability or mitigating its effects**

11 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 12 • Before dereferencing a pointer, ensure it is not equal to `NULL`.

13 **6.23.6 Implications for standardization**

14 In future standardization activities, the following items should be considered:

- 15 • Pointers could be checked for a `NULL` value before performing the access. This could be implemented
16 through a compiler option.

17 **6.23.7 Bibliography**

18 [None]

19 **6.24 Dangling Reference to Heap [XYK]**

20 **6.24.1 Description of application vulnerability**

21 A dangling reference is a reference to an object whose lifetime has ended due to explicit deallocation or the stack
22 frame in which the object resided has been freed due to exiting the dynamic scope. The memory for the object may
23 be reused; therefore, any access through the dangling reference may affect an apparently arbitrary location of
24 memory, corrupting data or code.

25 This description concerns the former case, dangling references to the heap. The description of dangling references
26 to stack frames is [DCM]. In many languages references are called pointers; the issues are identical.

27 A notable special case of using a dangling reference is calling a deallocator, for example, `free()`, twice on the
28 same pointer value. Such a "Double Free" may corrupt internal data structures of the heap administration, leading
29 to faulty application behaviour (such as infinite loops within the allocator, returning the same memory repeatedly as
30 the result of distinct subsequent allocations, or deallocating memory legitimately allocated to another request since
31 the first `free()` call, to name but a few), or it may have no adverse effects at all.

32 Memory corruption through the use of a dangling reference is among the most difficult of errors to locate.

33 With sufficient knowledge about the heap management scheme (often provided by the OS (Operating System) or
34 run-time system), use of dangling references is an exploitable vulnerability, since the dangling reference provides a
35 method with which to read and modify valid data in the designated memory locations after freed memory has been
36 re-allocated by subsequent allocations.

1 **6.24.2 Cross reference**

2 CWE:

3 415. Double Free (Note that Double Free (415) is a special case of Use After Free (416))

4 416. Use After Free

5 MISRA C 2004: 17.1-6

6 MISRA C++ 2008: 0-3-1, 7-5-1, 7-5-2, 7-5-3, and 18-4-1

7 CERT/CC guidelines: MEM01-C, MEM30-C, and MEM31.C

8 **6.24.3 Mechanism of failure**

9 The lifetime of an object is the portion of program execution during which storage is guaranteed to be reserved for
10 it. An object exists and retains its last-stored value throughout its lifetime. If an object is referred to outside of its
11 lifetime, the behaviour is undefined. Explicit deallocation of heap-allocated storage ends the lifetime of the object
12 residing at this memory location (as does leaving the dynamic scope of a declared variable). The value of a pointer
13 becomes indeterminate when the object it points to reaches the end of its lifetime. Such pointers are called
14 dangling references.

15 The use of dangling references to previously freed memory can have any number of adverse consequences —
16 ranging from the corruption of valid data to the execution of arbitrary code, depending on the instantiation and
17 timing of the deallocation causing all remaining copies of the reference to become dangling, of the system's reuse
18 of the freed memory, and of the subsequent usage of a dangling reference.

19 Like memory leaks and errors due to double de-allocation, the use of dangling references has two common and
20 sometimes overlapping causes:

- 21 • An error condition or other exceptional circumstances.
- 22 • Developer confusion over which part of the program is responsible for freeing the memory.

23 If a pointer to previously freed memory is used, it is possible that the referenced memory has been reallocated.
24 Therefore, assignment using the original pointer has the effect of changing the value of an unrelated variable. This
25 induces unexpected behaviour in the affected program. If the newly allocated data happens to hold a class
26 description, in an object-oriented language for example, various function pointers may be scattered within the heap
27 data. If one of these function pointers is overwritten with an address of malicious code, execution of arbitrary code
28 can be achieved.

29 **6.24.4 Applicable language characteristics**

30 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 31 • Languages that permit the use of pointers and that permit explicit deallocation by the developer or provide
32 for alternative means to reallocate memory still pointed to by some pointer value.

33 **6.24.5 Avoiding the vulnerability or mitigating its effects**

34 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 35 • Languages that permit definitions of constructs that can be parameterized without enforcing the
36 consistency of the use of the parameters at compile time.
- 37 • Use an implementation that checks whether a pointer is used that designates a memory location that has
38 already been freed.
- 39 • Use a coding style that does not permit deallocation.
- 40 • In complicated error conditions, be sure that clean-up routines respect the state of allocation properly. If the
41 language is object-oriented, ensure that object destructors delete each chunk of memory only once.
42 Ensuring that all pointers are set to NULL once the memory they point to have been freed can be an
43 effective strategy. The utilization of multiple or complex data structures may lower the usefulness of this
44 strategy.

- 1 • Use a static analysis tool that is capable of detecting some situations when a pointer is used after the
2 storage it refers to is no longer a pointer to valid memory location.
3 • Allocating and freeing memory in different modules and levels of abstraction burdens the programmer with
4 tracking the lifetime of that block of memory. This may cause confusion regarding when and if a block of
5 memory has been allocated or freed, leading to programming defects such as double-free vulnerabilities,
6 accessing freed memory, or dereferencing NULL pointers or pointers that are not initialized. To avoid
7 these situations, it is recommended that memory be allocated and freed at the same level of abstraction,
8 and ideally in the same code module.

9 **6.24.6 Implications for standardization**

10 In future standardization activities, the following items should be considered:

- 11 • Implementations of the free function could tolerate multiple frees on the same reference/pointer or frees of
12 memory that was never allocated.
13 • Language specifiers should design generics in such a way that any attempt to instantiate a generic with
14 constructs that do not provide the required capabilities results in a compile-time error.
15 • For properties that cannot be checked at compile time, language specifiers should provide an assertion
16 mechanism for checking properties at run-time. It should be possible to inhibit assertion checking if
17 efficiency is a concern.
18 • A storage allocation interface should be provided that will allow the called function to set the pointer used to
19 NULL after the referenced storage is deallocated.

20 **6.24.7 Bibliography**

21 [None]

22 **6.25 Templates and Generics [SYM]**

23 **6.25.1 Description of application vulnerability**

24 Many languages provide a mechanism that allows objects and/or functions to be defined parameterized by type
25 and then instantiated for specific types. In C++ and related languages, these are referred to as “templates”, and in
26 Ada and Java, “generics”. To avoid having to keep writing ‘templates/generics’, in this section these will simply be
27 referred to collectively as generics.

28 Used well, generics can make code clearer, more predictable and easier to maintain. Used badly, they can have
29 the reverse effect, making code difficult to review and maintain, leading to the possibility of program error.

30 **6.25.2 Cross reference**

31 JSF AV Rules: 101, 102, 103, 104, and 105

32 MISRA C 2004: 14-7-2, 14-8-1, and 14-8-2

33 MISRA C++ 2008: 14-6-1, 14-6-2, 14-7-1 to 14-7-3, 14-8-1, and 14-8-2

34

35 **6.25.3 Mechanism of failure**

36 The value of generics comes from having a single piece of code that supports some behaviour in a type
37 independent manner. This simplifies development and maintenance of the code. It should also assist in the
38 understanding of the code during review and maintenance, by providing the same behaviour for all types with which
39 it is instantiated.

40 Problems arise when the use of a generic actually makes the code harder to understand during review and
41 maintenance, by not providing consistent behaviour.

1 In most cases, the generic definition will have to make assumptions about the types it can legally be instantiated
 2 with. For example, a sort function requires that the elements to be sorted can be copied and compared. If these
 3 assumptions are not met, the result is likely to be a compiler error. For example if the sort function is instantiated
 4 with a user defined type that doesn't have a relational operator. Where 'misuse' of a generic leads to a compiler
 5 error, this can be regarded as a development issue, and not a software vulnerability.

6 Confusion, and hence potential vulnerability, can arise where the instantiated code is apparently illegal, but doesn't
 7 result in a compiler error. For example, a generic class defines a set of members, a subset of which rely on a
 8 particular property of the instantiation type (e.g., a generic container class with a sort member function, only the
 9 sort function relies on the instantiating type having a defined relational operator). In some languages, such as C++,
 10 if the generic is instantiated with a type that doesn't meet all the requirements but the program never subsequently
 11 makes use of the subset of members that rely on the property of the instantiating type, the code will compile and
 12 execute (e.g., the generic container is instantiated with a user defined class that doesn't define a relational
 13 operator, but the program never calls the sort member of this instantiation). When the code is reviewed the generic
 14 class will appear to reference a member of the instantiating type that doesn't exist.

15 The problem as described in the two prior paragraphs can be reduced by a language feature (such as the *concepts*
 16 language feature being designed by the C++ committee).

17 Similar confusion can arise if the language permits specific elements of a generic to be explicitly defined, rather
 18 than using the common code, so that behaviour is not consistent for all instantiations. For example, for the same
 19 generic container class, the sort member normally sorts the elements of the container into ascending order. In
 20 languages such as C++, a 'special case' can be created for the instantiation of the generic with a particular type.
 21 For example, the sort member for a 'float' container may be explicitly defined to provide different behaviour, say
 22 sorting the elements into descending order. Specialization that doesn't affect the apparent behaviour of the
 23 instantiation is not an issue. Again, for C++, there are some irregularities in the semantics of arrays and pointers
 24 that can lead to the generic having different behaviour for different, but apparently very similar, types. In such
 25 cases, specialization can be used to enforce consistent behaviour.

26 **6.25.4 Applicable language characteristics**

27 This vulnerability is intended to be applicable to languages with the following characteristics:

- 28 • Languages that permit definitions of objects or functions to be parameterized by type, for later instantiation
 29 with specific types, e.g.:
 - 30 ○ templates: C++
 - 31 ○ generics: Ada, Java

32 **6.25.5 Avoiding the vulnerability or mitigating its effects**

33 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 34 • Document the properties of an instantiating type necessary for the generic to be valid.
- 35 • If an instantiating type has the required properties, the whole of the generic should be ensured to be valid,
 36 whether actually used in the program or not.
- 37 • Preferably avoid, but at least carefully document, any 'special cases' where the generic instantiated with a
 38 specific type doesn't behave as it does for other types.

39 **6.25.6 Implications for standardization**

40 In future standardization activities, the following items should be considered:

- 41 • Language specifiers should standardize on a common, uniform terminology to describe generics/templates
 42 so that programmers experienced in one language can reliably learn and refer to the type system of
 43 another language that has the same concept, but with a different name.
- 44 • Language specifiers should design generics in such a way that any attempt to instantiate a generic with
 45 constructs that do not provide the required capabilities results in a compile-time error.

- 1 • Language specifiers should provide an assertion mechanism for checking properties at run-time, for those
2 properties that cannot be checked at compile time.. It should be possible to inhibit assertion checking if
3 efficiency is a concern.

4 **6.25.7 Bibliography**

5 [None]

6 **6.26 Inheritance [RIP]**

7 **6.26.1 Description of application vulnerability**

8 Inheritance, the ability to create enhanced and/or restricted object classes based on existing object classes can
9 introduce a number of vulnerabilities, both inadvertent and malicious. Because Inheritance allows the overriding of
10 methods of the parent class and because object oriented systems are designed to separate and encapsulate code
11 and data, it can be difficult to determine where in the hierarchy an invoked method is actually defined. Also, since
12 an overriding method does not need to call the method in the parent class that has been overridden, essential
13 initialization and manipulation of class data may be bypassed. This can be especially dangerous during constructor
14 and destructor methods.

15 Languages that allow multiple inheritance add additional complexities to the resolution of method invocations.
16 Different object brokerage systems may resolve the method identity to different classes, based on how the
17 inheritance tree is traversed.

18 **6.26.2 Cross reference**

19 JSF AV Rules: 86 to 97

20 MISRA C++ 2008: 0-1-12, 8-3-1, 10-1-1 to 10-1-3, and 10-3-1 to 10-3-3

21 **6.26.3 Mechanism of failure**

22 The use of inheritance can lead to an exploitable application vulnerability or negatively impact system safety in
23 several ways:

- 24 • Execution of malicious redefinitions - this can occur through the insertion of a class into the class hierarchy
25 that overrides commonly called methods in the parent classes.
- 26 • Accidental redefinition - where a method is defined that inadvertently overrides a method that has already
27 been defined in a parent class.
- 28 • Accidental failure of redefinition - when a method is incorrectly named or the parameters are not defined
29 properly, and thus does not override a method in a parent class.
- 30 • Breaking of class invariants - this can be cause by redefining methods that initialize or validate class data
31 without including that initialization or validation in the overriding methods.

32 These vulnerabilities can increase dramatically as the complexity of the hierarchy increases, especially in the use
33 of multiple inheritance.

34 **6.26.4 Applicable language characteristics**

35 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 36 • Languages that allow single and multiple inheritances.

37 **6.26.5 Avoiding the vulnerability or mitigating its effects**

38 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 1 • Avoid the use of multiple inheritance whenever possible.
- 2 • Provide complete documentation of all encapsulated data, and how each method affects that data for each
- 3 object in the hierarchy.
- 4 • Inherit only from trusted sources, and, whenever possible, check the version of the parent classes during
- 5 compilation and/or initialization.
- 6 • Provide a method that provides versioning information for each class.

7 **6.26.6 Implications for standardization**

8 In future standardization activates, the following items should be considered:

- 9 • Language specification should include the definition of a common versioning method.
- 10 • Compilers should provide an option to report the class in which a resolved method resides.
- 11 • Runtime environments should provide a trace of all runtime method resolutions.

12 **6.26.7 Bibliography**

- 13 [1] P. V. Bhansali, A systematic approach to identifying a safe subset for safety-critical software, ACM SIGSOFT
14 Software Engineering Notes, v.28 n.4, July 2003
- 15 [2] Ghassan, A., & Alkadi, I. (2003). Application of a Revised DIT Metric to Redesign an OO Design. *Journal of*
16 *Object Technology* , 127-134.
- 17 [3] Subramanian, S., Tsai, W.-T., & Rayadurgam, S. (1998). Design Constraint Violation Detection in Safety-Critical
18 Systems. The 3rd IEEE International Symposium on High-Assurance Systems Engineering , 109 - 116.

20 **6.27 Initialization of Variables [LAV]**

21 **6.27.1 Description of application vulnerability**

22 Reading a variable that has not been assigned a value appropriate to its type can cause unpredictable execution in
23 the block that uses the value of the variable, and has the potential to export bad values to callers, or cause out-of-
24 bounds memory accesses.

25 Uninitialized variable usage is frequently not detected until after testing and often when the code in question is
26 delivered and in use, because happenstance will provide variables with adequate values (such as default data
27 settings or accidental left-over values) until some other change exposes the defect.

28 Variables that are declared during module construction (by a class constructor, instantiation, or elaboration) may
29 have alternate paths that can read values before they are set. This can happen in straight sequential code but is
30 more prevalent when concurrency or co-routines are present, with the same impacts described above.

31 Another vulnerability occurs when compound objects are initialized incompletely, as can happen when objects are
32 incrementally built, or fields are added under maintenance.

33 When possible and supported by the language, whole-structure initialization is preferable to field-by-field
34 initialization statements, and named association is preferable to positional, as it facilitates human review and is less
35 susceptible to failures under maintenance. For classes, the declaration and initialization may occur in separate
36 modules. In such cases it must be possible to show that every field that needs an initial value receives that value,
37 and to document ones that do not require initial values.

38 **6.27.2 Cross reference**

- 39 CWE:
40 457. Use of Uninitialized Variable
41 JSF AV Rules: 71, 143, and 147
42 MISRA C 2004: 9.1, 9.2, and 9.3

1 CERT/CC guidelines: DCL14-C and EXP33-C
 2 MISRA C++ 2008: 8-5-1

3 **6.27.3 Mechanism of failure**

4 Uninitialized objects may have illegal values, legal but wrong values, or legal and dangerous values. Wrong values
 5 could cause unbounded branches in conditionals or unbounded loop executions, or could simply cause wrong
 6 calculations and results.

7 There is a special case of pointers or access types. When such a type contains null values, a bound violation and
 8 hardware exception can result. When such a type contains plausible but meaningless values, random data reads
 9 and writes can collect erroneous data or can destroy data that is in use by another part of the program; when such
 10 a type is an access to a subprogram with a plausible (but wrong) value, then either a bad instruction trap may occur
 11 or a transfer to an unknown code fragment can occur. All of these scenarios can result in unbounded behaviours.

12 Uninitialized variables are difficult to identify and use for attackers, but can be arbitrarily dangerous in safety
 13 situations.

14 **6.27.4 Applicable Language Characteristics**

15 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 16 • Languages that permit variables to be read before they are assigned.

18 **6.27.5 Avoiding the vulnerability or mitigating its effects**

19 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 20 • The general problem of showing that all objects are initialized is intractable; hence developers must
 21 carefully structure programs to show that all variables are set before first read on every path throughout the
 22 subprogram. Where objects are visible from many modules, it is difficult to determine where the first read
 23 occurs, and identify a module that must set the value before that read. When concurrency, interrupts and
 24 coroutines are present, it becomes especially imperative to identify where early initialization occurs and to
 25 show that the correct order is set via program structure, not by timing, OS precedence, or chance.
- 26 • The simplest method is to initialize each object at elaboration time, or immediately after subprogram
 27 execution commences and before any branches. If the subprogram must commence with conditional
 28 statements, then the programmer is responsible to show that every variable declared and not initialized
 29 earlier is initialized on each branch.
- 30 • Applications can consider defining or reserving fields or portions of the object to only be set when fully
 31 initialized.
- 32 • It should be possible to use static analysis tools to show that all objects are set before use in certain
 33 specific cases, but as the general problem is intractable, programmers should keep initialization algorithms
 34 simple so that they can be analyzed.
- 35 • When declaring and initializing the object together, if the language does not require that the compiler
 36 statically verify that the declarative structure and the initialization structure match, use static analysis tools
 37 to help detect any mismatches.
- 38 • When setting compound objects, if the language provides mechanisms to set all components together, use
 39 those in preference to a sequence of initializations as this helps coverage analysis; otherwise use tools that
 40 perform such coverage analysis and document the initialization. Do not perform partial initializations unless
 41 there is no choice, and document any deviations from 100% initialization.
- 42 • Where default assignments of multiple components are performed, explicit declaration of the component
 43 names and/or ranges helps static analysis and identification of component changes during maintenance.
- 44 • Some languages have named assignments that can be used to build reviewable assignment structures that
 45 can be analyzed by the language processor for completeness. Languages with positional notation only can
 46 use comments and secondary tools to help show correct assignment.

1 **6.27.6 Implications for standardization**

2 In future standardization activates, the following items should be considered:

- 3 • Some languages have ways to determine if modules and regions are elaborated and initialized and to raise
4 exceptions if this does not occur. Languages that do not could consider adding such capabilities.
- 5 • Languages could consider setting aside fields in all objects to identify if initialization has occurred,
6 especially for security and safety domains.
- 7 • Languages that do not support whole-object initialization could consider adding this capability.

8 **6.27.7 Bibliography**

9 [None]

10 **6.28 Wrap-around Error [XYY]**

11 **6.28.1 Description of application vulnerability**

12 Wrap-around errors can occur whenever a value is incremented past the maximum or decremented past the
13 minimum value representable in its type and if so specified by the language semantics "wraps around" to either a
14 very small, negative, or undefined value. Using shift operations as a surrogate for multiply or divide may produce a
15 similar error.

16 **6.28.2 Cross reference**

17 CWE:
18 128. Wrap-around Error
19 JSF AV Rules: 164 and 15
20 MISRA C 2004: 10.1 to 10.6, 12.8 and 12.11
21 MISRA C++ 2008: 2-13-3, 5-0-3 to 5-0-10, and 5-19-1
22 CERT/CC guidelines: INT30-C, INT32-C, and INT34-C

23 **6.28.3 Mechanism of failure**

24 Due to how arithmetic is performed by computers, if a variable is incremented past the maximum value
25 representable in its type, the system may fail to provide an overflow indication to the program. One of the most
26 common processor behaviour is to "wrap" to a very large negative value, or set a condition flag for overflow or
27 underflow, or saturate at the largest representable value.

28 Shift operations may also produce values that cannot be easily predicted as a result of the different representations
29 of negative integers on various hardware, and, when treating signed quantities, of the differences in behaviour
30 between logical shifts and arithmetic shifts (the particular effect of filling with the sign bit).

31 Wrap-around often generates an unexpected negative value; this unexpected value may cause a loop to continue
32 for a long time (because the termination condition requires a value greater than some positive value) or an array
33 bounds violation. A wrap-around can sometimes trigger buffer overflows that can be used to execute arbitrary
34 code.

35 **6.28.4 Applicable language characteristics**

36 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 37 • Languages that do not trigger an exception condition when a wrap-around error occurs.
- 38 • Languages that do not fully specify the distinction between arithmetic and logical shifts.

1 **6.28.5 Avoiding the vulnerability or mitigating its effects**

2 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 3 • Determine applicable upper and lower bounds for the range of all variables and use language mechanisms
- 4 or static analysis to determine that values are confined to the proper range.
- 5 • Analyze the software using static analysis looking for unexpected consequences of arithmetic operations.
- 6 • Avoid using shift operations as a surrogate for multiplication and division. Most compilers will use the
- 7 correct operation in the appropriate fashion when it is applicable.

8 **6.28.6 Implications for standardization**

9 In future standardization activities, the following items should be considered:

- 10 • Language standards developers should consider providing facilities to specify either an error, a saturated
- 11 value, or a modulo result when numeric overflow occurs.

12 **6.28.7 Bibliography**

13 [None]

14 **6.29 Sign Extension Error [XZI]**

15 **6.29.1 Description of application vulnerability**

16 Extending a signed variable that holds a negative value may produce an incorrect result.

17 **6.29.2 Cross reference**

18 CWE:
 19 194. Incorrect Sign Extension
 20 MISRA C++ 2008: 5-0-4
 21 CERT/CC guidelines: INT13-C

22 **6.29.3 Mechanism of failure**

23 Converting a signed data type to a larger data type or pointer can cause unexpected behaviour due to the
 24 extension of the sign bit. A negative data element that is extended with an unsigned extension algorithm will
 25 produce an incorrect result. For instance, this can occur when a signed character is converted to a type short or a
 26 signed integer (32-bit) is converted to an integer type long (64-bit). Sign extension errors can lead to buffer
 27 overflows and other memory based problems. This can occur unexpectedly when moving software designed and
 28 tested on a 32-bit architecture to a 64-bit architecture computer.

29 **6.29.4 Applicable language characteristics**

30 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 31 • Languages that are weakly typed due to their lack of enforcement of type classifications and interactions.
- 32 • Languages that explicitly or implicitly allow applying unsigned extension operations to signed entities or
- 33 vice-versa.

34 **6.29.5 Avoiding the vulnerability or mitigating its effects**

35 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 1 • Use a sign extension library, standard function, or appropriate language-specific coding methods to extend
2 signed values.
- 3 • Use static analysis tools to help locate situations in which the conversion of variables might have
4 unintended consequences.

5 **6.29.6 Implications for standardization**

6 In future standardization activities, the following items should be considered:

- 7 • Language definitions should define implicit and explicit conversions in a way that prevents alteration of the
8 mathematical value beyond traditional rounding rules.

9 **6.29.7 Bibliography**

10 [None]

11 **6.30 Operator Precedence/Order of Evaluation [JCW]**

12 **6.30.1 Description of application vulnerability**

13 Each language provides rules of precedence and associativity, for each expression that operands bind to which
14 operators. These rules are also known as “grouping” or “binding”.

15 Experience and experimental evidence shows that developers can have incorrect beliefs about the relative
16 precedence of many binary operators. See, *Developer beliefs about binary operator precedence*. C Vu, 18(4):14-
17 21, August 2006

18 **6.30.2 Cross reference**

19 JSF AV Rules: 204 and 213

20 MISRA C 2004: 12.1, 12.2, 12.5, 12.6, 13.2, 19.10, 19.12, and 19.13

21 MISRA C++ 2008: 4-5-1, 4-5-2, 4-5-3, 5-0-1, 5-0-2, 5-2-1, 5-3-1, 16-0-6, 16-3-1, and 16-3-2

22 CERT/CC Guidelines: EXP00-C

23 **6.30.3 Mechanism of failure**

24 In C and C++, the bitwise operators (bitwise logical and bitwise shift) are sometimes thought of by the programmer
25 having similar precedence to arithmetic operations, so just as one might correctly write “ $x - 1 == 0$ ” (“ x minus
26 one is equal to zero”), a programmer might erroneously write “ $x \& 1 == 0$ ”, mentally thinking “ x anded-with 1 is
27 equal to zero”, whereas the operator precedence rules of C and C++ actually bind the expression as “compute
28 $1==0$, producing ‘false’ i.e. zero, then bitwise-and the result with x ”, producing (a constant) zero, contrary to the
29 programmer’s intent.

30 Examples from an opposite extreme can be found in programs written in APL, which is noteworthy for the absence
31 of any distinctions of precedence. One commonly made mistake is to write “ $a * b + c$ ”, intending to produce “ a
32 times b plus c ”, whereas APL’s uniform right-to-left associativity produces “ b plus c , times a ”.

33 **6.30.4 Applicable language characteristics**

34 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 35 • Languages whose precedence and associativity rules are sufficiently complex that developers do not
36 remember them.

1 **6.30.5 Avoiding the vulnerability or mitigating its effects**

2 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 3 • Adopt programming guidelines (preferably augmented by static analysis). For example, consider the rules
- 4 itemized above from JSF C++, CERT/CC or MISRA C.
- 5 • Use parentheses around binary operator combinations that are known to be a source of error (e.g., mixed
- 6 arithmetic/bitwise and bitwise/relational operator combinations).
- 7 • Break up complex expressions and use temporary variables to make the order clearer.

8 **6.30.6 Implications for standardization**

9 In future standardization activities, the following items should be considered:

- 10 • Language definitions should avoid providing precedence or a particular associativity for operators that are
- 11 not typically ordered with respect to one another in arithmetic, and instead require full parenthesization to
- 12 avoid misinterpretation.
- 13 • Language designers should develop an ISO standard containing one (or at most a few) accepted
- 14 precedence orders.

15 **6.30.7 Bibliography**

16 "Developer beliefs about binary operator precedence" by Derek M. Jones CVu Vol. 18, No. 4. (August 2006), pp.

17 14-21.

18 **6.31 Side-effects and Order of Evaluation [SAM]**

19 **6.31.1 Description of application vulnerability**

20 Some programming languages allow subexpressions to cause side-effects (such as assignment, increment, or

21 decrement). For example, some programming languages permit such side-effects, and if, within one expression

22 (such as "i = v[i++]"), two or more side-effects modify the same object, undefined behaviour results.

23 Some languages allow subexpressions to be evaluated in an unspecified ordering. If these subexpressions contain

24 side-effects, then the value of the full expression can be dependent upon the order of evaluation. Furthermore, the

25 objects that are modified by the side-effects can receive values that are dependent upon the order of evaluation.

26 If a program contains these unspecified or undefined behaviours, testing the program and seeing that it yields the

27 expected results may give the false impression that the expression will always yield the expected result.

28 **6.31.2 Cross reference**

29 JSF AV Rules: 157, 158, 166, 204, 204.1, and 213

30 MISRA C 2004: 12.1-12.5

31 MISRA C++ 2008: 5-0-1

32 CERT/CC Guidelines: EXP10-C, EXP30-C

33

34 **6.31.3 Mechanism of failure**

35 When subexpressions with side effects are used within an expression, the unspecified order of evaluation can

36 result in a program producing different results on different platforms, or even at different times on the same

37 platform. For example, consider

38
$$a = f(b) + g(b);$$

1 where f and g both modify b . If $f(b)$ is evaluated first, then the b used as a parameter to $g(b)$ may be a different
2 value than if $g(b)$ is performed first. Likewise, if $g(b)$ is performed first, $f(b)$ may be called with a different value
3 of b .

4 Other examples of unspecified order, or even undefined behaviour, can be manifested, such as

5 `a = f(i) + i++;`

6 or

7 `a[i++] = b[i++];`

8 Parentheses around expressions can assist in removing ambiguity about grouping, but the issues regarding side-
9 effects and order of evaluation are not changed by the presence of parentheses; consider

10 `j = i++ * i++;`

11 where even if parentheses are placed around the `i++` subexpressions, undefined behaviour still remains. (All
12 examples use the syntax of C or Java for brevity; the effects can be created in any language that allows functions
13 with side-effects in the places where C allows the increment operations.)

14 The unpredictable nature of the calculation means that the program cannot be tested adequately to any degree of
15 confidence. A knowledgeable attacker can take advantage of this characteristic to manipulate data values
16 triggering execution that was not anticipated by the developer.

17 **6.31.4 Applicable language characteristics**

18 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 19 • Languages that permit expressions to contain subexpressions with side effects.
- 20 • Languages whose subexpressions are computed in an unspecified ordering.

22 **6.31.5 Avoiding the vulnerability or mitigating its effects**

23 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 24 • Make use of one or more programming guidelines which (a) prohibit these unspecified or undefined
25 behaviours, and (b) can be enforced by static analysis. (See JSF AV and MISRA rules in Cross reference
26 section [SAM])
- 27 • Keep expressions simple. Complicated code is prone to error and difficult to maintain.

28 **6.31.6 Implications for standardization**

29 In future standardization activities, the following items should be considered:

- 30 • In developing new or revised languages, give consideration to language features that will eliminate or
31 mitigate this vulnerability, such as pure functions.

32 **6.31.7 Bibliography**

33 [None]

1 **6.32 Likely Incorrect Expression [KOA]**

2 **6.32.1 Description of application vulnerability**

3 Certain expressions are symptomatic of what is likely to be a mistake made by the programmer. The statement is
 4 not wrong, but it is unlikely to be right. The statement may have no effect and effectively is a null statement or may
 5 introduce an unintended side-effect. A common example is the use of = in an `if` expression in C where the
 6 programmer meant to do an equality test using the `==` operator. Other easily confused operators in C are the
 7 logical operators such as `&&` for the bitwise operator `&`, or vice versa. It is legal and possible that the programmer
 8 intended to do an assignment within the `if` expression, but due to this being a common error, a programmer doing
 9 so would be using a poor programming practice. A less likely occurrence, but still possible is the substitution of `==`
 10 for = in what is supposed to be an assignment statement, but which effectively becomes a null statement. These
 11 mistakes may survive testing only to manifest themselves in deployed code where they may be maliciously
 12 exploited.

13 **6.32.2 Cross reference**

14 CWE:

- 15 480. Use of Incorrect Operator
- 16 481. Assigning instead of Comparing
- 17 482. Comparing instead of Assigning
- 18 570. Expression is Always False
- 19 571. Expression is Always True

20 JSF AV Rules: 160 and 166

21 MISRA 2004: 12.3, 12.4, 12.13, 13.1, 13.7, and 14.2

22 MISRA 2008: 0-1-9, 5-0-1, 6-2-1, and 6-5-2

23 CERT/CC guidelines: MSC02-C and MSC03-C

24 **6.32.3 Mechanism of failure**

25 Some of the failures are simply a case of programmer carelessness. Substitution of = instead of == in a Boolean
 26 test is easy to do and most C and C++ programmers have made this mistake at one time or another. Other
 27 instances can be the result of intricacies of the language definition that specifies what part of an expression must
 28 be evaluated. For instance, having an assignment expression in a Boolean statement is likely making an
 29 assumption that the complete expression will be executed in all cases. However, this is not always the case as
 30 sometimes the truth-value of the Boolean expression can be determined after only executing some portion of the
 31 expression. For instance:

```
32     if ((a == b) | (c = (d-1)))
```

33 There is no guarantee which of the two subexpressions `(a == b)` or `(c=(d-1))` will be executed first. Should
 34 `(a==b)` be determined to be true, then there is no need for the subexpression `(c=(d-1))` to be executed and as
 35 such, the assignment `(c=(d-1))` will not occur.

36 Embedding expressions in other expressions can yield unexpected results. Increment and decrement operators
 37 (`++` and `--`) can also yield unexpected results when mixed into a complex expression.

38 Incorrect calculated results can lead to a wide variety of erroneous program execution

39 **6.32.4 Applicable language characteristics**

40 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 41 • All languages are susceptible to likely incorrect expressions.

1 **6.32.5 Avoiding the vulnerability or mitigating its effects**

2 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 3 • Simplify expressions.
- 4 • Do not use assignment expressions as function parameters. Sometimes the assignment may not be
5 executed as expected. Instead, perform the assignment before the function call.
- 6 • Do not perform assignments within a Boolean expression. This is likely unintended, but if not, then move
7 the assignment outside of the Boolean expression for clarity and robustness.
- 8 • On some rare occasions, some statements intentionally do not have side effects and do not cause control
9 flow to change. These should be annotated through comments and made obvious that they are
10 intentionally no-ops with a stated reason. If possible, such reliance on null statements should be avoided.
11 In general, except for those rare instances, all statements should either have a side effect or cause control
12 flow to change.

13 **6.32.6 Implications for standardization**

14 In future standardization activities, the following items should be considered:

- 15 • Languages should consider providing warnings for statements that are unlikely to be right such as
16 statements without side effects. A null (no-op) statement may need to be added to the language for those
17 rare instances where an intentional null statement is needed. Having a null statement as part of the
18 language will reduce confusion as to why a statement with no side effects is present in code.
- 19 • Languages should consider not allowing assignments used as function parameters.
- 20 • Languages should consider not allowing assignments within a Boolean expression.
- 21 • Language definitions should avoid situations where easily confused symbols (e.g. = and ==, or ; and :, or
22 != and /=) are legal in the same context. For example, = is not generally legal in an `if` statement in Java
23 because it does not normally return a boolean value.

24 **6.32.7 Bibliography**

25 [None]

26 **6.33 Dead and Deactivated Code [XYQ]**

27 **6.33.1 Description of application vulnerability**

28 Dead and Deactivated code (the distinction is addressed in 6.33.3) is code that exists in the executable, but which
29 can never be executed, either because there is no call path that leads to it (e.g., a function that is never called), or
30 the path is semantically infeasible (e.g., its execution depends on the state of a conditional that can never be
31 achieved).

32 Dead and Deactivated code is undesirable because it indicates the possibility of a coding error and because it may
33 provide a "jump" target for an intrusion.

34 Also covered in this vulnerability is code which is believed to be dead, but which is inadvertently executed.

35 **6.33.2 Cross reference**

36 CWE:

- 37 561. Dead Code
- 38 570. Expression is Always False
- 39 571. Expression is Always True
- 40 JSF AV Rules: 127 and 186
- 41 MISRA C 2004: 14.1 and 2.4
- 42 MISRA C++ 2008: 0-1-1 to 0-1-10, 2-7-2, and 2-7-3
- 43 CERT/CC guidelines: MSC07-C and MSC12-C

1 DO178B/C

2 6.33.3 Mechanism of failure

3 DO-178B defines Dead and Deactivated code as:

- 4 • Dead code – Executable object code (or data) which... cannot be executed (code) or used (data) in an
- 5 operational configuration of the target computer environment and is not traceable to a system or software
- 6 requirement.
- 7 • Deactivated code – Executable object code (or data) which by design is either (a) not intended to be
- 8 executed (code) or used (data), for example, a part of a previously developed software component, or (b) is
- 9 only executed (code) or used (data) in certain configurations of the target computer environment, for
- 10 example, code that is enabled by a hardware pin selection or software programmed options.

11 Dead code is code that exists in an application, but which can never be executed, either because there is no call
12 path to the code (e.g., a function that is never called) or because the execution path to the code is semantically
13 infeasible, e.g., in

```
14     integer i = 0;
15     if ( i == 0)
16         then fun_a();
17         else fun_b();
```

18 fun_b is dead code, as only fun_a can ever be executed.

19 The presence of dead code is not in itself an error, but begs the question why is it there? Is its presence an
20 indication that the developer believed it to be necessary, but some error means it will never be executed? Or is
21 there a legitimate reason for its presence, for example:

- 22 • Defensive code, only executed as the result of a hardware failure.
- 23 • Code that is part of a library not required in this application.
- 24 • Diagnostic code not executed in the operational environment.

25 Such code may be referred to as “deactivated”. That is, dead code that is there by intent.

26 There is a secondary consideration for dead code in languages that permit overloading of functions and other
27 constructs that use complex name resolution strategies. The developer may believe that some code is not going to
28 be used (deactivated), but its existence in the program means that it appears in the namespace, and may be
29 selected as the best match for some use that was intended to be of an overloading function. That is, although the
30 developer believes it is never going to be used, in practice it is used in preference to the intended function.

31 6.33.4 Applicable language characteristics

32 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 33 • Languages that allow code to exist in the executable that can never be executed.

34 6.33.5 Avoiding the vulnerability or mitigating its effects

35 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 36 • The developer should endeavour to remove, as a first resort and as far as practical, dead code from an
- 37 application.
- 38 • When a developer identifies code that is dead because a conditional always evaluates to the same value,
- 39 this could be indicative of an earlier bug and additional testing may be needed to ascertain why the same
- 40 value is occurring.

- 1 • The developer should identify any dead code in the application, and provide a justification (if only to
2 themselves) as to why it is there.
- 3 • The developer should also ensure that any code that was expected to be unused is actually recognized as
4 dead code.
- 5 • The developer should apply standard branch coverage measurement tools and ensure by 100% coverage
6 that all branches are neither dead nor deactivated

7 **6.33.6 Implications for standardization**

8 [None]

9 **6.33.7 Bibliography**

10 [None]

11 **6.34 Switch Statements and Static Analysis [CLL]**

12 **6.34.1 Description of application vulnerability**

13 Many programming languages provide a construct, such as a `switch` statement, that chooses among multiple
14 alternative control flows based upon the evaluated result of an expression. The use of such constructs may
15 introduce application vulnerabilities if not all possible cases appear within the switch or if control unexpectedly flows
16 from one alternative to another.

17 **6.34.2 Cross reference**

18 JSF AV Rules: 148, 193, 194, 195, and 196
19 MISRA C 2004: 15.2, 15.3, and 15.5
20 MISRA C++ 2008: 6-4-3, 6-4-5, 6-4-6, and 6-4-8
21 CERT/CC guidelines: MSC01-C

22 **6.34.3 Mechanism of failure**

23 The fundamental challenge when using a `switch` statement is to make sure that all possible cases are, in fact,
24 dealt with.

25 **6.34.4 Applicable language characteristics**

26 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 27 • Contain a construct, such as a `switch` statement, that provides a selection among alternative control flows
28 based on the evaluation of an expression.

29
30 **6.34.5 Avoiding the vulnerability or mitigating its effects**

31 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 32 • Switch on an expression that has a small number of potential values that can be statically enumerated. In
33 languages that provide them, a variable of an enumerated type is to be preferred because its possible set
34 of values is known statically and is small in number (as compared, for example, to the value set of an
35 integer variable). In languages that don't provide enumerated types, a tightly constrained integer sub-type
36 might be a good alternative. Where it is practical to statically enumerate the switched type, it is preferable
37 to omit the default case, because the static analysis is simplified and because maintainers can better
38 understand the intent of the original programmer. When one must switch on some other form of type, it is
39 necessary to have a default case, preferably to be regarded as a serious error condition.

- 1 • Avoid “flowing through” from one case to another. Even if correctly implemented, it is difficult for reviewers
2 and maintainers to distinguish whether the construct was intended or is an error of omission. (Using
3 multiple labels on individual alternatives is not a violation of this guideline, though.) In cases where flow-
4 through is necessary and intended, an explicitly coded branch may be preferable in order to clearly mark
5 the intent. Providing comments regarding intention can be helpful to reviewers and maintainers.
6 • Perform static analysis to determine if all cases are, in fact, covered by the code. (Note that the use of a
7 default case can hamper the effectiveness of static analysis since the tool cannot determine if omitted
8 alternatives were or were not intended for default treatment.)
9 • Other means of mitigation include manual review, bounds testing, tool analysis, verification techniques, and
10 proofs of correctness.

11 **6.34.6 Implications for standardization**

12 In future standardization activities, the following items should be considered:

- 13 • Language specifications could require compilers to ensure that a complete set of alternatives is provided in
14 cases where the value set of the switch variable can be statically determined.

15 **6.34.7 Bibliography**

16 Hatton 14: `Switch` statements

17 **6.35 Demarcation of Control Flow [EOJ]**

18 **6.35.1 Description of application vulnerability**

19 Some programming languages explicitly mark the end of an `if` statement or a loop, whereas other languages mark
20 only the end of a block of statements. Languages of the latter category are prone to oversights by the programmer,
21 causing unintended sequences of control flow.

22 **6.35.2 Cross reference**

23 JSF AV Rules: 59 and 192

24 MISRA C 2004: 14.8, 14.9, 14.10, and 19.5

25 MISRA C++ 2008: 6-3-1, 6-4-1, 6-4-2, 6-4-3, 6-4-8, 6-5-1, 6-5-6, 6-6-1 to 6-6-5, and 16-0-2

26 Hatton 18: Control flow – `if` structure

27 **6.35.3 Mechanism of failure**

28 Programmers may rely on indentation to determine inclusion of statements within constructs. Testing of the
29 software may not reveal that statements thought to be included in an `if-then`, `if-then-else`, or loops that are not in
30 reality a part of it the statement. Moreover, for a nested `if-then-else` statement the programmer may be confused
31 about which “`if`” statement controls the “`else`” part directly. This can lead to unexpected results.

32 **6.35.4 Applicable language characteristics**

33 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 34 • Languages that contain loops and conditional statements that are not explicitly terminated by an “`end`”
35 construct.

36 **6.35.5 Avoiding the vulnerability or mitigating its effects**

37 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 1 • Adopt a convention for marking the closing of a construct that can be checked by a tool, to ensure that
- 2 program structure is apparent.
- 3 • Adopt programming guidelines (preferably augmented by static analysis). For example, consider the rules
- 4 itemized above from JSF AV, MISRA C, MISRA C++ or Hatton.
- 5 • Other means of assurance might include proofs of correctness, analysis with tools, verification techniques,
- 6 etc.
- 7 • Pretty-printers and syntax-aware editors may be helpful in finding such problems, but sometimes disguise
- 8 them.
- 9 • Include a final else statement at the end of `if...else-if` constructs to avoid confusion.
- 10 • Always enclose the body of statements of an `if`, `while`, `for`, etc. in braces (“{ }”) or other demarcation
- 11 indicators appropriate to the language used.

12 6.35.6 Implications for standardization

13 In future standardization activities, the following items should be considered:

- 14 • Specifiers of languages should consider adding a mode that strictly enforces compound conditional and
- 15 looping constructs with explicit termination, such as “end if” or a closing bracket.
- 16 • Specifiers of languages might consider explicit termination of loops and conditional statements.
- 17 • Specifiers might consider features to terminate named loops and conditionals and determine if the structure
- 18 as named matches the structure as inferred.

19 6.35.7 Bibliography

20 Hatton 18: Control flow – `if` structure

21 6.36 Loop Control Variables [TEX]

22 6.36.1 Description of application vulnerability

23 Many languages support a looping construct whose number of iterations is controlled by the value of a loop control

24 variable. Looping constructs provide a method of specifying an initial value for this loop control variable, a test that

25 terminates the loop and the quantity by which it should be decremented/incremented on each loop iteration.

26 In some languages it is possible to modify the value of the loop control variable within the body of the loop.

27 Experience shows that such value modifications are sometimes overlooked by readers of the source code,

28 resulting in faults being introduced.

29 6.36.2 Cross reference

30 JSF AV Rule: 201

31 MISRA C 2004: 13.6

32 MISRA C++ 2008: 6-5-1 to 6-5-6

33 6.36.3 Mechanism of failure

34 Readers of source code often make assumptions about what has been written. A common assumption is that a

35 loop control variable is a constant since such variables are not usually modified in the body of the associated loop.

36 A reader of the source may incorrectly assume that a loop control variable is not modified in the body of its loop

37 and write (incorrect) code based on this assumption.

38 6.36.4 Applicable language characteristics

39 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 1 • Languages that permit a loop control variable to be modified in the body of its associated loop.

2 **6.36.5 Avoiding the vulnerability or mitigating its effects**

3 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 4 • Not modifying a loop control variable in the body of its associated loop body.
 5 • Some languages, such as C and C++ do not explicitly specify which of the variables appearing in a loop
 6 header is the loop control variable. MISRA-C [15], MISRA C++ [20], and Jones have proposed algorithms
 7 for deducing which, if any, of these variables is the loop control variable in the programming languages C
 8 and C++ (these algorithms could also be applied to other languages that support a C-like for-loop).

9 **6.36.6 Implications for standardization**

10 In future standardization activities, the following items should be considered:

- 11 • Language designers should consider the addition of an identifier type for loop control that cannot be
 12 modified by anything other than the loop control construct.

13 **6.36.7 Bibliography**

14 Loops and their control variables: Discussion and proposed guidelines, Derek M. Jones, February 2006.

15 **6.37 Off-by-one Error [XZH]**

16 **6.37.1 Description of application vulnerability**

17 A program uses an incorrect maximum or minimum value that is 1 more or 1 less than the correct value. This
 18 usually arises from one of a number of situations where the bounds as understood by the developer differ from the
 19 design, such as:

- 20 • Confusion between the need for < and <= or > and >= in a test.
 21 • Confusion as to the index range of an algorithm, such as: beginning an algorithm at 1 when the underlying
 22 structure is indexed from 0; beginning an algorithm at 0 when the underlying structure is indexed from 1 (or
 23 some other start point); or using the length of a structure as its bound instead of the sentinel values.
 24 • Failing to allow for storage of a sentinel value, such as the NULL string terminator that is used in the C and
 25 C++ programming languages.

26 These issues arise from mistakes in mapping the design into a particular language, in moving between languages
 27 (such as between languages where all arrays start at 0 and other languages where arrays start at 1), and when
 28 exchanging data between languages with different default array bounds.

29 The issue also can arise in algorithms where relationships exist between components, and the existence of a
 30 bounds value changes the conditions of the test.

31 The existence of this possible flaw can also be a serious security hole as it can permit someone to surreptitiously
 32 provide an unused location (such as 0 or the last element) that can be used for undocumented features or hidden
 33 channels.

34 **6.37.2 Cross reference**

35 CWE:

36 193. Off-by-one Error

1 **6.37.3 Mechanism of failure**

2 An off-by-one error could lead to:

- 3 • an out-of bounds access to an array (buffer overflow),
- 4 • incomplete comparisons or calculation mistakes,
- 5 • a read from the wrong memory location, or
- 6 • an incorrect conditional.

7 Such incorrect accesses can cause cascading errors or references to illegal locations, resulting in potentially
8 unbounded behaviour.

9 Off-by-one errors are not often exploited in attacks because they are difficult to identify and exploit externally, but
10 the cascading errors and boundary-condition errors can be severe.

11 **6.37.4 Applicable language characteristics**

12 As this vulnerability arises because of an algorithmic error by the developer, it can in principle arise in any
13 language; however, it is most likely to occur when:

- 14 • The language relies on the developer having implicit knowledge of structure start and end indices (e.g.,
15 knowing whether arrays start at 0 or 1 – or indeed some other value).
- 16 • Where the language relies upon explicit bounds values to terminate variable length arrays.

17 **6.37.5 Avoiding the vulnerability or mitigating its effects**

18 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 19 • A systematic development process, use of development/analysis tools and thorough testing are all
20 common ways of preventing errors, and in this case, off-by-one errors.
- 21 • Where references are being made to structure indices and the languages provide ways to specify the
22 whole structure or the starting and ending indices explicitly (e.g., Ada provides xxx'First and xxx'Last for
23 each dimension), these should be used always. Where the language doesn't provide these, constants can
24 be declared and used in preference to numeric literals.
- 25 • Where the language doesn't encapsulate variable length arrays, encapsulation should be provided through
26 library objects and a coding standard developed that requires such arrays to only be used via those library
27 objects, so the developer does not need to be explicitly concerned with managing bounds values.
28

29 **6.37.6 Implications for standardization**

30 In future standardization activities, the following items should be considered:

- 31 • Languages should provide encapsulations for arrays that:
 - 32 ○ Prevent the need for the developer to be concerned with explicit bounds values.
 - 33 ○ Provide the developer with symbolic access to the array start, end and iterators.

34 **6.37.7 Bibliography**

35 [None]

1 **6.38 Structured Programming [EWD]**

2 **6.38.1 Description of application vulnerability**

3 Programs that have a convoluted control structure are likely to be more difficult to be human readable, less
4 understandable, harder to maintain, more difficult to modify, harder to statically analyze, and more difficult to match
5 the allocation and release of resources.

6

7 **6.38.2 Cross reference**

8 JSF AV Rules: 20, 113, 189, 190, and 191

9 MISRA C 2004: 14.4,14.5, and 20.7

10 MISRA C++ 2008: 6-6-1, 6-6-2, 6-6-3, and 17-0-5

11 CERT/CC guidelines: SIG32-C

12

13 **6.38.3 Mechanism of failure**

14 Lack of structured programming can lead to:

- 15 • Memory or resource leaks.
- 16 • Error prone maintenance.
- 17 • Design that is difficult or impossible to validate.
- 18 • Source code that is difficult or impossible to statically analyze.

19 **6.38.4 Applicable language characteristics**

20 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 21 • Languages that allow leaving a loop without consideration for the loop control.
- 22 • Languages that allow local jumps (`goto` statement).
- 23 • Languages that allow non-local jumps (`set jmp/long jmp` in the C programming language).
- 24 • Languages that support multiple entry and exit points from a function, procedure, subroutine or method.

25 **6.38.4 Avoiding the vulnerability or mitigating its effects**

26 Use only those features of the programming language that enforce a logical structure on the program. The
27 program flow follows a simple hierarchical model that employs looping constructs such as `for`, `repeat`, `do`, and
28 `while`.

29 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 30 • Avoid using language features such as `goto`.
- 31 • Avoid using language features such as `continue` and `break` in the middle of loops.
- 32 • Avoid using language features that transfer control of the program flow via a jump.
- 33 • Avoid multiple exit points to a function/procedure/method/subroutine.
- 34 • Avoid multiple entry points to a function/procedure/method/subroutine.

35 **6.38.5 Implications for standardization**

36 In future standardization activities, the following items should be considered:

- 37 • Languages should support and favor structured programming through their constructs to the extent
38 possible.

1 **6.38.6 Bibliography**

2 Holtzmann-1

3 **6.39 Passing Parameters and Return Values [CSJ]**4 **6.39.1 Description of application vulnerability**

5 Nearly every procedural language provides some method of process abstraction permitting decomposition of the
 6 flow of control into routines, functions, subprograms, or methods. (For the purpose of this description, the term
 7 subprogram will be used.) To have any effect on the computation, the subprogram must change data visible to the
 8 calling program. It can do this by changing the value of a non-local variable, changing the value of a parameter, or,
 9 in the case of a function, providing a return value. Because different languages use different mechanisms with
 10 different semantics for passing parameters, a programmer using an unfamiliar language may obtain unexpected
 11 results.

12 **6.39.2 Cross reference**

13 JSF AV Rules: 116, 117, and 118
 14 MISRA C 2004: 16.1, 16.2, 16.3, 16.4, 16.5, 16.6, 16.7, and 16.9
 15 MISRA C++ 2008: 0-3-2, 7-1-2, 8-4-1, 8-4-2, 8-4-3, and 8-4-4
 16 CERT/CC guidelines: EXP12-C and DCL33-C
 17

18 **6.39.3 Mechanism of failure**

19 The mechanisms for parameter passing include: *call by reference*, *call by copy*, and *call by name*. The last is so
 20 specialized and supported by so few programming languages that it will not be treated in this description.

21 In call by reference, the calling program passes the addresses of the arguments to the called subprogram. When
 22 the subprogram references the corresponding formal parameter, it is actually sharing data with the calling program.
 23 If the subprogram changes a formal parameter, then the corresponding actual argument is also changed. If the
 24 actual argument is an expression or a constant, then the address of a temporary location is passed to the
 25 subprogram; this may be an error in some languages.

26 In call by copy, the called subprogram does not share data with the calling program. Instead, formal parameters act
 27 as local variables. Values are passed between the actual arguments and the formal parameters by copying. Some
 28 languages may control changes to formal parameters based on labels such as *in*, *out*, or *inout*. There are three
 29 cases to consider: *call by value* for *in* parameters; *call by result* for *out* parameters and function return values;
 30 and *call by value-result* for *inout* parameters. For call by value, the calling program evaluates the actual
 31 arguments and copies the result to the corresponding formal parameters that are then treated as local variables by
 32 the subprogram. For call by result, the values of the locals corresponding to formal parameters are copied to the
 33 corresponding actual arguments. For call by value-result, the values are copied in from the actual arguments at the
 34 beginning of the subprogram's execution and back out to the actual arguments at its termination.

35 The obvious disadvantage of call by copy is that extra copy operations are needed and execution time is required
 36 to produce the copies. Particularly if parameters represent sizable objects, such as large arrays, the cost of call by
 37 copy can be high. For this reason, many languages also provide the call by reference mechanism. The
 38 disadvantage of call by reference is that the calling program cannot be assured that the subprogram hasn't
 39 changed data that was intended to be unchanged. For example, if an array is passed by reference to a subprogram
 40 intended to sum its elements, the subprogram could also change the values of one or more elements of the array.
 41 However, some languages enforce the subprogram's access to the shared data based on the labeling of actual
 42 arguments with modes—such as *in*, *out*, or *inout* or by constant pointers.

43 Another problem with call by reference is unintended aliasing. It is possible that the address of one actual argument
 44 is the same as another actual argument or that two arguments overlap in storage. A subprogram, assuming the two
 45 formal parameters to be distinct, may treat them inappropriately. For example, if one codes a subprogram to swap

1 two values using the exclusive-or method, then a call to `swap(x, x)` will zero the value of `x`. Aliasing can also
 2 occur between arguments and non-local objects. For example, if a subprogram modifies a non-local object as a
 3 side-effect of its execution, referencing that object by a formal parameter will result in aliasing and, possibly,
 4 unintended results.

5 Some languages provide only simple mechanisms for passing data to subprograms, leaving it to the programmer to
 6 synthesize appropriate mechanisms. Often, the only available mechanism is to use call by copy to pass small
 7 scalar values or pointer values containing addresses of data structures. Of course, the latter amounts to using call
 8 by reference with no checking by the language processor. In such cases, subprograms can pass back pointers to
 9 anything whatsoever, including data that is corrupted or absent.

10 Some languages use call by copy for small objects, such as scalars, and call by reference for large objects, such
 11 as arrays. The choice of mechanism may even be implementation-defined. Because the two mechanisms produce
 12 different results in the presence of aliasing, it is very important to avoid aliasing.

13 An additional problem may occur if the called subprogram fails to assign a value to a formal parameter that the
 14 caller expects as an output from the subprogram. In the case of call by reference, the result may be an uninitialized
 15 variable in the calling program. In the case of call by copy, the result may be that a legitimate initialization value
 16 provided by the caller is overwritten by an uninitialized value because the called program did not make an
 17 assignment to the parameter. This error may be difficult to detect through review because the failure to initialize is
 18 hidden in the subprogram.

19 An additional complication with subprograms occurs when one or more of the arguments are expressions. In such
 20 cases, the evaluation of one argument might have side-effects that result in a change to the value of another or
 21 unintended aliasing. Implementation choices regarding order of evaluation could affect the result of the
 22 computation. This particular problem is described in Side-effects and Order of Evaluation section [SAM].

23 6.39.4 Applicable language characteristics

24 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 25 • Languages that provide mechanisms for defining subprograms where the data passes between the calling
 26 program and the subprogram via parameters and return values. This includes methods in many popular
 27 object-oriented languages.

28 6.39.5 Avoiding the vulnerability or mitigating its effects

29 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 30 • Use available mechanisms to label parameters as constants or with modes like `in`, `out`, or `inout`.
- 31 • When a choice of mechanisms is available, pass small simple objects using call by copy.
- 32 • When a choice of mechanisms is available and the computational cost of copying is tolerable, pass larger
 33 objects using call by copy.
- 34 • When the choice of language or the computational cost of copying forbids using call by copy, then take
 35 safeguards to prevent aliasing:
 - 36 ○ Minimize side-effects of subprograms on non-local objects; when side-effects are coded, ensure
 37 that the affected non-local objects are not passed as parameters using call by reference.
 - 38 ○ To avoid unintentional aliasing, avoid using expressions or functions as actual arguments; instead
 39 assign the result of the expression to a temporary local and pass the local.
 - 40 ○ Utilize tooling or other forms of analysis to ensure that non-obvious instances of aliasing are
 41 absent.
 - 42 ○ Perform reviews or analysis to determine that called subprograms fulfill their responsibilities to
 43 assign values to all output parameters.

44 6.39.6 Implications for standardization

45 In future standardization activities, the following items should be considered:

- 1 • Programming language specifications could provide labels—such as `in`, `out`, and `inout`—that control the
2 subprogram's access to its formal parameters, and enforce the access.

3 6.39.7 Bibliography

- 4 [1] Robert W. Sebesta, Concepts of Programming Languages, 8th edition, ISBN-13: 978-0-321-49362-0, ISBN-10:
5 0-321-49362-1, Pearson Education, Boston, MA, 2008
6 [2] Carlo Ghezzi and Mehdi Jazayeri, Programming Language Concepts, 3rd edition, ISBN-0-471-10426-4, John
7 Wiley & Sons, 1998

8 6.40 Dangling References to Stack Frames [DCM]

9 6.40.1 Description of application vulnerability

10 Many languages allow treating the address of a local variable as a value stored in other variables. Examples are
11 the application of the address operator in C or C++, or of the 'Access or 'Address attributes in Ada. In some
12 languages, this facility is also used to model the call-by-reference mechanism by passing the address of the actual
13 parameter by-value. An obvious safety requirement is that the stored address shall not be used after the lifetime of
14 the local variable has expired. Technically, the stack frame, in which the local variable lived, has been popped and
15 memory may have been reused for a subsequent call. Therefore, the invalidity of the stored address is very difficult
16 to decide. This situation can be described as a “dangling reference to the stack”.

17 6.40.2 Cross reference

18 CWE:
19 562. Return of Stack Variable Address
20 JSF AV Rule: 173
21 MISRA C 2004: 17.6 and 21.1
22 MISRA C++ 2008: 0-3-1, 7-5-1, 7-5-2, and 7-5-3
23 CERT/CC guidelines: EXP35-C and DCL30-C

24 6.40.3 Mechanism of failure

25 The consequences of dangling references to the stack come in two variants: a deterministically predictable variant,
26 which therefore can be exploited, and an intermittent, non-deterministic variant, which is next to impossible to elicit
27 during testing. The following code sample illustrates the two variants; the behaviour is not language-specific:

```
28     struct s { ... };
29     typedef struct s array_type[1000];
30     array_type* ptr;
31     array_type* F()
32     {
33         struct s Arr[1000];
34         ptr = &Arr;      // Risk of variant 1;
35         return &Arr;    // Risk of variant 2;
36     }
37     ...
38     struct s secret;
39     array_type* ptr2;
40     ptr2 = F();
41     secret = (*ptr2)[10]; // Fault of variant 2
42     ...
43     secret = (*ptr)[10]; // Fault of variant 1
```

44 The risk of variant 1 is the assignment of the address of `Arr` to a pointer variable that survives the lifetime of `Arr`.
45 The fault is the subsequent use of the dangling reference to the stack, which references memory since altered by

1 other calls and possibly validly owned by other routines. As part of a call-back, the fault allows systematic
 2 examination of portions of the stack contents without triggering an array-bounds-checking violation. Thus, this
 3 vulnerability is easily exploitable. As a fault, the effects can be most astounding, as memory gets corrupted by
 4 completely unrelated code portions. (A life-time check as part of pointer assignment can prevent the risk. In many
 5 cases, e.g., the situations above, the check is statically decidable by a compiler. However, for the general case, a
 6 dynamic check is needed to ensure that the copied pointer value lives no longer than the designated object.)

7 The risk of variant 2 is an idiom “seen in the wild” to return the address of a local variable in order to avoid an
 8 expensive copy of a function result, as long as it is consumed before the next routine call occurs. The idiom is
 9 based on the ill-founded assumption that the stack will not be affected by anything until this next call is issued. The
 10 assumption is false, however, if an interrupt occurs and interrupt handling employs a strategy called “stack
 11 stealing”, i.e., using the current stack to satisfy its memory requirements. Thus, the value of `ARR` can be overwritten
 12 before it can be retrieved after the call on `F`. As this fault will only occur if the interrupt arrives after the call has
 13 returned but before the returned result is consumed, the fault is highly intermittent and next to impossible to re-
 14 create during testing. Thus, it is unlikely to be exploitable, but also exceedingly hard to find by testing. It can begin
 15 to occur after a completely unrelated interrupt handler has been coded or altered. Only static analysis can relatively
 16 easily detect the danger (unless the code combines it with risks of variant 1). Some compilers issue warnings for
 17 this situation; such warnings need to be heeded, and some forms of static analysis are effective in identifying such
 18 problems.

19 **6.40.4 Applicable language characteristics**

20 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 21 • The address of a local entity (or formal parameter) of a routine can be obtained and stored in a variable or
 22 can be returned by this routine as a result.
- 23 • No check is made that the lifetime of the variable receiving the address is no larger than the lifetime of the
 24 designated entity.

25 **6.40.5 Avoiding the vulnerability or mitigating its effects**

26 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 27 • Do not use the address of locally declared entities as storable, assignable or returnable value (except
 28 where idioms of the language make it unavoidable).
- 29 • Where unavoidable, ensure that the lifetime of the variable containing the address is completely enclosed
 30 by the lifetime of the designated object.
- 31 • Never return the address of a local variable as the result of a function call.

32 **6.40.6 Implications for standardization**

33 In future standardization activities, the following items should be considered:

- 34 • Do not provide means to obtain the address of a locally declared entity as a storable value; or
- 35 • Define implicit checks to implement the assurance of enclosed lifetime expressed in subsection 5 of this
 36 vulnerability. Note that, in many cases, the check is statically decidable, for example, when the address of
 37 a local entity is taken as part of a return statement or expression.

38 **6.40.7 Bibliography**

39 [None]

1 6.41 Subprogram Signature Mismatch [OTR]

2 6.41.1 Description of application vulnerability

3 If a subprogram is called with a different number of parameters than it expects, or with parameters of different types
 4 than it expects, then the results will be incorrect. Depending on the language, the operating environment, and the
 5 implementation, the error might be as benign as a diagnostic message or as extreme as a program continuing to
 6 execute with a corrupted stack. The possibility of a corrupted stack provides opportunities for penetration.

7 6.41.2 Cross reference

8 CWE:

9 628. Function Call with Incorrectly Specified Arguments

10 686. Function Call with Incorrect Argument Type

11 683. Function Call with Incorrect Order of Arguments

12 JSF AV Rule: 108

13 MISRA C 2004: 8.1, 8.2, 8.3, 16.1, 16.3, 16.4, 16.5, 16.6

14 MISRA C++ 2008: 0-3-2, 3-2-1, 3-2-2, 3-2-3, 3-2-4, 3-3-1, 3-9-1, 8-3-1, 8-4-1, and 8-4-2

15 CERT/CC guidelines: DCL31-C, and DCL35-C

16

17 6.41.3 Mechanism of failure

18 When a subprogram is called, the actual arguments of the call are pushed on to the execution stack. When the
 19 subprogram terminates, the formal parameters are popped off the stack. If the number and type of the actual
 20 arguments do not match the number and type of the formal parameters, then the push and the pop will not be
 21 commensurable and the stack will be corrupted. Stack corruption can lead to unpredictable execution of the
 22 program and can provide opportunities for execution of unintended or malicious code.

23 The compilation systems for many languages and implementations can check to ensure that the list of actual
 24 parameters and any expected return match the declared set of formal parameters and return value (the
 25 *subprogram signature*) in both number and type. (In some cases, programmers should observe a set of
 26 conventions to ensure that this is true.) However, when the call is being made to an externally compiled
 27 subprogram, an object-code library, or a module compiled in a different language, the programmer must take
 28 additional steps to ensure a match between the expectations of the caller and the called subprogram.

29 6.41.4 Applicable language characteristics

30 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 31 • Languages that do not require their implementations to ensure that the number and types of actual
 32 arguments are equal to the number and types of the formal parameters.
- 33 • Implementations that permit programs to call subprograms that have been externally compiled (without a
 34 means to check for a matching subprogram signature), subprograms in object code libraries, and any
 35 subprograms compiled in other languages.

36 6.41.5 Avoiding the vulnerability or mitigating its effects

37 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 38 • Take advantage of any mechanism provided by the language to ensure that parameter signatures match.
- 39 • Avoid any language features that permit variable numbers of actual arguments without a method of
 40 enforcing a match for any instance of a subprogram call.
- 41 • Take advantage of any language or implementation feature that would guarantee matching the
 42 subprogram signature in linking to other languages or to separately compiled modules.
- 43 • Intensively review subprogram calls where the match is not guaranteed by tooling.

1 **6.41.6 Implications for standardization**

2 In future standardization activates, the following items should be considered:

- 3 • Language specifiers could ensure that the signatures of subprograms match within a single compilation
- 4 unit and could provide features for asserting and checking the match with externally compiled
- 5 subprograms.

6 **6.41.7 Bibliography**

7 [None]

8 **6.42 Recursion [GDL]**

9 **6.42.1 Description of application vulnerability**

10 Recursion is an elegant mathematical mechanism for defining the values of some functions. It is tempting to write
 11 code that mirrors the mathematics. However, the use of recursion in a computer can have a profound effect on the
 12 consumption of finite resources, leading to denial of service.

13 **6.42.2 Cross reference**

14 CWE:

- 15 674. Uncontrolled Recursion
- 16 JSF AV Rule: 119
- 17 MISRA C 2004: 16.2
- 18 MISRA C++ 2008: 7-5-4
- 19 CERT/CC guidelines: MEM05-C

20 **6.42.3 Mechanism of failure**

21 Recursion provides for the economical definition of some mathematical functions. However, economical definition
 22 and economical calculation are two different subjects. It is tempting to calculate the value of a recursive function
 23 using recursive subprograms because the expression in the programming language is straightforward and easy to
 24 understand. However, the impact on finite computing resources can be profound. Each invocation of a recursive
 25 subprogram may result in the creation of a new stack frame, complete with local variables. If stack space is limited
 26 and the calculation of some values will lead to an exhaustion of resources resulting in the program terminating.

27 In calculating the values of mathematical functions the use of recursion in a program is usually obvious, but this is
 28 not true in the general case. For example, finalization of a computing context after treating an error condition might
 29 result in recursion (e.g., attempting to "clean up" by closing a file after an error was encountered in closing the
 30 same file). Although such situations may have other problems, they typically do not result in exhaustion of
 31 resources but may otherwise result in a denial of service.

32 **6.42.4 Applicable language characteristics**

33 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 34 • Any language that permits the recursive invocation of subprograms.

35 **6.42.5 Avoiding the vulnerability or mitigating its effects**

36 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 37 • Minimize the use of recursion.

- 1 • Converting recursive calculations to the corresponding iterative calculation. In principle, any recursive
2 calculation can be remodeled as an iterative calculation which will have a smaller impact on some
3 computing resources but which may be harder for a human to comprehend. The cost to human
4 understanding must be weighed against the practical limits of computing resource.
- 5 • In cases where the depth of recursion can be shown to be statically bounded by a tolerable number, then
6 recursion may be acceptable, but should be documented for the use of maintainers.

7 It should be noted that some languages or implementations provide special (more economical) treatment of a form
8 of recursion known as *tail-recursion*. In this case, the impact on computing economy is reduced. When using such
9 a language, tail recursion may be preferred to an iterative calculation.

10 6.42.6 Implications for standardization

11 [None]

12 6.42.7 Bibliography

13 [None]

14 6.43 Returning Error Status [NZN]

15 6.43.1 Description of application vulnerability

16 Unpredicted error conditions—perhaps from hardware (such as an I/O device error), perhaps from software (such
17 as heap exhaustion)—sometimes arise during the execution of code. Programming languages provide a
18 surprisingly wide variety of mechanisms to deal with such errors. The choice of a mechanism that doesn't match
19 the programming language can lead to errors in the execution of the software or unexpected termination of the
20 program. This could lead to a simple decrease in the robustness of a program or it could be exploited in a denial of
21 service attack.

22 6.43.2 Cross reference

23 JSF AV Rules: 115 and 208
24 MISRA C 2004: 16.10
25 MISRA C++ 2008: 15-3-2 and 19-3-1
26 CERT/CC guidelines: DCL09-C, ERR00-C, and ERR02-C

27 6.43.3 Mechanism of failure

28 Even in the best-written programs, error conditions sometimes arise. Some errors occur because of defects in the
29 software itself, but some result from external conditions in hardware, such as errors in I/O devices, or in the
30 software system, such as exhaustion of heap space. If left untreated, the effect of the error might result in
31 termination of the program or continuation of the program with incorrect results. To deal with the situation,
32 designers of programming languages have equipped their languages with different mechanisms to detect and treat
33 such errors. These mechanisms are typically intended to be used in specific programming idioms. However, the
34 mechanisms differ among languages. A programmer expert in one language might mistakenly use an inappropriate
35 idiom when programming in a different language with the result that some errors are left untreated, leading to
36 termination or incorrect results. Attackers can exploit such weaknesses in denial of service attacks.

37 In general, languages make no distinction between dealing with programming errors (like an access to protected
38 memory), unexpected hardware errors (like device error), expected but unusual conditions (like end of file), and
39 even usual conditions that fail to provide the typical result (like an unsuccessful search). This description will use
40 the term "error" to apply to all of the above. The description applies equally to error conditions that are detected via
41 hardware mechanisms and error conditions that are detected via software during execution of a subprogram (such
42 as an inappropriate parameter value).

1 6.43.4 Applicable language characteristics

2 Different programming languages provide remarkably different mechanisms for treating errors. In languages that
3 provide a number of error detection and treatment mechanisms, it becomes a design issue to match the
4 mechanism to the condition. This section will describe the mechanisms that are provided in widely used languages.

5 The simplest case is the set of languages that provide no special mechanism for the notification and treatment of
6 unusual conditions. In such languages, error conditions are signaled by the value of an auxiliary status variable,
7 sometimes a subprogram parameter. The programming language C standard library functions use a variant of this
8 approach; the error status is provided as the return value and sometimes in an additional global error value.
9 Obviously, in such languages, it is imperative to check and act upon the status variable after every call to a
10 subprogram that might provide an error indication. If error conditions can occur in an asynchronous manner, it is
11 necessary to provide means to check for errors in a systematic and periodic manner.

12 Some languages permit the passing of a label parameter. If an error is encountered, the subprogram returns to the
13 indicated label rather than to the point at which it was called. Similarly some languages accept the name of a
14 subprogram to be used to handle errors. In either case, it is imperative to provide labeled code or a subprogram to
15 deal with all possible error situations.

16 The approaches described above have the disadvantage that error checking must be provided at every call to a
17 subprogram. This can clutter the code immensely to deal with situations that may occur rarely. For this reason,
18 some languages provide an exception mechanism that automatically transfers control when an error is
19 encountered. This has the potential advantage of allowing error treatment to be factored into distinct error handlers,
20 leaving the main execution path to deal with the usual results. The disadvantages, of course, are that the language
21 design is complicated and the programmer must deal with the conceptually more complex problem of providing
22 error handlers that are removed from the immediate context of a specific call to a subprogram. Furthermore,
23 different languages provide exception-handling mechanisms that differ in the manner in which various design
24 issues are treated:

- 25 • How is the occurrence of an exception bound to a particular handler?
- 26 • What happens when no handler is local to an exception occurrence? Is the exception propagated in some
27 manner or is it lost?
- 28 • What happens after an exception handler executes? Is control returned to the point before the call or after
29 the call, or is the calling routine terminated in some way? If the calling routine is terminated, is there some
30 provision for finalization, such as closing files or releasing resources?
- 31 • Are programmers permitted to define additional exceptions?
- 32 • Does the language provide default handlers for some exceptions or must the programmer explicitly provide
33 for all of them?
- 34 • Can predefined exceptions be raised explicitly by a subprogram?
- 35 • Under what circumstances can error checking be disabled?

36 6.43.5 Avoiding the vulnerability or mitigating its effects

37 Given the variety of error handling mechanisms, it is difficult to write general guidelines. However, dealing with
38 exception handlers can stress the capability of many static analysis tools and can, in some cases, reduce the
39 effectiveness of their analysis. Therefore, for situations where the highest of reliability is required, the application
40 should be designed so that exception handling is not used at all. In the more general case, exception-handling
41 mechanisms should be reserved for truly unexpected situations and other situations (possibly hardware arithmetic
42 overflow) where no other mechanism is available. Situations which are merely unusual, like end of file, should be
43 treated by explicit testing—either prior to the call which might raise the error or immediately afterward.

44 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 45 • Checking error return values or auxiliary status variables following a call to a subprogram is mandatory
46 unless it can be demonstrated that the error condition is impossible.
- 47 • In dealing with languages where untreated exceptions can be lost (e.g., an exception that goes untreated
48 within an Ada task), it is mandatory to deal with the exception in the local context before it is lost.

- 1 • When execution within a particular context is abandoned due to an exception, it is important to finalize the
- 2 context by closing open files, releasing resources and restoring any invariants associated with the context.
- 3 • It is often not appropriate to repair an error condition and retry the operation. In such cases, one often
- 4 treats a symptom but not the underlying problem. It is usually a better solution to finalize and terminate the
- 5 current context and retreat to a context where the situation is known.
- 6 • Error checking provided by the language, the software system, or the hardware should never be disabled in
- 7 the absence of a conclusive analysis that the error condition is rendered impossible.
- 8 • Because of the complexity of error handling, careful review of all error handling mechanisms is appropriate.
- 9 • In applications with the highest requirements for reliability, defense-in-depth approaches are often
- 10 appropriate, for example, checking and handling errors thought to be impossible.

11 **6.43.6 Implications for standardization**

12 In future standardization activities, the following items should be considered:

- 13 • A standardized set of mechanisms for detecting and treating error conditions should be developed so that
- 14 all languages to the extent possible could use them. This does not mean that all languages should use the
- 15 same mechanisms as there should be a variety (e.g., label parameters, auxiliary status variables), but each
- 16 of the mechanisms should be standardized.

17 **6.43.7 Bibliography**

- 18 [1] Robert W. Sebesta, Concepts of Programming Languages, 8th edition, ISBN-13: 978-0-321-49362-0, ISBN-10:
- 19 0-321-49362-1, Pearson Education, Boston, MA, 2008
- 20 [2] Carlo Ghezzi and Mehdi Jazayeri, Programming Language Concepts, 3rd edition, ISBN-0-471-10426-4, John
- 21 Wiley & Sons, 1998

22 **6.44 Termination Strategy [REU]**

23 **6.44.1 Description of application vulnerability**

24 Expectations that a system will be dependable are based on the confidence that the system will operate as

25 expected and not fail in normal use. The dependability of a system and its fault tolerance can be measured

26 through the component part's reliability, availability, safety and security. Reliability is the ability of a system or

27 component to perform its required functions under stated conditions for a specified period of time [IEEE 1990

28 glossary]. Availability is how timely and reliable the system is to its intended users. Both of these factors matter

29 highly in systems used for safety and security. In spite of the best intentions, systems may encounter a failure,

30 either from internally poorly written software or external forces such as power outages/variability, floods, or other

31 natural disasters. The reaction to a fault can affect the performance of a system and in particular, the safety and

32 security of the system and its users.

33 When the software does not terminate in the planned mechanism, safety or security is compromised, as failing in

34 an unspecified way interferes with the alternative recovery features. In safety-related systems the results can be

35 catastrophic: for other systems the result can mean failure of the complete system

36 **6.44.2 Cross reference**

- 37 JSF AV Rule: 24
- 38 MISRA C 2004: 20.11
- 39 MISRA C++ 2008: 0-3-2, 15-5-2, 15-5-3, and 18-0-3
- 40 CERT/CC guidelines: ERR04-C, ERR06-C and ENV32-C

41 **6.44.3 Mechanism of failure**

42 The reaction to a fault in a system can depend on the criticality of the part in which the fault originates. When a

43 program consists of several tasks, each task may be critical, or not. If a task is critical, it may or may not be

1 restartable by the rest of the program. Ideally, a task that detects a fault within itself should be able to halt leaving
 2 its resources available for use by the rest of the program, halt clearing away its resources, or halt the entire
 3 program. The latency of task termination and whether tasks can ignore termination signals should be clearly
 4 specified. Having inconsistent reactions to a fault can potentially be a vulnerability.

5 When a fault is detected, there are many ways in which a system can react. The quickest and most noticeable way
 6 is to fail hard, also known as fail fast or fail stop. The reaction to a detected fault is to immediately halt the system.
 7 Alternatively, the reaction to a detected fault could be to fail soft. The system would keep working with the faults
 8 present, but the performance of the system would be degraded. Systems used in a high availability environment
 9 such as telephone switching centers, e-commerce, etc. would likely use a fail soft approach. What is actually done
 10 in a fail soft approach can vary depending on whether the system is used for safety critical or security critical
 11 purposes. For fail-safe systems, such as flight controllers, traffic signals, or medical monitoring systems, there
 12 would be no effort to meet normal operational requirements, but rather to limit the damage or danger caused by the
 13 fault. A system that fails securely, such as cryptologic systems, would maintain maximum security when a fault is
 14 detected, possibly through a denial of service.

15 **6.44.4 Applicable language characteristics**

16 This vulnerability description is intended to be applicable to all languages.

17 **6.44.5 Avoiding the vulnerability or mitigating its effects**

18 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 19 • A strategy for fault handling should be decided. Consistency in fault handling should be the same with
 20 respect to critically similar parts.
- 21 • A multi-tiered approach of fault prevention, fault detection and fault reaction should be used.
- 22 • System-defined components that assist in uniformity of fault handling should be used when available. For
 23 one example, designing a "runtime constraint handler" (as described in ISO/IEC TR 24731-1) permits the
 24 application to intercept various erroneous situations and perform one consistent response, such as flushing
 25 a previous transaction and re-starting at the next one.
- 26 • When there are multiple tasks, a fault-handling policy should be specified whereby a task may
 - 27 ○ halt, and keep its resources available for other tasks (perhaps permitting restarting of the faulting
 28 task)
 - 29 ○ halt, and remove its resources (perhaps to allow other tasks to use the resources so freed, or to
 30 allow a recreation of the task)
 - 31 ○ halt, and signal the rest of the program to likewise halt.

32 **6.44.6 Implications for standardization**

33 In future standardization activities, the following items should be considered:

- 34 • Languages should consider providing a means to perform fault handling. Terminology and the means
 35 should be coordinated with other languages.

36 **6.44.7 Bibliography**

37 [None]

38 **6.45 Extra Intrinsic [LRM]**

39 **6.45.1 Description of application vulnerability**

40 Most languages define intrinsic procedures, which are easily available, or always "simply available", to any
 41 translation unit. If a translator extends the set of intrinsics beyond those defined by the standard, and the standard

1 specifies that intrinsics are selected before procedures of the same signature defined by the application, a different
2 procedure may be unexpectedly used when switching between translators.

3 **6.45.2 Cross reference**

4 **6.45.3 Mechanism of failure**

5 Most standard programming languages define a set of intrinsic procedures which may be used in any application.
6 Some language standards allow a translator to extend this set of intrinsic procedures. Some language standards
7 specify that intrinsic procedures are selected ahead of an application procedure of the same signature. This may
8 cause a different procedure to be used when switching between translators.

9 For example, most languages provide a routine to calculate the square root of a number, usually named `sqrt()`. If
10 a translator also provided, as an extension, a cube root routine, say named `cbrt()`, that extension may override
11 an application defined procedure of the same signature. If the two different `cbrt()` routines chose different branch
12 cuts when applied to complex arguments, the application could unpredictably go wrong.

13 If the language standard specifies that application defined procedures are selected ahead of intrinsic procedures of
14 the same signature, the use of the wrong procedure may mask a linking error.

15 **6.45.4 Applicable language characteristics**

16 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 17
 - Any language where translators may extend the set of intrinsic procedures and where intrinsic procedures
18 are selected ahead of application defined (or external library defined) procedures of the same signature.

19 **6.45.5 Avoiding the vulnerability or mitigating its effects**

20 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 21
 - Use whatever language features are available to mark a procedure as language defined or application
22 defined.
 - Be aware of the documentation for every translator in use and avoid using procedure signatures matching
23 those defined by the translator as extending the standard set.
24

25 **6.45.6 Implications for standardization**

26 In future standardization activities, the following items should be considered:

- 27
 - Clearly state whether translators can extend the set of intrinsic procedures or not.
 - Clearly state what the precedence is for resolving collisions.
 - Clearly provide ways to mark a procedure signature as being the intrinsic or an application provided
29 procedure.
 - Require that a diagnostic is issued when an application procedure matches the signature of an intrinsic
31 procedure.
32

33 **6.45.7 Bibliography**

34 **6.46 Type-breaking Reinterpretation of Data [AMV]**

35 **6.46.1 Description of application vulnerability**

36 In most cases, objects in programs are assigned locations in processor storage to hold their value. If the same
37 storage space is assigned to more than one object—either statically or temporarily—then a change in the value of

1 one object will have an effect on the value of the other. Furthermore, if the representation of the value of an object
 2 is reinterpreted as being the representation of the value of an object with a different type, unexpected results may
 3 occur.

4 **6.46.2 Cross reference**

5 JSF AV Rules 153 and 183
 6 MISRA 2004: 18.2, 18.3, and 18.4
 7 MISRA C++ 2008: 4-5-1 to 4-5-3, 4-10-1, 4-10-2, and 5-0-3 to 5-0-9
 8 CERT/CC guidelines: MEM08-C

9 **6.46.3 Mechanism of failure**

10 Sometimes there is a legitimate need for applications to place different interpretations upon the same stored
 11 representation of data. The most fundamental example is a program loader that treats a binary image of a program
 12 as data by loading it, and then treats it as a program by invoking it. Most programming languages permit type-
 13 breaking reinterpretation of data, however, some offer less error prone alternatives for commonly encountered
 14 situations.

15 Type-breaking reinterpretation of representation presents obstacles to human understanding of the code, the ability
 16 of tools to perform effective static analysis, and the ability of code optimizers to do their job.

17 Examples include:

- 18 • Providing alternative mappings of objects into blocks of storage performed either statically (e.g., Fortran
 19 `COMMON`) or dynamically (e.g., pointers).
- 20 • Union types, particularly unions that do not have a discriminant stored as part of the data structure.
- 21 • Operations that permit a stored value to be interpreted as a different type (e.g., treating the representation
 22 of a pointer as an integer).

23 In all of these cases accessing the value of an object may produce an unanticipated result.

24 A related problem, the aliasing of parameters, occurs in languages that permit call by reference because
 25 supposedly distinct parameters might refer to the same storage area, or a parameter and a non-local object might
 26 refer to the same storage area. That vulnerability is described in Passing Parameters and Return Values [CSJ].

27 **6.46.4 Applicable language characteristics**

28 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 29 • A programming language that permits multiple interpretations of the same bit pattern.

30 **6.46.5 Avoiding the vulnerability or mitigating its effects**

31 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 32 • Programmers should avoid reinterpretation performed as a matter of convenience; for example, using an
 33 integer pointer to manipulate character string data should be avoided. When type-breaking reinterpretation
 34 is necessary, it should be carefully documented in the code. However this vulnerability cannot be
 35 completely avoided because some applications view stored data in alternative ways.
- 36 • When using union types it is preferable to use discriminated unions. This is a form of a union where a
 37 stored value indicates which interpretation is to be placed upon the data. Some languages (e.g., variant
 38 records in Ada) enforce the view of data indicated by the value of the discriminant. If the language does not
 39 enforce the interpretation (e.g., equivalence in Fortran and union in C and C++), then the code should
 40 implement an explicit discriminant and check its value before accessing the data in the union, or use some
 41 other mechanism to ensure that correct interpretation is placed upon the data value.

- 1 • Operations that reinterpret the same stored value as representing a different type should be avoided. It is
 2 easier to avoid such operations when the language clearly identifies them. For example, the name of Ada's
 3 `Unchecked_Conversion` function explicitly warns of the problem. A much more difficult situation occurs
 4 when pointers are used to achieve type reinterpretation. Some languages perform type-checking of
 5 pointers and place restrictions on the ability of pointers to access arbitrary locations in storage. Others
 6 permit the free use of pointers. In such cases, code must be carefully reviewed in a search for unintended
 7 reinterpretation of stored values. Therefore it is important to explicitly comment the source code where
 8 *intended* reinterpretations occur.
- 9 • Static analysis tools may be helpful in locating situations where unintended reinterpretation occurs. On the
 10 other hand, the presence of reinterpretation greatly complicates static analysis for other problems, so it
 11 may be appropriate to segregate intended reinterpretation operations into distinct subprograms.

12 6.46.6 Implications for standardization

13 In future standardization activities, the following items should be considered:

- 14 • Because the ability to perform reinterpretation is sometimes necessary, but the need for it is rare,
 15 programming language designers might consider putting caution labels on operations that permit
 16 reinterpretation. For example, the operation in Ada that permits unconstrained reinterpretation is called
 17 `Unchecked_Conversion`.
- 18 • Because of the difficulties with indiscriminated unions, programming language designers might consider
 19 offering union types that include distinct discriminants with appropriate enforcement of access to objects.
 20

21 6.46.7 Bibliography

- 22 [1] Robert W. Sebesta, Concepts of Programming Languages, 8th edition, ISBN-13: 978-0-321-49362-0, ISBN-10:
 23 0-321-49362-1, Pearson Education, Boston, MA, 2008
 24 [2] Carlo Ghezzi and Mehdi Jazayeri, Programming Language Concepts, 3rd edition, ISBN-0471-10426-4, John
 25 Wiley & Sons, 1998

26 6.47 Memory Leak [XYL]

27 6.47.1 Description of application vulnerability

28 A memory leak occurs when software does not release allocated memory after it ceases to be used. Repeated
 29 occurrences of a memory leak can consume considerable amounts of available memory. A memory leak can be
 30 exploited by attackers to generate denial-of-service by causing the program to execute repeatedly a sequence that
 31 triggers the leak. Moreover, a memory leak can cause any long-running critical program to shutdown prematurely.

32 6.47.2 Cross reference

33 CWE:

34 401. Failure to Release Memory Before Removing Last Reference (aka 'Memory Leak')

35 JSF AV Rule: 206

36 MISRA C 2004: 20.4

37 CERT/CC guidelines: MEM00-C and MEM31-C

38 6.47.3 Mechanism of failure

39 As a process or system runs, any memory taken from dynamic memory and not returned or reclaimed (by the
 40 runtime system or a garbage collector) after it ceases to be used, may result in future memory allocation requests
 41 failing for lack of free space. Alternatively, memory claimed and returned can cause the heap to fragment, which
 42 will eventually result in an inability to take the necessary size storage. Either condition will result in a memory
 43 exhaustion exception, and program termination or a system crash.

1 If an attacker can determine the cause of an existing memory leak, the attacker may be able to cause the
2 application to leak quickly and therefore cause the application to crash.

3 **6.47.4 Applicable language characteristics**

4 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 5 • Languages that support mechanisms to dynamically allocate memory and reclaim memory under program
6 control.

7 **6.47.5 Avoiding the vulnerability or mitigating its effects**

8 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 9 • Use of garbage collectors that reclaim memory that will never be used by the application again. Some
10 garbage collectors are part of the language while others are add-ons
- 11 • Allocating and freeing memory in different modules and levels of abstraction may make it difficult for
12 developers to match requests to free storage with the appropriate storage allocation request. This may
13 cause confusion regarding when and if a block of memory has been allocated or freed, leading to memory
14 leaks. To avoid these situations, it is recommended that memory be allocated and freed at the same level
15 of abstraction, and ideally in the same code module.
- 16 • Storage pools are a specialized memory mechanism where all of the memory associated with a class of
17 objects is allocated from a specific bounded region. When used with strong typing one can ensure a strong
18 relationship between pointers and the space accessed such that storage exhaustion in one pool does not
19 affect the code operating on other memory.
- 20 • Memory leaks can be eliminated by avoiding the use of dynamically allocated storage entirely, or by doing
21 initial allocation exclusively and never allocating once the main execution commences. For safety-critical
22 systems and long running systems, the use of dynamic memory is almost always prohibited, or restricted to
23 the initialization phase of execution.
- 24 • Use static analysis that is capable of detecting when allocated storage is no longer used and has not been
25 freed (for reuse).

26 **6.47.6 Implications for standardization**

27 In future standardization activities, the following items should be considered:

- 28 • Languages can provide syntax and semantics to guarantee program-wide that dynamic memory is not
29 used (such as the configuration `pragmas` feature offered by some programming languages).
- 30 • Languages can document or can specify that implementations must document choices for dynamic
31 memory management algorithms, to help designers decide on appropriate usage patterns and recovery
32 techniques as necessary.

33 **6.47.7 Bibliography**

34 [None]

35 **6.48 Argument Passing to Library Functions [TRJ]**

36 **6.48.1 Description of application vulnerability**

37 Libraries that supply objects or functions are in most cases not required to check the validity of parameters passed
38 to them. In those cases where parameter validation is required there might not be adequate parameter validation.

39 **6.48.2 Cross reference**

40 CWE:

- 1 114. Process Control
- 2 JSF AV Rules 16, 18, 19, 20, 21, 22, 23, 24, and 25
- 3 MISRA C 2004: 20.2, 20.3, 20.4, 20.6, 20.7, 20.8, 20.9, 20.10, 20.11, and 20.12
- 4 MISRA C++ 2008: 17-0-1, 17-0-5, 18-0-2, 18-0-3, 18-0-4, 18-2-1, 18-7-1 and 27-0-1
- 5 CERT/CC guidelines: INT03-C and STR07-C

6 **6.48.3 Mechanism of failure**

7 When calling a library, either the calling function or the library may make assumptions about parameters. For
8 example, it may be assumed by a library that a parameter is non-zero so division by that parameter is performed
9 without checking the value. Sometimes some validation is performed by the calling function, but the library may
10 use the parameters in ways that were unanticipated by the calling function resulting in a potential vulnerability.
11 Even when libraries do validate parameters, their response to an invalid parameter is usually undefined and can
12 cause unanticipated results.

13 **6.48.4 Applicable language characteristics**

14 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 15
- 16 • Languages providing or using libraries that do not validate the parameters accepted by functions, methods
17 and objects.
- 18

19 **6.48.5 Avoiding the vulnerability or mitigating its effects**

20 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 21 • Libraries should be defined so that as many parameters as possible are validated.
- 22 • Libraries should be defined to validate any values passed to the library before the value is used.
- 23 • Develop wrappers around library functions that check the parameters before calling the function.
- 24 • Demonstrate statically that the parameters are never invalid.
- 25 • Use only libraries known to have been developed with consistent and validated interface requirements.

26 It is noted that several approaches can be taken, some work best if used in conjunction with each other.

27 **6.48.6 Implications for standardization**

28 In future standardization activities, the following items should be considered:

- 29 • All languages that define a support library should consider removing most if not all cases of undefined
30 behaviour from the library sections.
- 31 • Libraries should be defined so that all parameters are validated.

32 **6.48.7 Bibliography**

33 Holtzmann-7

34 **6.49 Dynamically-linked Code and Self-modifying Code [NYY]**

35 **6.49.1 Description of application vulnerability**

36 Code that is dynamically linked may be different from the code that was tested. This may be the result of replacing
37 a library with another of the same name or by altering an environment variable such as `LD_LIBRARY_PATH` on
38 UNIX platforms so that a different directory is searched for the library file. Executing code that is different than that
39 which was tested may lead to unanticipated errors or intentional malicious activity.

1 On some platforms, and in some languages, instructions can modify other instructions in the code space.
 2 Historically self-modifying code was needed for software that was required to run on a platform with very limited
 3 memory. It is now primarily used (or misused) to hide functionality of software and make it more difficult to reverse
 4 engineer or for specialty applications such as graphics where the algorithm is tuned at runtime to give better
 5 performance. Self-modifying code can be difficult to write correctly and even more difficult to test and maintain
 6 correctly leading to unanticipated errors.

7 **6.49.2 Cross reference**

8 JSF AV Rule: 2

9 **6.49.3 Mechanism of failure**

10 Through the alteration of a library file or environment variable, the code that is dynamically linked may be different
 11 from the code which was tested resulting in different functionality.

12 On some platforms, a pointer-to-data can erroneously be given an address value that designates a location in the
 13 instruction space. If subsequently a modification is made through that pointer, then an unanticipated behaviour can
 14 result.

15 **6.49.4 Applicable language characteristics**

16 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 17 • Languages that allow a pointer-to-data to be assigned an address value that designates a location in the
- 18 instruction space.
- 19 • Languages that allow execution of code that exists in data space.
- 20 • Languages that permit the use of dynamically linked or shared libraries.
- 21 • Languages that execute on an OS that permits program memory to be both writable and executable.

22 **6.49.5 Avoiding the vulnerability or mitigating its effects**

23 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 24 • Verify that the dynamically linked or shared code being used is the same as that which was tested.
- 25 • Do not write self-modifying code except in extremely rare instances. Most software applications should
- 26 never have a requirement for self-modifying code.
- 27 • In those extremely rare instances where its use is justified, self-modifying code should be very limited and
- 28 heavily documented.

29 **6.49.6 Implications for standardization**

30 In future standardization activities, the following items should be considered:

- 31 • Languages should consider providing a means so that a program can either automatically or manually
- 32 check that the digital signature of a library matches the one in the compile/test environment.

33 **6.49.7 Bibliography**

34 [None]

1 **6.50 Library Signature [NSQ]**

2 **6.50.1 Description of application vulnerability**

3 Programs written in modern languages may use libraries written in other languages than the program
4 implementation language. If the library is large, the effort of adding signatures for all of the functions use by hand
5 may be tedious and error-prone. Portable cross-language signatures will require detailed understanding of both
6 languages, which a programmer may lack.

7 Integrating two or more programming languages into a single executable relies upon knowing how to interface the
8 function calls, argument list and global data structures so the symbols match in the object code during linking.

9 Byte alignment can be a source of data corruption if memory boundaries between the programming languages are
10 different. Each language may also align structure data differently.

11 **6.50.2 Cross reference**

12 MISRA C 2004: 1.3
13 MISRA C++ 2008: 1-0-2

14 **6.50.3 Mechanism of failure**

15 When the library and the application in which it is to be used are written in different languages, the specification of
16 signatures is complicated by inter-language issues.

17 As used in this vulnerability description, the term library includes the interface to the operating system, which may
18 be specified only for the language used to code the operating system itself. In this case, any program written in
19 any other language faces the inter-language interoperability issue of creating a fully-functional signature.

20 When the application language and the library language are different, then the ability to specify signatures
21 according to either standard may not exist, or be very difficult. Thus, a translator-by-translator solution may be
22 needed, which maximizes the probability of incorrect signatures (since the solution must be recreated for each
23 translator pair). Incorrect signatures may or may not be caught during the linking phase.

24 **6.50.4 Applicable language characteristics**

25 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 26 • Languages that do not specify how to describe signatures for subprograms written in other languages.

27 **6.50.5 Avoiding the vulnerability or mitigating its effects**

28 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 29 • Use tools to create the signatures.
- 30 • Avoid using translator options or language features to reference library subprograms without proper
31 signatures.

32 **6.50.6 Implications for standardization**

33 In future standardization activities, the following items should be considered:

- 34 • Provide correct linkage even in the absence of correctly specified procedure signatures. (Note that this
35 may be very difficult where the original source code is unavailable.)
- 36 • Provide specified means to describe the signatures of subprograms.

1 6.50.7 Bibliography

2 [None]

3 6.51 Unanticipated Exceptions from Library Routines [HJW]

4 6.51.1 Description of application vulnerability

5 A library in this context is taken to mean a set of software routines produced outside the control of the main
6 application developer, usually by a third party, and where the application developer may not have access to the
7 source. In such circumstances the application developer has limited knowledge of the library functions, other than
8 from their behavioural interface.

9 Whilst the use of libraries can present a number of vulnerabilities, the focus of this vulnerability is any undesirable
10 behaviour that a library routine may exhibit, in particular the generation of unexpected exceptions.

11 6.51.2 Cross reference

12 JSF AV Rule: 208

13 MISRA C 2004: 3.6, 20.3

14 MISRA C++ 2008: 15-3-1, 15-3-2, 17-0-4

15 6.51.3 Mechanism of failure

16 In some languages, unhandled exceptions lead to implementation dependant behaviour. This can include
17 immediate termination, without for example, releasing previously allocated resources. If a library routine throws an
18 unanticipated exception, this undesirable behaviour may result.

19 It should be noted that the considerations of [NZN], Returning Error Status, are also relevant here.

20 6.51.4 Applicable language characteristics

21 This vulnerability description is intended to be applicable to languages with the following characteristics:

- 22 • Languages that can link previously developed library code (where the developer and compiler don't have
- 23 access to the library source).
- 24 • Languages that permit exceptions to be thrown but do not require handlers for them.

25 6.51.5 Avoiding the vulnerability or mitigating its effects

26 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 27 • All library calls should be wrapped within a 'catch-all' exception handler (if the language supports such a
- 28 construct), so that any unanticipated exceptions can be caught and handled appropriately. This wrapping
- 29 may be done for each library function call or for the entire behaviour of the program, e.g. having the
- 30 exception handler in main for C++. However, note that the later isn't a complete solution, as static objects
- 31 are constructed before main is entered and are destroyed after it has been exited. Consequently, MISRA
- 32 C++ [20] bars class constructors and destructors from throwing exceptions (unless handled locally).
- 33 • An alternative approach would be to use only library routines for which all possible exceptions are
- 34 specified.

35 6.51.6 Implications for standardization

36 In future standardization activities, the following items should be considered:

- 1 • Languages that provide exceptions should provide a mechanism for catching all possible exceptions (for
2 example, a 'catch-all' handler). The behaviour of the program when encountering an unhandled exception
3 should be fully defined.
4 • Languages should provide a mechanism to determine which exceptions might be thrown by a called library
5 routine.

6 **6.51.7 Bibliography**

7 [None]

1 **7. Application Vulnerabilities**

2 **7.1 Adherence to Least Privilege [XYN]**

3 **7.1.1 Description of application vulnerability**

4 Failure to adhere to the principle of least privilege amplifies the risk posed by other vulnerabilities.

5 **7.1.2 Cross reference**

6 CWE:
7 250. Design Principle Violation: Failure to Use Least Privilege
8 CERT/CC guidelines: POS02-C

9 **7.1.3 Mechanism of failure**

10 This vulnerability type refers to cases in which an application grants greater access rights than necessary.
11 Depending on the level of access granted, this may allow a user to access confidential information. For example,
12 programs that run with root privileges have caused innumerable Unix security disasters. It is imperative that you
13 carefully review privileged programs for all kinds of security problems, but it is equally important that privileged
14 programs drop back to an unprivileged state as quickly as possible in order to limit the amount of damage that an
15 overlooked vulnerability might be able to cause. Privilege management functions can behave in some less-than-
16 obvious ways, and they have different quirks on different platforms. These inconsistencies are particularly
17 pronounced if you are transitioning from one non-root user to another. Signal handlers and spawned processes run
18 at the privilege of the owning process, so if a process is running as root when a signal fires or a sub-process is
19 executed, the signal handler or sub-process will operate with root privileges. An attacker may be able to leverage
20 these elevated privileges to do further damage. To grant the minimum access level necessary, first identify the
21 different permissions that an application or user of that application will need to perform their actions, such as file
22 read and write permissions, network socket permissions, and so forth. Then explicitly allow those actions while
23 denying all else.

24 **7.1.4 Avoiding the vulnerability or mitigating its effects**

25 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 26 • Very carefully manage the setting, management and handling of privileges. Explicitly manage trust zones in
- 27 the software.
- 28 • Follow the principle of least privilege when assigning access rights to entities in a software system.

29 **7.1.5 Bibliography**

30 [None]

31 **7.2 Privilege Sandbox Issues [XYO]**

32 **7.2.1 Description of application vulnerability**

33 A variety of vulnerabilities occur with improper handling, assignment, or management of privileges. These are
34 especially present in sandbox environments, although it could be argued that any privilege problem occurs within
35 the context of some sort of sandbox.

36 **7.2.2 Cross reference**

37 CWE:

- 1 266. Incorrect Privilege Assignment
- 2 267. Privilege Defined With Unsafe Actions
- 3 268. Privilege Chaining
- 4 269. Privilege Management Error
- 5 270. Privilege Context Switching Error
- 6 272. Least Privilege Violation
- 7 273. Failure to Check Whether Privileges were Dropped Successfully
- 8 274. Failure to Handle Insufficient Privileges
- 9 276. Insecure Default Permissions
- 10 CERT/CC guidelines: POS36-C

11 **7.2.3 Mechanism of failure**

12 The failure to drop system privileges when it is reasonable to do so is not an application vulnerability by itself. It
13 does, however, serve to significantly increase the severity of other vulnerabilities. According to the principle of least
14 privilege, access should be allowed only when it is absolutely necessary to the function of a given system, and only
15 for the minimal necessary amount of time. Any further allowance of privilege widens the window of time during
16 which a successful exploitation of the system will provide an attacker with that same privilege.

17 Many situations could lead to a mechanism of failure:

- 18 • A product could incorrectly assign a privilege to a particular entity.
- 19 • A particular privilege, role, capability, or right could be used to perform unsafe actions that were not
20 intended, even when it is assigned to the correct entity. (Note that there are two separate sub-categories
21 here: privilege incorrectly allows entities to perform certain actions; and the object is incorrectly accessible
22 to entities with a given privilege.)
- 23 • Two distinct privileges, roles, capabilities, or rights could be combined in a way that allows an entity to
24 perform unsafe actions that would not be allowed without that combination.
- 25 • The software may not properly manage privileges while it is switching between different contexts that cross
26 privilege boundaries.
- 27 • A product may not properly track, modify, record, or reset privileges.
- 28 • In some contexts, a system executing with elevated permissions will hand off a process/file etc. to another
29 process/user. If the privileges of an entity are not reduced, then elevated privileges are spread throughout
30 a system and possibly to an attacker.
- 31 • The software may not properly handle the situation in which it has insufficient privileges to perform an
32 operation.
- 33 • A program, upon installation, may set insecure permissions for an object.

34 **7.2.4 Avoiding the vulnerability or mitigating its effects**

35 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 36 • The principle of least privilege when assigning access rights to entities in a software system should be
37 followed. The setting, management and handling of privileges should be managed very carefully. Upon
38 changing security privileges, one should ensure that the change was successful.
- 39 • Consider following the principle of separation of privilege. Require multiple conditions to be met before
40 permitting access to a system resource.
- 41 • Trust zones in the software should be explicitly managed. If at all possible, limit the allowance of system
42 privilege to small, simple sections of code that may be called atomically.
- 43 • As soon as possible after acquiring elevated privilege to call a privileged function such as `chroot()`, the
44 program should drop root privilege and return to the privilege level of the invoking user.
- 45 • In newer Windows implementations, make sure that the process token has the `SeImpersonatePrivilege`.

46 **7.2.5 Bibliography**

47 [None]

1 **7.3 Executing or Loading Untrusted Code [XYS]**

2 **7.3.1 Description of application vulnerability**

3 Executing commands or loading libraries from an untrusted source or in an untrusted environment can cause an
4 application to execute malicious commands (and payloads) on behalf of an attacker.

5 **7.3.2 Cross reference**

6 CWE:

7 114. Process Control

8 CERT/CC guidelines: PRE09-C, ENV02-C, and ENV03-C

9 **7.3.3 Mechanism of failure**

10 Process control vulnerabilities take two forms:

- 11 • An attacker can change the command that the program executes so that the attacker explicitly controls
12 what the command is.
- 13 • An attacker can change the environment in which the command executes so that the attacker implicitly
14 controls what the command means.

15 Considering only the first scenario, the possibility that an attacker may be able to control the command that is
16 executed, process control vulnerabilities occur when:

- 17 • Data enters the application from a source that is not trusted.
- 18 • The data is used as or as part of a string representing a command that is executed by the application.
- 19 • By executing the command, the application gives an attacker a privilege or capability that the attacker
20 would not otherwise have.

21 **7.3.4 Avoiding the vulnerability or mitigating its effects**

22 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 23 • Libraries that are loaded should be well understood and come from a trusted source with a digital
24 signature. The application can execute code contained in native libraries, which often contain calls that
25 are susceptible to other security problems, such as buffer overflows or command injection.
- 26 • All native libraries should be validated.
- 27 • Determine if the application requires the use of the native library. It can be very difficult to determine what
28 these libraries actually do, and the potential for malicious code is high.
- 29 • To help prevent buffer overflow attacks, validate all input to native calls for content and length.
- 30 • If the native library does not come from a trusted source, review the source code of the library. The library
31 should be built from the reviewed source before using it.

32 **7.3.5 Bibliography**

33 [None]

34 **7.4 Unspecified Functionality [BVQ]**

35 **7.4.1 Description of application vulnerability**

36 *Unspecified functionality* is code that may be executed, but whose behaviour does not contribute to the
37 requirements of the application. While this may be no more than an amusing 'Easter Egg', like the flight simulator in
38 a spreadsheet, it does raise questions about the level of control of the development process.

1 In a security-critical environment particularly, the developer of an application could include a 'trap-door' to allow
2 illegitimate access to the system on which it is eventually executed, irrespective of whether the application has
3 obvious security requirements.

4 **7.4.2 Cross reference**

- 5 JSF Rule: 127
- 6 MISRA C 2004: 2.2, 2.3, 2.4, and 14.1
- 7 XYQ: Dead and Deactivated code.

8 **7.4.3 Mechanism of failure**

9 Unspecified functionality is not a software vulnerability per se, but more a development issue. In some cases,
10 unspecified functionality may be added by a developer without the knowledge of the development organization. In
11 other cases, typically Easter Eggs, the functionality is unspecified as far as the user is concerned (nobody buys a
12 spreadsheet expecting to find it includes a flight simulator), but is specified by the development organization. In
13 effect they only reveal a subset of the program's behaviour to the users.

14 In the first case, one would expect a well managed development environment to discover the additional
15 functionality during validation and verification. In the second case, the user is relying on the supplier not to release
16 harmful code.

17 In effect, a program's requirements are 'the program should behave in the following manner and do nothing
18 else'. The 'and do nothing else' clause is often not explicitly stated, and can be difficult to demonstrate.

19 **7.4.4 Avoiding the vulnerability or mitigating its effects**

20 End users can avoid the vulnerability or mitigate its ill effects in the following ways:

- 21 • Programs and development tools that are to be used in critical applications should come from a developer
22 who uses a recognized and audited development process for the development of those programs and
23 tools. For example: ISO 9001 or CMMI@.
- 24 • The development process should generate documentation showing traceability from source code to
25 requirements, in effect answering 'why is this unit of code in this program?'. Where unspecified functionality
26 is there for a legitimate reason (e.g., diagnostics required for developer maintenance or enhancement), the
27 documentation should also record this. It is not unreasonable for customers of bespoke critical code to ask
28 to see such traceability as part of their acceptance of the application

29 **7.4.5 Bibliography**

30 [None]

31 **7.5 Distinguished Values in Data Types [KLK]**

32 **7.5.1 Description of application vulnerability**

33 Sometimes, in a type representation, certain values are distinguished as not being members of the type, but rather
34 as providing auxiliary information. Examples include special characters used as string terminators, distinguished
35 values used to indicate out of type entries in SQL (Structured Query Language) database fields, and sentinels used
36 to indicate the bounds of queues or other data structures. When the usage pattern of code containing distinguished
37 values is changed, it may happen that the distinguished value happens to coincide with a legitimate in-type value.
38 In such a case, the value is no longer distinguishable from an in-type value and the software will no longer produce
39 the intended results.

1 7.5.2 Cross reference

2 CWE

3 20: Improper input validation

4 137: Representation errors

5 JSF AV Rule: 151

6 7.5.3 Mechanism of failure

7 A "distinguished value" or a "magic number" in the representation of a data type might be used to represent out-of-
8 type information. Some examples include the following:

- 9
- 10 • The use of a special code, e.g., "00", to indicate the termination of a coded character string.
 - 11 • The use of a special value, e.g., "999...9", as the indication that the actual value is either not known or is invalid.

12 If the use of the software is later generalized, the once-special value can become indistinguishable from valid data.
13 Note that the problem may occur simply if the pattern of usage of the software is changed from that anticipated by
14 the software's designers. It may also occur if the software is reused in other circumstances.

15 An example of a change in the pattern of usage is this: An organization logs visitors to its buildings by recording
16 their names and national identity numbers or social security numbers in a database. Of course, some visitors
17 legitimately don't have or don't know their social security number, so the receptionists enter numbers to "make the
18 computer happy." Receptionists at one building have adopted the convention of using the code "555-55-5555" to
19 designate children of employees. Receptionists at another building have used the same code to designate foreign
20 nationals. When the databases are merged, the children are reclassified as foreign nationals or vice-versa
21 depending on which set of receptionists are using the newly merged database.

22 An example of an unanticipated change due to reuse is this: Suppose a software component analyzes radar data,
23 recording data every degree of azimuth from 0 to 359. Packets of data are sent to other components for
24 processing, updating displays, recording, and so on. Since all degree values are non-negative, a distinguished
25 value of -1 is used as a signal to stop processing, compute summary data, close files, and so on. Many of the
26 components are to be reused in a new system with a new radar analysis component. However the new component
27 represents direction by numbers in the range -180 degrees to 179 degrees. When an azimuth value of -1 is
28 provided, the downstream components will interpret that as the indication to stop processing. If the magic value is
29 changed to, say, -999, the software is still at risk of failing when future enhancements (say, counting accumulated
30 degrees on complete revolutions) bring -999 into the range of valid data.

31 Distinguished values should be avoided. Instead, the software should be designed to use distinct variables to
32 encode the desired out-of-type information. For example, the length of a character string might be encoded in a
33 dope vector and validity of data entries might be encoded in distinct Boolean values.

34 7.5.4 Avoiding the vulnerability or mitigating its effects

35 End users can avoid the vulnerability or mitigate its ill effects in the following ways:

- 36
- 37 • Use auxiliary variables (perhaps enclosed in variant records) to encode out-of-type information.
 - 38 • Use enumeration types to convey category information. Do not rely upon large ranges of integers, with
39 distinguished values having special meanings.
 - Use named constants to make it easier to change distinguished values.

40 7.5.5 Bibliography

41 [None]

1 **7.6 Memory Locking [XZX]**

2 **7.6.1 Description of application vulnerability**

3 Sensitive data stored in memory that was not locked or that has been improperly locked may be written to swap
4 files on disk by the virtual memory manager.

5 **7.6.2 Cross reference**

6 CWE:
7 591. Sensitive Data Storage in Improperly Locked Memory
8 CERT/CC guidelines: MEM06-C

9 **7.6.3 Mechanism of failure**

10 Sensitive data that is not kept cryptographically secure may become visible to an attacker by any of several
11 mechanisms. Some operating systems may write memory to swap or page files that may be visible to an attacker.
12 Some operating systems may provide mechanisms to examine the physical memory of the system or the virtual
13 memory of another application. Application debuggers may be able to stop the target application and examine or
14 alter memory.

15 **7.6.4 Avoiding the vulnerability or mitigating its effects**

16 In almost all cases, these attacks require elevated or appropriate privilege.

17 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 18 • Remove debugging tools from production systems.
- 19 • Log and audit all privileged operations.
- 20 • Identify data that needs to be protected and use appropriate cryptographic and other data obfuscation
- 21 techniques to avoid keeping plaintext versions of this data in memory or on disk.
- 22 • If the operating system allows, clear the swap file on shutdown.

23 **Note:** Several implementations of the POSIX `mlock()` and the Microsoft Windows `VirtualLock()` functions
24 will prevent the named memory region from being written to a swap or page file. However, such usage is not
25 portable.

26 Systems that provide a "hibernate" facility (such as laptops) will write all of physical memory to a file that may be
27 visible to an attacker on resume.

28 **7.6.5 Bibliography**

29 [None]

30 **7.7 Resource Exhaustion [XZP]**

31 **7.7.1 Description of application vulnerability**

32 The application is susceptible to generating and/or accepting an excessive number of requests that could
33 potentially exhaust limited resources, such as memory, file system storage, database connection pool entries, or
34 CPU. This could ultimately lead to a denial of service that could prevent any other applications from accessing
35 these resources.

1 7.7.2 Cross reference

2 CWE:
3 400. Resource Exhaustion

4 7.7.3 Mechanism of failure

5 There are two primary failures associated with resource exhaustion. The most common result of resource
6 exhaustion is denial of service. In some cases an attacker or a defect may cause a system to fail in an unsafe or
7 insecure fashion by causing an application to exhaust the available resources.

8 Resource exhaustion issues are generally understood but are far more difficult to prevent. Taking advantage of
9 various entry points, an attacker could craft a wide variety of requests that would cause the site to consume
10 resources. Database queries that take a long time to process are good *DoS* (Denial of Service) targets. An
11 attacker would only have to write a few lines of Perl code to generate enough traffic to exceed the site's ability to
12 keep up. This would effectively prevent authorized users from using the site at all.

13 Resources can be exhausted simply by ensuring that the target machine must do much more work and consume
14 more resources in order to service a request than the attacker must do to initiate a request. Prevention of these
15 attacks requires either that the target system either recognizes the attack and denies that user further access for a
16 given amount of time or uniformly throttles all requests in order to make it more difficult to consume resources more
17 quickly than they can again be freed. The first of these solutions is an issue in itself though, since it may allow
18 attackers to prevent the use of the system by a particular valid user. If the attacker impersonates the valid user, he
19 may be able to prevent the user from accessing the server in question. The second solution is simply difficult to
20 effectively institute and even when properly done, it does not provide a full solution. It simply makes the attack
21 require more resources on the part of the attacker.

22 The final concern that must be discussed about issues of resource exhaustion is that of systems which "fail open."
23 This means that in the event of resource consumption, the system fails in such a way that the state of the system
24 — and possibly the security functionality of the system — are compromised. A prime example of this can be found
25 in old switches that were vulnerable to "macof" attacks (so named for a tool developed by Dugsong). These attacks
26 flooded a switch with random IP and MAC address combinations, therefore exhausting the switch's cache, which
27 held the information of which port corresponded to which MAC addresses. Once this cache was exhausted, the
28 switch would fail in an insecure way and would begin to act simply as a hub, broadcasting all traffic on all ports and
29 allowing for basic sniffing attacks.

30 7.7.4 Avoiding the vulnerability or mitigating its effects

31 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 32 • Implement throttling mechanisms into the system architecture. The best protection is to limit the amount of
33 resources that an application can cause to be expended. A strong authentication and access control model
34 will help prevent such attacks from occurring in the first place. The authentication application should be
35 protected against denial of service attacks as much as possible. Limiting the database access, perhaps by
36 caching result sets, can help minimize the resources expended. To further limit the potential for a denial of
37 service attack, consider tracking the rate of requests received from users and blocking requests that
38 exceed a defined rate threshold.
- 39 • Ensure that applications have specific limits of scale placed on them, and ensure that all failures in
40 resource allocation cause the application to fail safely.

41 7.7.5 Bibliography

42 [None]

1 7.8 Injection [RST]

2 7.8.1 Description of application vulnerability

3 Injection problems span a wide range of instantiations. The basic form of this weakness involves the software
 4 allowing injection of additional data in input data in order to alter the control flow of the process. Command
 5 injection problems are a subset of injection problem, in which the process can be tricked into calling external
 6 processes of an attacker's choice through the injection of command syntax into the input data. Multiple
 7 leading/internal/trailing special elements injected into an application through input can be used to compromise a
 8 system. As data is parsed, improperly handled multiple leading special elements may cause the process to take
 9 unexpected actions that result in an attack. Software may allow the injection of special elements that are non-
 10 typical but equivalent to typical special elements with control implications. This frequently occurs when the product
 11 has protected itself against special element injection. Software may allow inputs to be fed directly into an output file
 12 that is later processed as code, e.g., a library file or template. Line or section delimiters injected into an application
 13 can be used to compromise a system.

14 Many injection attacks involve the disclosure of important information -- in terms of both data sensitivity and
 15 usefulness in further exploitation. In some cases injectable code controls authentication; this may lead to a remote
 16 vulnerability. Injection attacks are characterized by the ability to significantly change the flow of a given process,
 17 and in some cases, to the execution of arbitrary code. Data injection attacks lead to loss of data integrity in nearly
 18 all cases as the control-plane data injected is always incidental to data recall or writing. Often the actions performed
 19 by injected control code are not logged.

20 SQL injection attacks are a common instantiation of injection attack, in which SQL commands are injected into
 21 input in order to effect the execution of predefined SQL commands. Since SQL databases generally hold sensitive
 22 data, loss of confidentiality is a frequent problem with SQL injection vulnerabilities. If poorly implemented SQL
 23 commands are used to check user names and passwords, it may be possible to connect to a system as another
 24 user with no previous knowledge of the password. If authorization information is held in a SQL database, it may be
 25 possible to change this information through the successful exploitation of the SQL injection vulnerability. Just as it
 26 may be possible to read sensitive information, it is also possible to make changes or even delete this information
 27 with a SQL injection attack.

28 Injection problems encompass a wide variety of issues — all mitigated in very different ways. The most important
 29 issue to note is that all injection problems share one thing in common — they allow for the injection of control data
 30 into the user controlled data. This means that the execution of the process may be altered by sending code in
 31 through legitimate data channels, using no other mechanism. While buffer overflows and many other flaws involve
 32 the use of some further issue to gain execution, injection problems need only for the data to be parsed. Many
 33 injection attacks involve the disclosure of important information in terms of both data sensitivity and usefulness in
 34 further exploitation. In some cases injectable code controls authentication, this may lead to a remote vulnerability.

35 7.8.2 Cross reference

36 CWE:

- 37 74. Failure to Sanitize Data into a Different Plane ('Injection')
- 38 76. Failure to Resolve Equivalent Special Elements into a Different Plane
- 39 78. Failure to Sanitize Data into an OS Command (aka 'OS Command Injection')
- 40 90. Failure to Sanitize Data into LDAP Queries (aka 'LDAP Injection')
- 41 91. XML Injection (aka Blind XPath Injection)
- 42 92. Custom Special Character Injection
- 43 95. Insufficient Control of Directives in Dynamically Code Evaluated Code (aka 'Eval Injection')
- 44 97. Failure to Sanitize Server-Side Includes (SSI) Within a Web Page
- 45 98. Insufficient Control of Filename for Include/Require Statement in PHP Program (aka 'PHP File Inclusion')
- 46 99. Insufficient Control of Resource Identifiers (aka 'Resource Injection')
- 47 144. Failure to Sanitize Line Delimiters
- 48 145. Failure to Sanitize Section Delimiters
- 49 161. Failure to Sanitize Multiple Leading Special Elements
- 50 163. Failure to Sanitize Multiple Trailing Special Elements
- 51 165. Failure to Sanitize Multiple Internal Special Elements

- 1 166. Failure to Handle Missing Special Element
- 2 167. Failure to Handle Additional Special Element
- 3 168. Failure to Resolve Inconsistent Special Elements
- 4 564. SQL Injection: Hibernate
- 5 CERT/CC guidelines: FIO30-C

6 7.8.3 Mechanism of failure

7 A software system that accepts and executes input in the form of operating system commands (e.g., `system()`,
8 `exec()`, `open()`) could allow an attacker with lesser privileges than the target software to execute commands with
9 the elevated privileges of the executing process. Command injection is a common problem with wrapper programs.
10 Often, parts of the command to be run are controllable by the end user. If a malicious user injects a character (such
11 as a semi-colon) that delimits the end of one command and the beginning of another, he may then be able to insert
12 an entirely new and unrelated command to do whatever he pleases.

13 Dynamically generating operating system commands that include user input as parameters can lead to command
14 injection attacks. An attacker can insert operating system commands or modifiers in the user input that can cause
15 the request to behave in an unsafe manner. Such vulnerabilities can be very dangerous and lead to data and
16 system compromise. If no validation of the parameter to the `exec` command exists, an attacker can execute any
17 command on the system the application has the privilege to access.

18 There are two forms of command injection vulnerabilities. An attacker can change the command that the program
19 executes (the attacker explicitly controls what the command is). Alternatively, an attacker can change the
20 environment in which the command executes (the attacker implicitly controls what the command means). The first
21 scenario where an attacker explicitly controls the command that is executed can occur when:

- 22 • Data enters the application from an untrusted source.
- 23 • The data is part of a string that is executed as a command by the application.
- 24 • By executing the command, the application gives an attacker a privilege or capability that the attacker
25 would not otherwise have.

26 Eval injection occurs when the software allows inputs to be fed directly into a function (e.g., "eval") that dynamically
27 evaluates and executes the input as code, usually in the same interpreted language that the product uses. Eval
28 injection is prevalent in handler/dispatch procedures that might want to invoke a large number of functions, or set a
29 large number of variables.

30 A PHP file inclusion occurs when a PHP product uses `require` or `include` statements, or equivalent statements,
31 that use attacker-controlled data to identify code or *HTML* (HyperText Markup Language) to be directly processed
32 by the PHP interpreter before inclusion in the script.

33 A resource injection issue occurs when the following two conditions are met:

- 34 • An attacker can specify the identifier used to access a system resource. For example, an attacker might be
35 able to specify part of the name of a file to be opened or a port number to be used.
- 36 • By specifying the resource, the attacker gains a capability that would not otherwise be permitted. For
37 example, the program may give the attacker the ability to overwrite the specified file, run with a
38 configuration controlled by the attacker, or transmit sensitive information to a third-party server. Note:
39 Resource injection that involves resources stored on the file system goes by the name path manipulation
40 and is reported in separate category. See Path Traversal [EWR] description for further details of this
41 vulnerability. Allowing user input to control resource identifiers may enable an attacker to access or modify
42 otherwise protected system resources.

43 Line or section delimiters injected into an application can be used to compromise a system. As data is parsed, an
44 injected/absent/malformed delimiter may cause the process to take unexpected actions that result in an attack. One
45 example of a section delimiter is the boundary string in a multipart MIME message. In many cases, doubled line
46 delimiters can serve as a section delimiter.

1 7.8.4 Avoiding the vulnerability or mitigating its effects

2 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 3 • Assume all input is malicious. Use an appropriate combination of black lists and white lists to ensure only
- 4 valid, expected and appropriate input is processed by the system.
- 5 • Narrowly define the set of safe characters based on the expected values of the parameter in the request.
- 6 • Developers should anticipate that delimiters and special elements would be injected/removed/manipulated
- 7 in the input vectors of their software system and appropriate mechanisms should be put in place to handle
- 8 them.
- 9 • Implement SQL strings using prepared statements that bind variables. Prepared statements that do not
- 10 bind variables can be vulnerable to attack.
- 11 • Use vigorous white-list style checking on any user input that may be used in a SQL command. Rather than
- 12 escape meta-characters, it is safest to disallow them entirely since the later use of data that have been
- 13 entered in the database may neglect to escape meta-characters before use.
- 14 • Follow the principle of least privilege when creating user accounts to a SQL database. Users should only
- 15 have the minimum privileges necessary to use their account. If the requirements of the system indicate that
- 16 a user can read and modify their own data, then limit their privileges so they cannot read/write others' data.
- 17 • Assign permissions to the software system that prevents the user from accessing/opening privileged files.
- 18 • To avert eval injections, refactor your code so that it does not need to use eval().

19 7.8.5 Bibliography

20 [None]

21 7.9 Cross-site Scripting [XYT]

22 7.9.1 Description of application vulnerability

23 *Cross-site scripting* (XSS) occurs when dynamically generated web pages display input, such as login information
 24 that is not properly validated, allowing an attacker to embed malicious scripts into the generated page and then
 25 execute the script on the machine of any user that views the site. If successful, cross-site scripting vulnerabilities
 26 can be exploited to manipulate or steal cookies, create requests that can be mistaken for those of a valid user,
 27 compromise confidential information, or execute malicious code on the end user systems for a variety of nefarious
 28 purposes.

29 7.9.2 Cross reference

30 CWE:

- 31 79. Failure to Preserve Web Page Structure ('Cross-site Scripting')
- 32 80. Failure to Sanitize Script-Related HTML Tags in a Web Page (Basic XSS)
- 33 81. Failure to Sanitize Directives in an Error Message Web Page
- 34 82. Failure to Sanitize Script in Attributes of IMG Tags in a Web Page
- 35 83. Failure to Sanitize Script in Attributes in a Web Page
- 36 84. Failure to Resolve Encoded URI Schemes in a Web Page
- 37 85. Doubled Character XSS Manipulations
- 38 86. Invalid Characters in Identifiers
- 39 87. Alternate XSS Syntax

40 7.9.3 Mechanism of failure

41 Cross-site scripting (XSS) vulnerabilities occur when an attacker uses a web application to send malicious code,
 42 generally JavaScript, to a different end user. When a web application uses input from a user in the output it
 43 generates without filtering it, an attacker can insert an attack in that input and the web application sends the attack
 44 to other users. The end user trusts the web application, and the attacks exploit that trust to do things that would not

1 normally be allowed. Attackers frequently use a variety of methods to encode the malicious portion of the tag, such
2 as using Unicode, so the request looks less suspicious to the user.

3 XSS attacks can generally be categorized into two categories: stored and reflected. Stored attacks are those where
4 the injected code is permanently stored on the target servers in a database, message forum, visitor log, and so
5 forth. Reflected attacks are those where the injected code takes another route to the victim, such as in an email
6 message, or on some other server. When a user is tricked into clicking a link or submitting a form, the injected code
7 travels to the vulnerable web server, which reflects the attack back to the user's browser. The browser then
8 executes the code because it came from a 'trusted' server. For a reflected XSS attack to work, the victim must
9 submit the attack to the server. This is still a very dangerous attack given the number of possible ways to trick a
10 victim into submitting such a malicious request, including clicking a link on a malicious Web site, in an email, or in
11 an inter-office posting.

12 XSS flaws are very common in web applications, as they require a great deal of developer discipline to avoid them
13 in most applications. It is relatively easy for an attacker to find XSS vulnerabilities. Some of these vulnerabilities can
14 be found using scanners, and some exist in older web application servers. The consequence of an XSS attack is
15 the same regardless of whether it is stored or reflected.

16 The difference is in how the payload arrives at the server. XSS can cause a variety of problems for the end user
17 that range in severity from an annoyance to complete account compromise. The most severe XSS attacks involve
18 disclosure of the user's session cookie, which allows an attacker to hijack the user's session and take over their
19 account. Other damaging attacks include the disclosure of end user files, installation of Trojan horse programs,
20 redirecting the user to some other page or site, and modifying presentation of content.

21 Cross-site scripting (XSS) vulnerabilities occur when:

- 22 • Data enters a Web application through an untrusted source, most frequently a web request. The data is
23 included in dynamic content that is sent to a web user without being validated for malicious code.
- 24 • The malicious content sent to the web browser often takes the form of a segment of JavaScript, but may
25 also include HTML, Flash or any other type of code that the browser may execute. The variety of attacks
26 based on XSS is almost limitless, but they commonly include transmitting private data like cookies or other
27 session information to the attacker, redirecting the victim to web content controlled by the attacker, or
28 performing other malicious operations on the user's machine under the guise of the vulnerable site.

29 Cross-site scripting attacks can occur wherever an untrusted user has the ability to publish content to a trusted web
30 site. Typically, a malicious user will craft a client-side script, which — when parsed by a web browser — performs
31 some activity (such as sending all site cookies to a given e-mail address). If the input is unchecked, this script will
32 be loaded and run by each user visiting the web site. Since the site requesting to run the script has access to the
33 cookies in question, the malicious script does also. There are several other possible attacks, such as running
34 "Active X" controls (under Microsoft Internet Explorer) from sites that a user perceives as trustworthy; cookie theft is
35 however by far the most common. All of these attacks are easily prevented by ensuring that no script tags — or for
36 good measure, HTML tags at all — are allowed in data to be posted publicly.

37 Specific instances of XSS are:

- 38 • 'Basic' XSS involves a complete lack of cleansing of any special characters, including the most
39 fundamental XSS elements such as "<", ">", and "&".
- 40 • A web developer displays input on an error page (e.g., a customized 403 Forbidden page). If an attacker
41 can influence a victim to view/request a web page that causes an error, then the attack may be successful.
- 42 • A Web application that trusts input in the form of HTML IMG tags is potentially vulnerable to XSS attacks.
43 Attackers can embed XSS exploits into the values for IMG attributes (e.g., SRC) that is streamed and then
44 executed in a victim's browser. Note that when the page is loaded into a user's browser, the exploit will
45 automatically execute.
- 46 • The software does not filter "javascript:" or other *URI*'s (Uniform Resource Identifier) from dangerous
47 attributes within tags, such as *onmouseover*, *onload*, *onerror*, or *style*.
- 48 • The web application fails to filter input for executable script disguised with URI encodings.
- 49 • The web application fails to filter input for executable script disguised using doubling of the involved
50 characters.

- 1 • The software does not strip out invalid characters in the middle of tag names, schemes, and other
2 identifiers, which are still rendered by some web browsers that ignore the characters.
- 3 • The software fails to filter alternate script syntax provided by the attacker.

4 Cross-site scripting attacks may occur anywhere that possibly malicious users are allowed to post unregulated
5 material to a trusted web site for the consumption of other valid users. The most common example can be found in
6 bulletin-board web sites that provide web based mailing list-style functionality. The most common attack performed
7 with cross-site scripting involves the disclosure of information stored in user cookies. In some circumstances it may
8 be possible to run arbitrary code on a victim's computer when cross-site scripting is combined with other flaws.

9 **7.9.4 Avoiding the vulnerability or mitigating its effects**

10 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 11 • Carefully check each input parameter against a rigorous positive specification (white list) defining the
12 specific characters and format allowed.
- 13 • All input should be sanitized, not just parameters that the user is supposed to specify, but all data in the
14 request, including hidden fields, cookies, headers, the *URL* (Uniform Resource Locator) itself, and so forth.
- 15 • A common mistake that leads to continuing XSS vulnerabilities is to validate only fields that are expected
16 to be redisplayed by the site.
- 17 • Data is frequently encountered from the request that is reflected by the application server or the
18 application that the development team did not anticipate. Also, a field that is not currently reflected may be
19 used by a future developer. Therefore, validating ALL parts of the HTTP request is recommended.

20 **7.9.5 Bibliography**

21 [None]

22 **7.10 Unquoted Search Path or Element [XZQ]**

23 **7.10.1 Description of application vulnerability**

24 Strings injected into a software system that are not quoted can permit an attacker to execute arbitrary commands.

25 **7.10.2 Cross reference**

26 CWE:

27 428. Unquoted Search Path or Element

28 CERT/CC guidelines: ENV04-C

29 **7.10.3 Mechanism of failure**

30 The mechanism of failure stems from missing quoting of strings injected into a software system. By allowing white-
31 spaces in identifiers, an attacker could potentially execute arbitrary commands. This vulnerability covers
32 "C:\Program Files" and space-in-search-path issues. Theoretically this could apply to other operating systems
33 besides Windows, especially those that make it easy for spaces to be in filenames or folders names.

34 **7.10.4 Avoiding the vulnerability or mitigating its effects**

35 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 36 • Software should quote the input data that can be potentially executed on a system.
- 37 • Use a programming language that enforces the quoting of strings.

1 **7.10.5 Bibliography**

2 [None]

3 **7.11 Improperly Verified Signature [XZR]**

4 **7.11.1 Description of application vulnerability**

5 The software does not verify, or improperly verifies, the cryptographic signature for data. By not adequately
6 performing the verification step, the data being received should not be trusted and may be corrupted or made
7 intentionally incorrect by an adversary.

8 **7.11.2 Cross reference**

9 CWE:

10 347. Improperly Verified Signature

11 **7.11.3 Mechanism of failure**

12 Data is signed using techniques that assure the integrity of the data. There are two ways that the integrity can be
13 intentionally compromised. The exchange of the cryptologic keys may have been compromised so that an attacker
14 could provide encrypted data that has been altered. Alternatively, the cryptologic verification could be flawed so
15 that the encryption of the data is flawed which again allows an attacker to alter the data.

16 **7.11.4 Avoiding the vulnerability or mitigating its effects**

17 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 18
- Use data signatures to the extent possible to help ensure trust in data.
 - Use built-in verifications for data.
- 19

20 **7.11.5 Bibliography**

21 [None]

22 **7.12 Discrepancy Information Leak [XZL]**

23 **7.12.1 Description of application vulnerability**

24 A discrepancy information leak is an information leak in which the product behaves differently, or sends different
25 responses, in a way that reveals security-relevant information about the state of the product, such as whether a
26 particular operation was successful or not.

27 **7.12.2 Cross reference**

28 CWE:

29 203. Discrepancy Information Leaks
30 204. Response Discrepancy Information Leak
31 206. Internal Behavioural Inconsistency Information Leak
32 207. External Behavioral Inconsistency Information Leak
33 208. Timing Discrepancy Information Leak

1 **7.12.3 Mechanism of failure**

2 A response discrepancy information leak occurs when the product sends different messages in direct response to
3 an attacker's request, in a way that allows the attacker to learn about the inner state of the product. The leaks can
4 be inadvertent (bug) or intentional (design).

5
6 A behavioural discrepancy information leak occurs when the product's actions indicate important differences based
7 on (1) the internal state of the product or (2) differences from other products in the same class. Attacks such as OS
8 fingerprinting rely heavily on both behavioural and response discrepancies. An internal behavioural inconsistency
9 information leak is the situation where two separate operations in a product cause the product to behave differently
10 in a way that is observable to an attacker and reveals security-relevant information about the internal state of the
11 product, such as whether a particular operation was successful or not. An external behavioural inconsistency
12 information leak is the situation where the software behaves differently than other products like it, in a way that is
13 observable to an attacker and reveals security-relevant information about which product is being used, or its
14 operating state.

15
16 A timing discrepancy information leak occurs when two separate operations in a product require different amounts
17 of time to complete, in a way that is observable to an attacker and reveals security-relevant information about the
18 state of the product, such as whether a particular operation was successful or not.

19 **7.12.4 Avoiding the vulnerability or mitigating its effects**

20 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 21 • Compartmentalize the system to have "safe" areas where trust boundaries can be unambiguously drawn.
- 22 • Do not allow sensitive data to go outside of the trust boundary and always be careful when interfacing with
- 23 a compartment outside of the safe area.

24 **7.12.5 Bibliography**

25 [None]

26 **7.13 Sensitive Information Uncleared Before Release [XZK]**

27 **7.13.1 Description of application vulnerability**

28 The software does not fully clear previously used information in a data structure, file, or other resource, before
29 making that resource available to another party that did not have access to the original information.

30 **7.13.2 Cross reference**

31 CWE:
32 226. Sensitive Information Uncleared Before Release
33 CERT/CC guidelines: MEM03-C

34 **7.13.3 Mechanism of failure**

35 This typically involves memory in which the new data occupies less memory than the old data, which leaves
36 portions of the old data still available ("memory disclosure"). However, equivalent errors can occur in other
37 situations where the length of data is variable but the associated data structure is not. This can overlap with
38 cryptographic errors and cross-boundary cleansing info leaks.

39 Dynamic memory managers are not required to clear freed memory and generally do not because of the additional
40 runtime overhead. Furthermore, dynamic memory managers are free to reallocate this same memory. As a result,
41 it is possible to accidentally leak sensitive information if it is not cleared before calling a function that frees dynamic
42 memory. Programmers should not and can not rely on memory being cleared during allocation.

1 7.13.4 Avoiding the vulnerability or mitigating its effects

2 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 3 • Use library functions and or programming language features (such as destructors or finalization
- 4 procedures) that provide automatic clearing of freed buffers and or the functionality to clear buffers.

5 7.13.5 Bibliography

6 [None]

7 7.14 Path Traversal [EWR]

8 7.14.1 Description of application vulnerability

9 The software constructs a path that contains relative traversal sequence such as ".." or an absolute path sequence
 10 such as "/path/here." Attackers run the software in a particular directory so that the hard link or symbolic link used
 11 by the software accesses a file that the attacker has under their control. In doing this, the attacker may be able to
 12 escalate their privilege level to that of the running process.

13 7.14.2 Cross reference

14 CWE:

- 15 22. Path Traversal
- 16 24. Path Traversal: - '../filedir'
- 17 25. Path Traversal: '/../filedir'
- 18 26. Path Traversal: '/dir../filename'
- 19 27. Path Traversal: 'dir/./../filename'
- 20 28. Path Traversal: '..filename'
- 21 29. Path Traversal: '\.filename'
- 22 30. Path Traversal: 'dir\..filename'
- 23 31. Path Traversal: 'dir\..filename'
- 24 32. Path Traversal: '...' (Triple Dot)
- 25 33. Path Traversal: '....' (Multiple Dot)
- 26 34. Path Traversal: '.../'
- 27 35. Path Traversal: '.../.../'
- 28 37. Path Traversal: '/absolute/pathname/here'
- 29 38. Path Traversal: '\absolute\pathname\here'
- 30 39. Path Traversal: 'C:dirname'
- 31 40. Path Traversal: '\\UNC\share\name\' (Windows UNC Share)
- 32 61. UNIX Symbolic Link (Symlink) Following
- 33 62. UNIX Hard Link
- 34 64. Windows Shortcut Following (.LNK)
- 35 65. Windows Hard Link
- 36 CERT/CC guidelines: FIO02-C

37 7.14.3 Mechanism of failure

38 There are two primary ways that an attacker can orchestrate an attack using path traversal. In the first, the attacker
 39 alters the path being used by the software to point to a location that the attacker has control over. Alternatively, the
 40 attacker has no control over the path, but can alter the directory structure so that the path points to a location that
 41 the attacker does have control over.

42 For instance, a software system that accepts input in the form of: '..filename', '\.filename', '/directory/./filename',
 43 'directory/./../filename', '..filename', '\.filename', '\directory\..filename', 'directory\..\filename', '...', '....' (multiple
 44 dots), '.../' or '.../.../' without appropriate validation can allow an attacker to traverse the file system to access an
 45 arbitrary file. Note that '..' is ignored if the current working directory is the root directory. Some of these input forms

1 can be used to cause problems for systems that strip out '..' from input in an attempt to remove relative path
2 traversal.

3 There are several common ways that an attacker can point a file access to a file the attacker has under their
4 control. A software system that accepts input in the form of '/absolute/pathname/here' or '\absolute\pathname\here'
5 without appropriate validation can also allow an attacker to traverse the file system to unintended locations or
6 access arbitrary files. An attacker can inject a drive letter or Windows volume letter ('C:dirname') into a software
7 system to potentially redirect access to an unintended location or arbitrary file. A software system that accepts input
8 in the form of a backslash absolute path without appropriate validation can allow an attacker to traverse the file
9 system to unintended locations or access arbitrary files. An attacker can inject a Windows UNC share
10 ('\\UNC\share\name') into a software system to potentially redirect access to an unintended location or arbitrary file.
11 A software system that allows UNIX symbolic links (symlink) as part of paths whether in internal code or through
12 user input can allow an attacker to spoof the symbolic link and traverse the file system to unintended locations or
13 access arbitrary files. The symbolic link can permit an attacker to read/write/corrupt a file that they originally did not
14 have permissions to access. Failure for a system to check for hard links can result in vulnerability to different types
15 of attacks. For example, an attacker can escalate their privileges if he/she can replace a file used by a privileged
16 program with a hard link to a sensitive file, for example, `etc/passwd`. When the process opens the file, the
17 attacker can assume the privileges of that process.

18 A software system that allows Windows shortcuts (.LNK) as part of paths whether in internal code or through user
19 input can allow an attacker to spoof the symbolic link and traverse the file system to unintended locations or access
20 arbitrary files. The shortcut (file with the .lnk extension) can permit an attacker to read/write a file that they originally
21 did not have permissions to access.

22 Failure for a system to check for hard links can result in vulnerability to different types of attacks. For example, an
23 attacker can escalate their privileges if he/she can replace a file used by a privileged program with a hard link to a
24 sensitive file (e.g., `etc/passwd`). When the process opens the file, the attacker can assume the privileges of that
25 process or possibly prevent a program from accurately processing data in a software system.

26 7.14.4 Avoiding the vulnerability or mitigating its effects

27 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 28 • Assume all input is malicious. Attackers can insert paths into input vectors and traverse the file system.
- 29 • Use an appropriate combination of black lists and white lists to ensure only valid and expected input is
30 processed by the system.
- 31 • A sanitizing mechanism can remove characters such as '.' and ';' which may be required for some exploits.
32 An attacker can try to fool the sanitizing mechanism into "cleaning" data into a dangerous form. Suppose
33 the attacker injects a '.' inside a filename (e.g., "sensitiveFile") and the sanitizing mechanism removes the
34 character resulting in the valid filename, "sensitiveFile". If the input data are now assumed to be safe, then
35 the file may be compromised.
- 36 • Files can often be identified by other attributes in addition to the file name, for example, by comparing file
37 ownership or creation time. Information regarding a file that has been created and closed can be stored
38 and then used later to validate the identity of the file when it is reopened. Comparing multiple attributes of
39 the file improves the likelihood that the file is the expected one.
- 40 • Follow the principle of least privilege when assigning access rights to files.
- 41 • Denying access to a file can prevent an attacker from replacing that file with a link to a sensitive file.
- 42 • Ensure good compartmentalization in the system to provide protected areas that can be trusted.
- 43 • When two or more users, or a group of users, have write permission to a directory, the potential for sharing
44 and deception is far greater than it is for shared access to a few files. The vulnerabilities that result from
45 malicious restructuring via hard and symbolic links suggest that it is best to avoid shared directories.
- 46 • Securely creating temporary files in a shared directory is error prone and dependent on the version of the
47 runtime library used, the operating system, and the file system. Code that works for a locally mounted file
48 system, for example, may be vulnerable when used with a remotely mounted file system.
- 49 • The mitigation should be centered on converting relative paths into absolute paths and then verifying that
50 the resulting absolute path makes sense with respect to the configuration and rights or permissions. This
51 may include checking "whitelists" and "blacklists", authorized super user status, access control lists, etc.

1 **7.14.5 Bibliography**

2 [None]

3 **7.15 Missing Required Cryptographic Step [XZS]**4 **7.15.1 Description of application vulnerability**5 Cryptographic implementations should follow the algorithms that define them exactly, otherwise encryption can be
6 faulty.7 **7.15.2 Cross reference**

8 CWE:

9 325. Missing Required Cryptographic Step

10 **7.15.3 Mechanism of failure**11 Not following the algorithms that define cryptographic implementations exactly can lead to weak encryption. This
12 could be the result of many factors such as a programmer missing a required cryptographic step or using weak
13 randomization algorithms.14 **7.15.4 Avoiding the vulnerability or mitigating its effects**

15 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 16
- Implement cryptographic algorithms precisely.
 - Use system functions and libraries rather than writing the function.

18 **7.15.5 Bibliography**

19 [None]

20 **7.16 Insufficiently Protected Credentials [XYM]**21 **7.16.1 Description of application vulnerability**22 This weakness occurs when the application transmits or stores authentication credentials and uses an insecure
23 method that is susceptible to unauthorized interception and/or retrieval.24 **7.16.2 Cross reference**

25 CWE:

26 256. Plaintext Storage of a Password

27 257. Storing Passwords in a Recoverable Format

28 **7.16.3 Mechanism of failure**29 Storing a password in plaintext may result in a system compromise. Password management issues occur when a
30 password is stored in plaintext in an application's properties or configuration file. A programmer can attempt to
31 remedy the password management problem by obscuring the password with an encoding function, such as Base64
32 encoding, but this effort does not adequately protect the password. Storing a plaintext password in a configuration
33 file allows anyone who can read the file access to the password-protected resource. Developers sometimes
34 believe that they cannot defend the application from someone who has access to the configuration, but this attitude

1 makes an attacker's job easier. Good password management guidelines require that a password never be stored
2 in plaintext.

3
4 The storage of passwords in a recoverable format makes them subject to password reuse attacks by malicious
5 users. If a system administrator can recover the password directly or use a brute force search on the information
6 available to him, he can use the password on other accounts.

7 The use of recoverable passwords significantly increases the chance that passwords will be used maliciously. In
8 fact, it should be noted that recoverable encrypted passwords provide no significant benefit over plain-text
9 passwords since they are subject not only to reuse by malicious attackers but also by malicious insiders.

10 **7.16.4 Avoiding the vulnerability or mitigating its effects**

11 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 12 • Avoid storing passwords in easily accessible locations.
- 13 • Never store a password in plaintext.
- 14 • Ensure that strong, non-reversible encryption is used to protect stored passwords.
- 15 • Consider storing cryptographic hashes of passwords as an alternative to storing in plaintext.

16 **7.16.5 Bibliography**

17 [None]

18 **7.17 Missing or Inconsistent Access Control [XZN]**

19 **7.17.1 Description of application vulnerability**

20 The software does not perform access control checks in a consistent manner across all potential execution paths.

21 **7.17.2 Cross reference**

22 CWE:
23 285. Missing or Inconsistent Access Control
24 CERT/CC guidelines: FIO06-C

25 **7.17.3 Mechanism of failure**

26 For web applications, attackers can issue a request directly to a page (URL) that they may not be authorized to
27 access. If the access control policy is not consistently enforced on every page restricted to authorized users, then
28 an attacker could gain access to and possibly corrupt these resources.

29 **7.17.4 Avoiding the vulnerability or mitigating its effects**

30 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 31 • For web applications, make sure that the access control mechanism is enforced correctly at the server side
32 on every page. Users should not be able to access any information simply by requesting direct access to
33 that page, if they do not authorization. Ensure that all pages containing sensitive information are not
34 cached, and that all such pages restrict access to requests that are accompanied by an active and
35 authenticated session token associated with a user who has the required permissions to access that page.

36 **7.17.5 Bibliography**

37 [None]

1 7.18 Authentication Logic Error [XZO]

2 7.18.1 Description of application vulnerability

3 The software does not properly ensure that the user has proven their identity.

4 7.18.2 Cross reference

5 CWE:

- 6 287. Improper Authentication
- 7 288. Authentication Bypass by Alternate Path/Channel
- 8 289. Authentication Bypass by Alternate Name
- 9 290. Authentication Bypass by Spoofing
- 10 294. Authentication Bypass by Capture-replay
- 11 301. Reflection Attack in an Authentication Protocol
- 12 302. Authentication Bypass by Assumed-Immutable Data
- 13 303. Improper Implementation of Authentication Algorithm
- 14 305. Authentication Bypass by Primary Weakness

15 7.18.3 Mechanism of failure

16 There are many ways that an attacker can potentially bypass the validation of a user. Some of the ways are means
17 of impersonating a legitimate user while others are means of bypassing the authentication mechanisms that are in
18 place. In either case, a user who should not have access to the software system gains access.

19 Authentication bypass by alternate path or channel occurs when a product requires authentication, but the product
20 has an alternate path or channel that does not require authentication. Note that this is often seen in web
21 applications that assume that access to a particular *CGI* (Common gateway Interface) program can only be
22 obtained through a "front" screen, but this problem is not just in web applications.

23
24 Authentication bypass by alternate name occurs when the software performs authentication based on the name of
25 the resource being accessed, but there are multiple names for the resource, and not all names are checked.

26
27 Authentication bypass by capture-replay occurs when it is possible for a malicious user to sniff network traffic and
28 bypass authentication by replaying it to the server in question to the same effect as the original message (or with
29 minor changes). Messages sent with a capture-relay attack allow access to resources that are not otherwise
30 accessible without proper authentication. Capture-replay attacks are common and can be difficult to defeat without
31 cryptography. They are a subset of network injection attacks that rely on listening in on previously sent valid
32 commands, then changing them slightly if necessary and resending the same commands to the server. Since any
33 attacker who can listen to traffic can see sequence numbers, it is necessary to sign messages with some kind of
34 cryptography to ensure that sequence numbers are not simply doctored along with content.

35
36 Reflection attacks capitalize on mutual authentication schemes in order to trick the target into revealing the secret
37 shared between it and another valid user. In a basic mutual-authentication scheme, a secret is known to both a
38 valid user and the server; this allows them to authenticate. In order that they may verify this shared secret without
39 sending it plainly over the wire, they utilize a Diffie-Hellman-style scheme in which they each pick a value, then
40 request the hash of that value as keyed by the shared secret. In a reflection attack, the attacker claims to be a valid
41 user and requests the hash of a random value from the server. When the server returns this value and requests its
42 own value to be hashed, the attacker opens another connection to the server. This time, the hash requested by the
43 attacker is the value that the server requested in the first connection. When the server returns this hashed value, it
44 is used in the first connection, authenticating the attacker successfully as the impersonated valid user.

45
46 Authentication bypass by assumed-immutable data occurs when the authentication scheme or implementation
47 uses key data elements that are assumed to be immutable, but can be controlled or modified by the attacker, e.g.,
48 if a web application relies on a cookie "Authenticated=1".

49
50 Authentication logic error occurs when the authentication techniques do not follow the algorithms that define them

1 exactly and so authentication can be jeopardized. For instance, a malformed or improper implementation of an
2 algorithm can weaken the authorization technique.

3
4 An authentication bypass by primary weakness occurs when the authentication algorithm is sound, but the
5 implemented mechanism can be bypassed as the result of a separate weakness that is primary to the
6 authentication error.

7 **7.18.4 Avoiding the vulnerability or mitigating its effects**

8 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 9 • Funnel all access through a single choke point to simplify how users can access a resource. For every
10 access, perform a check to determine if the user has permissions to access the resource. Avoid making
11 decisions based on names of resources (e.g., files) if those resources can have alternate names.
- 12 • Canonicalize the name to match that of the file system's representation of the name. This can sometimes
13 be achieved with an available API (e.g. in Win32 the `GetFullPathName` function).
- 14 • Utilize some sequence or time stamping functionality along with a checksum that takes this into account in
15 order to ensure that messages can be parsed only once.
- 16 • Use different keys for the initiator and responder or of a different type of challenge for the initiator and
17 responder.

18 **7.18.5 Bibliography**

19 [None]

20 **7.19 Hard-coded Password [XYP]**

21 **7.19.1 Description of application vulnerability**

22 Hard coded passwords may compromise system security in a way that cannot be easily remedied. It is never a
23 good idea to hardcode a password. Not only does hard coding a password allow all of the project's developers to
24 view the password, it also makes fixing the problem extremely difficult. Once the code is in production, the
25 password cannot be changed without patching the software. If the account protected by the password is
26 compromised, the owners of the system will be forced to choose between security and availability.

27 **7.19.2 Cross reference**

28 CWE:
29 259. Hard-Coded Password

30 **7.19.3 Mechanism of failure**

31 The use of a hard-coded password has many negative implications -- the most significant of these being a failure of
32 authentication measures under certain circumstances. On many systems, a default administration account exists
33 which is set to a simple default password that is hard-coded into the program or device. This hard-coded password
34 is the same for each device or system of this type and often is not changed or disabled by end users. If a malicious
35 user comes across a device of this kind, it is a simple matter of looking up the default password (which is likely
36 freely available and public on the Internet) and logging in with complete access. In systems that authenticate with a
37 back-end service, hard-coded passwords within closed source or drop-in solution systems require that the back-
38 end service use a password that can be easily discovered. Client-side systems with hard-coded passwords
39 present even more of a threat, since the extraction of a password from a binary is exceedingly simple. If hard-
40 coded passwords are used, it is almost certain that unauthorized users will gain access through the account in
41 question.

1 **7.19.4 Avoiding the vulnerability or mitigating its effects**

2 Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

- 3
- 4 • Rather than hard code a default username and password for first time logins, utilize a "first login" mode that requires the user to enter a unique strong password.
 - 5 • For front-end to back-end connections, there are three solutions that may be used.
 - 6 1. Use of generated passwords that are changed automatically and must be entered at given time intervals by a system administrator. These passwords will be held in memory and only be valid for the time intervals.
 - 7 2. The passwords used should be limited at the back end to only performing actions for the front end, as opposed to having full access.
 - 8 3. The messages sent should be tagged and checksummed with time sensitive values so as to prevent replay style attacks.
- 9
10
11
12

13 **7.19.5 Bibliography**

14 [None]

1 **Annex A**
2 (informative)
3 **Guideline Recommendation Factors**

4 **A. Guideline Recommendation Factors**

5 **A.1 Factors that need to be covered in a proposed guideline recommendation**

6 These are needed because circumstances might change, for instance:

- 7 • Changes to language definition.
8 • Changes to translator behaviour.
9 • Developer training.
10 • More effective recommendation discovered.

11 **A.1.1 Expected cost of following a guideline**

12 How to evaluate likely costs.

13 **A.1.2 Expected benefit from following a guideline**

14 How to evaluate likely benefits.

15 **A.2 Language definition**

16 Which language definition to use. For instance, an ISO/IEC Standard, Industry standard, a particular
17 implementation.

18 Position on use of extensions.

19 **A.3 Measurements of language usage**

20 Occurrences of applicable language constructs in software written for the target market.

21 How often do the constructs addressed by each guideline recommendation occur?

22 **A.4 Level of expertise**

23 How much expertise, and in what areas, are the people using the language assumed to have?

24 Is use of the alternative constructs less likely to result in faults?

25 **A.5 Intended purpose of guidelines**

26 For instance: How the listed guidelines cover the requirements specified in a safety critical standard.

27 **A.6 Constructs whose behaviour can vary**

28 The different ways in which language definitions specify behaviour that is allowed to vary between implementations
29 and how to go about documenting these cases.

1 **A.7 Example guideline proposal template**

2 **A.7.1 Coding Guideline**

3 Anticipated benefit of adhering to guideline

- 4 • Cost of moving to a new translator reduced.
- 5 • Probability of a fault introduced when new version of translator used reduced.
- 6 • Probability of developer making a mistake is reduced.
- 7 • Developer mistakes more likely to be detected during development.
- 8 • Reduction of future maintenance costs.
- 9

Annex B (informative) Guideline Selection Process

1
2
3
4

B. Guideline Selection Process

6 It is possible to claim that any language construct can be misunderstood by a developer and lead to a failure to
7 predict program behaviour. A cost/benefit analysis of each proposed guideline is the solution adopted by this
8 Technical Report.

9 The selection process has been based on evidence that the use of a language construct leads to unintended
10 behaviour (i.e., a cost) and that the proposed guideline increases the likelihood that the behaviour is as intended
11 (i.e., a benefit). The following is a list of the major source of evidence on the use of a language construct and the
12 faults resulting from that use:

- 13 • a list of language constructs having undefined, implementation-defined, or unspecified behaviours,
- 14 • measurements of existing source code. This usage information has included the number of occurrences of
15 uses of the construct and the contexts in which it occurs,
- 16 • measurement of faults experienced in existing code,
- 17 • measurements of developer knowledge and performance behaviour.

18 The following are some of the issues that were considered when framing guidelines:

- 19 • An attempt was made to be generic to particular kinds of language constructs (i.e., language independent),
20 rather than being language specific.
- 21 • Preference was given to wording that is capable of being checked by automated tools.
- 22 • Known algorithms for performing various kinds of source code analysis and the properties of those
23 algorithms (i.e., their complexity and running time).

B.1 Cost/Benefit Analysis

25 The fact that a coding construct is known to be a source of failure to predict correct behaviour is not in itself a
26 reason to recommend against its use. Unless the desired algorithmic functionality can be implemented using an
27 alternative construct whose use has more predictable behaviour, then there is no benefit in recommending against
28 the use of the original construct.

29 While the cost/benefit of some guidelines may always come down in favor of them being adhered to (e.g., don't
30 access a variable before it is given a value), the situation may be less clear-cut for other guidelines. Providing a
31 summary of the background analysis for each guideline will enable development groups.

32 Annex A provides a template for the information that should be supplied with each guideline.

33 It is unlikely that all of the guidelines given in this Technical Report will be applicable to all application domains.

B.2 Documenting of the selection process

35 The intended purpose of this documentation is to enable third parties to evaluate:

- 36 • the effectiveness of the process that created each guideline,
- 37 • the applicability of individual guidelines to a particular project.

Annex C (informative)

Template for use in proposing programming language vulnerabilities

C. Skeleton template for use in proposing programming language vulnerabilities

C.1 6.<x> <short title> [<unique immutable identifier>]

No text should appear here—in the space between 6.x and 6.x.1. This is simply an explanation of the header.

Notes on template header. The number "x" depends on the order in which the vulnerabilities are listed in Clause 6. It will be assigned by the editor. The "short title" should be a noun phrase summarizing the description of the application vulnerability. The "unique immutable identifier" is intended to provide an enduring identifier for the vulnerability description, even if their order is changed in the document. No additional text should appear here.

C.1.0 6.<x>.0 Status and history

The header will be removed before publication.

This temporary section will hold the edit history for the vulnerability along with the current status of the vulnerability.

C.1.1 6.<x>.1 Description of application vulnerability

Replace this with a brief description of the application vulnerability. It should be a short paragraph.

C.1.2 6.<x>.2 Cross reference

CWE: Replace this with the CWE identifier. At a later date, other cross-references may be added.

C.1.3 6.<x>.3 Mechanism of failure

Replace this with a brief description of the mechanism of failure. This description provides the link between the programming language vulnerability and the application vulnerability. It should be a short paragraph.

C.1.4 6.<x>.4 Applicable language characteristics

This vulnerability description is applicable to languages with the following characteristics:

Replace this with a bullet list summarizing the pertinent range of characteristics of languages for which this discussion is applicable. This list is intended to assist readers attempting to apply the guidance to languages that have not been treated in the language-specific annexes.

C.1.5 6.<x>.5 Avoiding the vulnerability or mitigating its effects

Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:

Replace this with a bullet list summarizing various ways in which programmers can avoid the programming language vulnerability, break the chain of causation to the application vulnerability, or contain the bad effects of the application vulnerability. Begin with the more direct, concrete, and effective means and then progress to the more indirect, abstract, and probabilistic means.

1 **C.1.6 6.<x>.6 Implications for standardization**

2 *Recommendations for other working groups will be recorded here. For example, we might record suggestions for*
3 *changes to language standards or API standards.*

4 **C.1.7 6.<x>.7 Bibliography**

5 *Insert numbered references for other documents cited in your description. These will eventually be collected into an*
6 *overall bibliography for the TR. So, please make the references complete. Someone will eventually have to*
7 *reformat the references into an ISO-required format, so please err on the side of providing too much information*
8 *rather than too little. Here [1] is an example of a reference:*

9 *[1] Greg Hoglund, Gary McGraw, *Exploiting Software: How to Break Code*, ISBN-0-201-78695-8, Pearson*
10 *Education, Boston, MA, 2004*

1 **Annex D**
 2 (informative)
 3 **Template for use in proposing application vulnerabilities**
 4

5 **D. Skeleton template for use in proposing application vulnerabilities**

6 **D.1 7.<x> <short title> [<unique immutable identifier>]**

7 *Notes on template header. The number "x" depends on the order in which the vulnerabilities are listed in Clause 7.*
 8 *It will be assigned by the editor. The "short title" should be a noun phrase summarizing the description of the*
 9 *application vulnerability. The "unique immutable identifier" is intended to provide an enduring identifier for the*
 10 *vulnerability description, even if their order is changed in the document. No additional text should appear here.*

11 **D.1.0 7.<x>.0 Status and history**

12 *The header will be removed before publication.*

13 *This temporary section will hold the edit history for the vulnerability. With the current status of the vulnerability.*

14 **D.1.1 7.<x>.1 Description of application vulnerability**

15 *Replace this with a brief description of the application vulnerability. It should be a short paragraph.*

16 **D.1.2 7.<x>.2 Cross reference**

17 *CWE: Replace this with the CWE identifier. At a later date, other cross-references may be added.*

18 **D.1.3 7.<x>.3 Mechanism of failure**

19 *Replace this with a brief description of the mechanism of failure. It should be a short paragraph.*

20 **D.1.4 7.<x>.4 Avoiding the vulnerability or mitigating its effects**

21 *Software developers can avoid the vulnerability or mitigate its ill effects in the following ways:*

22 *Replace this with a bullet list summarizing various ways in which programmers can avoid the application*
 23 *vulnerability, break the chain of causation to the application vulnerability, or contain the bad effects of the*
 24 *application vulnerability. Begin with the more direct, concrete, and effective means and then progress to the more*
 25 *indirect, abstract, and probabilistic means.*

26 **D.1.5 7.<x>.5 Bibliography**

27 *Insert numbered references for other documents cited in your description. These will eventually be collected into an*
 28 *overall bibliography for the TR. So, please make the references complete. Someone will eventually have to*
 29 *reformat the references into an ISO-required format, so please err on the side of providing too much information*
 30 *rather than too little. Here [1] is an example of a reference:*

31 *[1] Greg Hoglund, Gary McGraw, Exploiting Software: How to Break Code, ISBN-0-201-78695-8, Pearson*
 32 *Education, Boston, MA, 2004*

Annex E (informative) Vulnerability Outline

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55

E. Vulnerability Outline

- E.1. Human Factors
 - E.1.1. [BRS] Obscure Language Features
- E.2. Environment
 - E.2.1. [XYN] Adherence to Least Privilege
 - E.2.2. [XYO] Privilege Sandbox Issues
 - E.2.3. Interactions with environment
 - E.2.3.1. [XYS] Executing or Loading Untrusted Code
- E.3. Core Language Issues
 - E.3.1. [BQF] Unspecified Behaviour
 - E.3.2. [EWF] Undefined Behaviour
 - E.3.3. [FAB] Implementation-defined Behaviour
 - E.3.4. [MEM] Deprecated Language Features
 - E.3.5. [BVQ] Unspecified Functionality
- E.4. Pre-processor
 - E.4.1. [NMP] Pre-processor Directives
- E.5. Declarations and Definitions
 - E.5.1. [NAI] Choice of Clear Names
 - E.5.2. [AJN] Choice of Filenames and other External Identifiers
 - E.5.3. [XYR] Unused Variable
 - E.5.4. [YOW] Identifier Name Reuse
- E.6. Types
 - E.6.1. Representation
 - E.6.1.1. [IHN] Type System
 - E.6.1.2. [STR] Bit Representations
 - E.6.1.3. [KLK] Distinguished Values in Data Types
 - E.6.2. Constants
 - E.6.3. Floating-point
 - E.6.3.1. [PLF] Floating-point Arithmetic
 - E.6.4. Enumerated Types
 - E.6.4.1. [CCB] Enumerator Issues
 - E.6.5. Integers
 - E.6.5.1. [FLC] Numeric Conversion Errors
 - E.6.6. Characters and strings
 - E.6.6.1 [CJM] String Termination
 - E.6.7. Arrays
 - E.6.7.1. [XYX] Boundary Beginning Violation
 - E.6.7.2. [XYZ] Unchecked Array Indexing
 - E.6.7.3. [XYW] Unchecked Array Copying
 - E.6.7.4. [XZB] Buffer Overflow
 - E.6.8. Structures and Unions
 - E.6.9. Pointers
 - E.6.9.1. [HFC] Pointer Casting and Pointer Type Changes
 - E.6.9.2. [RVG] Pointer Arithmetic
 - E.6.9.3. [XYH] Null Pointer Dereference
 - E.6.9.4. [XYK] Dangling Reference to Heap
- E.7. Templates/Generics
 - E.7.1. [SYM] Templates and Generics
 - E.7.2. [RIP] Inheritance
- E.8. Initialization
 - E.8.1. [LAV] Initialization of Variables
- E.9. Type Conversions/Limits

- 1 E.9.1. [XYY] Wrap-around Error
- 2 E.9.2. [XZI] Sign Extension Error
- 3 E.10. Operators/Expressions
- 4 E.10.1. [JCW] Operator Precedence/Operator Precedence
- 5 E.10.2. [SAM] Side-effects and Order of Evaluation
- 6 E.10.3. [KOA] Likely Incorrect Expressions
- 7 E.10.4. [XYQ] Dead and Deactivated Code
- 8 E.11. Control Flow
- 9 E.11.1. Conditional Statements
- 10 E.11.1.1. [CLL] Switch Statements and Static Analysis
- 11 E.11.1.2. [EOJ] Demarcation of Control Flow
- 12 E.11.2. Loops
- 13 E.11.2.1. [TEX] Loop Control Variables
- 14 E.11.2.2. [XZH] Off-by-one Error
- 15 E.11.3. Subroutines (Functions, Procedures, Subprograms)
- 16 E.11.3.1. [EWD] Structured Programming
- 17 E.11.3.2. [CSJ] Passing Parameters and Return Values
- 18 E.11.3.3. [DCM] Dangling References to Stack Frames
- 19 E.11.3.4. [OTR] Subprogram Signature Mismatch
- 20 E.11.3.5. [GDL] Recursion
- 21 E.11.3.7. [NZN] Returning Error Status
- 22 E.11.4. Termination Strategy
- 23 E.11.4.1. [REU] Termination Strategy
- 24 E.12. Internal interfaces
- 25 E.12.1 [LRM] Extra Intrinsic
- 26 E.13. External interfaces
- 27 E.13.1. Memory Management
- 28 E.13.1.1. [AMV] Type-breaking Reinterpretation of Data
- 29 E.13.1.2. [XYL] Memory Leak
- 30 E.13.1.3. [XZX] Memory Locking
- 31 E.13.1.4. [XZP] Resource Exhaustion
- 32 E.13.2. Input
- 33 E.13.2.1. [RST] Injection
- 34 E.13.2.2. [XYT] Cross-site Scripting
- 35 E.13.2.3. [XZQ] Unquoted Search Path or Element
- 36 E.13.2.4. [XZR] Improperly Verified Signature
- 37 E.13.2.5. [XZL] Discrepancy Information Leak
- 38 E.13.3. Output
- 39 E.13.3.1. [XZK] Sensitive Information Uncleared Before Use
- 40 E.13.4. Libraries
- 41 E.13.4.1. [TRJ] Argument Passing to Library Functions
- 42 E.13.4.2. [NYY] Dynamically-linked Code and Self-modifying Code
- 43 E.13.4.3. [NSQ] Library Signature
- 44 E.13.4.4. [HJW] Unanticipated Exceptions from Library Routines
- 45 E.13.5. Files
- 46 E.13.5.1. [EWR] Path Traversal
- 47 E.14. Miscellaneous
- 48 E.14.1. [XZS] Missing Required Cryptographic Step
- 49 E.14.2. Authentication
- 50 E.14.2.1. [XYM] Insufficiently Protected Credentials
- 51 E.14.2.2. [XZN] Missing or Inconsistent Access Control
- 52 E.14.2.3. [XZO] Authentication Logic Error
- 53 E.14.2.4. [XYP] Hard-coded Password

1 **Annex F**
 2 (informative)
 3 **Language Specific Vulnerability Template**
 4

5 **F. Skeleton template for use in proposing language specific information for**
 6 **vulnerabilities**

7 Every vulnerability description of Clause 6 of the main document should be addressed in the annex in the same
 8 order even if there is simply a notation that it is not relevant to the language in question.

9 **F.1 Identification of standards**

10 This section should list the relevant language standards.

11 **F.2 General Terminology**

12 This section should provide an overview of general terminology if desired.

13 **F.3 <language annex>.3.<x> <short title><unique immutable identifier>**

14 *The "short title" and "unique immutable identifier" are copied from 6.x*

15 **F.3.0 <language annex>.3.<x>.0 Status and history**

16 *The header will be removed before publication.*

17 *This temporary section will hold the edit history and status for this language specific vulnerability.*

18 *[The following descriptions should assume that the reader has read the corresponding material in the main body of*
 19 *the document. They should explain differences in terminology and how the general concepts relate to the specific*
 20 *language but should not repeat the general material. The text should be the minimum necessary to describe the*
 21 *relationship to the language, supplemented with examples.]*

22 **F.3.1 <language annex>.3.<x>.1 Language-specific terminology**

23 *This sub-section should deal with terminology differences between the body of the document and the terminology*
 24 *used in the language standard.*

25 **F.3.2 <language annex>.3.<x>.2 Description of application vulnerability**

26 *Replace this with a brief description and examples of how the general description relates to the language. The*
 27 *description should describe why the language is susceptible and provide examples of the vulnerability in the*
 28 *language if possible.*

29 **F.3.3 <language annex>.3.<x>.3 Mechanism of failure**

30 *Replace this with a section describing the mechanism of failure in terms relevant to the language being described.*

31 **F.3.3 <language annex>.3.<x>.4 Avoiding the vulnerability or mitigating its effects in <language>**

32 This vulnerability can be avoided or mitigated in <language> in the following ways:

1 *Replace this with a bullet list summarizing the ways in which the vulnerability can be avoided or mitigated in <*
2 *language>. The bullet points should only apply to <language>. This is the spot for usage guidelines.*

3 **F.3.4 <language annex>.3.<x>.5 Implications for standardization in <language>**

4 *Optionally, replace with a bullet list summarizing various ways that standardization for <language> has assisted in*
5 *mitigating the vulnerability. This is the opportunity to talk about actions that the language committee has already*
6 *taken, such as deprecating features or providing alternatives. There is also the possibility of talking about future*
7 *plans.*

8 **F.3.5 <language annex>.3.<x>.6 Bibliography**

9 *Insert numbered references for other documents specific to <language>. These will eventually be collected into an*
10 *overall bibliography for the TR. So, please make the references complete. Someone will eventually have to*
11 *reformat the references into an ISO-required format, so please err on the side of providing too much information*

12 *rather than too little. Here [1] is an example of a reference:*

13 *[1] Greg Hoglund, Gary McGraw, Exploiting Software: How to Break Code, ISBN-0-201-78695-8, Pearson*
14 *Education, Boston, MA, 2004*

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31

Bibliography

- [1] ISO/IEC Directives, Part 2, *Rules for the structure and drafting of International Standards*, 2001
- [2] ISO/IEC TR 10000-1, *Information technology — Framework and taxonomy of International Standardized Profiles — Part 1: General principles and documentation framework*
- [3] ISO 10241, *International terminology standards — Preparation and layout*
- [4] ISO/IEC TR 15942:2000, "Information technology - Programming languages - Guide for the use of the Ada programming language in high integrity systems"
- [5] Joint Strike Fighter Air Vehicle: C++ Coding Standards for the System Development and Demonstration Program. Lockheed Martin Corporation. December 2005.
- [6] ISO/IEC 9899:1999, *Programming Languages – C*
- [7] ISO/IEC 1539-1:2004, *Programming Languages – Fortran*
- [8] ISO/IEC 8652:1995/Cor 1:2001/Amd 1:2007, Information technology -- *Programming languages – Ada*
- [9] ISO/IEC 15291:1999, Information technology - Programming languages - Ada Semantic Interface Specification (ASIS)
- [10] Software Considerations in Airborne Systems and Equipment Certification. Issued in the USA by the Requirements and Technical Concepts for Aviation (document RTCA SC167/DO-178B) and in Europe by the European Organization for Civil Aviation Electronics (EUROCAE document ED-12B).December 1992.
- [11] IEC 61508: Parts 1-7, Functional safety: safety-related systems. 1998. (Part 3 is concerned with software).
- [12] ISO/IEC 15408: 1999 Information technology. Security techniques. Evaluation criteria for IT security.
- [13] J Barnes, High Integrity Software - the SPARK Approach to Safety and Security. Addison-Wesley. 2002.
- [14] R. Seacord, Preliminary draft of the CERT C Programming Language Secure Coding Standard. ISO/IEC JTC 1/SC 22/OWGV N0059, April 2007.
- [15] Motor Industry Software Reliability Association. *Guidelines for the Use of the C Language in Vehicle Based Software*, 2004 (second edition)².
- [16] ISO/IEC TR24731-1, *Extensions to the C Library, — Part I: Bounds-checking interfaces*
- [17] Steve Christy, *Vulnerability Type Distributions in CVE*, V1.0, 2006/10/04
- [18] Douglas Gregor, Jaakko Jarbvi, Jeremy Siek, *Concepts: Linguistic Support for Generic Programming in C++*
- [19] Gabriel Dos Reis and Bjarne Stroustrup, *Specifying C++ Concepts* POPL06. January 2006

² The first edition should not be used or quoted in this work.

- 1 [20] Motor Industry Software Reliability Association. *Guidelines for the Use of the C++ Language in critical*
2 *systems*, June 2008
- 3 [21] ISO/IEC TR 24718: 2004, *Guide for the use of the Ada Ravenscar Profile in high integrity systems*
- 4 [22] L. Hatton, *Safer C: developing software for high-integrity and safety-critical systems*. McGraw-Hill 1995