

从 *Lessons Learned from* *C/C++ Defects* 缺陷中 学习 C/C++

刘新浙 刘玲 王超 李敬娜 等 编著

这是一本

在研究大量C/C++程序Bug基础上集结而成的书！

这是一本

汇集众多一线C/C++编程人员智慧的书！

这是一本

让您学好C/C++，绕过编程陷阱和障碍的必备案头书！

从 *Lessons Learned from*
C/C++ Defects
缺陷中
学习
C/C++

刘新浙 刘玲 王超 李敬娜 等 编著

人民邮电出版社
北京

23126
205

图书在版编目 (C I P) 数据

从缺陷中学习C/C++ / 刘新浙等编著. — 北京: 人民邮电出版社, 2013. 9
ISBN 978-7-115-32159-6

I. ①从… II. ①刘… III. ①C语言—程序设计
IV. ①TP312

中国版本图书馆CIP数据核字(2013)第151124号

内 容 提 要

C/C++是广泛用于系统和应用软件开发的语言,也是使用最为广泛的编程语言。C/C++易学难用,尤其C++,经过20多年的发展,已经变得非常复杂,给编程人员带来了很大的挑战。那么如何在工程项目中用好C/C++语言、如何绕过Bug构建稳定可靠的生产系统、如何以最快速度全面了解C/C++编程中的陷阱和障碍,编写出健壮可靠的代码呢?本书将通过102个案例,帮助程序员尽快从新手成长为专家。案例涵盖基础问题、编译问题、库函数问题、文件处理、类和对象、内存使用、多线程问题、性能问题等。读者每掌握一个案例就掌握了一个或几个知识点,就能避免一类问题。由于是从大量编程中总结出来的具体Bug案例中学习,这种学习方式更直接,让人印象更深刻。本书将为你成为C和C++高手、编写出完美的程序助一臂之力。

本书适合程序员、测试人员以及C和C++初学者使用,也可以作为各大专院校和培训学校的教学用书。

-
- ◆ 编 著 刘新浙 刘 玲 王 超 李敬娜等
责任编辑 张 涛
责任印制 程彦红 焦志炜
 - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街14号
邮编 100061 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京鑫正大印刷有限公司印刷
 - ◆ 开本: 720×960 1/16
印张: 11.25
字数: 176千字 2013年9月第1版
印数: 1-3500册 2013年9月北京第1次印刷
-

定价: 39.00元

读者服务热线: (010)67132692 印装质量热线: (010)67129223
反盗版热线: (010)67171154

本书作者

主要编著人员:

刘新浙、刘玲、王超、李敬娜、李爱华、陈足先。

其他编著人员:

谭淑丹、曹恒智、付英。

素材提供者:

刘新浙、刘玲、刘晓俊、王艳、李爱华、黄元君、谭淑丹、杨晓霞、王竟时、王超、李颖、曹恒智、李敬娜、陈足先、刘云卿、刘霏暄、邵翔、陶旭颖。

致 谢

本书在编写过程中得到了很多人的帮助。

除本书作者外，原淘宝北京搜索与广告算法测试小组的很多同事都参与了本书的审查工作，提出了文字等方面的一些改进建议，在此一并表示深深的感谢。在本书编写过程中，淘宝公司技术研发部尤其搜索技术部的一些同事，也给予了一些中肯的意见和反馈，在此表示衷心的感谢。

另外，要特别感谢淘宝公司北京测试部门负责人刘立川先生，没有他的鼓励和督促，作者可能就不会在两年多的写作中坚持下来，这本书可能就不会出版面世。

在本书即将出版时，我们非常荣幸地邀请到了著名计算机专家潘爱民老师为本书做推荐序。潘老师在百忙之中抽出时间通读本书，并作序，还给予了非常专业的修改意见，这让本书受益良多。

最后，感谢人民邮电出版社的编辑，在本书出版过程中出谋划策，给予了很专业的建议，并促成了本书的最终顺利出版。

编者

序

C++是易学难用的语言。经过 20 多年的发展，C++已经变得极为复杂，很多语言特性看似优美，但给工程上带来了很大的挑战。因此，如何在工程项目中用好 C++ 语言，并不像通过一本教程来学会写 C++ 程序那么简单，这是一个不断积累经验和吸取教训的过程。工程上推荐使用的往往是 C++ 大量语言特性的一个子集，原因是为了避免在工程项目中埋下各种陷阱。

譬如，Google 是一家重度使用 C++ 的公司，它有许多开源软件（如 Android）使用了 C++ 语言。Google 公开了一份“Google C++ Style Guide——Google C++ 编程风格指南”，阐明了在 Google 的项目中如何有效地使用 C++，应遵守哪些规范，以避免各种可能的陷阱。Google 的这份指南值得工程线上的 C++ 程序员认真阅读。阿里巴巴也是一家重度使用 C++ 的公司，内部有大量的工程用到了 C++ 语言。虽然阿里巴巴集团没有统一的 C++ 编程指导，但各个团队都在摸索并制定出符合自己需要的编程规范，处于快速积累经验的过程中。很高兴淘宝广告技术部的测试团队很早就意识到了 C++ 语言在工程中的诸多陷阱，并有意识地将其收集和整理出来，这就是你看到的这本书。

《从缺陷中学习 C/C++》这本书收集并整理了 102 个实例，这些实例都来自于工程一线实践，虽然大多数看起来像是初学者犯的低级错误，但实质上有一定的代表性，有的错误根源是对 C++ 机制的不理解或者编译过程中的副作用，或者是对 C++ 标准库的实现依赖。通过阅读这些实例，你可以对 C++ 有更细致的理解。

当刘新浙邀请我为本书作序时，我欣然答应，并有幸成为这本书的第一个读者。当时觉得难能可贵的一点是，这样一本书并非是深入研究 C++ 工程多年的 C++ 高手

在指点技术，而是一群懂 C++ 的测试工程师在跟开发工程师探讨如何避免 C++ 语言本身的各种缺陷。这个角度提炼出来的知识，真正是千金难买，因为他们是从“地雷阵”上走过来分享教训的。

我认真阅读了所有的实例，颇有收获，也愿意推荐给每一位在工程线上从事 C++ 编程工作的工程师。若能够结合本书与“Google C++ Style Guide——Google C++ 编程风格指南”一起阅读，我相信一定会有相得益彰之成效。

潘爱民

于杭州

前 言

这是一本在研究大量 C/C++ 程序 Bug 基础上集结而成的书！

这是一本汇集众多一线 C/C++ 编程人员智慧的书！

这是一本让您学好 C/C++，绕过编程陷阱和障碍的必备案头书！

为什么写这样一本书

• 在不同的项目或产品中，不同的开发人员重复着同样的 Bug，甚至同一个人重复相同的 Bug。如果将时间周期拉得更长一些看：一个程序员，从刚毕业参加工作到具备丰富编程经验，从一个新手到成为专家，在这个过程中，每个人都在重复着前人走过的弯路，重复着同样的编程错误。测试人员在日常工作中积累了大量验证 Bug 方面的经验，这些 Bug 是有价值的，总结出来可以让更多人受益。

• C/C++ 是软件/互联网行业最常用的编程语言之一，相对其他语言学习难度高，从新手到专家往往需要多年的磨练。另一方面，C/C++ 开发的系统往往更容易产生严重的生产事故，一旦出现事故，定位问题根源也比较困难。所以，每一个程序员掌握扎实的 C/C++ 基础知识，对于构建稳定可靠的生产系统非常重要。我们希望通过这本书帮助 C/C++ 程序员以最快速度全面了解 C/C++ 编程中的陷阱，编写健壮可靠的代码，从而达到提高软件质量、减少生产故障、提高工作效率的目的。

相对其他 C/C++ 类书籍，本书有以下特点和优势：

- 从具体 Bug 中学习

全书由 102 个案例组成，每个案例分析一个 Bug。读者掌握了一个案例就是掌握了一个知识点，就能避免一类问题。由于是从具体 Bug 案例中学习，这种学习方式更直接，更让人印象深刻。普通的 C/C++ 编程书，即便看完后，写代码时也无法避免 Bug，这是因为，书虽然看完了，知识也理解了，但你并不知道哪里有陷阱。

- 来源于工程实战的宝贵经验

本书中的所有案例都来自于软件/互联网行业开发生产中遇到的实际问题，都是在前人的错误和弯路中总结出来的实战经验，堪称 C/C++ 编程方面的“干货”。

- 学习起来更有成就感

本书由一个个案例组成，在每个案例中，先给出错误代码示例，然后描述 Bug 的现象和后果，随后对该 Bug 进行具体分析，最后给出解决方案及建议。这种案例分析式的组织方式，会引导读者先对案例中提出的问题进行思考，当看到案例分析和解决方案时常常有恍然大悟的感觉，让学习过程变得简单，并充满乐趣。

- 更宽的知识面

一个 C/C++ 程序员即使工作多年，由于受所接触项目和技术方向的限制，视野（C/C++ 编程中值得注意的知识点）往往是有限的。这本书中的案例收集于大量工程项目，几乎涵盖了 C/C++ 编程中的方方面面，远超出一般程序员所能接触的范围。掌握了这本书中的内容，能避免大多数问题。

本书适用范围

• 本书适合已经写过一些 C/C++ 代码、期望尽快积累实战经验的 C/C++ 程序员阅读学习。本书也适合打算提高代码编写和代码阅读分析能力的软件测试人员，书中的每个案例都可应用于白盒测试中。

本书代码运行环境

• 本书中的所有案例代码都是针对 Linux C/C++ 环境，在 Redhat Linux 环境下编译（GCC, G++）测试通过。代码可以通过人民邮电出版社网站（www.ptpress.com.cn）下载。

最后，由于本书作者在 C/C++ 编程方面的经验和技能有限，书中可能会有一些描述不够清晰或不正确的地方，读者如有发现请及时告诉我们，我们会再进行修正。我们的联系邮箱是 cppbugbook@gmail.com，编辑联系邮箱是 zhangtao@ptpress.com.cn。在此表示感谢！

期望这本书能真真切切地帮助到他人。

编者

目 录

第 1 章 基础问题	1
1.1 运算符优先级引起的问题 ●*	1
1.2 不加括号的宏定义引起的错误 ●*	2
1.3 污染环境的宏定义 ●*	3
1.4 多语句宏定义使用错误 ●*●*	4
1.5 char 转为 int 时高位符号扩展的问题 ●*	6
1.6 int 转为 char 时的数据损失 ●*●*	7
1.7 非法的数组下标 ●*●*	9
1.8 有符号 int 与无符号 int 比较的后果 ●*●*	10
1.9 有符号的困惑 ●*●*	11
1.10 整除的精度问题 ●*	13
1.11 浮点数比较的精度问题 ●*●*	14
1.12 最小负整数取相反数溢出 ●*●*	15
1.13 临时变量溢出 ●*●*	16
1.14 size_t 导致的死循环 ●*	17

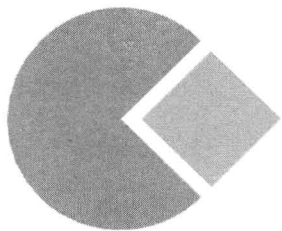
1.15	误用 short 引起缓冲区溢出	●*	18
1.16	区分 continue 和 return	●*	20
1.17	指针常量和常量指针的区别	●*	21
1.18	字符数组和指针不总是等价的	●*	23
1.19	结构体成员变量初始化的隐患	●*	24
1.20	返回值非 void 的函数没有返回值	●*●*	25
1.21	cin>>和 getline 混用导致的奇怪问题	●*●*	27
1.22	小结		29
第 2 章 编译问题			30
2.1	动态链接库加载错误版本	●*	30
2.2	相同名称静态库的链接顺序	●*	32
2.3	使用命名空间来区分不同 cpp 中的同名类	●*●*	33
2.4	C++模板编译时依赖名称查找	●*	34
2.5	违背 ODR 原则可能会带来的意想不到的问题	●*●*●*	36
2.6	变量共用内存时使用 O2 优化编译	●*●*●*●*	38
2.7	小结		40
第 3 章 库函数问题			41
3.1	sprintf 函数引起的缓冲区溢出	●*●*	41
3.2	snprintf 函数 format 参数的问题	●*	43
3.3	错误使用 snprintf 函数返回值	●*●*	44
3.4	字符串复制不完整	●*●*	45

3.5	string 类的 c_str 方法使用不当	46
3.6	string 类的 “[]” 操作符使用不正确	48
3.7	不正确的字符串比较	49
3.8	strncpy 函数没有复制结束符	51
3.9	调用 memcpy 函数前未初始化缓冲区	52
3.10	误用 sizeof 操作符取字符串长度	53
3.11	string 类 find 函数返回值判定	54
3.12	stringstream 的清空	56
3.13	调用 strtptime 函数前需初始化 tm	57
3.14	使用 feof 函数的陷阱	59
3.15	for 循环中调用 vector 容器 insert 函数	60
3.16	multiset 容器 erase 函数的误用	62
3.17	慎用容器类 erase 函数的返回值	63
3.18	for 循环中调用 vector 容器 erase 函数	65
3.19	getopt 函数参数问题	67
3.20	不用 errno 判断系统调用是否成功	69
3.21	strcat 函数造成的段错误	70
3.22	危险的 strdup 函数	71
3.23	小结	73
第 4 章 文件处理		74
4.1	程序异常退出时未关闭已打开文件	74

4.2	目录打开后未关闭	●*	75
4.3	写文件没有调用 <code>fflush</code>	●*	76
4.4	读文件 <code>fread</code> 的返回值不能忽略	●*●*	77
4.5	<code>getline()</code> 使用不当导致的死循环	●*●*	78
4.6	未重置流状态导致读文件错误	●*●*	80
4.7	小结		82
第 5 章 类和对象			83
5.1	对象的浅复制	●*●*●*	83
5.2	构造函数中的操作符重载	●*●*●*	85
5.3	拷贝构造函数不能模板化	●*●*●*	87
5.4	析构函数未捕获异常引发 <code>coredump</code>	●*●*	89
5.5	构造函数中抛出异常引起内存泄露	●*●*	91
5.6	多态性未生效	●*	93
5.7	基类成员函数被隐藏	●*●*●*	94
5.8	匿名对象引起的内存泄露	●*	96
5.9	基类非虚析构函数引发内存泄露	●*●*●*	97
5.10	删除 <code>void*</code> 指针引发内存泄露	●*●*●*	100
5.11	STL 容器不会自动释放指针指向的对象	●*●*●*	102
5.12	静态成员类内初始化	●*	104
5.13	<code>union</code> 作为类的成员时需要构造函数	●*●*	105
5.14	成员函数尾部缺失 <code>const</code> 标注	●*●*●*	107

5.15	使用 <code>memset</code> 初始化 <code>class</code>	●*●*●*	109
5.16	<code>dynamic_cast</code> 转换失败返回 <code>NULL</code>	●*●*	110
5.17	小结		113
第 6 章 内存使用			114
6.1	数组越界	●*	114
6.2	数组定义和值初始化形式混淆	●*	115
6.3	数组传参时的 <code>sizeof</code>	●*	116
6.4	临时对象的生存期	●*●*	117
6.5	变量的作用域	●*	119
6.6	指针变量的传值和传址	●*	120
6.7	指针赋值和指针赋址的混淆	●*	121
6.8	指针释放后再次使用	●*	122
6.9	重复申请内存未释放	●*	124
6.10	<code>delete</code> 与 <code>delete[]</code> 的区别	●*●*	126
6.11	函数中途退出忘记释放内存	●*●*	126
6.12	二维数组的内存泄露	●*●*	127
6.13	临时变量内存不能返回	●*●*	128
6.14	正确使用引用参数和引用返回值	●*●*	129
6.15	试图产生的指针很可能不存在	●*●*	130
6.16	结构体成员内存对齐问题	●*●*	131
6.17	<code>String</code> 对象何时需 <code>delete</code>	●*	134

6.18	小结	135
第 7 章	多线程问题	136
7.1	局部静态变量非线程安全 ●*●*	136
7.2	string 类 append 操作非线程安全 ●*●*●*	138
7.3	中途退出造成的线程阻塞 ●*●*●*	142
7.4	结构体位域成员线程安全问题 ●*●*●*●*	144
7.5	多线程写文件引发内容被覆盖 ●*●*●*	146
7.6	线程未 join 引起的内存泄露 ●*●*●*	148
7.7	小结	150
第 8 章	性能问题	151
8.1	strlen 用作循环条件影响性能 ●*	151
8.2	STL 容器 list 使用时忌频繁调用 size() ●*●*	152
8.3	误用 clear 回收 vector 内存 ●*●*●*	154
8.4	calloc 在 glibc 高版本下性能劣于低版本下 ●*●*●*●*	156
8.5	小结	157
第 9 章	C/C++编程中其他问题	158
9.1	中文截断成乱码 ●*●*●*	158
9.2	不必要的类型转换 ●*●*	159
9.3	不确定的函数参数赋值 ●*●*●*	161
9.4	epoll 边沿触发模式下的陷阱 ●*●*●*●*	163
9.5	小结	166



第1章 基础问题

本章从基础问题讲起，涉及运算符优先级、宏定义、类型转换（显式和隐式）、指针、数组等 C/C++ 编程的一些基本概念和知识点。这些问题虽然简单，但如果不引起重视，也会给编程带来很多麻烦。

1.1

运算符优先级引起的问题



- 代码示例

```
//To get 2*n+1
int func ( int n )
{
    return n << 1 + 1;
}
```

- 现象&后果

上述代码中的函数 `func` 本意是期望计算 $2*n+1$ ，但程序实际运行结果是 $4*n$ 。

- Bug 分析

这段代码使用左移 1 位来代替乘以 2 的运算，是很好的方法，但是编程者弄错了运算符“<<”和“+”的优先级。C/C++语言规定运算符“+”的优先级高于运算符“<<”，因此，上述语句“`return n<<1+1`”等同于“`return n<<(1+1)`”，所以，会先进行加法运算，再进行左移运算，得到结果 $4*n$ 。

修正的方法是在表达式中添加必要的括号。

- 正确代码

```
//To get 2*n+1
int func ( int n )
{
    return (n << 1) + 1;
}
```

1.2 不加括号的宏定义引起的错误



- 代码示例

```
#define PERIMETER(X,Y) 2*X+2*Y
int main(void)
{
    int length = 5;
    int width = 2;
    int high = 8;
    int result = 0;
    result = PERIMETER(length, width) * high;
    printf("result=%d \n", result);
}
```

- 现象&后果

上述代码是期望通过一个宏先计算一个矩形的周长，然后再乘以高，结果应该为 112，但实际计算结果为 42，与预期不符。

- Bug 分析

上述代码中，语句“`result = PERIMETER(length, width) * high`”本意是希望先进行宏定义部分的运算，然后再乘以变量 `high`，但是，由于宏替换是在预编译阶段进行的，宏替换本身只是文本替换，上述语句在宏替换后变成了“`result = 2 * length + 2 * width * high`”，按照运算符优先级规则会先做乘法运算再做加法运算，故导致运算结果不符合预期。因此，用于表达式的宏，在宏定义时给整体语句加上小括号是比较保险的做法。