Xpert.press

Die Reihe **Xpert.press** vermittelt Professionals in den Bereichen Softwareentwicklung, Internettechnologie und IT-Management aktuell und kompetent relevantes Fachwissen über Technologien und Produkte zur Entwicklung und Anwendung moderner Informationstechnologien.

Ralf Schneeweiß

Moderne C++ Programmierung

Klassen, Templates, Design Patterns

Mit 19 Abbildungen und 397 Listings

Springer

Ralf Schneeweiß
Gölzstraße 8
72072 Tübingen
ralf.schneeweiss@oop-trainer.de

Bibliografische Information der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.ddb.de abrufbar.

ISSN 1439-5428
ISBN-10 3-540-22281-2 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-22281-1 Springer Berlin Heidelberg New York

Springer ist ein Unternehmen von Springer Science+Business Media
springer.de

Satz: ptp, Berlin
Herstellung: LE-TEX, Jelonek, Schmidt & Vöckler GbR, Leipzig
Umschlaggestaltung: KünkelLopka Werbeagentur, Heidelberg
Gedruckt auf säurefreiem Papier 33/3142 YL – 5 4 3 2 1 0

Vorwort

Um im Jahr 2006 noch ein Buch über C++ zu schreiben, bedurfte es meinerseits einiger Vorüberlegungen, bis ich mich zu diesem Schritt entscheiden konnte. Da ist zum einen die aktuelle Veränderung in der Nutzung von Programmiersprachen in der Anwendungsentwicklung, die ein deutliches Zeichen für die Zukunft von Java und anderen moderneren Programmiersprachen ist und ein ebensolches Zeichen für das allmähliche Verschwinden von C++ aus dieser Domäne. Zum anderen ist da ein scheinbares Defizit in der Sprache C++ selbst, die sich ein Stück weit gegen die unmittelbare Anwendung neuer Technologietrends zu sträuben scheint. Die enorme Komplexität der Sprache ist dafür sicher ein Grund unter anderen. Was sind also meine Gründe, doch ein Buch über diese alte und noch in vielen Bereichen unersetzliche Sprache zu schreiben?

In den letzten Jahren fand eine rasante Entwicklung bei den Mobile Devices, den Embedded Systems und den vermittelnden Systemen in Netzen statt. Diese Entwicklung dauert auch heute noch an. Es entsteht viel mehr neue Hardware als in den neunziger Jahren und es entsteht die dazugehörige Betriebssystemlandschaft. Gerade für diese neuen Systeme braucht man heute C++ als systemnahe und deterministische Programmiersprache. Da heute aufgrund der besseren Portierbarkeit mancher gängigen Compiler C++ auch auf fast allen neuen Plattformen zur Verfügung steht, erfährt die Sprache genau dort eine späte Verbreitung. Häufig werden Systeme und Teile davon in Crosscompileumgebungen entwickelt, die Windows oder Linux als Entwicklungsbasis definieren und ein Embedded System als Zielsystem haben. Diese aktuelle Situation ist es, die C++ meiner Meinung nach noch zu einem starken Entwicklungspotential verhilft, denn sie stellt an den Entwickler dieser neuen Domäne viel höhere Anforderungen als an den reinen Anwendungsentwickler, der es in den neunziger Jahren gewohnt war, C++ zusammen mit der Unterstützung mächtiger Frameworks einzusetzen.

Während in den Neunzigern die verwendeten Methoden lange Zeit auf der klassischen Objektorientierung stagnierten, brachte der ANSI/ISO-Standard Ansätze mit, die in eine ganz andere Richtung weisen, als es die traditionellen C++-Techniken tun. Techniken, deren methodische Anwendung in der Praxis noch längst nicht vollständig erprobt und ausgeschöpft ist und die sich mit Portierbarkeit, Wartbarkeit, Wiederverwendbarkeit und Reduktion von Komplexität befassen. Diese neuen Methoden und die dazugehörige Syntax so darzustellen, dass sie einerseits in einem theoretischen Kontext stehen und andererseits praktisch anwendbar werden, das war mein Ziel für dieses Buch.

Danksagung

Zur Fertigstellung dieses Buches haben einige Freunde beigetragen, denen ich an dieser Stelle ganz ausdrücklich meinen Dank aussprechen möchte. Da ist Marcel Rüdinger, der mit dem Angebot, die Schreibarbeit am Sitz seiner Firma Datronic Portugal in Madeira zu erledigen, das ganze Projekt ans Laufen brachte. Sein Inselleben brachte mich dazu, in See zu stechen. Er ist ein Mensch, der sich Ziele setzen kann, und möglicherweise färbte das ein wenig auf mich ab. In der langwierigen Arbeit mit dem Text und den Beispielen war mir Ekaterina Myrsova eine gleichmütige Begleiterin. Sie gab mir die Kraft zum Durchhalten. Sascha Morgenstern übernahm mit viel Geduld die inhaltliche Korrektur des Textes und der Beispiele und stand für Diskussionen zur Verfügung. Ebenso möchte ich Tobias Geiger und Robert Fichtner danken, die mit ihren kritischen Anmerkungen gegen Ende des Buchprojekts wesentliche Vorschläge zur Verbesserung der Textstruktur lieferten. Nicht zuletzt stand mir bei der Perfektionierung des LaTeX-Satzes Annett Eichstädt bei.

Inhaltsverzeichnis

1 Einführung

Dieses Buch verfolgt zwei Zielrichtungen. Es wendet sich einerseits an Entwickler, die einen fundierten Einstieg in die C++-Programmiersprache suchen und sie in Projekten praktisch einsetzen wollen. Andererseits ist dieses Buch eine Art thematisches Nachschlagewerk, das es dem Kenner der Sprache ermöglicht, bestimmte Bereiche nachzuarbeiten und zu vertiefen. Insbesondere derjenige, der sich mit den dahinter liegenden Leitideen befassen möchte, sollte bei diesem Buch auf seine Kosten kommen.

Vor allem der erste Teil des Buchs ist in der Form eines Lehrbuchs geschrieben. Die einzelnen Unterkapitel der Sprachbeschreibung bauen inhaltlich aufeinander auf und führen den interessierten Leser schrittweise an die Sprache C++ heran. Der zweite Teil des Buchs greift verschiedene moderne Techniken der C++-Entwicklung auf und erläutert diese in einer Form, die nicht auf den Zusammenhang mit den Nachbarkapiteln angewiesen ist. Es wird weniger auf die Zusammenhänge der Techniken eingegangen als vielmehr auf die Alleinstellungsmerkmale von diesen. Das Buch ist mit zahlreichen Listings ausgestattet, die das im Text Gesagte veranschaulichen und einen Anknüpfungspunkt für eigene Experimente abgeben. Deshalb finden Sie die meisten größeren Listings als Quellcode zum Download auf der Internetseite:

```
http://www.oop-trainer.de/publications/de/moderncpp
```

Die Sprache C++ hat nun schon eine beachtliche Entwicklungszeit hinter sich. In dieser Zeit übernahm C++ vieles der sich parallel entwickelnden Programmierparadigmen, wie der Objektorientierung und später der generischen und generativen Programmierung. Nicht alles, was auf diesem zeitlichen Weg entstand, war tragfähig. Manches verschwand nach wenigen Jahren schon völlig aus dem Methodenschatz des Entwicklers, wenn sich herausstellte, dass eine Methode nicht mit anderen integrierbar oder auch nur fehlerhaft war[1]. C++ hatte Anteil an diesen Entwicklungen und nahm vieles auf, so dass heute die Sprache Altes und Neues wie keine andere Sprache vereint. Nicht alles davon ist nach dem heutigen Kenntnisstand sinnvoll. Anderes ist wiederum sehr zukunftsweisend, wird aber noch wenig eingesetzt. In den nachfolgenden Kapiteln soll Altes und Neues angesprochen werden. Dabei wird sich der Text am ANSI/ISO-Standard von 1998 orientieren, obwohl der Autor weiß, dass alle breit eingesetzten Compiler den Standard bis dato nicht vollständig unterstützen. Manche von ihnen weichen sogar grob von ihm ab und unterstützen alte AT&T-Standards oder Mischformen zwischen AT&T C++ und ANSI C++. Da das Buch die praktische Anwendung von C++ beschreiben

[1] Wie zum Beispiel die private oder die geschützte Sichtbarkeit in der Vererbung.

soll, kann diese Tatsache nicht ignoriert werden. Es wird also an den wesentlichen Stellen auch Hinweise auf die unterschiedlichen Compiler geben, und es werden Methoden beschrieben, wie kritische Stellen bezüglich einer Unverträglichkeit unterschiedlicher Standards im Code zu vermeiden sind. Es werden also auch die gängigen Produktionscompiler GNU C++, MS VC++, Borland C++ und Metrowerks C++ an entsprechender Stelle angesprochen.

C++ ist eine objektorientierte Programmiersprache. Trotz der Kritik an ihr bezüglich der Umsetzung des OO-Konzepts ist sie vielleicht auch die OO-Sprache, die diesem Paradigma einen breiten Durchbruch in der Akzeptanz ermöglicht hat – schließlich war es C++, das in den 90er Jahren prozedurale Compilersprachen wie Pascal und C in der PC-Programmierung ablöste und damit selbst eine enorme Verbreitung erfuhr. Es ist deshalb für ein solches Buch wie das vorliegende notwendig, auf die objektorientierten Techniken einzugehen. Es kann jedoch keine durchgehende Methodik des objektorientierten Entwurfs geben, zumal der syntaktische Umfang der Sprache C++ den textuellen Raum des Buches zur Genüge in Anspruch nimmt. Die syntaktischen und technischen Mittel, mit denen C++ die objektorientierte Modellierung unterstützt, werden ausgiebig behandelt. Der Leser sollte also schon Kenntnisse über die Objektorientierte Programmierung mitbringen oder sich parallel erarbeiten, wenn er dieses Buch liest. Neben der Objektorientierung wurde Mitte der 90er Jahre ein weiteres Paradigma in C++ integriert: die Methodik der generischen Programmierung. Seit einigen Jahren wird sie auch der aspektorientierte Ansatz genannt. Dazu war die neue Technik der parametrisierbaren Klassen, der „Templates“, notwendig. Große Bereiche der Standardbibliothek wurden seit der Integration der Templates in C++ auf diese Technik umgeschrieben, da sie große Vorteile gegenüber der objektorientierten aufweist, wenn es darum geht, wiederverwendbaren Code zu schreiben. Dem fortschreitenden Einsatz der generischen Programmierung in C++-Code wird im zweiten Teil des Buchs Rechnung getragen, indem diese Technik in verschiedenen Anwendungsbereichen facettenreich dargestellt wird.

Einen großen Teil im Buch nimmt also die Templateprogrammierung und deren methodische Anwendung ein. Gerade in diesem Bereich hat sich mit der ANSI-Standardisierung 1998 am meisten getan. Dabei werden bis heute, acht Jahre nach dem Standard, die Techniken nur wenig genutzt. Häufig wird nicht erkannt, was in diesen neuen Modellierungstechniken steckt, oder man hat eine gewisse Angst davor, dass die negativen Folgen des Einsatzes die positiven überwiegen könnten. Dazu ist es notwendig die positiven und negativen Konsequenzen des Einsatzes parametrisierbarer Klassen kennenzulernen, um Entscheidungen für eigene Projekte fällen zu können.

Ein durchgängiger Schwerpunkt liegt in dem Buch auf dem Schreiben von portierbarem Code. Das heißt in der Praxis etwas mehr als nur das Einhalten des Standards. Es wurde ja schon angesprochen, dass die gängigen Compi-

ler nur unzureichend standardkonform sind. Deshalb ist es sinnvoll, zwar standardkonform zu programmieren, aber nur solche Standardkonstrukte zu verwenden, die von allen Compilern identisch verstanden werden. Die in diesem Sinne wesentlichen Problemkonstrukte werden Sie in diesem Buch finden.

Abschließend soll noch gesagt werden, dass die Themen des Buchs vollkommen systemunabhängig sind. Die Programmierbeispiele und der Inhalt des Textes sind mit wenigen Ausnahmen nicht spezifisch für die Programmierung des einen oder anderen konkreten Betriebssystems. Es sind unabhängige Techniken, wie sie der ANSI/ISO C++-Standard von 1998 vorsieht. Die Inhalte können also auf beliebige Projekte angewandt werden, sei es in der Entwicklung einer Anwendung für einen Windows Desktop oder in der Embedded-Entwicklung. Das Verständnis der Leitideen des ANSI/ISO C++-Standards ist heute auf fast allen Plattformen, auf denen C++-Entwicklung durchgeführt wird, eine wichtige Grundlage für ein erfolgreiches Erreichen des Projektziels. Dieses Verständnis möchte das Buch geben.

2 Die Sprache C++

C++ ist inzwischen eine relativ alte Programmiersprache, die viele Brüche in ihrer Geschichte erlebt hat. Vieles, was man heute in C++ findet, ist älteren Programmiermethoden und Ideen zu verdanken, die heute nicht mehr unbedingt aktuell sind. Modernere objektorientierte Sprachen, wie zum Beispiel JAVA, sind in ihrer syntaktischen Struktur weit weniger komplex, da sie weniger unterschiedlichen Paradigmen[1] folgen müssen. In den folgenden Abschnitten soll ein Überblick über die verschiedenen Phasen der Entwicklung von C++ gegeben werden. Damit sollen auch die unterschiedlichen Ideen umrissen werden, die bis heute in der Sprache C++ stecken, ihre Grundlage bilden und sie einerseits mächtig und andererseits an manchen Stellen auch unübersichtlich machen. Für den C++-Entwickler stellt sich immer die schwierige Frage der Auswahl geeigneter Methoden für seine aktuelle Arbeit. Dabei muss zwischen konkurrierenden Leitideen ausgewählt werden und es müssen solche verworfen werden, die nicht mehr aktuell sind.

2.1 Geschichte und Paradigmenwandel

C++ entstand in den frühen 80er Jahren aus der Programmiersprache C heraus. Die Sprache wurde in den Bell Laboratories, die später von AT&T übernommen wurden, entwickelt. Federführender Ideengeber und Entwickler war Bjarne Stroustrup, der diese Rolle über lange Jahre beibehielt. Aus C wurde zunächst ein „C mit Klassen“ entwickelt, um die damals neuen Methoden der Objektorientierung den C-Programmierern zugänglich zu machen. Später, 1983, wurde erstmals der Name „C++“ für die neue Sprache verwendet. Er setzt sich zusammen aus dem Namen „C“ und dem Inkrementoperator „++“ aus C, um anzudeuten, dass C++ über C quasi einen Schritt hinausgeht. Die Objektorientierte Programmierung steckte zu damaliger Zeit noch in den Kinderschuhen. Die Methodenausstattung dieser neuen Leitidee war nach heutigem Maßstab noch etwas dürftig und erst in Entwicklung begriffen. Die Sprachkonstrukte von C++ wurden natürlich an den Stand der damaligen Sofwaretechnik angepasst. Mit der fortschreitenden Technik der OO-Programmierung bekam C++ auch neue sprachliche Strukturen. Die Firma AT&T, die die Bell Laboratories übernahm, standardisierte C++ mit einem eigenen Firmenstandard, der von fremden Compilerherstellern übernommen werden konnte. Die meisten Hersteller von C++-Entwicklungssystemen

[1] Paradigma: Leitidee.

machten sich diese AT&T-Standards zunutze und unterstützten sie. Natürlich boten die verschiedenen Hersteller auch spezifische Features ihrer Compiler an, die über den jeweils gültigen Stand der AT&T-Quasi-Standards hinausgingen. Die wesentlichen Standards von AT&T waren durch die Compiler 1.0, 1.1, 2.0, 2.1 und 3.0 vorgegeben. Die jeweilige aktuelle Standardimplementierung wird bis heute *cfront* genannt. Ein cfront-Compiler ist also konform zum aktuellen Standard. Seit Anfang der 90er Jahre wurde die Standardisierung durch das American National Standardisation Institute angestrebt.

Das für C++ zuständige Komitee brauchte allerdings bis 1998, um einen Standard zu verabschieden. In der Folge der ANSI-Standardisierung wurde der Standard auch zur internationalen ISO-Norm erklärt. Was war nun in den verschiedenen Standardisierungsschritten mit C++ passiert?

In den AT&T Standards der 1er Versionen bekam C++ vor allem die Sprachmerkmale, die zur Kapselung von Daten in Klassen notwendig sind. In der objektorientierten Entwicklung der damaligen Zeit bis 1986 standen in besonderer Weise die Gedanken der Datenkapselung und der Funktionsüberladung im Vordergrund. Vererbung wurde eher zur Anhäufung von Daten und Funktionalität verwendet, um eine Art Wiederverwendung gemeinsamer Funktionalität zu erreichen. Die dafür geschaffenen Sprachstrukturen bestehen in C++ bis heute, auch wenn wir die Vererbung nicht mehr in dem damaligen Sinne anwenden, denn sie hat sich als Sackgasse erwiesen. Dass die Vererbung in C++ eine Sichtbarkeit besitzt[2] und diese auch noch privat ist, hat den beschriebenen Hintergrund.

Mit der AT&T-Version 2.0 fanden die virtuellen Funktionen Einzug in die Sprache C++ und mit ihnen das Konzept der Polymorphie. Man kann daher sagen, dass im Grunde genommen erst diese Version der Sprache dem ähnlich wurde, was wir heute als C++ kennen. Mit der Polymorphie als zentralem Konzept der Objektorientierung wurde C++ überhaupt erst zu einer OO-Sprache. Vorher war das, was man C++ nannte, ein erweitertes C mit etwas besserer Typenkontrolle, anderen Kommentarzeichen und der Möglichkeit Daten zu verstecken. Jetzt erst konnten OO-Entwürfe mit C++ verwirklicht werden, die auf die späte Bindung von Methoden aufbauten und damit den Funktionsfokus zugunsten des Objektfokus aufgaben. OO war in C++ angekommen, wenngleich es an der einen oder anderen Stelle noch etwas hakte. So konnten virtuelle Methoden in Basisklassen deklariert werden, um die Polymorphiebasis für abgeleitete Klassen bereitzustellen. Diese virtuellen Methoden mussten formal aber immer implementiert werden, was inhaltlich natürlich wenig sinnvoll war und den einen oder anderen Entwickler noch zu Fehlentscheidungen verführte („Was die Sprache erzwingt, kann ja so falsch nicht sein! Oder?“).

Erst 1992 in der AT&T-Version 3.0 bekam C++ den OO-Rundschliff. Dazu gehörten in allererster Linie die rein virtuellen Funktionen ohne Implementierung. Dieses sehnlichst erwartete und in Newsgroups vorher schon disku-

[2] Siehe Abschnitt 2.2.21 auf Seite 76

tierte Sprachfeature machte nun allen Entwicklern deutlich, dass es Klassen ohne eigene Funktionalität geben kann. Klassen, die eine reine Schnittstellenfunktion haben. Es gab auch noch Versionen bis 3.3, die aber hauptsächlich Elemente einführten, die mit dem OO-Paradigma nichts mehr zu tun hatten. So wurden die Templates und das Exception Handling in den 3er Versionen von C++ aufgenommen. Anfang der 90er Jahre gab es in der Entwicklergemeinde eine Diskussion um die Anwendung der neuen Templatetechnik. So gab es schon Stimmen, die die Templates befürworteten, da sie doch den Compiler in die Lage versetzten, Code zu erzeugen, statt nur zu übersetzen. Andere befürchteten unbeherrschbare Nebeneffekte – z. B. Aufblähung des erzeugten Codes – und eine zunehmende Komplexität der Sprache. Man war schließlich zu einem gesicherten Wissensstand über die Objektorientierung gelangt, da sollte das Instrumentarium nicht durch völlig fremde Konzepte aufgeweicht oder gestört werden. Insbesondere am Design der für C++ längst überfälligen Containerbibliothek entzündete sich die Diskussion. Dynamische Datencontainer werden in C++ benötigt um 1-zu-n-Relationen zwischen Klassen und Objekten zu modellieren. In den C++-Standardbibliotheken der AT&T-Standards waren keine Containerbibliotheken enthalten, weshalb Compilerhersteller eigene herstellerspezifische Bibliotheken auslieferten. Software, die diese herstellerspezifischen Implementierungen nutzte, war natürlich nicht mehr portierbar, was dazu führte, dass häufig für jede spezielle Software extra Container implementiert wurden. In dieser Zeit entstanden sehr viele – sehr viele! – verschiedene, mehr oder weniger gut funktionierende Container (in jedem Softwareprojekt die eigene verkettete Liste). Solche Container können mit reinen OO-Methoden entwickelt werden, oder aber auch mit den generischen Techniken, die durch die Templates in die Sprache eingeführt wurden. Den Streit entschied A. Stephanov[3] mit seiner Standard Template Library zugunsten der generischen Technologie. Diese Bibliothek beruht ganz auf dem Einsatz der C++-Templates und war frei verfügbar. Die STL ist gegenüber entsprechenden OO-Containerbibliotheken von überzeugender Eleganz und Flexibilität. Ein weiterer Vorteil ist die gute Beherrschbarkeit der Laufzeitaspekte beim Einsatz der Bibliothek. Dies alles zusammengenommen führte zunächst dazu, dass die generischen Templatetechniken anerkannt wurden.

Eine weitere Folge war die Aufnahme der STL in den ANSI/ISO-Standard. Die Arbeiten um die Standardisierung veränderten die STL nicht wesentlich. Die STL ist seitdem in leicht angepasster Form Bestandteil der C++-Standardbibliothek. Außerdem verdrängten generische Technologien ihre OO-Pendants aus einigen Bereichen der Bibliothek. Mit gewisser Berechtigung kann man sagen, dass mit der Integration der Templatetechnologie in C++ nun Methoden zur Verfügung stehen, die die lang versprochene Wiederverwendbarkeit Wirklichkeit werden lassen. Die OO-Methoden sind dazu nur marginal in der Lage. Es kam auch das Konzept der Namensräume mit

[3] Der Autor der STL, Alexander Stephanov, wurde durch Hewlett Packard und SGI für deren Entwicklung unterstützt. Über Server bei HP stand die STL schon 1994 zum Download bereit.

in den Standard. Außerdem wurden neue Typenkonvertierungen definiert. Die Templatetechniken wurden mit dem ANSI/ISO-Standard weiter verfeinert und die syntaktischen Möglichkeiten nahmen zu. Sie haben mit dem Standard eine Komplexität erreicht, die das bloße Erlernen der Syntaxregeln enorm erschweren. Dazu kommt noch, dass mit der Verabschiedung des Standards praktisch kein Hersteller in der Lage war, einen Compiler mit korrekter Standardkonformität zu liefern. Lange Zeit stand praktisch nur ein Experimentalcompiler, der Comeau-C++-Compiler, zur Verfügung. Der erste nahezu standardkonforme Produktivcompiler war der Metrowerks Code Warrior, der vor allem für die Macintosh-Plattform verwendet wurde. Eine ähnlich gute Unterstützung des Standards erreichte der GCC-C++-Compiler mit seinen 3er Versionen. Erst mit den stabilisierten Versionen dieses Compilers 2005 stand eine standardkonforme Entwicklungsplattform in der Breite zur Verfügung. Es kommt noch eine weitere Schwierigkeit für die Anwendung des Standards hinzu: in den ANSI/ISO-Standard wurde Exception Handling als integraler Bestandteil aufgenommen. Dabei wurde der `new`-Operator zur Speicherallokation neu definiert. In den AT&T-Versionen von C++ liefert der `new`-Operator einen Nullzeiger zurück, wenn die Allokation fehlschlägt. In der ANSI-Version wirft `new` eine Exception, um einen Fehlschlag anzuzeigen. Diese Änderung ist derart tiefgreifend, dass eine Abwärtskompatibilität standardkonformer Compiler nicht mehr gegeben ist. Wenn Softwareprojekte alte Codeteile verwenden, kann damit der Standard nicht zur Anwendung kommen, da sonst die alte Fehlerbehandlung deaktiviert wird und es zu undefinierten Laufzeitzuständen kommen kann – fast schon zwangsläufig kommen muss. Alter und neuer Code vertragen sich nicht. Alter Code muss zunächst umgeschrieben werden, um mit neuem standardkonformen Code zu harmonisieren[4]. Viele erfolgreiche Bibliotheken sind nach einem AT&T- oder verwandten Standard geschrieben und werden weiter verwendet[5]. Ein weiteres Problem besteht in den Gewohnheiten der Entwickler. Um standardkonform zu entwickeln müssen sie Exception Handling als integralen Bestandteil ihrer Software begreifen, was einen Bruch mit alten Gewohnheiten und ganz neue Programmstrukturen erfordert. Diese Gründe zusammengenommen verhindern acht Jahre nach der Verabschiedung des Standards immer noch seine breite Akzeptanz.

Heute ist schon die nächste Version des ISO-Standards in Diskussion. Dabei wird auch die Aufnahme von freien Bibliotheken, wie z. B. der boost-Library diskutiert. Zu hoffen bleibt allerdings, dass Verhaltensweisen des Sprachkerns nicht mehr redefiniert werden, wie es bei der letzten Standardisierung geschah, da das einer zügigen Annahme eines Standards erneut im Wege stehen könnte.

[4] Es ist im ANSI/ISO-Standard zwar ein Migrationspfad für die alte Anwendung des new-Operators vorgesehen – die „nothrow"-Variante –, jedoch ist eine Portierung alten AT&T-konformen Codes deutlich aufwändiger als die reine Anpassung der Verwendungen des Operators new.

[5] Ein prominentes Beispiel dafür ist sicher die MFC.

2.2 Grundlagen

C++ ist eine reine Compilersprache. Der Quelltext muss durch einen Compiler in ausführbaren Code übersetzt werden, damit ein Programm ablaufen kann. Der Quelltext wird für den Ablaufvorgang der Software nicht mehr gebraucht. Dieses traditionelle Konzept teilt C++ mit älteren prozeduralen Sprachen wie Fortran, Pascal und natürlich auch C. Von C hat C++ die typische Zusammenarbeit zwischen Compiler und Linker geerbt. Nicht jede Compilersprache braucht einen separaten Linker. Das C-Konzept der zweistufigen Bearbeitung bringt jedoch eine sehr flexible Möglichkeit zur Modularisierung von Softwareprojekten mit sich. In der ersten Stufe übersetzt der Compiler ein Quelltextmodul zu einer Objektdatei. In der zweiten Stufe bindet der Linker verschiedene Objektdateien – und damit die Funktionen der verschiedenen Module – zur lauffähigen Software zusammen. C++ wird ebenso durch einen Compiler zu Objektdateien übersetzt, die später durch einen Linker gebunden werden. Dadurch lässt C++ große Freiheiten bei der Aufteilung eines Softwareprojektes in Module. Technisch ist die Flexibilität der Modularisierung eines Softwareprojektes sehr hoch. Es bleibt dem Entwickler überlassen, wie er mit dieser Flexibilität umgeht. Im Abschnitt 6.1 auf Seite 347 wird auf die Modularisierung näher eingegangen. Zunächst sollen jedoch einfache Beispiele gegeben werden, bei welchen die Modulaufteilung noch keine Rolle spielt.

Ein einfaches erstes Beispiel zum Einsatz des Compilers

Fangen wir also mit einem ganz einfachen, für einen Anfang mit einer Programmiersprache sehr typischen „Hello world“ an:

Listing 2.1. Erstes Beispiel

```
#include <iostream>

int main()
{
  std::cout << "Hello world" << std::endl;
  return 0;
}
```

Wenn man mit einem beliebigen Texteditor diesen Text erfasst und unter einem beliebigen Namen mit der Endung `.cc` oder `.cpp` abspeichert, hat man schon einen echten C++-Quelltext. Schreiben Sie zum Beispiel „`prg1.cpp`“. Der Quelltext stellt ein einfaches Programm dar, das die Ausgabe „Hello world“ auf die Standardausgabe schreiben soll. Das Beispiel kann also unter einer textorientierten Shell zum Laufen gebracht werden. Sie brauchen eine beliebige Textshell unter UNIX, Linux oder MacOS. Wenn Sie mit Win-

dows arbeiten, nehmen Sie einfach die DOS-Eingabeaufforderung. Natürlich muss ein C++-Compiler auf dem System installiert sein, damit das Programm übersetzt werden kann. Wichtig ist auch, dass der Compiler aus der Kommandoumgebung, in der Sie arbeiten, erreichbar ist. Auch bei einem installierten Compilersystem müssen nicht immer alle wichtigen Pfade systemweit gesetzt sein. Manche Hersteller legen dazu eine kleine Batch-Datei ihrem Compiler bei, damit die Systempfade damit gesetzt werden können. Dazu muss die Batch natürlich innerhalb der Kommandoumgebung aufgerufen werden. Für manche Compiler muss man eine solche Batch-Datei auch selbst schreiben. Da es von Hersteller zu Hersteller sehr unterschiedlich sein kann, was vorhanden ist und was nicht, und da es darüber hinaus noch auf jedem Betriebssystem andere Verfahren der Batcherstellung gibt, kann dieses Installationsproblem an dieser Stelle nicht hinreichend behandelt werden. Dazu nur drei Regeln (wobei `compiler_root` für das Verzeichnis steht, in dem Ihr Entwicklungssystem installiert wurde):

1. Der Compiler und der Linker sind ausführbare Tools, die im Suchpfad der Umgebung gefunden werden müssen. Normalerweise `compiler_root/bin`.
2. Der Compiler muss das Verzeichnis mit den Includedateien finden. Normalerweise `compiler_root/include`.
3. Der Linker muss das Verzeichnis mit den Bibliotheksdateien finden. Normalerweise `compiler_root/lib`.

Manchmal gehen Compilerhersteller etwas unterschiedliche Wege, um ihren Tools die typischen Orte ihres Arbeitsmaterials mitzuteilen. Das können Umgebungsvariablen sein oder auch Konfigurationsdateien im Compilerverzeichnis. Wenn Sie noch nie mit einem C- oder C++-Compiler zu tun hatten, sollten Sie die erste Installation am besten von jemandem durchführen lassen, der es mindestens schon einmal gemacht hat. Gehen wir also von einem installierten und in Ihrer Kommandoumgebung bekannten Entwicklungssystem aus. Im Allgemeinen ruft man den Compiler auf und übergibt ihm den Quelltextnamen in der Kommandozeile. Also:

```
g++ prg1.cpp
```

für den GNU-Compiler unter Linux oder UNIX.

```
bcc32 prg1.cpp
```

für den Borland-C++-Compiler unter Windows.

```
cl /GX prg1.cpp
```

für den MS-Visual C++-Compiler unter Windows.

Auf manchen Systemen wie zum Beispiel QNX heißt der Aufruf des C++-Compilers `CC`. Also: `CC prg1.cpp`.

Dabei sollte nun auf den Unices eine Datei `a.out` entstanden sein. Bei den Windows-Compilern müsste eine Exedatei `prg1.exe` entstanden sein. Diese Dateien sind ausführbar und können direkt von der Kommandozeile aus gestartet werden.

Also:

```
./a.out
```

für Unix und Linux bzw.

```
prg1
```

für Windows.

Wir haben den Aufruf des Compilers kennen gelernt, der eine ausführbare Datei produziert. Ganz ist dieser Satz nicht richtig, denn irgendwie muss ja noch der in den vorangegangenen Abschnitten beschriebene Linker mit von der Partie sein. In unserem ersten Beispiel haben wir den Linker mit dem Compiler aufgerufen. Die genannten Compilerimplementierungen vereinfachen den Einstieg dadurch, dass bei einem direkten Aufruf des Compilers nach dessen Durchlauf noch der Linker gestartet wird, um die Arbeit zu Ende zu bringen. Wenn man die beiden Arbeitsschritte voneinander entkoppeln möchte – oder muss –, kann man das dem Compiler mit einem Kommandozeilenargument mitteilen. Die entsprechenden Outputs des reinen Compileprozesses finden sich meistens auch bei einem integrierten Compiler-Linker-Lauf im aktuellen Verzeichnis. Sie tragen die Endung `.obj` bei den meisten Windows-Compilern und `.o` bei den Compilern der Unices. Auch diese Zwischenprodukte werden für den Ablauf des Programms nicht mehr gebraucht.

Um dieses erste Programmbeispiel nicht ganz unerklärt zu lassen, sollen die einzelnen Zeilen noch kurz besprochen und entsprechende Verweise auf spätere Kapitel eingefügt werden. Mit der Zeile „`#include <iostream>`“ haben wir eine typische Präprozessoranweisung, wie sie im Abschnitt 2.2.2 erklärt ist. Die Includeanweisung fügt die Datei „`iostream`“ in den von uns geschriebenen Programmquelltext ein. Wo er die Datei `iostream` findet? Sie befindet sich im Allgemeinen im Includeverzeichnis des Compilers. Was es mit dieser Datei auf sich hat, wird im Abschnitt 5.2.1 auf Seite 280 beschrieben. An dieser Stelle sei nur gesagt, dass darin die Definitionen für `std::cout` und `std::endl` zu finden sind. Die Zeile „`int main()`“ und der dazugehörige Block, der durch die geschweiften Klammern begrenzt wird, definiert die Hauptfunktion des Programms. Die Funktion heißt `main`, das Wörtchen `int` zeigt an, dass die Hauptfunktion einen ganzzahligen Wert zurück gibt (in diesem Fall an das aufrufende System), und die einfachen Klammern zeigen an, dass es sich bei `main` überhaupt um eine Funktion handelt. Dazu mehr im Abschnitt 2.2.16 auf Seite 49. Jedenfalls springt die Programmausführung vom Betriebssystem aus zuallererst in die Funktion `main()` und führt die darin enthaltenen Anweisungen aus. Im Beispielprogramm sind das nur zwei. Die Ausgabe auf der Konsole „`std::cout << "Hello world" << std::endl;`“ und die Returnanweisung der Funktion „`return 0;`“, die vereinbarungsgemäß etwas – hier die Null – zurückliefert. Vereinbarungsgemäß deshalb, weil die Funktion `int main()` ja

einen ganzzahligen Rückgabewert in ihrer Deklaration verspricht. Momentan sehen wir von dieser Werterückgabe an das Betriebssystem nichts.

Da jetzt die grundsätzliche Funktionsweise des Compiler-Linker-Gespanns geklärt ist, soll nun der Compilerlauf in seinen Einzelheiten genauer betrachtet werden. Das nächste Kapitel beschreibt die sprachlichen Möglichkeiten, den Quelltext selbst mit Hilfe des Compilers zu manipulieren.

2.2.1 Bezeichner in C++

Es gibt in C++ Schlüsselworte und Operatoren, die den Sprachumfang darstellen. Alles andere in einem Programm muss man selbst definieren und benennen. Auch wenn noch nicht eingeführt wurde, was die Dinge für eine Bedeutung haben, so sollen sie hier doch schon einmal aufgezählt werden: Bezeichnen muss man Konstanten, Variablen, Funktionen, Namensräume, Strukturen, Enums, Unions, Klassen und Klassenmethoden. Welche Namen dafür gewählt werden, ist weitgehend dem Programmierer überlassen. Natürlich sollte er solche Namen finden, die im jeweiligen Zusammenhang einen „Sinn" ergeben. Ganz frei ist er aber beim Erfinden der Bezeichner nicht. Es gibt ein paar Regeln, die für alle Bezeichner gelten, die in einem Programm eingeführt werden können:

1. Bei Bezeichnern in C++ werden Groß- und Kleinschreibung unterschieden. Das heißt, sie sind case-sensitiv. Die beiden Bezeichner `Prozentwert` und `prozentwert` bezeichnen also unterschiedliche Dinge.
2. Bezeichner bestehen aus Buchstaben und können auch Ziffern enthalten.
3. Beginnen muss ein Bezeichner immer mit einem Buchstaben des lateinischen ASCII-Zeichensatzes oder mit dem Unterstrich _. Der ASCII-Zeichensatz umfasst das Alphabet ohne nationale Sonderzeichen. Also `a`, `b`, `c`, `d`, `e`, `f`, `g`, `h`, `i`, `j`, `k`, `l`, `m`, `n`, `o`, `p`, `q`, `r`, `s`, `t`, `u`, `v`, `w`, `x`, `y`, `z` und deren groß geschriebene Varianten `A` – `Z`. Die deutschen Sonderzeichen `ä`, `ö`, `ü` und `ß` sind beispielsweise nicht erlaubt. Das gleiche gilt für die dänischen, die isländischen usw.
4. Erst nach dem ersten Zeichen dürfen auch Ziffern verwendet werden. Also `0`, `1`, `2`, `3`, `4`, `5`, `6`, `7`, `8` und `9`. Der Bezeichner `a42` ist ein gültiger Bezeichner. Dagegen ist `42a` nicht gültig.
5. Alle Zeichen in einem Bezeichner sind signifikant. Sie werden also zur Erkennung und Unterscheidung herangezogen. Es gibt keine Größenbegrenzung[6].

[6] Das war in älteren C++-Dialekten einmal anders. Dort galten 32 oder 256 Stellen im Bezeichner als signifikant. Alles was darüber hinausging wurde einfach abgeschnitten. Auch heute sind noch solche Compiler im breiten Einsatz, die nach diesem alten Verfahren arbeiten. Natürlich gelten auch bei Compilern, die dem ANSI/ISO-Standard entsprechen, technische Grenzen. Ein Bezeichner sollte also nicht den ganzen Hauptspeicher füllen.

6. Schlüsselworte dürfen nicht als Bezeichner verwendet werden.
7. Es gibt noch einige andere durch den Compiler vordefinierte Bezeichner, die nicht verwendet werden dürfen. So sind beispielsweise einige Konstanten oder Makros durch den Compiler vordefiniert.

Bei der Wahl von geeigneten Bezeichnern müssen die vorangegangenen Regeln beachtet werden. Die Regeln sind allerdings so geartet, dass die Einhaltung einigermaßen intuitiv erfolgen kann. Es kommt selten vor, dass man mit ihnen in Konflikt gerät.

2.2.2 Der Präprozessor

Der Präprozessor ist eine Phase des Compilers, die Textersetzungen am Quelltext durchführt. Diese erste Phase hat mit dem eigentlichen Parsen des C++-Codes noch nichts zu tun. Sie ist dem eigentlichen Compilevorgang zeitlich vorgeschaltet und man kann mit ihr den Quelltext für den eigentlichen Compiliervorgang vorbereiten. So kommt es beispielsweise oft vor, dass man Code portabel auf mehreren Plattformen lauffähig hält, indem man Teile, die nur für eine bestimmte Plattform Gültigkeit besitzen, durch Präprozessoranweisungen für eben diese Plattform inkludiert. Andere Teile können durch den Präprozessor herausgeschnitten werden. Es gibt viele Gründe dafür, warum man Textmanipulationen am Quelltext durch den Compiler durchführen lässt.

Nur die Präprozessorphase kann man für solche Textoperationen nutzen. Man programmiert sie mit speziellen Anweisungen, den sogenannten Präprozessordirektiven[7]. Die Präprozessorphase gibt es auch in der Sprache C und sie wurde von dort nach C++ übernommen (die Präprozessorbefehle unterscheiden sich auch zwischen den zwei Sprachen kaum). Man erkennt die Präprozessordirektiven an der Raute # vor dem Schlüsselwort. Es gibt nicht viele verschiedene Präprozessordirektiven. Sie lassen sich in der folgenden Liste zusammenfassen:

```
#define
#undef
#if
#ifdef
#ifndef
#elif
#else
#endif
```

[7] In den meisten Fällen ist der Präprozessor ein eigenes Tool, das durch den Compiler selbst aufgerufen wird. Daher kann man einen solchen Präprozessorlauf auch ohne einen ganzen Compilerlauf anstoßen, um die Ergebnisse bestimmter Präprozessordirektiven zu überprüfen.

```
#include
#line
#error
#pragma
```

Dazu kommen noch die beiden Operatoren # und ##, die für die Bildung von symbolischen Namen gebraucht werden.

Die Präprozessordirektive #define

Eine sehr einfache Anwendung des Präprozessors findet sich in der Definition von Konstanten durch die #define-Anweisung.

```
#define MAXIMUM    1000000
```

Diese Zeile stellt für den Präprozessor die Anweisung dar, alle Zeichenketten im nachfolgenden Text, die „MAXIMUM“ lauten, durch die Zeichenkette „1000000“ zu ersetzen. Damit wird eine symbolische Konstante geschaffen, die an beliebig vielen Stellen des Quelltextes verwendet werden kann. Vor dem eigentlichen Parsevorgang des Compilers ersetzt der Präprozessor die symbolische Konstante MAXIMUM durch die numerische 1000000. Die #define-Anweisung kann noch wesentlich mehr, als nur Konstanten zu definieren. Mit ihr lassen sich Makros definieren, die im Abschnitt 2.2.19 auf Seite 58 erklärt werden.

Die Präprozessordirektive #include ..

Ebenso leicht ist die #include-Anweisung zu verstehen. In fast allen C- und C++-Programmen findet man am Anfang der Quellcodedateien diese Art der Präprozessordirektiven:

```
#include <iostream>
```

Mit dieser Anweisung fügt der Präprozessor die Textdatei, die in den spitzen Klammern genannt ist, in den aktuellen Code ein. Im Folgenden wird diese Textdatei „Includedatei“ oder „Headerdatei“ genannt. Es wird dabei einfach der Text der Quellcodedatei um den Text der Includedatei[8] an der aktuellen Stelle erweitert. Es ist wirklich nichts Weiteres. Dem Compiler ist es an der Stelle der Includeanweisung vollkommen egal, was in der einzufügenden Includedatei steht. Die Wirkungsweise ist einfach auf die Einfügung des Textes beschränkt, der in der aufgeführten Datei enthalten ist. Spätere Compilephasen nach dem Präprozessorlauf überprüfen das Eingefügte dann syntaktisch. Die spitzen Klammern, die den Dateinamen umschließen, haben die Wirkung, dass der Compiler in einem vorher eingestellten Pfad nach der

[8] Das ist im Prinzip eine normale Quellcodedatei. Allerdings enthält sie für gewöhnlich nur bestimmte Definitionen. In folgenden Abschnitten wird näher darauf eingegangen werden.

angegebenen Datei sucht. In diesem Pfad stehen für gewöhnlich die Bibliotheksheaderdateien. Beim Einsatz mehrerer Bibliotheken kann man den Pfad erweitern. Das geschieht in Abhängigkeit vom eingesetzten Entwicklungssystem in Konfigurationsdateien oder in Umgebungsvariablen. Dem Compiler selbst kann man auch einen Parameter übergeben, der den genannten Pfad setzt. Dieser Kommandozeilenparameter beginnt für gewöhnlich mit `-I`. Möchte man eine Headerdatei vom gleichen Verzeichnis einfügen, aus dem auch die Quellcodedatei stammt, in der die Includeanweisung steht, so verwendet man Anführungszeichen.

```
#include "Definitionen.h"
```

Von diesem Wurzelverzeichnis des aktuellen Compiliervorgangs kann man auch relative Pfadangaben machen.

```
#include "../../mylibs/include/Definitionen.h"
```

Die Headerdateien haben für gewöhnlich die Endung `.h`, `.hpp` oder `.hxx`. Die Endung `.h` war schon in der Sprache C üblich und wurde für C++ beibehalten. Manchmal möchte man die Dateien als C++-Quelltexte kenntlich machen und verwendet dabei eine der beiden anderen Varianten. Notwendig ist das nicht, dem Compiler ist es egal.

In ANSI-C haben auch alle Bibliotheksheaderdateien die Endung `.h`. In den alten AT&T-Standards für C++ wurden die Dateiendungen beibehalten. Erst mit der ANSI/ISO-Standardisierung wurden die Dateiendungen für Bibliotheksheader gestrichen. Das hat den Grund, dass man C++ für so viele Systeme einsetzbar gestalten wollte wie möglich. Dabei kann man sich nicht immer auf das Vorhandensein von Dateisystemen verlassen, die Dateierweiterungen kennen.

Die Präprozessordirektiven `#if`, `#else`, `#elif` und `#endif`

Die Anweisungen `#if`, `#else`, `#elif` und `#endif` werden zur so genannten bedingten Compilierung verwendet. Mit ihnen findet ein „Zurechtschneiden" des Quelltextes im Präprozessorlauf statt. Bevor der Quelltext der eigentlichen syntaktischen Überprüfung unterzogen wird, werden Teile aktiviert bzw. deaktiviert, indem sie eingefügt oder ausgeschnitten werden. Dazu werden mit Hilfe dieser Präprozessordirektiven Konstanten abgefragt, die zur Compilezeit schon feststehen.

Listing 2.2. Verzweigungen im Präprozessorlauf

```
#if   DEBUG == 1
  void TraceOut( char *txt )
  {
    std::cerr << txt << std::endl;
  }
```

```
#elif DEBUG == 2
  void TraceOut( char *txt )
  {
    whiteToFile( "TRACES.TXT"', txt );
  }
#else
  inline void TraceOut( char * )
  {
  }
#endif
```

In dem vorangegangenen Beispiel werden drei Quellcodeabschnitte zur Auswahl gegeben. Je nach dem Wert der Konstanten `DEBUG` wird einer der drei Abschnitte gewählt und übersetzt. Die anderen Abschnitte sondert der Präprozessor schon vor dem eigentlichen Übersetzungsvorgang aus.

Die Anwendung der Präprozessoranweisung `#if` zieht nicht zwingend `#else` oder `#elif` nach sich. Beide sind optional.

Die Präprozessordirektiven `#ifdef` und `#ifndef`

Die beiden Anweisungen `#ifdef` und `#ifndef` sind erweiterte Versionen der Anweisung `#if`. Sie fragen ab, ob ein Bezeichner durch den Präprozessor definiert wurde oder nicht. Man kann sie auch mit der Grundform umschreiben.

Listing 2.3. Verzweigungen im Präprozessorlauf anhand einer Definition

```
#ifdef BIGENDIAN
 #define HIGHBYTE RIGHTBYTE
 #define LOWBYTE  LEFTBYTE
#else // LITTLEENDIAN
 #define HIGHBYTE LEFTBYTE
 #define LOWBYTE  RIGHTBYTE
#endif
```

In dem Beispiel in Listing 2.3 werden zwei Definitionen in Abhängigkeit einer schon existierenden Definition unterschiedlich definiert. Solch ein Beispiel ist durchaus realistisch. Es bleibt aber noch ganz im Bereich der Präprozessordirektiven. Mit der Anweisung `#if` kann man das Beispiel folgendermaßen umformulieren:

Listing 2.4. Die Anwendung des Präprozessoroperators `defined`

```
#if defined(BIGENDIAN)
// ...
```

Das nächste Beispiel präpariert ein Stückchen Code unterschiedlich, je nachdem, ob es von einem C- oder einem C++-Compiler verarbeitet wird.

Listing 2.5. Beispiel: bedingte Compilierung

```
#ifdef __cplusplus
extern "C" {
#endif

int f()
{
  // ...
}

#ifdef __cplusplus
}
#endif
```

Auch dieses ist ein sehr realistisches Beispiel. Bei C++-Compilern ist die symbolische Konstante `__cplusplus` vordefiniert (siehe Abschnitt 2.2.2 auf Seite 20). Das führt in dem Beispiel dazu, dass für einen C++-Compilevorgang zwei kleine Codeschnipselchen mehr compiliert werden als für C[9].

Etwas anders ausgedrückt, kann man den Effekt auch so erreichen:

Listing 2.6. Beispiel: bedingte Compilierung

```
#ifndef EXTC
 #ifdef __cplusplus
  #define EXTC extern "C"
 #else
  #define EXTC
 #endif
#endif // EXTC

EXTC int f()
{
  // ...
}
```

Die Anweisung `#ifndef` negiert die Bedeutung von `#ifdef` einfach. Das `n` steht für „not“. Das Beispiel in Listing 2.6 kann auch mit `#if` reformuliert werden:

Listing 2.7. `#ifndef` anders ausgedrückt

```
#if !defined(EXTC)
 #if defined(__cplusplus)
  #define EXTC extern "C"
```

[9] Die hier gezeigten bedingten Codefragmente `extern „C“` mit den dazugehörigen geschweiften Klammern werden für die Integration von C- und C++-Code gebraucht.

```
 #else
  #define EXTC
 #endif
#endif // EXTC
...
```

Ein weiteres Beispiel demonstriert, wie man zur Compilezeit feststellen kann, ob für ein 16-Bit- oder für ein 32-Bit-System compiliert wird, und legt dann die maximale Größe eines Puffers fest:

Listing 2.8. Detektion zur Compilezeit

```
#if sizeof(void*)==2 // 16 Bit
 #define MAXBUFFERSIZE 32000
#else
 #define MAXBUFFERSIZE 10000000L
#endif
```

Dazu wird der Operator `sizeof` verwendet. Dieser Operator wird im Abschnitt 2.2.4 auf Seite 24 erklärt und liefert immer ein konstantes Ergebnis.

Die Direktive `#undef` macht eine Definition rückgängig. Man hebt mit der Anweisung `#undef` Dinge auf, die mit `#define` definiert wurden. Das sind symbolische Konstanten und Makros. Anzuwenden ist `#undef` wie eine `#define`-Anweisung.

```
   #undef Bezeichner
```

Die Präprozessordirektive `#error`

Die Direktive `#error` führt zu einem sofortigen Abbruch des Compilevorgangs und gibt eine Fehlermeldung aus.

Listing 2.9. Abbruch des Compilerlaufs bei Fehlerfall

```
#if   sizeof(void*)==2 // 16 Bit
 #define MAXBUFFERSIZE 32000
#elif sizeof(void*)==4 // 32 Bit
 #define MAXBUFFERSIZE 10000000L
#else
 #error Kein 16 oder 32-Bit System!
#endif
```

Diese `#error`-Direktive wird dazu verwendet, den Compilevorgang abzubrechen, wenn nicht die benötigten Bedingungen für eine erfolgreiche Übersetzung detektiert werden. Der Text in der Argumentenzeile der Anweisung wird dabei durch den Compiler auf der Standardausgabe ausgegeben.

Die Präprozessoranweisung #pragma

Die Anweisung #pragma ist zwar selbst eine Standarddirektive, ihre Funktion ist es jedoch, nichtstandardisierte Einstellungen spezifischer Compiler vorzunehmen. Die Argumente der Direktive sind also jeweils herstellerspezifisch. Wenn ein Compiler die Argumente einer #pragma-Direktive nicht erkennt, muss er sie ignorieren. Das ist eine Standardvorgabe. Welche Einstellungen durch Pragmas beeinflusst werden können, ist absolut abhängig von den technischen Notwendigkeiten eines Zielcompilers oder vom Erfindungsreichtum des Compilerherstellers. So erzwingt die folgende Pragmaanweisung beim MS-Visual C++ 6.0-Compiler, dass die Warnung mit der Nummer 164 als Fehler behandelt wird.

```
#pragma warning( error : 164 )
```

Für den gleichen Compiler hat die folgende Pragmaanweisung die Bedeutung, dass das Funktioninlining auf eine bestimmte Aufruftiefe begrenzt wird. Ruft also eine Inline-Funktion eine andere Inline-Funktion, und so weiter, werden diese Funktionen nur bis zur der Aufrufebene inline expandiert, die in der Pragmaanweisung genannt wurde.

```
#pragma inline_depth( [0... 255] )
```

Diese Beispiele sollen zeigen, wie compilerabhängig Pragmaanweisungen sind. Will man sie für einen bestimmten Compiler verstehen, muss man seine Dokumentation zu Rate ziehen. Jeder Compiler hat einen großen Satz von möglichen Einstellungen und oft auch bestimmte Funktionen, die nur bei diesem anzutreffen sind.

So erzeugt beispielsweise die Quelltextzeile

```
#pragma keeka
```

beim Borland C++-Compiler Version 5.5.1 für Windows die folgende Ausgabe auf der Konsole:

```
Borland C++ 5.5.1 for Win32 Copyright (c) 1993, 2000 Borland
keeka.cpp:
 /\__/\
| -  - |
* ^--^ *
 x=--=x
  |    |
|       |
| |  | |     .
| |__| |\___//
 vv  vv *---*

Keeka: Simply the best darn cat in the Universe.
```

Andere Compiler ignorieren dieses Pragma völlig.

Vordefinierte Makros und symbolische Konstanten

Bestimmte Makros und Konstanten sind durch den Compiler bereits vordefiniert. Sie können bei der Definition eigener Makros behilflich sein, ermöglichen die programmtechnische Aufnahme bestimmter Daten, die mit dem Compilevorgang in Verbindung stehen und unterstützen auch die Implementierung von Fehleranalyse- und Abbruchcodes.

Die standardmäßig vordefinierten Makros und Konstanten sind:

`__LINE__`	Zeilennummer als Dezimalkonstante.
`__FILE__`	Dateiname als Stringliteral.
`__DATE__`	Datum als Stringliteral im Format `mm dd yyyy`.
`__TIME__`	Zeit als Stringliteral im Format `hh:mm:ss`.
`__cplusplus`	In ANSI/ISO C++ von 1998 mit 199711L definiert. Wird in zukünftigen Standards wahrscheinlich entsprechend höhere Werte annehmen.

Neben den standardisierten Makros und Konstanten gibt es eine Unmenge von Vordefinitionen, die herstellerspezifisch sind. Jeder Compiler bringt einen Satz von Makros und Konstanten mit, der auf sein bestimmtes Einsatzgebiet und auf seine spezielle Technik abgestimmt ist. Diese Vordefinitionen sind nur in sehr seltenen Fällen kompatibel.

Die Präprozessoroperatoren # und

Die beiden Operatoren werden gebraucht um symbolische Konstanten oder Makronamen durch den Präprozessor generieren zu lassen. Dabei wandelt # eine Zahl in einen Text und ## fügt zwei Texte aneinander.

Mit dem Makro `#define STR(n) #n` lassen sich durch den Präprozessor Literalkonstanten, die Zahlen repräsentieren, in einen Text konvertieren. Der Ausdruck `STR(42)` wird durch den Compiler in den Text „42“ übersetzt.

Listing 2.10. Die Anwendung des Präprozessoroperators ##

```
#include <iostream>

#define SOURCE "Datei: "##__FILE__
#define TRACE( msg ) std::cerr << \
(SOURCE##" => "##msg)<<std::endl;

int main()
{
  TRACE( "Eine Debugmeldung" );

  return 0;
}
```

2.2.3 Variablen

Alle Daten, die in einem C++-Programm verwaltet werden und damit im Speicher gehalten werden, müssen in C++ typisiert sein. Das heißt, dass man sich vorher über die Art der Daten im Klaren sein muss, bevor man mit ihnen umgeht. Für die Daten werden dann Platzhalter definiert. Im einfachsten Fall – die komplizierteren folgen in den Absätzen weiter hinten in diesem Buch – sind es Variablen, die Daten aufnehmen können.

Sehen wir uns also hier die Variablendeklaration in C++ an. Es können in C++ nur Variablen verwendet werden, die vorher deklariert wurden. Variablen als Platzhalter für irgendwelche Daten, die erst zur Laufzeit feststehen, müssen dem Typ ihrer Daten entsprechen. Es gibt keine Variablen, die wie in BASIC verschiedene Datentypen, also zum Beispiel ganzzahlige Werte oder Texte, aufnehmen können. Die Datentypen, die solchen Variablen zugrunde liegen, nennt man variante Typen[10]. Eine Variable in C++ wird also zuerst mit deklariert – mit dem nötigen Datentyp – und dann benutzt:

```
int i;
```

Die Variable `i` hat den Datentyp `int`, der ganzzahlige Werte in einem bestimmten Wertebereich erlaubt. Der Bezeichner `i` erlaubt nun mit der Variablen umzugehen. Man kann dieser Variablen zum Beispiel einen Wert zuweisen:

```
i = 7;
```

Der Begriff „Variable“ steht wie in der Algebra dafür, dass der Platzhalter mit einer bestimmten Bezeichnung für verschiedene Werte steht. Allerdings gibt es einen großen Unterschied zur Algebra. In einem C++-Programm kann eine Variable zu einem Zeitpunkt immer nur genau einen Wert aufnehmen. In der Algebra kann eine Variable durchaus für mehrere mögliche Werte stehen. Die Mathematik steht außerhalb von Zeit und Raum. In einem Programm sieht es anders aus. Ein Programm läuft in der Zeit ab und ändert dabei seinen internen Zustand. Es werden also die Variablen während der Ablaufzeit des Programms geändert. Zu einem bestimmten Zeitpunkt hat eine Variable also immer genau einen bestimmten Wert.

Technisch gesehen ist eine Variable ein Speicherplatz, der Werte eines definierten Typs aufnehmen kann. Der Typ bestimmt die Größe des Speicherplatzes und auch die Weise, wie der Speicherplatz interpretiert wird. Letzteres muss man so verstehen, dass der Speicher, der ja nur eine Aneinanderreihung von Bytes darstellt, alle beliebigen Werte aufnehmen kann, die man in Bytes speichert. In jede Zelle (Byte) passen ganzzahlige Werte von 0 bis 255. Wenn

[10] So etwas kann zwar auch mit C++ erreicht werden. Dazu muss allerdings erst das Klassenkonzept verstanden werden. Deshalb stelle ich mich zunächst auf den Standpunkt, dass es variante Datentypen in C++ nicht gibt.

beispielsweise Text oder Fließkommazahlen abgespeichert werden sollen, so muss es ein Verfahren geben, die zu speichernde Information in eben diesen Werten abzulegen, die man in Bytes speichern kann. Hinter einem Fließkommatyp, wie zum Beispiel `double`, steht also die Speichergröße, die ein solcher Zahlentyp im Speicher braucht (acht Bytes), und das Verfahren, wie eine solche Kommazahl in dem Speicher untergebracht wird. Eine Zuweisung, wie sie an der Variablen `i` demonstriert wurde, ordnet die Daten in einer dem Typ inneliegenden Weise und legt sie im Speicher ab.

Im Gegensatz zu C können in C++ fast an jeder Stelle Variablen deklariert werden[11]. Wenn man mehrere Variablen des gleichen Typs braucht, kann man die Variablenbezeichner hinter dem Typ durch Komma trennen.

```
int a, b, c;
```

Diese Deklaration legt drei Variablen des Typs `int` an. Außerdem können Variablen an der Stelle der Deklaration gleich mit einem Wert initialisiert werden.

```
int a = 1, b = 2, c = 3;
```

Die Variablen, die innerhalb von Funktionen deklariert werden, liegen auf dem Stack. Das ist ein Speicherbereich, der unter anderem für diese lokalen Daten zur Verfügung steht. Der Stack und die möglichen Orte der Variablendeklaration sind in Abschnitt 2.2.23 auf Seite 94 beschrieben.

2.2.4 Standarddatentypen

C++ ist eine typisierende Programmiersprache. Das heißt, dass alle Daten, die in einem C++-Programm verwaltet werden, einen Datentyp besitzen. Ein Datentyp beschreibt, um welche Art von Daten es sich handelt. Diese Aussage geht durchaus über das hinaus, was in diesem Abschnitt besprochen wird. Hier werden die Grundbausteine der Datentypen besprochen: die einfachen numerischen Typen, die man auch als *Standarddatentypen* bezeichnet. In späteren Abschnitten und Kapiteln werden noch ganz andere Aspekte des Datentyps in der Objektorientierten Programmierung beschrieben.

Zunächst aber zu den Datentypen, die durch die Sprache C++ fertig mitgeliefert werden: die Standarddatentypen. Es gibt vier vorzeichenbehaftete und vier vorzeichenlose ganzzahlige Typen. Die Grundtypen `int`, `short`, `long` sind vorzeichenbehaftet. Man kann das Vorzeichen auch dadurch zum Ausdruck bringen, dass man ihnen das Schlüsselwort `signed` voranstellt. Der vierte vorzeichenbehaftete ganzzahlige Typ ist `signed char`. Von diesen Typen lassen sich die vorzeichenlosen Varianten bilden, indem man ihnen das

[11] In C müssen sie immer am Anfang eines Blocks vor den Anweisungen stehen.