

C

Een inleiding in de programmeertaal C

Jesse op den Brouw

Eerste druk

©2020 Jesse op den Brouw, Den Haag

Versie: 0.1

Datum: 14 april 2020

DE HAAGSE
HOGESCHOOL

De auteur kan niet aansprakelijk worden gesteld voor enige schade, in welke vorm dan ook, die voortvloeit uit informatie in dit boek. Evenmin kan de auteur aansprakelijk worden gesteld voor enige schade die voortvloeit uit het gebruik, het onvermogen tot gebruik of de resultaten van het gebruik van informatie in dit boek.

Dit boek mag overal voor gebruikt worden, commercieel en niet-commercieel, mits de wijzigingen worden gedeeld en naam van de originele auteur wordt gemeld. De broncode van dit boek is beschikbaar op https://github.com/jesseopdenbrouw/book_c.



C van Jesse op den Brouw is in licentie gegeven volgens een [Creative Commons Naamsvermelding-NietCommercieel-GelijkDelen 3.0 Nederland-licentie](#).

Suggesties en/of opmerkingen over dit boek kunnen worden gestuurd naar:
J.E.J.opdenBrouw@hhs.nl.

Voorwoord

Een van de beste boeken over C is *The C Programming Language* van Brian Kernighan en Dennis Ritchie [1]. Het boek is kort en bondig maar laat toch ruimte om de tekst te ondersteunen met voorbeelden. Helaas zijn veel van de voorbeelden lastig te volgen voor de beginnende programmeur. Zoals de schrijvers zelf opmerken:

The book is not an introductory programming manual; it assumes some familiarity with basic programming concepts like variables, assignment statements, loops, and functions. [...] C is not a big language, and it is not well served by a big book.

Dit boek is bedoeld als inleiding op de programmeertaal C en is niet toereikend om alle facetten van de taal te leren. Dat is ook niet de bedoeling van dit boek. We laten alleen maar zien wat gebruikelijk is bij het ontwikkelen van C-programma's.

Veel boeken over C beschrijven het ontwikkelen van programma's op het Unix operating system en afgeleide varianten zoals Linux, FreeBSD en Mac OS-X. Dat is echter niet de werkwijze van veel programmeurs. Veel software wordt ontwikkeld met een Integrated Development Environment (IDE). Bekende IDE's zijn Visual Studio, Code:Blocks en Xcode. Het is natuurlijk niet mogelijk om alle mogelijkheden van zulke IDE's te beschrijven (of af te beelden).

Dit boek is opgemaakt in L^AT_EX [2] (L^AT_EX-engine = pdfL^AT_EX 1.40.20). L^AT_EX leent zich uitstekend voor het opmaken van lopende tekst, tabellen, figuren, programmacode, vergelijkingen en referenties. De gebruikte L^AT_EX-distributie is TexLive uit 2019 [3]. Als editor is TexStudio [4] gebruikt. Tekst is gezet in Charter [5], een van de standaard fonts in L^AT_EX. De keuze hiervoor is dat het een prettig te lezen lettertype is, een *slanted* letterserie heeft en een bijbehorende wiskundige tekenset heeft. Code is opgemaakt in Nimbus Mono [6] met behulp van de *listings*-package [7]. Voor het tekenen van array's en pointers is *TikZ/PGF* gebruikt [8]. Alle figuren zijn door de auteur zelf ontwikkeld, behalve de logo's van Creative Commons en De Haagse Hogeschool.

Natuurlijk zullen docenten ook nu opmerken dat dit boek niet aan al hun verwachtingen voldoet. Dat zal altijd wel zo blijven. Daarom wordt de broncode van vrij beschikbaar. Eenieder die dit wil, kan de inhoud vrijelijk aanpassen, mits wijzigingen gedeeld worden. Hiermee wordt dit boek een levend document dat nooit af is. De broncode van dit boek is beschikbaar op https://github.com/jesseopdenbrouw/book_c.

Leeswijzer

Hoofdstuk 1

Studiewijzer

In onderstaande tabel wordt een overzicht gegeven van de stof die een student bij een eerste introductie onderwezen zou moeten krijgen. Uiteraard is iedereen vrij om zelf de onderwerpen te kiezen.

Tekst

Verantwoording inhoud

Dit boek voldoet alleen aan aandachtspunt 4,01 van de basis Basic of Knowledge and Skills (BoKS) Elektrotechniek. De BoKS is te vinden via [9].

Website

Op de website <http://ds.opdenbrouw.nl> zijn slides, practicumopdrachten en aanvullende informatie te vinden. De laatste versie van dit boek wordt hierop gepubliceerd. Er zijn ook voorbeeldprojecten voor Visual Studio en Code::Blocks te vinden.

Dankbetuigingen

Dit boek had niet tot stand kunnen komen zonder hulp van een aantal mensen. Ik wil collega Harry Broeders van Hogeschool Rotterdam bedanken voor zijn geweldige bijdrage. Niet alleen op technisch gebied, maar ook taalkundig en op de indeling van dit boek. Collega Ben Kuiper heeft een waardevolle bijdrage geleverd op technisch en taalkundig gebied.

De ASCII-tabel op pagina 84 is ontwikkeld door Victor Eijkhout. Deze tabel kan gevonden worden op <http://www.ctan.org/tex-archive/info/ascii-chart>.

Jesse op den Brouw, 2020.

Inhoudsopgave

Voorwoord	iii
1 Un tour de C	1
1.1 De computer	2
1.2 De compiler	3
1.3 Bibliotheken en functies	4
1.4 Ontwikkelssystemen	4
1.5 Afdrukken op het scherm	5
1.6 Invoer van het toetsenbord	7
2 Variabelen, datatypes en expressies	9
2.1 Variabelen	9
2.2 Datatypes	10
2.3 Constanten	12
2.4 Expressies	15
2.4.1 Rekenkundige operatoren	15
2.4.2 Relationale operatoren	16
2.4.3 Logische operatoren	16
2.4.4 Bitsgewijze operatoren	17
2.5 Datatypeconversie	19
2.6 Overflow	20
2.7 Vaste-lenge datatypes	20
2.8 Overige operatoren	21
2.8.1 Grootte van een variabele	21
2.8.2 Verhogen of verlagen met 1	22
2.8.3 Toekenningsoperatoren	22
2.8.4 Conditionele expressie	23
2.8.5 Komma-operator	24
2.8.6 Nog enkele operatoren	24
2.9 Voorrangsregels van operatoren	24
3 Flow control	27
4 Functies	29
5 Array's	31

5.1	Declaratie en selectie	31
5.2	Initialisatie	32
5.3	Eendimensionale array	33
5.4	Aantal elementen van een array bepalen	34
5.5	Tweedimensionale array	35
5.6	Afdrukken van array	36
5.7	Array's en functies	37
5.8	Array's vergelijken	39
5.9	Strings	40
5.9.1	Stringfuncties	41
6	Pointers	43
6.1	Pointers naar enkelvoudige datatypes	44
6.2	De NULL-pointer	46
6.3	Pointer naar void	47
6.4	Afdrukken van pointers	47
6.5	Pointers naar array's	48
6.6	Strings	50
6.7	Rekenen met pointers	51
6.8	Relatie tussen pointers en array's	52
6.9	Pointers als functie-argumenten	53
6.10	Pointer als return-waarde	56
6.11	Pointers naar pointers	57
6.12	Array van pointers	60
6.13	Pointers naar een array van pointers	61
6.14	Argumenten meegeven aan een C-programma	62
6.15	Pointers naar functies	64
6.16	Pointers naar vaste adressen	67
6.17	Dynamische geheugenallocatie	67
6.18	Subtiele verschillen en complexe declaraties	69
7	Structures	71
8	Invoer en uitvoer	73
8.1	Inlezen van een getal met scanf	73
8.1.1	Problemen met scanf	73
	Bibliografie	81
A	De ASCII-tabel	83
B	Tutorial Visual Studio	85
C	Voorrangsregels van operatoren	91
	Index	93

1

Un tour de C

C is al een oude taal. De taal is rond 1970 ontworpen door Dennis Ritchie¹ en is dus al zo'n 50 jaar oud. Hij was bezig met het schrijven van een besturingssysteem (Engels: operating system)² en zocht naar een mogelijkheid om dit te schrijven op een hoger niveau dat gebruikelijk voor die tijd. In die tijd werden besturingssystemen geschreven in assembly, een taal die dicht de hardware van de computer ligt. Programma's schrijven in assembly is tijdrovend en gevoelig voor fouten van de programmeur.

C is een zogenoemde derde-generatie-taal (3GL) en zorgt ervoor dat programma's gestructureerd kunnen worden geschreven zonder dat de programmeur kennis hoeft te hebben van de hardware waarop het programma draait. Toch biedt de taal constructies om die hardware in te stellen, een voorwaarde voor het schrijven van een operating system. Daarom is de taal geliefd bij programmeurs van kleine computersystemen waarop geen operating system draait, de zogenoemde *bare metal systems*. Het is vaak ook de enige taal die gebruikt kan worden op dit soort systemen, naast assembly.

C is geschreven voor ervaren programmeurs die behoefte hebben voor het schrijven van compacte programma's, dat te zien is in de vele, soms onduidelijke, taalconstructies. Het is eigenlijk niet geschikt voor beginnende programmeurs. Het is zeker mogelijk om programma's te schrijven die niet correct werken, maar met enige discipline kan ook de minder ervaren programmeur de taal prima gebruiken.

C is een algemeen bruikbare taal (*general purpose language*) en is dus niet voor een specifiek doeleind ontworpen (behalve voor het schrijven van operating systems). Zo kunnen met C wiskundige berekeningen worden uitgevoerd maar de programmeur moet zelf alles schrijven. Er zijn wel *functies* beschikbaar voor bijvoorbeeld sinus, cosinus en tangens. Er

¹ Dennis MacAlister Ritchie (1941 – 2011). Hij was de ontwerper van de programmeertaal C en was een van de ontwerpers van Unix. Bekende afgeleiden van Unix zijn Linux en FreeBSD.

² Een besturingssysteem is een programma dat draait op een computer en zorgt voor het beschikbaar stellen van *resources* aan de gebruiker (lees: programma's). Zo zorgt een besturingssysteem ervoor dat de bestanden op de harde schijf netjes worden opgeslagen en beschikbaar zijn. Daarnaast zorgt een besturingssysteem ervoor dat programma's netjes naast elkaar kunnen draaien. Drie bekende besturingssystemen zijn Windows, Linux en OS-X.

zijn echter talen die dit veel beter ondersteunen. Ook het verwerken van bestanden is mogelijk maar het manipuleren van *strings* (een rij karakters) moet door de programmeur zelf worden uitgewerkt. Ook hier zijn diverse functies beschikbaar die het programmeerwerk enigszins verlichten.

C is een *imperatieve* programmeertaal, een van de bekende programmeerparadigma's³. C moet concurreren met vele andere talen zoals C++, C#, Python en Java. C++ is ontwikkeld door Bjarne Stroustrup als “de betere C” en ondersteunt *objectgeoriënteerd* programmeren. C# is ontwikkeld door MicroSoft en is de *de facto* programmeertaal op Windows-systemen. Python is ontwikkeld door het Centrum voor Wiskunde en Informatica van de Universiteit van Amsterdam.

1.1 De computer

Een *computer* is een elektronisch, digitaal systeem dat *data* verwerkt aan de hand van een lijst *instructies*. Het Engelse woord voor verwerken is “to process”. En daar komt de naam vandaan van een belangrijk onderdeel van de computer: de *processor*.

De instructies die de processor kan uitvoeren zijn zeer eenvoudig. Voorbeelden zijn “tel op” en “trek af” en “bepaal of het ene getal kleiner is dan het andere getal”. Een processor kan niet in één keer “doe dit tien keer en bereken de wortel van diverse getallen en druk af op het scherm” uitvoeren. Als we dat willen, dan moeten we dit opdelen in de eenvoudige instructies die de processor wel kan verwerken. Alle instructies bij elkaar wordt een *programma* genoemd en de verzamelnaam voor alle programma's wordt *software* genoemd.

Een computer heeft naast de processor *geheugen* om de data en instructies op te slaan. Er zijn twee soorten geheugens: ROM (Read Only Memory) is een geheugen waarvan de inhoud niet gewijzigd kan worden; RAM (Random Access Memory) is geheugen waarvan de inhoud wel gewijzigd kan worden. Over het algemeen wordt het programma in de ROM opgeslagen en wordt de data in de RAM opgeslagen.

Naast ROM en RAM heeft de computer nog invoer- en uitvoermogelijkheden, anders is er geen communicatie met de buitenwereld mogelijk. De verzamelnaam is *I/O* dat “input” en “output” betekent. Het is best lastig om de invoer en uitvoer te realiseren. Daarom wordt bij de gangbare computers een stuk software geladen (of het is al aanwezig in de ROM) om dat voor de gebruiker (en programmeur) te vereenvoudigen. Die software wordt een *besturingssysteem* genoemd, maar gangbaar is om de Engelse naam Operating System te gebruiken. Bekende besturingssystemen zijn Windows, Mac OS-X en Linux.

Niet alle gegevens zijn in het geheugen van de computer aanwezig. Een computer heeft (vaak) *secundair geheugen*. Voorbeelden van secundair geheugen zijn harddisk (harde schijf), USB-sticks en SD-cards. De informatie op deze geheugendragers blijft aanwezig ook al wordt de computer uitgezet en weer aangezet. Om de informatie beschikbaar te maken worden de gegevens in *bestanden* (Engels: file) gezet. De manier waarop dat wordt gerealiseerd wordt een *bestandssysteem* (Engels: file system) genoemd. Bekende bestandssystemen zijn NTFS op Windows en ext3 op Linux. Het besturingssysteem zorgt ervoor dat de bestanden op ordentelijke wijze kunnen worden benaderd.

³ Een programmeerparadigma is een manier van programmeren en een wijze waarop een programma wordt vormgegeven.

In de praktijk komen twee soorten computers tegen: de algemeen bruikbare computers waarop een besturingssysteem draait (PC, laptops) en dat grote hoeveelheid aan verschillende programma's kan uitvoeren en de zogenoemde *embedded systems*. Een embedded system is meestal geschikt om één taak uit te voeren en is voorzien van een kleine processor met geheugen en I/O-faciliteiten. Al deze componenten zijn op één ic (Integrated Circuit of chip genoemd) geplaatst. We noemen zo'n processorsysteem een *microcontroller*. Voorbeelden van embedded systems zijn (de besturingen van) wasmachines, koelkasten, televisies, thermostaten, horloges en bloeddrukmeters. De markt voor embedded systems is trouwens vele malen groter dan de markt voor algemene bruikbare computers.

De microcontroller heeft meestal geen besturingssysteem. Dat noemen we *bare metal* systemen. Dat betekent dat de programmeur alles zelf moet ontwikkelen. Gelukkig leveren fabrikanten ontwikkelsystemen om programma's voor microcontrollers te produceren. Dat worden *compiler suites* of *toolchains* genoemd.

1.2 De compiler

C is een zogenoemde derde-generatie-taal. De broncode van een C-programma is voor een mens gewoon te lezen. Deze broncode kan niet direct door de processor worden uitgevoerd. De *C-compiler* vertaalt de broncode naar instructies die de processor wel kan uitvoeren. Dat vertalen wordt *compileren* genoemd. De instructies worden in bitpatronen in het geheugen opgeslagen. De bitpatronen noemen we *machinecode*. Elk type processor heeft zijn eigen verzameling bitpatronen, dat wordt de *Instruction Set Architecture* genoemd. De ISA van een Intel-processor is dus anders dan die van een ATmega-processor, maar we kunnen wel zeggen dat elke processor ongeveer dezelfde soort instructies bevat. De C-compiler moet weten voor welke processor een C-programma wordt vertaald.

De C-programma kan niet worden uitgevoerd door de computer. We moeten eerst een programma compileren zodat instructies worden gegenereerd die de processor wel kan uitvoeren. Het vertaalde programma wordt een *uitvoerbaar bestand* of *executable* genoemd. Een uitvoerbaar bestand wordt net als een C-programma opgeslagen op de harddisk. Bij het ontwikkelen van programma's voor microcontrollers wordt gebruik gemaakt van PC of laptop met een besturingssysteem en een toolchain. We noemen zo'n compiler een *cross compiler*. De uitvoerbare instructies worden via de PC in het geheugen van de microcontroller *geprogrammeerd*.

In het licht van het compilatieproces noemen we nog twee begrippen die we vaak zullen gebruiken: *compile-time* en *runtime*. Met compile-time bedoelen we dat iets tijdens compileren van een C-programma bekend moet zijn. Een voorbeeld hiervan is het bepalen van de grootte van een variabele. Met runtime bedoelen we dat tijdens het uitvoeren van een programma iets wordt bepaald, bijvoorbeeld hoe vaak iets moet worden uitgerekend.

Hoewel een derde-generatie-taal suggereert dat we een C-programma door verschillende compilers kunnen laten vertalen en op verschillende computersystemen (lees: processoren) kunnen uitvoeren, moeten we helaas opmerken dat dat niet het geval is. Het is zeker mogelijk om een programma te schrijven dat door de Microsoft C-compiler en door de GNU C-compiler kan worden vertaald. Maar er zijn ook veel verschillen. Zo kan de grootte van gegevens door verschillende compilers op verschillende wijze worden geïnterpreteerd. Ook het bestandssysteem is op computers verschillend. Windows en Linux doen dat op hun eigen

wijze. Het is dus best lastig om een programma te schrijven dat op beide besturingssystemen zonder problemen draait. En wat te denken van een microcontroller dat niet eens een bestandssysteem heeft. Daar heeft het bewerken van bestanden geen zinnige betekenis.

1.3 Bibliotheken en functies


Stel dat we een klein stukje C-programma hebben geschreven dat iets voor ons uitrekent, bijvoorbeeld het gemiddelde van een aantal getallen. We willen dit stukje programma vaker gebruiken in een groot C-programma. Dan plaatsen we de berekening in een *functie*. We kunnen nu in ons C-programma de functie meerdere malen *aanroepen*. De functie hoeft dus maar één keer geprogrammeerd te worden om meerdere keren gebruikt te worden.

Er zijn ook functies die al geschreven zijn, zoals het afdrukken van tekst op het scherm. We hoeven dus niet zelf een functie te schrijven die dat voor ons doet. Als deze functies zijn ondergebracht in *bibliotheken* (Engels: library). Als tijdens het compileren zo'n functie nodig is, dan zoekt de compiler in de bibliotheek of de functie beschikbaar is. De compiler “plakt” dan de functie bij het programma. Dat plakken wordt *linken* genoemd.

Bij de C-compiler wordt een zeer uitgebreide bibliotheek geleverd, de zogenoemde *standard library*. De standard library bevat functies voor het afdrukken van tekst en gegevens, het bewerken van bestanden en het manipuleren van *strings* (een stukje tekst opgeslagen in het geheugen). Daarnaast wordt ook de *mathematical library* meegeleverd waarin functies voor het berekenen van sinus, cosinus, logaritmen en e-machten zijn opgeslagen. Tijdens het compileren moet opgegeven worden dat moet worden gezocht in de bibliotheken.

1.4 Ontwikkelssystemen

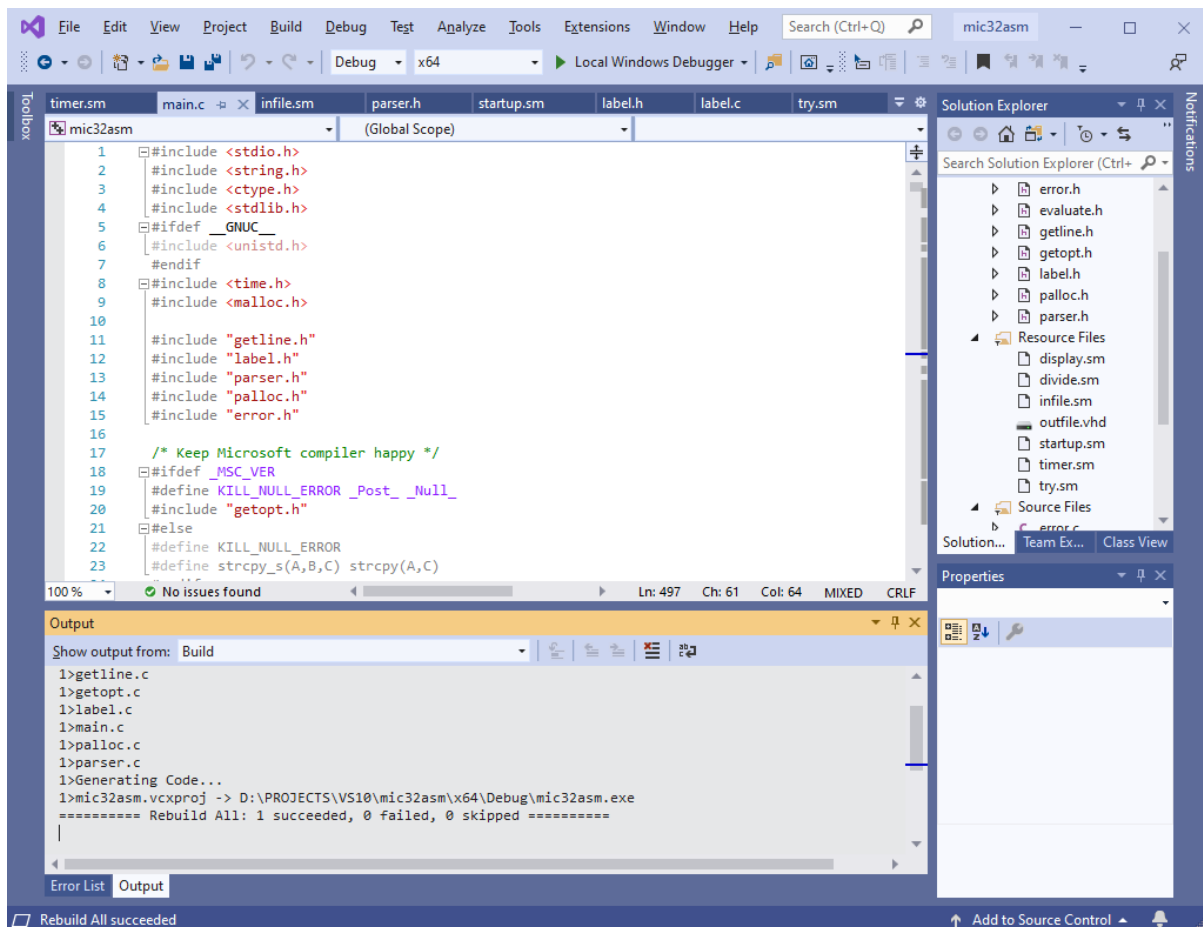
De meeste boeken gaan uit van een *command line interface* op een Unix-derivaat. Met behulp van een *editor* wordt een programma ingevoerd en wordt op de commandoregel de C-compiler gestart. Daarna wordt het programma (ook via de commandoregel) gestart. Natuurlijk is het ook mogelijk om alles met behulp van de command line te doen. Deze werkwijze wordt veel gebruikt op Linux-systemen maar kan ook op Windows worden gebruikt. Een voorbeeld is te zien in onderstaande figuur.

 Command Prompt	
C:\Users\C> notepad mooi.c	(start Notepad)
C:\Users\C> gcc -o mooi.exe mooi.c	(start C-compiler)
C:\Users\C> .\mooi.exe	(start uitvoerbaar programma)
C is een mooie taal	(de uitvoer op het scherm)
C:\Users\C>	

Dit is echter niet de werkwijze van veel programmeurs. Gelukkig zijn er goede ontwikkel-systemen (IDE: Integrated Development Environment) die het programmeerwerk verlichten. Bekende systemen zijn Microsoft Visual Studio, Code::Blocks en Apple's Xcode. Zulke ontwikkel-systemen zorgen ervoor dat de programmeur gemakkelijk het programma kan invoeren, de compiler kan starten en het programma kan *debuggen*. Dat laatste is vaak nodig

omdat blijkt dat (de uitvoer) een programma niet verloopt zoals de programmeur voor ogen had. Met debuggen wordt het programma stap voor stap doorlopen en kan de programmeur (of is het debugger) de inhoud van *variabelen* bekijken. Ook kan de programmeur bepalen of de *statements* (opdrachten voor de computer) op de juiste volgorde worden uitgevoerd.

Een voorbeeld van Microsoft Visual Studio is te zien in figuur 1.1. Visual Studio ondersteunt het ontwikkelen van software voor Windows-computers met behulp van een groot aantal talen: C, C++, C#, Python. Het is zelfs mogelijk om software te ontwikkelen voor Linux-systemen. Een korte introductie wordt gegeven in bijlage B.



Figuur 1.1: Voorbeeld van Microsoft Visual Stdio.

We zullen in de overige paragrafen in sneltreinvaart enkelen concepten van C bespreken. In de volgende hoofdstukken wordt de taal verder uitgediept.

1.5 Afdrukken op het scherm

Laten we beginnen met een eenvoudig voorbeeld. Het programma in listing 1.1 drukt de regel De som van 3 en 7 is 10 op het scherm af. We zullen het programma stap voor stap doorlopen.

In regel 1 wordt een zogenoemde *header-bestand* geladen, in dit geval het bestand `stdio.h`. We leggen zo meteen uit waarom dat nodig is.

```

1 #include <stdio.h>
2
3 int main(void) {
4
5     int a = 3;
6     int b = 7;
7
8     int som;
9
10    som = a+b;
11
12    printf("De_som_van_%d_en_%d_is_%d\n", a, b, som);
13
14    return 0;
15 }

```

Listing 1.1: Eerste C-programma.

In regel 3 wordt kenbaar gemaakt dat het programma de functie `main` heeft. Een C-programma heeft de *altijd* de functie `main`. Het (gecompileerde) programma wordt hier gestart.

In regel 5 en 6 worden twee *variabelen* gedeclareerd, de variabelen `a` en `b`. Een variabele kan door het programma gebruikt worden om gegevens te bewerken. Declaratie wil zeggen dat een variabele kenbaar worden gemaakt. Bij de declaratie wordt opgegeven wat het type is van de variabele. Dit gebeurt middels het *keyword* `int`, dat betekent dat `a` en `b` alleen geheeltallige getallen kan opslaan (Engels: integer). Aan de variabelen worden waarden toegekend. We noemen dit *initialisatie* van variabelen.

In regel 8 wordt de variabele `som` gedeclareerd zonder initialisatie. Dat betekent dat op dat moment de waarde (of inhoud) van de variabele *onbekend* is. Vervolgens wordt in regel 10 de som van `a` en `b` berekend middels de *optel-operator* `+` en wordt het resultaat *toegekend* aan variabele `som`. Vanaf regel 10 is de waarde van `som` dus 10.

In regel 12 wordt de functie `printf` aangeroepen. Deze functie is al geschreven is zit in de standard library. De functie krijgt vier *argumenten* mee, gegevens die in de functie verwerkt worden. Het eerste argument is een *string*. Een string is een stukje tekst, maar we spreken meestal van een rij karakters. Daarna volgen de (waarden) van de variabelen `a`, `b` en `som`.

Het eerste argument van `printf` wordt een *format string* genoemd. Dat ligt niet in C-taal vast maar wordt algemeen gebruikt. De format string bevat karakters, zogenoemde *format specifications* en een *escape sequence*. Een format specification begint met een procentteken gevolgd door een letter. De functie `printf` gebruikt deze format specifications om de gegevens af te drukken. De eerste `%d` zorgt ervoor dat variabele `a` wordt afgedrukt als een decimaal geheel getal. Op overeenkomstige wijze worden ook `b` en `som` afgedrukt. Aan het einde van de format string is een escape sequence te zien, in dit geval `\n`. Dit zorgt ervoor dat een volgende afdruk wordt begonnen aan het begin van de volgende regel (Engels: newline).

In regel 14 wordt met het keyword `return` aangegeven dat het programma wordt afgesloten. Dat behoeft enige uitleg. Een (gecompileerd) programma wordt gestart door het besturingssysteem. Aan het einde wordt het programma afgesloten. Het besturingssysteem “ruimt” het programma op en zorgt ervoor dat het gebruikte geheugen weer vrijgegeven wordt voor volgende programma’s. We kunnen aan het besturingssysteem een getal teruggeven, in dit geval 0. Het is aan het besturingssysteem om hier wat mee te doen. Gebruikelijk is om 0 terug te geven als alles goed verlopen is. Een ander getal dan 0 geeft over het algemeen aan dat er iets fout gegaan is.

Dat de `return`-waarde een geheel getal moet zijn, kunnen we zien aan de definitie van de functie `main`. We zien in regel 3 dat `main` een geheel getal teruggeeft (keyword `int`) en geen argumenten meekrijgt (keyword `void`).

Hoe weet de C-compiler nou hoe de functie `printf` moet worden aangeroepen? Dat wordt geregeld met een *function prototype*. We hoeven dat zelf niet op te geven want dat is al gedaan en is te vinden in het header-bestand `stdio.h`. De eerste regel geeft dus aan dat dit bestand geladen moet worden. De tekens `<` en `>` geven aan dat gezocht moet worden op een bepaalde plek in het bestandssysteem. We hoeven dat verder niet te weten.

1.6 Invoer van het toetsenbord

We kunnen het eerste programma interessanter maken door aan de gebruiker te vragen om twee gehele getallen in te voeren. Het programma is te zien in listing 1.2. We springen eerst naar regel 14 waar de functie `scanf` ervoor zorgt dat een geheel getal wordt ingelezen van het toetsenbord en in variabele `a` wordt gezet. De format specifier `%d` hadden we al eerder gezien, maar de constructie `&a` is nieuw. De ampersand (`&`) zorgt ervoor dat aan de functie `scanf` het *adres* van variabele `a` meegegeven wordt. Het adres van de variabele is de plek waar de variabele in het geheugen ligt. Op deze manier kan `scanf` de informatie op je juiste plek zetten. In regel 17 wordt hetzelfde gedaan voor variabele `b`. We drukken voor het inlezen nog even netjes af hoe de gebruiker moet handelen.

OPMERKING

De Microsoft C-compiler bestempelt `scanf` als “onveilig”. Een compilatie met `scanf` zal eindigen met een foutmelding. In plaats daarvan moet de functie `scanf_s` worden gebruikt. Helaas ondersteunen andere compilers deze functie niet. Dat zal resulteren in een programma dat niet door iedere compiler kan worden vertaald. Om het probleem te omzeilen hebben we gebruik gemaakt van een *macro*. In regel 4 definiëren we dat het woord `scanf` vervangen moet worden door `scanf_s`. Op andere compilers, bijvoorbeeld de GNU-C compiler, moet deze regel verwijderd worden. Overigens wordt op vele fora gewaarschuwd voor de onveiligheid van `scanf` en worden alternatieven gegeven. Wij gebruiken `scanf` hier wel *for the sake of simplicity*. Het is beter om `scanf` te vermijden.

```

1 #include <stdio.h>
2
3 /* Make Visual Studio happy */
4 #define scanf scanf_s
5
6 int main(void) {
7
8     int a;
9     int b;
10
11     int som;
12
13     printf("Geef_eeen_getal:_");
14     scanf("%d", &a);
15
16     printf("Geef_nog_eeen_getal:_");
17     scanf("%d", &b);
18
19     som = a + b;
20
21     printf("De_som_van_%d_en_%d_is_%d\n", a, b, som);
22
23     return 0;
24 }

```

Listing 1.2: *Invoer van de gebruiker opvragen.*

2

Variabelen, datatypes en expressies

Een C-programma bestaat uit variabelen en functies. We gebruiken variabelen om data te bewerken, zoals berekenen van nieuwe waarden of testen of twee variabelen aan elkaar gelijk zijn. De compiler moet notie hebben welke variabelen in een programma gebruikt worden en hoe groot de variabelen zijn. De meeste moderne computersystemen werken met 64-bits eenheden en dat betekent dat een variabele een bepaald bereik heeft; niet elk getal kunnen we in een variabele opslaan. We gebruiken functies om bewerkingen op de variabelen te beschrijven. We hebben al één functie gezien: `main`. Functies worden in detail besproken in hoofdstuk 4.

2.1 Variabelen

Een *variabele* is een object waaraan een waarde kan worden toegekend. Een variabele moet aan het C-programma kenbaar gemaakt worden. We noemen dat de *declaratie* van de variabele. Bij de declaratie wordt het *type* van de variabele opgegeven. Zo geeft de declaratie

```
int a;
```

aan dat in het programma de variabele `a` wordt gebruikt van het type `int`. Een `int` is een geheeltallig getal; het Engelse woord hiervoor is *integer*. Getallen met een komma, zoals 3,14159, kunnen we niet in een `int` opslaan. We kunnen nu in het programma de waarde van de variabele vastleggen met een *toekenning*. In het Engels is dat een *assignment*. Als we aan `a` de waarde 2 willen toekennen dan schrijven we

```
a = 2;
```

De C-compiler zorgt ervoor dat in het geheugen van de computer ruimte wordt gereserveerd om de variabele op te slaan. Als we in een programma de variabele `a` gebruiken dan weet de compiler op welk geheugenadres hij moet zoeken. We mogen de variabele onbeperkt aantal keer aanpassen. Zo kunnen we schrijven:

```

1 int a;
2 ...
3 a = 2;
4 ...
5 a = -3;
6 ...
7 a = 5;

```

Listing 2.1: Meerdere toekenningen aan een variabele.

De naam van de variabelen mag 31 karakters lang zijn. Er is geen noodzaak (en zelfs *bad practice*) om namen uit één karakters te laten bestaan. Hoofdletters en kleine letters zijn *niet* gelijk aan elkaar dus `count` is niet hetzelfde als `COUNT`. Dit wordt kast-ongevoelig genoemd. Overigens accepteren C-compilers langere namen dan 31 karakters.

Er zijn enkele regels aan de namen van variabelen:

- Moet beginnen met een letter;
- Mag alleen letters, cijfers en de underscore bevatten;
- Mag geen keyword zijn;
- Mag geen bekende functienaam zijn;¹
- Mag niet eerder gedeclareerd zijn;

Enkele goede voorbeelden:

```
Count val23 loop_counter ary2to4
```

Enkele foute voorbeelden:

```
_loop 50cent int printf
```

De meeste C-compilers accepteren de underscore als eerste karakter. Dit wordt echter afgeraden omdat veel *systeemroutines* de underscore als eerste karakter gebruiken en dat kan problemen geven in de latere stadia van de compilatie.

Bij de declaratie van variabelen mogen ook gelijk waarden worden toegekend, op voorpraak dat de waarden constant zijn. Een voorbeeld is hieronder te zien:

```
int a = 2;
int b = 3+5;
```

2.2 Datatypes

Een variabele is van een bepaald *datatype*. We hebben er al één gezien, de `int`. Een datatype heeft een bepaalde grootte. Zo bestaat een `int` *meestal* uit 32 bits. Er zijn echter

¹ Dat mag wel, maar dan kunnen we de functie niet in een programma gebruiken. We kunnen dus schrijven `int printf;` maar dan kan de `printf`-functie niet aangeroepen worden.

ook systemen en C-compilers waar een `int` uit 16 bits bestaat. Dat een `int` uit 32 bits bestaat, betekent dat niet elk willekeurig getal aan de `int` kunnen worden toegekend. Er is een minimale en maximale waarde. Voor de `int` is de minimale waarde -2147483648 en de maximale waarde $+2147483647$.

Nu is de grootte van een `int` niet altijd noodzakelijk of toereikend in een bepaalde situatie. We kunnen dan kiezen voor een ander type. C kent een aantal geheeltallige datatypes van diverse grootten. Deze zijn weergegeven in tabel 2.1.

Tabel 2.1: De datatypes die beschikbaar zijn in C.

type	bits	bereik
<code>char</code>	8	$-128 \text{ — } +127$
<code>short int</code>	16	$-32768 \text{ — } +32767$
<code>int</code>	16	$-32768 \text{ — } +32767$
	32	$-2147483648 \text{ — } +2147483647$
<code>long int</code>	32	$-2147483648 \text{ — } +2147483647$
<code>long long int</code>	64	$-9223372036854775808 \text{ — } +9223372036854775807$

Een `char` wordt gebruikt om een karakter op te slaan en is 8 bits groot. De interpretatie van het opgeslagen karakter is afhankelijk van de computer waarom het programma draait. Merk op dat een `char` zowel negatieve als positieve waarden kan bevatten. De meest gebruikte codering is de ASCII-code. Om het karakter 'A' op te slaan schrijven we:

```
char karak;
karak = 'A';    /* assign character A */
```

Nu is de C-compiler erg coulant in het toekennen van waarden. We mogen voor de toekenning ook een getal gebruiken. We kunnen het bovenstaande ook schrijven als:

```
char karak;
karak = 65;    /* assign character A */
```

Uiteraard moet de waarde passen en we merken terloops op dat negatieve waarden geen echte karakters voorstellen. Een `short int` is meestal 16 bits. Een `int` is 16 of 32 bits groot afhankelijk van de gebruikte C-compiler en de onderliggende hardware. Een `long int` is meestal 32 bits. De C-standaard legt vast dat:

$$\text{grootte short int} \leq \text{grootte int} \leq \text{grootte long int}$$

De keywords `short` en `long` worden *qualifiers* genoemd. Bij gebruik hiervan mag `int` weggelaten worden:

```
short sh;    /* short int */
long lo;    /* long int */
```

Als we aan een integer alleen maar positieve waarden of 0 toekennen dan kunnen we de qualifier `unsigned` bij declaratie opgeven. Het (positieve) bereik is dan ongeveer twee keer zo groot als bij de signed variant:

```
unsigned short int ush;    /* range 0 to 65535 */
```

Naast geheeltallige datatypes kent C ook een drietal *floating point* datatypes. Het Nederlandse woord hiervoor is *drijvende komma*. Hier ontstaat al gelijk de eerste verwarring: in C wordt de komma vervangen door de punt, zoals gebruikelijk is in Engelstalige literatuur. We schrijven dus 3.14 en niet 3,14.

Een `float` is een datatype met een grootte van 32 bits. Een `double` is een datatype met een grootte van 64 bits. We zouden verwachten dat de `long double` dan een datatype van 128 bits is, maar dat is niet altijd waar. De `long double` is soms ook 80 bits, bijvoorbeeld in Intel-processoren.

Tabel 2.2: De *floating point* datatypes die beschikbaar zijn in C.

type	bits	kleinste getal	grootste getal
<code>float</code>	32	$\approx 1,18 \times 10^{-38}$	$\approx 3,4 \times 10^{38}$
<code>double</code>	64	$\approx 2.2 \times 10^{-308}$	$\approx 1.8 \times 10^{308}$
<code>long double</code>	80/128	¹⁾	¹⁾

¹⁾ Afhankelijk van de implementatie 80 of 128 bits.

Voorbeelden van *floating point* variabelen:

```
float f;
double d;
```

Met betrekking tot de nauwkeurigheid kunnen we nog het volgende vermelden. Een `float` heeft een nauwkeurigheid van ongeveer 6 decimale cijfers. Het heeft dus geen zin om meer cijfers toe te voegen, de extra cijfers worden genegeerd. Een `double` heeft een nauwkeurigheid van ongeveer 16 decimale cijfers. Niet alle getallen kunnen exact in een *floating point*-variabele worden opgeslagen. Zo is $1/7$ niet exact te representeren. Bij veelvuldig rekenen met *floating point*-variabelen treedt er verlies op in de representatie. Als we bijvoorbeeld $1/3$ met 3 vermenigvuldigen, kan het zijn dat het resultaat 0,99999 is en niet 1,0.

Al deze datatypes worden *enkelvoudige datatypes* genoemd. Dat betekent dat de compiler ze ziet als één eenheid. Het is niet mogelijk om enkelvoudige datatypes te splitsen over meerdere andere datatypes.

Naast de genoemde datatypes bestaat er nog een datatype: de *pointer*. Een pointer is een variabele die het *adres* bevat van een (andere) variabele. Pointers worden in detail besproken in hoofdstuk 6.

2.3 Constanten

Een geheeltallige constante zoals 123 is van het type `int`. Door er een L achter te zetten wordt de constante een `long`. Dus 123L is een `long`. Als een constante te groot is voor een `int` wordt het automatisch gezien als een `long`. Een unsigned constante wordt aangegeven met een U aan het einde. Dat mag in combinatie met de L. Een *floating point*-constante bestaat uit cijfers, optioneel met een punt. Dus 1.23 is een *floating point*-constante. Om 10-machten aan te geven wordt de *E*-notatie gebruikt. Zo staat $125.0E-3$ voor 0,125. Een

floating point-constante is automatisch van het type `double`, tenzij er een `F` aan het einde staat, dan is de constante van het type `float`.

Geheeltallige constanten kunnen op drie manieren worden ingevoerd: als decimaal getal, als octaal getal of als hexadecimaal getal. Een constante die begint met `0` wordt gezien als octaal getal. Een octaal getal bestaat alleen uit de cijfers `1 t/m 7`. Een constante die begint met de *prefix* `0x` of `0X` wordt gezien als een hexadecimaal getal. Een constante die begint met de cijfers `1 t/m 9` wordt gezien als een decimaal getal.

```
int a = 0377;    /* octal 377 is decimal 255 */
int b = 0xa9;    /* hexadecimal A9 is decimal 169 */
int c = 127;     /* decimal 127 */
```

De C-standaard kent geen manier om binaire getallen in te voeren. Veel compilers ondersteunen dit toch door de prefix `0b` te gebruiken. Zo is de constante `0b10101001` gelijk aan `169`.

Een *karakterconstante* is een geheel getal, geschreven als een één karakter tussen apostrofes. Zo is `'A'` een karakterconstante. De interne representatie is een geheel getal dat overeen komt met de karakterset van de computer. Over het algemeen wordt de ASCII-code gebruikt en komt `'A'` overeen met de waarde `65`. De standaard ASCII-tabel is te vinden in bijlage [A](#).

Niet alle karakters kunnen zo worden ingevoerd. Een voorbeeld is het karakter `'` zelf. Om dit karakter in te voeren gebruiken we een *escape sequence*. Een escape sequence bestaat uit een *backslash* (`\`) en een karakter. We kunnen de apostrofe dus voorstellen met `'\''`.

De C-standaard kent een hele verzameling van dit soort escape sequences. Eén daarvan hebben we al meerdere malen gezien: de newline. Deze wordt aangegeven met `'\n'`. Een aantal is vermeld in tabel [2.3](#).

Tabel 2.3: Enkele escape sequences in C.

<code>\n</code>	newline	<code>\\</code>	backslash
<code>\r</code>	carriage return	<code>\'</code>	single quote
<code>\a</code>	bell	<code>\"</code>	double quote
<code>\b</code>	backspace	<code>\0</code>	null byte
<code>\f</code>	formfeed	<code>\t</code>	horizontal tab

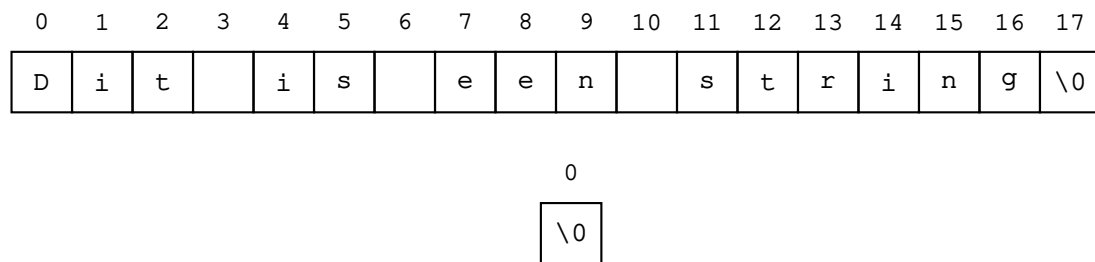
Andere karakters kunnen worden gevormd door de backslash, gevolgd door de letter `x` en één of twee hexadecimale cijfers. We kunnen de letter `A` dus ook schrijven als `'\x41'`.

Een *string constante*, of kortweg string, is een rij van karakters die wordt begonnen en afgesloten met aanhalingstekens. De rij mag ook leeg zijn. Dan spreken we dan van een lege string. Twee voorbeelden:

```
"Dit is een string"
"" /* The empty string*/
```

Let erop dat een string *geen* enkelvoudig datatype is zoals `int` en `double`. Een string is een rij karakters die in het geheugen liggen. Technisch gezien is een string een *array*. We kunnen dus niet de waarde van een string bepalen zoals dat wel mogelijk is van een `int` of een `char`. Aan het einde van een string wordt automatisch een nul-karakter geplaatst. Er

is dus altijd één karakter meer nodig dan het aantal karakters in een string. In figuur 2.1 zijn de twee strings afgebeeld.



Figuur 2.1: *Uitbeelding van twee strings.*

C kent geen ingebouwde operaties op strings, zoals inkorten, aan elkaar plakken, kopiëren of lengte bepalen. Dit moet allemaal door de programmeur zelf ontworpen worden. Gelukkig kent de standaard C-bibliotheek een groot aantal functies voor het bewerken van strings. We komen hierop terug in hoofdstuk 5.

Bij de declaratie van variabelen mogen ook gelijk waarden worden toegekend, op voorspraak dat de waarden constant zijn. De constante waarde mag ook berekend worden. Een aantal voorbeelden is hieronder te zien:

```
int a = 2;
char z = 'z';
short a = -10+5;
long l = 123L;
float tau = 2.0F*3.14159F;
double e = 2.718281828;
double googol = 1E100;
unsigned char = 128+127;
```

Let erop dat de constante wordt geconverteerd naar het type van de variabele. Sommige compilers geven een waarschuwing als zo'n conversie plaatsvindt.

```
int a = 25.6;    /* converted to 25 */
float b = 7;     /* converted to 7.0 */
```

Bij de declaratie van een variabele mag de qualifier `const` worden opgegeven dat inhoudt dat aan de variabele eenmalig een waarde wordt toegekend. De `const`-variabele kan daarna niet meer veranderen. Dat is bijzonder handig als we een constante waarde moeten gebruiken in een programma. Zo kunnen we na de declaratie

```
const int aantal = 10;
```

de variabele `aantal` gebruiken als vervanging voor het getal 10. Als later blijkt dat het aantal moet worden aangepast, dan hoeven we alleen maar de variabele `aantal` aan te passen.

Uiteraard moet de constante passen. Als we bijvoorbeeld

```
char ch = 65536;
```

uitvoeren zal de compiler een waarschuwing geven dat de constante niet past in een `char`.

2.4 Expressies

Een *expressie* is elk stukje programma dat een uitkomst oplevert. Zo is $2+2$ een expressie die de waarde 4 oplevert. We kunnen de uitkomst van een expressie toekennen aan een variabele:

```
a = 2 + 2;
```

Een expressie mag variabelen bevatten:

```
a = b - c;
```

Zelfs de regel

```
a;
```

is een expressie die 4 oplevert als de waarde van a 4 is. Ook het vergelijken van twee variabelen is een expressie en mag toegekend worden aan een variabele:

```
a = b == c;    /* note the == */
```

Als b gelijk is aan c dan wordt a gelijk aan 1. Als b ongelijk is aan c dan wordt a gelijk aan 0. We zullen het vergelijken van variabelen en constanten nader bekijken in hoofdstuk 3.

Een expressie mag willekeurig complex zijn, maar let op de voorrangsregels: vermenigvuldigen en delen gaan voor op optellen en aftrekken. Haakjes worden gebruikt de prioriteiten te veranderen:

```
a = (2+b)*5+c;
```

2.4.1 Rekenkundige operatoren

De *binaire* rekenkundige operatoren zijn $*$, $/$, $\%$ modulus operator, $+$ optellen en $-$ aftrekken. Binair wil in dit verband zeggen dat de operatoren op twee variabelen (of constanten of expressies) werken. De expressie

```
a % b
```

berekent de rest van de deling van a gedeeld door b . Dit wordt de *modulus operator* genoemd. Daarnaast kunnen $+$ en $-$ ook als *unaire operator* gebruikt worden. Een voorbeeld hiervan is

```
int a = -5;
```

Let goed op de prioriteiten van de operatoren. De unaire operatoren gaan voor op vermenigvuldigen ($*$), delen ($/$) en modulus ($\%$). De binaire optelling ($+$) en aftrekking ($-$) hebben en lagere prioriteit dan $*$, $/$ en $\%$. Haakjes kunnen de volgorde veranderen. Merk op dat de *associativiteit* van links naar rechts is. Dus de expressie

```
a / b / c
```

is equivalent aan

```
(a / b) / c
```

maar *niet* aan

`a / (b / c)`

Verder moet worden opgelet bij deling van geheeltallige getallen. Een deling rondt naar beneden af, dus de *fractie*, het deel van een getal na de komma, wordt weggelaten. Dat betekent dat

`1 / 3 * 3`

de waarde 0 oplevert. Eerst wordt $1/3$ berekend en de geheeltallige uitkomst hiervan is 0. Daarna wordt de uitkomst met 3 vermenigvuldigd. De uitkomst is nog steeds 0.

2.4.2 Relationale operatoren

De zes *relationele operatoren* zijn

`== != > >= < <=`

Hierin zijn `==` gelijk aan en `!=` ongelijk aan. Verder is `>` groter dan, `>=` is groter dan of gelijk aan, `<` is kleiner dan en `<=` is kleiner dan of gelijk aan. Deze hebben een lagere prioriteit dan de alle rekenkundige operatoren, dus `i < len-1` wordt gelezen als `i < (len-1)`. Overigens hebben `==` en `!=` een lagere prioriteit dan de overige vier. Een relationele operator levert de waarde 1 op als de vergelijking waar is en anders levert de relationele operator 0 op.

De relationele operatoren hebben vooral nut bij beslissingen. Zie listing 2.2.

```
1 int main(void) {  
2  
3     int a = 2, b = 3;  
4  
5     if (a<b) {  
6         printf("a_is_kleiner_dan_b\n");  
7     }  
8     return 0;  
9 }
```

Listing 2.2: Voorbeeld van een beslissing.

2.4.3 Logische operatoren

De `&&`- en `||`-operatoren zijn logische operatoren die expressies met elkaar verbinden. Ze hebben een relatie met de relationele operatoren. Zo wordt variabele `isdig` in de expressie

`isdig = ch>='0' && ch<='9';`

gelijk aan 1 als `ch` een cijferkarakter is, anders wordt `isdig` 0. In de expressie

`isbit = ch=='0' || ch=='1';`

wordt `isbit` gelijk aan 1 als `ch` een '0' of een '1' is, anders wordt `isbit` 0. Het berekenen van een dergelijke expressie wordt gestopt op het moment dat de uitkomst al duidelijk is. Dus de berekening van

```
i==5 && j>3
```

wordt gestopt als *i* ongelijk aan 5 is. De uitkomst staat dan namelijk al vast. Dit wordt *shortcut evaluation* genoemd.

De unaire *negatieoperator* `!` zet een expressie die 0 oplevert om in een 1 en een expressie die *niet-nul* oplevert wordt omgezet in een 0. Een typisch gebruik van `!` is te zien in [listing 2.3](#).

```
1 int main(void) {  
2  
3     int valid = a<5 || a>10;  
4  
5     if (!valid) { /* if valid is 0 */  
6         ...  
7     }  
8  
9     return 0;  
10 }
```

Listing 2.3: Voorbeeld van de negatieoperator.

Met al deze operatoren kunnen we willekeurig complexe expressies realiseren. Om bijvoorbeeld te bepalen of een jaartal een schrikkeljaar (Engels: leap year) is, gebruiken we de volgende expressie:

```
int leap = (year % 4 == 0 && year % 100 != 0) || year % 400 == 0;
```

We zullen het even uitleggen. Een schrikkeljaar is een jaar met als extra dag 29 februari. Zo'n jaartal is deelbaar door 4 en niet deelbaar door 100, of als het jaartal deelbaar is door 400, dan is het wel een schrikkeljaar. Dus 1984 is een schrikkeljaar, 2100 is geen schrikkeljaar maar 2000 weer wel. Deze expressie wordt vaak gebruikt bij het bepalen van een geldige datum.

2.4.4 Bitsgewijze operatoren

C biedt zes zogenoemde *bitsgewijze operatoren*. Ze kunnen alleen maar gebruikt worden bij geheeltallige variabele, constanten of expressies. Deze zijn:

<code>&</code>	AND
<code> </code>	OR
<code>^</code>	EXOR
<code><<</code>	naar links schuiven
<code>>></code>	naar rechts schuiven
<code>~</code>	one's complement (alle bits geïnverteerd)

De *waarheidstabellen* van de AND-, OR- en EXOR-functies zijn gegeven in [tabel 2.4](#). De AND-functie wordt gebruikt om een of meer bits uit een variabele te selecteren of op 0 te zetten. De OR-functie wordt gebruikt om bits in een variabele op 1 te zetten en de EXOR-functie wordt gebruikt om bits in een variabele te inverteren.

Tabel 2.4: De drie bitsgewijze operatoren.

(a) AND-functie.

0	0		0
0	1		0
1	0		0
1	1		1

(b) OR-functie.

0	0		0
0	1		1
1	0		1
1	1		1

(c) EXOR-functie.

0	0		0
0	1		1
1	0		1
1	1		0

De schuifoperator `<<` schuift de bits in een variabele een aantal plekken naar links. De vrijgekomen bits worden opgevuld met nullen. De bits die aan de linkerkant “eruit vallen” gaan verloren. De expressie

```
1 << 4
```

geeft als resultaat 16. De schuifoperator `>>` schuift de bits in een variabele een aantal plekken naar rechts. De vrijgekomen bits worden *bij unsigned variabelen* aangevuld met nullen. Bij *signed variabelen* worden ze aangevuld met de *tekenbit*. De bits die aan de rechterkant eruit vallen gaan verloren. De unaire operator `~` inverteert alle bits in een variabele. Een bit die 1 is wordt een 0 en een bit die 0 is wordt een 1. In veel boeken wordt dit *one's complement* genoemd.

De operatoren kunnen door elkaar gebruikt worden. Een voorbeeld is:

```
a = a & ~0xff;
```

Deze expressie zorgt ervoor dat de 8 minst significante bits allemaal 0 worden. In figuur 2.2 zijn drie voorbeelden te zien van bitmanipulaties met AND, OR en EXOR. Hiervoor zijn de gegevens uit tabel 2.4 gebruikt.

7								0	
1	0	1	0	1	0	1	0		
1	0	0	0	0	0	0	0	0	AND
1	0	0	0	0	0	0	0	0	
1	1	0	1	1	0	1	0		
1	1	0	0	1	1	0	0		OR
1	1	0	1	1	1	1	0		
1	1	0	1	1	0	1	0		
1	1	1	1	0	0	0	0		EXOR
0	0	1	0	1	0	1	0		

Figuur 2.2: Drie voorbeelden van AND, OR en EXOR.

De bitsgewijze operatoren zijn bijzonder handig bij het gebruik van microcontrollers. We geven ter illustratie een voorbeeld voor een ATmega-microcontroller. We willen testen of een schakelaar is ingedrukt. We lezen de stand van de schakelaars in via de (speciale) variabele `PINA`. Als de schakelaar op bit 7 (hoogste bit van `PINA`) is ingedrukt, dan inverteren we de stand van de led die is gekoppeld aan bit 0 van de (speciale) variabele `PORTB`.

```
1 #include <avr/io.h>          /* PINA, PORTB, ... */
2
3 int main(void) {
4
5     if (PINA & 0x80) {        /* if key is pressed ... */
6         PORTB = PORTB ^ 0x01; /* ... invert lower led */
7     }
8
9 }
```

Listing 2.4: *bla*

2.5 Datatypeconversie

Het omzetten of converteren van het ene datatype naar het andere datatype kan op twee manieren worden gerealiseerd: automatisch door de C-compiler of expliciet door een *type cast* in het programma. We kunnen redelijkerwijs verwachten dat een “kleiner” type wordt omgezet naar een “groter” type. We noemen dat *promotie*. De regels voor automatische conversie zijn complex; we geven hier slecht een korte opsomming:

- Als een van de twee operanden een `long double` is, converteer de andere naar `long double`;
- Als een van de twee operanden een `double` is, converteer de andere naar `double`;
- Als een van de twee operanden een `float` is, converteer de andere naar `float`;
- Converteer `char` en `short` naar `int`;
- Als een van de twee operanden een `long` is, converteer de andere naar `long`.

De regels voor unsigned variabelen zijn nog lastiger, zeker als ze gecombineerd worden met signed integers. Het beste is om deze *mixed types* te vermijden.

Een conversie kan ook expliciet worden opgegeven. We noemen dit een *type cast*. We geven dit aan door het type tussen haken te zetten:

```
i = (int) ch;    /* promote character to integer */
```

In dit geval geeft dat geen problemen want een karakter past in een integer. Maar de cast:

```
ch = (char) i;   /* demote integer to character */
```

kan voor problemen zorgen als de waarde van `i` te groot is om in een karakter te plaatsen. Er gaan dan bits verloren. Een mooi voorbeeld van een type case is het omrekenen van een

temperatuur in graden Celsius naar graden Fahrenheit. Stel we hebben twee `float`s voor de temperaturen. Dan kan de conversie berekend worden met:

```
f = (float)9/5 * c + 32;
```

We hebben hier de integer 9 gecast naar een `float` want anders gaat bij de deling 9/5 informatie verloren. De constante 32 hoeft niet gecast te worden omdat eerder al de getallen en variabelen naar een `float` gecast zijn. Het is overigens *good practice* om 32.0 te schrijven.

2.6 Overflow

Overflow is de situatie als een expressie een waarde oplevert die te groot of te klein is en niet in de variabele kan worden opgeslagen. De C-compiler genereert geen code om hierop te testen. Dat zou wel kunnen, maar dat zou de executietijd van programma's nadelig beïnvloeden. De ontwerper van het programma moet hier zelf op toezien dat een overflow-conditie niet kan voorkomen. De reden dat overflow voorkomt is dat datatypes (en dus variabelen) uit een eindig aantal bits bestaan. Zo bestaat een `unsigned char` uit 8 bits en het grootste getal dat kan worden opgeslagen is 255 (binair 1111111₂). Tellen we daar 1 bij op dat is het resultaat 256 (binair 10000000₂) maar dat kan niet in de variabele worden opgeslagen. Het resultaat is dat de overvloedige bits gewoonweg worden geschrapt.

Overflow bij floating point-getallen werkt anders. Dat heeft te maken met de gebruikte specificatie van de getallen. Naast representeerbare getallen kent de floating point-specificatie nog drie speciale getallen: $+\infty$ (oneindig positief), $-\infty$ (oneindig negatief) en NaN (Not A Number). Ze geeft de deling 1/0 als resultaat $+\infty$ ². De deling 0/0 resulteert in NaN.

2.7 Vaste-lenge datatypes

Dat de grootte van integers aan elkaar gerelateerd zijn, kan in programma's voor problemen leiden. We weten nu niet of een `int` nu 16 of 32 bits is. Dit heeft geleid tot een aantal nieuwe datatypes die in de C99-standaard zijn vastgelegd. De grootte van deze types zijn exact. Zo is een `int8_t` een integer van precies 8 bits. Naast de integer met teken zijn er ook enkele types zonder teken. Zo is een `uint16_t` een integer van precies 16 bits met een minimale waarde van 0 en een maximale waarde van 65535.

Voor "normaal" gebruik zijn deze datatypes niet echt nodig. Maar bij het schrijven van programma's op microcontrollers is het vaak de enige goede methode om zeker ervan te zijn dat een bepaalde grootte van een variabele gegarandeerd is. Dit is met name belangrijk bij het gebruik van zogenoemde *I/O-adressen*. Dat zijn speciale adressen in het geheugen van de microcontroller waarmee communicatie met de buitenwereld mogelijk is, zoals het inlezen van de stand van een schakelaar of het laten branden van een led.

In de tabellen 2.5 en 2.6 zijn de nieuwe datatypes voor signed en unsigned te zien. Voordat we ze kunnen gebruiken moeten we eerst het *header-bestand* `stdint.h` laden. Dit is te zien in listing 2.5.

² Volgens de wiskunde is het resultaat van een deling door 0 onbepaald. De specificatie definieert echter de deling als oneindig.

Tabel 2.5: De vaste-lengte datatypes die beschikbaar zijn in C (signed).

type	bits	kleinste getal	grootste getal
int8_t	8	−128	+127
int16_t	16	−32768	+32767
int32_t	32	−2147483648	+2147483647
int64_t	64	−9223372036854775808	+9223372036854775807

Tabel 2.6: De vaste-lengte datatypes die beschikbaar zijn in C (unsigned).

type	bits	grootste getal
uint8_t	8	255
uint16_t	16	65535
uint32_t	32	4294967295
uint64_t	64	18446744073709551615

```

1  /* Load new datatypes */
2  #include <stdint.h>
3
4  int main(void) {
5
6      uint8_t byte = 0;
7      ...
8      while (byte < 128) {
9          ...
10         byte = byte + 1;
11     }
12 }
```

Listing 2.5: Voorbeeld van het gebruik van vaste-lengte datatypes.

2.8 Overige operatoren

C kent veel operatoren. In deze paragraaf beschrijven we ze even kort. In de volgende hoofdstukken worden ze verder uitgelegd.

2.8.1 Grootte van een variabele

De grootte van een variabele of datatype in *bytes* kan berekend worden met de operator `sizeof`. Dit kan alleen *tijdens het compileren* van het C-programma gebeuren. Tijdens het uitvoeren van het programma zijn alle grootten berekend. Er is één restrictie in het gebruik van `sizeof`: bij het gebruik met een datatype *moeten* haakjes gebruikt worden. Zie listing 2.6. Deze operator heeft een hoge prioriteit. We zullen `sizeof` nader bespreken in hoofdstuk 5.

```

1 int main(void) {
2     int i;
3
4     if (sizeof i == sizeof (long)) {
5         printf("De_grootte_van_i_is_gelijk_aan_een_long\n");
6     }
7
8     return 0;
9 }

```

Listing 2.6: Gebruik van *sizeof*.

2.8.2 Verhogen of verlagen met 1

Een toekenning als

```
i = i + 1;
```

kan geschreven worden als

```
i++;
```

of als

```
++i;
```

Het verschil is dat bij `i++` eerst de waarde van `i` wordt gebruikt en daarna met 1 wordt opgehoogd (dit wordt *postfix* genoemd), bij `++i` wordt de waarde van `i` eerst opgehoogd en daarna gebruikt (dit wordt *prefix* genoemd). Stel dat `i` is 5 dan zorgt

```
k = i++;
```

dat `k` gelijk is aan 5 en `i` gelijk is aan 6. Bij

```
k = ++i;
```

zijn `k` en `i` gelijk aan 6. We zullen deze operatoren tegenkomen in hoofdstuk 5. Deze operatoren hebben een hoge prioriteit.

2.8.3 Toekenningsoperatoren

Een toekenning als

```
i = i * 3;
```

mag geschreven worden als

```
i *= 3;
```

Let daarbij op dat de expressie aan de rechterkant van de toekenningsoperator als één eenheid wordt gezien. Dus

```
i *= y + 1;
```

wordt uitgewerkt als

Tabel 2.7: Toekenningsoperatoren.

Operatie		Evaluatie
<code>a *= b</code>	lees als	<code>a = (a) * (b)</code>
<code>a /= b</code>	lees als	<code>a = (a) / (b)</code>
<code>a %= b</code>	lees als	<code>a = (a) % (b)</code>
<code>a += b</code>	lees als	<code>a = (a) + (b)</code>
<code>a -= b</code>	lees als	<code>a = (a) - (b)</code>
<code>a <<= b</code>	lees als	<code>a = (a) << (b)</code>
<code>a >>= b</code>	lees als	<code>a = (a) >> (b)</code>
<code>a &= b</code>	lees als	<code>a = (a) & (b)</code>
<code>a = b</code>	lees als	<code>a = (a) (b)</code>
<code>a ^= b</code>	lees als	<code>a = (a) ^ (b)</code>

```
i = i * (y + 1);
```

Zo'n beetje alle gangbare operatoren kunnen worden gebruikt. Een lijst is te vinden in tabel 2.7. Deze operatoren hebben een lage prioriteit. Merk op dat bij de kolom "Evaluatie" de expressies `a` en `b` omringd zijn door haakjes. Dat geeft ook gelijk aan dat deze expressies eerst geëvalueerd worden voordat de operatie en toekenning plaatsvindt.

2.8.4 Conditionele expressie

De *conditionele expressie* is een operator die aan de hand van de waarde van een variabele (of constante of expressie) een van de twee opgegeven expressies uitvoert. Als we de maximale waarde van `a` en `b` willen bepalen dan kunnen we schrijven:

```
max = (a>b) ? a : b;    /* max = maximum(a, b) */
```

Een typisch voorbeeld is het afdrukken van de meervoudsvorm van een woord (of niet). Stel dat we aan het einde van een programma afdrukken hoeveel documenten verwerkt zijn. Dan kunnen we met behulp van een conditionele expressie het woord `document` of `documenten` afdrukken. We doen dat zoals te zien is in de regels 11 en 12 in listing 2.7.

De eerste parameter (`%d`) is het aantal documenten dat verwerkt is (`nrdocs`). De tweede parameter is een string (`%s`). De conditionele expressie

```
(nrdocs == 1) ? " " : "en";
```

geeft de (lege) string `" "` als `nrdocs` gelijk is aan 1 of de string `"en"` als `nrdocs` ongelijk aan 1 is. De haken rond de expressie `nrdocs == 1` zijn eigenlijk niet nodig want de prioriteit van de conditionele expressie is erg laag, maar het komt de leesbaarheid wel ten goede.

```

1 #include <stdio.h>
2
3 int main(void) {
4
5     int nrdocs = 0;
6
7     ...
8     /*          +----- nrdcos          */
9     /*          |          +---- " " or "en" */
10    /*          v          v                */
11    printf("Er_zijn_%d_document%s_verwerkt\n", nrdocs,
12           (nrdocs == 1) ? " " : "en"));
13
14    return 0;
15 }

```

Listing 2.7: Het afdrukken van een meervoudsvorm.

2.8.5 Komma-operator

De komma-operator scheidt twee expressies die één statement worden gezien; We bespreken de komma-operator in hoofdstuk 3.

2.8.6 Nog enkele operatoren

De volgende operatoren worden de verderop in het boek besproken:

[] -> . * (*dereference*) & (*address*)

2.9 Voorrangsregels van operatoren

Bij het uitwerken van expressies worden voorrangsregels gehanteerd. C kent een groot aantal operatoren. We hebben de *meest voorkomende* operatoren op volgorde van prioriteit opgesomd in tabel 2.8. Merk op dat de toekenningoperator = ook in de lijst voorkomt. Een in C-programma's gebruikelijke constructie is het inlezen van een karakter én gelijk testen of het een newline-karakter betreft:

```
while ((ch = getchar()) != '\n') { ... }
```

Hier gebeuren eigenlijk drie dingen. Eerst wordt de *functie* `getchar` aangeroepen om een karakter van het toetsenbord te lezen. Daarna wordt dit karakter toegekend aan variabele `ch`. De haakjes zorgen voor de juiste prioriteit. De uitkomst van de toekenning is de geheeltallige waarde van de toekenning en dat is het ingelezen karakter. Daarna wordt getest of het ingelezen karakter gelijk is aan het newline-karakter.

Uit deze tabel zijn de operatoren voor array's ([]), pointers (* en &) en structures (-> en .) weggelaten. Deze operatoren worden behandeld in volgende hoofdstukken.

Veel van deze operatoren volgen uit de noodzaak (of is het drang) om programma's compact te schrijven. Doorgewinterde programmeurs kunnen op deze manier vlot, zij het onleesbare,

Tabel 2.8: Voorrangsregels van de meeste operatoren.

Operatie	Associativiteit
()	links naar rechts
! ~ + - ++ -- (type cast) sizeof	rechts naar links
* / %	links naar rechts
+ -	links naar rechts
<< >>	links naar rechts
< <= > >=	links naar rechts
== !=	links naar rechts
&	links naar rechts
^	links naar rechts
	links naar rechts
&&	links naar rechts
	links naar rechts
?:	rechts naar links
= += -= *= /= %= &= ^= = <<= >>=	rechts naar links
,	links naar rechts

code produceren. Een eerste gedachte is dat zulke constructies bijdragen aan compacte uitvoerbare bestanden. Maar de huidige generatie compilers is zeer goed in het herkennen van de taal en genereren al bijna-minimale uitvoerbare bestanden. Het is beter om eerst een leesbaar programma te schrijven dat correct werkt en daarna gaan nadenken of het slimmer kan. Donald Knuth schreef al eens [10]:

Premature optimization is the root of all evil.

De lezer is gewaarschuwd.

3

Flow control

```
1 #include <stdio.h>
2 #include <stdlib.h>
3
4 int main() {
5
6     int i, ione, j;
7
8     for (i=0, ione=1, j=9; i<10; i++, ione++, j--) {
9         printf("%2dx9_=%1d%1d\n", ione, i, j);
10    }
11    return 0;
12 }
```

Listing 3.1: Afdrukken van de tafel van 9 met behulp van de komma-operator.

4

Funcities

5

Array's

De enkelvoudige datatypes hebben een beperking in de zin dat een groot aantal variabelen niet gemakkelijk kan worden verwerkt. Stel dat we de som van een aantal variabelen willen bepalen. We moeten dan voor elke variabele een declaratie geven en in het programma moeten we de som bepalen door alle variabelen op te tellen. Willen nu meer variabelen optellen, dan moeten we door het hele programma aanpassingen maken. Dat is nog wel te doen als het aantal variabelen beperkt is maar naar mate dat aantal groeit, wordt het programma groter en complexer. We kunnen veel van dit soort vraagstukken elegant oplossen met behulp van een *array*.

We bespreken array's op vele manieren, onder andere hoe een array wordt gedeclareerd en het wordt gebruikt. We laten zien dat het mogelijk is om “langs een array te lopen” met behulp van een lus. We kunnen tweedimensionale array's gebruiken waarbij een array uit rijen en kolommen bestaat. We zien dat een *string* een array is van karakters en dat in C bewerkingen op strings mogelijk zijn. We kunnen een array als parameter aan een functie meegeven met de kanttekening dat de grootte van de array niet berekend kan worden in de functie. Maar er zijn ook zaken die we hier niet bespreken, namelijk de relatie tussen array's en pointers. Dat wordt besproken in hoofdstuk 6.

5.1 Declaratie en selectie

Een *array* is een manier om bij elkaar behorende gegevens onder één naam te groeperen en te bewerken. Een array is een *samengesteld datatype*, een datatype dat bestaat uit een aantal enkelvoudige datatypes. Zo zorgt de declaratie

```
int a[10];
```

ervoor dat we gebruik kunnen maken van de tien `int`-variabelen:

```
a[0]  a[1]  a[2]  a[3]  a[4]  a[5]  a[6]  a[7]  a[8]  a[9]
```

Bij de declaratie wordt het aantal *elementen* opgegeven. In het voorbeeld zijn dat er 10. De nummering van de elementen begint bij 0 en loopt tot het aantal elementen minus 1.

Tevens wordt bij de declaratie het datatype opgegeven. Dit mag elk gangbaar datatype zijn.

We kunnen een element selecteren door gebruik te maken van de blokhaken `[]` met daarin het elementnummer. Dit elementnummer mag ook berekend worden met een expressie als het resultaat van de expressie maar een integer is. Om het achtste element toe te kennen aan de variabele `i` gebruiken we de toekenning:

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

Om de variabele `i` toe te kennen aan het vijfde element gebruiken we de toekenning:

```
a[4] = i;
```

We mogen alleen de elementen gebruiken die gedeclareerd zijn, dus de expressies

```
i = a[-1]; j = a[99];
```

zijn niet geldig¹. Overigens voegen de meeste compilers geen extra programmatuur toe om deze grenzen te bewaken. De programmeur moet het zelf in de gaten houden.

Een prettige manier om een array voor te stellen is om de array als een rij vierkante hokjes te tekenen. Boven de hokjes schrijven we de elementnummers. In de hokjes schrijven we de waarden (of inhouden) van de elementen. Een voorbeeld is te zien in figuur 5.1.

	0	1	2	3	4	5	6	7	8	9
a	2	5	3	9	8	2	6	-1	10	0

Figuur 5.1: Uitbeelding van een array met tien elementen.

Overigens zijn de blokhaken `[]` als geheel een *operator* met een hoge prioriteit.

5.2 Initialisatie

De array `a` met declaratie

```
int a[10];
```

wordt, als *globale* variabele, automatisch geïnitieerd met nullen. Als *lokale* variabele is de inhoud onbekend. We kunnen de array bij declaratie gelijk initialiseren. Om de array in figuur 5.1 direct te initialiseren gebruiken we een rij getallen tussen accolades en gescheiden door komma's:

```
int a[10] = {2, 5, 3, 9, 8, 2, 6, -1, 10, 0};
```

Omdat de compiler de lengte van de lijst kan uitrekenen, mogen we het aantal element bij de declaratie weglaten. We kunnen dus ook schrijven:

```
int a[] = {2, 5, 3, 9, 8, 2, 6, -1, 10, 0};
```

¹ Technisch gezien zijn de elementnummers wel geldig. De compiler berekent via het elementnummer de plaats in het geheugen waar het element zich bevindt. Door gebruik te maken van *pointers* kan op zinnige wijze gebruik gemaakt worden van negatieve elementnummers.

Als de lijst korter is dan het aantal opgegeven elementen, dan worden de overige elementen met nullen geïnitieerd:

```
int a[10] = {2, 5 ,3 ,9, 8};    /* the rest are zero */
```

Willen we een lokale array met nullen initialiseren, dan kunnen we volstaan met één nul:

```
int a[10] = {0};    /* all elements are zero */
```

Een *constante array* moet altijd geïnitieerd worden en de elementen mogen niet gewijzigd worden:

```
const int a[] = {2, 5 ,3 ,9, 8, 2, 6, -1, 10, 0};
```

5.3 Eendimensionale array

De array *a* wordt na de declaratie

```
int a[10];
```

een *eendimensionale* array genoemd. Zoals al eerder is geschreven, mogen we het elementnummer van een array met een expressie uitrekenen, op voorwaarde dat het uitkomst een geheel getal is en binnen de grenzen van van array ligt. In listing 5.1 is een programma te zien dat een array vult met kwadraten van 0 t/m 9. We gebruiken *i* als lusvariabele en berekenen voor elk element het kwadraat. Daarna bepalen we de som van twee kwadraten door de juiste elementen uit de array bij elkaar op te tellen. Vervolgens drukken we de gegevens op het scherm af.

```
1 #include <stdio.h>
2
3 int main(void) {
4
5     int kwad[10], i;
6     int a = 3, b = 5, som;
7
8     for (i = 0; i < 10; i = i + 1) {
9         kwad[i] = i * i;
10    }
11
12    som = kwad[a] + kwad[b];
13
14    printf("De_som_van_de_kwadraten_%d_en_%d_is_%d\n", a, b, som);
15
16    return 0;
17 }
```

Listing 5.1: Afdrukken van de som van twee kwadraten kwadraten.

Een interessant geval is dat we de inhoud van een element van een array mogen gebruiken als een elementnummer van een andere array. Dit is te zien in listing 5.2. We declareren een array *rij* en initialiseren de array met de getallen waarvan we de kwadraten willen

optellen. Daarna vullen we de array `kwad` met de kwadraten zoals we dat al eerder deden. In de regels 13 t/m 15 worden een voor een de waarden uit array `rij` opgevraagd en gebruikt als elementnummer voor array `kwad`. Merk op dat `rij[i]` eerst wordt bepaald en de uitkomst daarvan wordt gebruikt om een element uit de array `kwad` te selecteren.

```
1 #include <stdio.h>
2
3 int main(void) {
4
5     int kwad[10], i, som = 0;
6
7     int rij[5] = { 2, 5, 8, 1, 6 };
8
9     for (i = 0; i < 10; i = i + 1) {
10         kwad[i] = i * i;
11     }
12
13     for (i = 0; i < 5; i = i + 1) {
14         som = som + kwad[rij[i]];
15     }
16
17     printf("De_som_van_de_kwadraten_is_%d\n", som);
18
19     return 0;
20 }
```

Listing 5.2: Afdrukken van de som van twee kwadraten kwadraten.

5.4 Aantal elementen van een array bepalen

We kunnen bij de declaratie van een array het aantal elementen opgeven, maar dat hoeft niet als de array gelijk geïnitieerd wordt. De declaratie

```
int a[] = {2, 5 ,3 ,9, 8, 2, 6, -1, 10, 0};
```

zorgt ervoor dat `a` uit 10 elementen bestaat. In een programma moeten we erop toezien dat we niet buiten de array komen. Als we de initialisatie groter of kleiner maken, dan verandert het aantal elementen en moet het programma hierop aangepast worden. Met behulp van de operator `sizeof` kunnen we het aantal elementen *tijdens compile time* uitrekenen. Let erop dat `sizeof` de grootte *in bytes* uitrekent. We moeten dus de grootte van de array delen door de grootte van één element. Het aantal elementen van een array wordt berekend met:

```
int siz = sizeof a / sizeof a[0];
```

Merk op dat `sizeof` een hogere prioriteit heeft dan `/` maar lager dan `[]`. We hoeven dus geen haakjes te zetten om `a[0]`. We kunnen het getal 5 regel 13 in listing 5.2 vervangen door een berekening van het aantal elementen. Dit is te zien in listing 5.3.


```

1     ...
2     for (i = 0; i < sizeof rij / sizeof rij[0]; i = i + 1) {
3         som = som + kwad[rij[i]];
4     }
5     ...

```

Listing 5.3: Bepalen van het aantal elementen in een array.

5.5 Tweedimensionale array

Een *tweedimensionale* array van 5×3 elementen wordt gedeclareerd met:

```
int matrix[5][3];
```

Het eerste elementnummer wordt het *rijnummer* genoemd en het tweede elementnummer wordt het *kolomnummer* genoemd. Het aantal rijen en kolommen mag natuurlijk gelijk zijn.

We kunnen de array tegelijk met de declaratie initialiseren. We gebruiken een dubbele array-initialisatie met accolades. De buitenste accolades geven de afbakening van de initialisatie aan. Binnen de accolades vormen we drie groepen die zijn afgebakend door accolades en gescheiden zijn door komma's.

```
int matrix[3][3] = { {1, 2, 3}, {4, 5, 6}, {7, 8, 9} };
```

We mogen hierbij het aantal rijen weglaten. De compiler berekent die automatisch. Als we een element uit de array willen gebruiken dan geven we eerst het rijnummer op en daarna het kolomnummer:

```
int i = matrix[2][1];    /* row 2, column 1 */
```

Een veelvoorkomende fout is om het rijnummer en kolomnummer te scheiden door een komma. Dat kan niet want de komma werkt als de komma-operator:

```
int i = matrix[2,1];    /* WRONG */
```

Let erop dat het verwisselen van het rijnummer en kolomnummer een ander element selecteert. Dus

```
matrix[2,1] en matrix[1,2]
```

zijn twee verschillende elementen.

We kunnen het aantal rijen en kolommen van een array bepalen met behulp van `sizeof`. Dit is te zien in listing 5.4. In regel 7 bepalen we het aantal rijen door de grootte van de gehele array te delen door de grootte van een rij. Dus

```
int rows = sizeof twodim / sizeof twodim[0];
```

kent het aantal rijen toe aan `rows`. In regel 8 bepalen we het aantal kolommen door de grootte van een rij te delen door de grootte van één element. Dus

```
int cols = sizeof twodim[0] / sizeof twodim[0][0];
```

kent het aantal kolommen toe aan `cols`. Daarna drukken we de twee variabelen af.

```
1 #include <stdio.h>
2
3 int main(void) {
4
5     int twodim[15][20];
6
7     int rows = sizeof twodim / sizeof twodim[0];
8     int cols = sizeof twodim[0] / sizeof twodim[0][0];
9
10    printf("Rows:_%d\n", rows);
11    printf("Cols:_%d\n", cols);
12
13    return 0;
14 }
```

Listing 5.4: Bepalen van het aantal rijen en kolommen.

5.6 Afdrukken van array

Een array kan niet als één eenheid worden afgedrukt. We moeten dat doen met behulp van een lus. Dit is te zien in listing 5.5. In regel 5 declareren en initialiseren we de array met in dit geval 12 elementen. Omdat de array kan wijzigen berekenen we in regel 7 het aantal elementen waaruit de array bestaat en drukken dat af in regel 9. Daarna selecteren we een voor een de elementen uit de array. Dat doen we met een `for`-lus. Binnen de lus, in regel 12, drukken we de waarde van het element af.

```
1 #include <stdio.h>
2
3 int main(void) {
4
5     int rij[] = { 2, 5, 8, 1, 6, 7, 3, 9, 0, 4, 7, 8}, i;
6
7     int len = sizeof rij / sizeof rij[0];
8
9     printf("De_array_heeft_%d_elementen:", len);
10
11    for (i = 0; i < len; i = i + 1) {
12        printf("%d%c", rij[i], i < len-1 ? '_' : '\n');
13    }
14
15    return 0;
16 }
```

Listing 5.5: Afdrukken van een array.

We zullen de aanroep van de `printf`-functie even uitleggen. Het eerste argument is de format string waarin twee format specifications staan: `%d` en `%c`. De eerste geeft aan dat een integer moet worden afgedrukt en de tweede geeft aan dat één karakter moet worden afgedrukt. Het tweede argument is `rij[i]`, het element uit de array. Het derde argument wordt gevormd door een *conditionele expressie*:

```
i < len-1 ? ' ' : '\n'
```

Dit werkt als volgt. Voor het vraagteken staat de relationele expressie `i < len-1`. Als de expressie waar is, dan wordt een *spatie* afgedrukt. Als de expressie onwaar is dan wordt een *newline* afgedrukt. Het gevolg is dat na een element een spatie wordt afgedrukt, behalve na het laatste element, dan wordt een *newline* afgedrukt. De conditionele expressie wordt besproken in paragraaf 2.8.4.

5.7 Array's en functies

Een array kan als argument van een functie dienen. Een functie kan *geen* array teruggeven. Dat heeft te maken met de manier waarop array's overgedragen worden. Bij enkelvoudige datatypes wordt een *kopie* gemaakt van de originele variabelen en die kopieën wordt doorgegeven. Bij overdragen van een array wordt geen kopie gemaakt, maar wordt de plaats in geheugen waar de array begint overgedragen². Omdat niet een complete kopie wordt gemaakt, kan de functie de grootte van de array niet uitrekenen; het kent immers de declaratie niet. Er moet dus altijd een parameter met de grootte van de array worden overgedragen.

In listing 5.6 is een functie te zien die de positie bepaalt van een getal dat in een array voorkomt. De waarde `-1` wordt teruggegeven als het getal niet in de array voorkomt. De aanroepende functie moet hierop testen. Als het getal meerdere keren voorkomt, wordt de eerste positie teruggegeven. Om aan te geven dat de eerste parameter een array is, wordt de declaratie `int ary[]` gegeven. De tweede parameter is de grootte van de array in elementen en de derde parameters is het getal dat gezocht wordt.

```
1 int findnumber(int ary[], int siz, int num) {
2     int pos;
3
4     for (pos = 0; pos < siz; pos = pos + 1) {
5         if (ary[pos] == num) {
6             return pos;
7         }
8     }
9     return -1;
10 }
```

Listing 5.6: Functie voor het vinden van een getal in een array.

² Dit is gedaan uit efficiëntieoverwegingen. Stel dat een array uit 10000 elementen bestaat, dan zou er een kopie moeten worden gemaakt van al die elementen. Niet alleen de geheugenruimte is een probleem, ook de snelheid waarmee de kopie wordt gemaakt vormt een obstakel. En wat te denken als een functie weer een functie aanroept met dezelfde array. Dan moet er weer een kopie gemaakt worden.

In listing 5.7 is te zien hoe de functie wordt aangeroepen. Zie regel 9. Als eerste argument wordt de *naam* van de array opgegeven. De compiler weet aan de hand van de functiedefinitie en de declaratie van de array dat het om een array gaat. Het tweede argument is het aantal elementen van de array dat berekend wordt met behulp van `sizeof`. Het derde argument is het getal waarnaar gezocht wordt.

```
1 #include <stdio.h>
2
3 int findnumber(int ary[], int siz, int num);
4
5 int main(void) {
6
7     int rij[] = { 4, 8, 3, 8, 2, 3, 4, 5, 1, 2 };
8
9     int plek = findnumber(rij, sizeof rij / sizeof rij[0], 5);
10
11     printf("Plek_%d\n", plek);
12
13     return 0;
14 }
```

Listing 5.7: Aanroepen van de functie *findnumber*.

Omdat de plaats van de array in het geheugen als argument wordt meegegeven, kunnen we de array *in de functie* wijzigen. De functie `patcharray` in listing 5.8 vervangt alle elementen die gelijk zijn aan parameter `what` met de parameter `with`. De functie geeft aan het einde het aantal vervangingen terug. Omdat er geen kopie wordt meegegeven maar het adres van het eerste element, kan de functie bij de originele array komen en zodoende de elementen veranderen.

```
1 int patcharray(int ary[], int siz, int what, int with) {
2
3     int pos;
4     int count = 0;
5
6     for (pos = 0; pos < siz; pos = pos + 1) {
7         if (ary[pos] == what) {
8             ary[pos] = with;
9             count = count + 1;
10        }
11    }
12    return count;
13 }
```

Listing 5.8: Functie om een getal te vervangen door een ander getal.

Doorgewinterde C-programmeurs zouden het trouwens zo willen oplossen (listing 5.9):

```

1 int patcharrayunreadable(int ary[], int siz, int what, int with) {
2     int count = 0;
3
4     while (--siz > 0) {
5         ary[siz] == what ? ary[siz] = with, count++ : 0;
6     }
7     return count;
8 }

```

Listing 5.9: *Functie om een getal te vervangen door een ander getal.*

5.8 Array's vergelijken

Een array is een samengesteld datatype en kan niet zonder meer met een andere array vergeleken worden. Het is verleidelijk om de twee namen van de array's te vergelijken zoals te zien is in listing 5.10. Door de manier waarop C met de namen omgaat zorgt ervoor dat niet de inhoud van de array's vergeleken worden, maar de adressen van het eerste element. Dit is vergelijkbaar met de array als argument van een functie.

```

1 #include <stdio.h>
2
3 int main(void) {
4
5     int ary1[10] = { 0 }, ary2[10] = { 0 };
6
7     printf("De_array's_zijn_%sgelijk\n", ary1 == ary2 ? "" : "on");
8 }

```

Listing 5.10: *Vergelijken van twee array's (foutief).*

De juiste manier om twee array's te vergelijken is om de elementen paarsgewijs te vergelijken. We doen dit met behulp van de functie `arraycompare`, te zien in listing 5.11.

```

1 int arraycompare(int ary1[], int ary2[], int siz) {
2
3     int i;
4
5     for (i = 0; i < siz; i++) {
6         if (ary1[i] != ary2[i]) {
7             return 0;
8         }
9     }
10    return 1;
11 }

```

Listing 5.11: *Vergelijken van twee array's.*

De twee array's en het aantal elementen worden als argumenten meegegeven. De array's moeten natuurlijk uit een gelijk aan elementen bestaan. Met een `for`-lus lopen we langs alle elementen van de array's. We vergelijken telkens twee elementen op dezelfde positie. Als twee elementen ongelijk zijn dan weten we dat de array's niet aan elkaar gelijk zijn en geven een 0 terug. Als alle elementen vergeleken zijn en we uit de `for`-lus komen dan weten we dat alle elementen gelijk zijn en geven we een 1 terug. In listing 5.12 is te zien hoe we de functie moeten aanroepen.

```

1 #include <stdio.h>
2
3 int arraycompare(int ary1[], int ary2[], int siz);
4
5 int main(void) {
6
7     int ary1[10] = { 0 }, ary2[10] = { 0 };
8
9     printf("De_array's_zijn_%sgelijk\n", arraycompare(ary1, ary2,
10                                     sizeof ary1 / sizeof ary1[0]) ? "" : "on");
11 }

```

Listing 5.12: Vergelijken van twee array's.

De standard library bevat een aantal functies waarmee we gemakkelijk array's kunnen vergelijken, onafhankelijk van het datatype. Dit wordt besproken in hoofdstuk 6.

5.9 Strings

Een *string* is een eendimensionale array van karakters afgesloten met een nul-karakter. In tegenstelling tot veel andere talen is een string in C dus *geen* enkelvoudig datatype. Dat betekent dat we niet op eenvoudige wijze strings kunnen manipuleren zoals het bepalen van de lengte, inkorten en veranderen. Gelukkig zijn er functies beschikbaar die het programmerwerk enigszins verlichten.

We definiëren een string door de rij karakters te beginnen en af te sluiten met aanhalingstekens. De declaratie

```
char str[] = "Ik ben een string";
```

initialiseert een string van 18 karakters. Het aantal elementen van de array wordt automatisch bepaald. In figuur 5.2 is te zien hoe de string in het geheugen ligt. Het einde van de string wordt gekenmerkt door een *nul-karakter* (`\0`). Een nul-karakter is een geheel getal, meestal een byte, waarvan alle bits 0 zijn. Er is dus altijd één karakter meer nodig dan de karakters die opgegeven zijn in de string.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
str	I	k		b	e	n		e	e	n		s	t	r	i	n	g	\0

Figuur 5.2: Uitbeelding van een stringarray.

Door middel van dit nul-karakter kan een programma het einde van de string vinden in het geheugen. Het berekenen van het aantal karakters kan heel eenvoudig met een lus uitgevoerd worden. Dit is te zien in listing 5.13. In regel 3 declareren en initialiseren we de `str` met een string constante. In de `while`-lus lopen we een voor een langs de elementen en zolang we het nul-karakter niet gevonden hebben, verhogen we een teller. Na het uitvoeren van de lus is het aantal karakters beschikbaar in variabele `len`.

```

1 #include <stdio.h>
2
3 int main() {
4
5     char str[] = "Hallo_wereld!";
6     int len = 0;
7
8     while (str[len] != '\0') {    /* while not end of string ... */
9         len = len + 1;           /* point to the next character */
10    }
11
12    printf("Lengte_is_%d\n", len);
13 }

```

Listing 5.13: Berekenen van de lengte van een string.

5.9.1 Stringfuncties

De standard library kent een groot aantal functies waarmee strings kunnen worden gemanipuleerd. Om ze te gebruiken moet het header-bestand `string.h` worden ingelezen. We hebben de vier belangrijkste functies op een rijtje gezet. Zie tabel 5.1.

Tabel 5.1: Enkele standaard functies voor stringmanipulaties.

<code>strlen(s)</code>	Geeft de lengte van <code>s</code> in karakters exclusief het nul-karakter.
<code>strcpy(to, from)</code>	Kopieert de string <code>from</code> naar string <code>to</code> .
<code>strcat(to, from)</code>	Voegt de string <code>from</code> toe aan het einde van string <code>to</code> .
<code>strcmp(s1, s2)</code>	Vergelijkt <code>s1</code> met <code>s2</code> .

De functie `strlen` geeft het aantal karakter in de string terug exclusief het nul-karakter. De functie `strcpy` kopieert de string `from` naar de string `to`. Let erop dat de ruimte van `to` groot genoeg is. De functie test dat niet en een *buffer overflow* is dan het gevolg. De functie `strcat` voegt de string `from` toe aan het einde van string `to`. In dit verband wordt vaak het Engelse woord *concatenation* gebruikt dat aaneenschakelen betekent. Ook hier moet erop gelet worden dat de ruimte in `to` groot genoeg is.

De functie `strcmp` vergelijkt `s1` met `s2` op *alfabetische ordening* met inachtneming van de karakterset van de computer. Alfabetische ordening wil min of meer zeggen: zoals de woorden in een woordenboek. Dat betekent dat de string "Jan" voorgaat op "Janus". De functie geeft `-1` terug als `s1` "kleiner" is dan `s2`, geeft `0` terug als de strings identiek

zijn en geeft 1 terug als `s1` “groter” is dan `s2`. Let erop dat de karakterset van de computer gebruikt wordt dus `"123"` is kleiner dan `"Jan"` en `"Janus"` is kleiner dan `"janus"`. Zie bijlage A over de ordening van de ASCII-karakterset.

Dat deze functies niet zo ingewikkeld zijn is te zien listing 5.14 waar we overigens gebruik hebben gemaakt van enkele juweeltjes van C-snelschrift. We dagen de lezer uit om de functies te analyseren.

```
1 int stringlength(char str[]) {
2     int len;
3
4     for (len = 0; str[len] != '\0'; len++);
5
6     return len;
7 }
8
9 void stringcopy(char to[], char from[]) {
10
11     int len;
12
13     for (len = 0; (to[len] = from[len]) != '\0'; len++);
14 }
15
16 int stringcompare(char str1[], char str2[]) {
17
18     int len;
19
20     for (len = 0; str1[len] == str2[len]; len++) {
21         if (str1[len] == '\0') {
22             return 0;
23         }
24     }
25     return str1[len] < str2[len] ? -1 : 1;
26 }
27
28 void stringconcat(char to[], char from[]) {
29     int i, j;
30
31     for (i = j = 0; to[i] != '\0'; i++);
32
33     while ((to[i++] = from[j++]) != '\0');
34 }
```

Listing 5.14: Een implementatie van de stringfuncties.

6

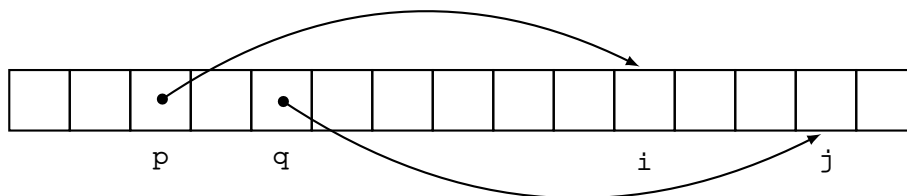
Pointers

Een pointer-variabele, of kortweg *pointer*, is een variabele waarvan de inhoud het adres is van een andere variabele. We zeggen dan ook wel dat de pointer *wijst* (Engels: “points to”) naar de andere variabele. Als een pointer naar een variabele wijst, is het mogelijk om via de pointer bij de variabele te komen.

Pointers zijn een krachtig middel om efficiënt gegevens te beheren en parameters over te dragen aan functies. Soms zijn pointers zelfs de enige manier voor het bewerken van data. Het is dan ook niet verwonderlijk dat in veel C-programma's pointers gebruikt worden.

Pointers worden samen met het `goto`-statement in verband gebracht met het schrijven van ondoorzichtige programma's. En zeker, onzorgvuldig gebruik van pointers komt de leesbaarheid en aantonen van correctheid van programma's niet ten goede. Maar met discipline kunnen programma's zeer efficiënt geschreven worden.

Een handige manier om pointers weer te geven, is door het geheugen van een computer voor te stellen als een rij vakjes. In figuur 6.1 is dat te zien. Elk vakje stelt een geheugenplaats voor¹. In de figuur zijn de variabelen *i* en *j* te zien. De twee pointers *p* en *q* wijzen respectievelijk naar variabele *i* en *j*. Dit is weergegeven met de twee pijlen. De inhoud van pointer *p* is dus het adres van variabele *i* en de inhoud van pointer *q* is het adres van variabele *j*.



Figuur 6.1: Uitbeelding van twee pointers naar variabelen in het geheugen wijzen.

¹ We gaan hier gemakshalve vanuit dat elke variabele precies één geheugenplaats in beslag neemt. In de praktijk bestaan variabelen en pointers meestal uit meer dan één geheugenplaats.

Een pointer kan wijzen naar een enkelvoudige variabele, een (element van een) array, een structure of (het begin van) een functie. Een pointer kan niet wijzen naar een constante, een uitdrukking en een register variabele.

Het is ook mogelijk om een pointer naar het datatype `void` te laten “wijzen”. Deze pointers worden generieke pointers genoemd. We zullen dit bespreken in paragraaf 6.3.

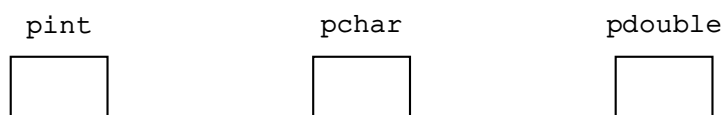
6.1 Pointers naar enkelvoudige datatypes

In listing 6.1 is de declaratie van enkele pointers van enkelvoudige datatypes te zien. Bij de declaratie moet het type variabele waarnaar de pointer wijst worden opgegeven. De asterisk (*) geeft aan dat het de declaratie van een pointer betreft.

```
1 int *pint;           /* pint is a pointer to an int */
2 char *pchar;         /* pchar is a pointer to a character */
3 double *pdouble;     /* pdouble is a pointer to a double */
```

Listing 6.1: Enkele declaraties van pointers.

We kunnen pointers uitbeelden door middel van vakjes, zoals te zien is in figuur 6.2. Elk vakje stelt een pointer voor. In beginsel hebben de pointers geen correcte inhoud. Er wordt dan wel gesproken dat de pointer nergens naar toe wijst, maar dat is feitelijk onjuist. Een pointer heeft altijd een adres als inhoud, maar het kan zijn dat de pointer niet naar een bekende variabele wijst.



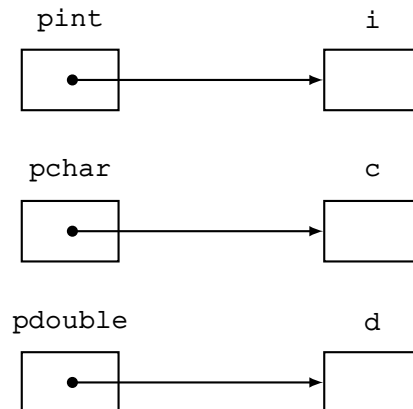
Figuur 6.2: Voorstelling van drie pointers in het geheugen.

Aan een pointer is het adres van een variabele van hetzelfde type toe te kennen. Hiervoor gebruiken we de *adres-operator* & (ampersand). In listing 6.2 is een aantal toekenningen van adressen te zien.

```
1 int i;               /* the variables */
2 char c;
3 double d;
4
5 int *pint;           /* the pointers */
6 char *pchar;
7 double *pdouble;
8
9 pint = &i;           /* pint points to variable i */
10 pchar = &c;          /* pchar points to variable c */
11 pdouble = &d;        /* pdouble points to variable d */
```

Listing 6.2: Enkele toekenningen van adressen aan pointers.

Nu de pointers geïnitieerd zijn, kunnen we een voorstelling maken van de relatie tussen de pointers en de variabelen. Dit is te zien in figuur 6.3. Pointer `pint` wijst naar variabele `i`, pointer `pchar` wijst naar variabele `c` en pointer `pdouble` wijst naar variabele `d`.



Figuur 6.3: Voorstelling van drie pointers die naar variabelen wijzen.

Via de pointers kunnen we de variabelen gebruiken. Stel dat we variabele `i` met één willen verhogen. Dat kunnen we doen door gebruik te maken van de *indirectie* of *dereferentie* operator `*`. Deze operator heeft voorrang op de rekenkundige operatoren.

```
1 *pint = *pint + 1;           /* increment i by one */
2 *pchar = *pchar + 1;        /* next character in ASCII table */
3 *pdouble = *pdouble + 1.0;   /* add 1.0 to double */
```

Listing 6.3: Gebruik van een pointer bij een toekenning.

We mogen de dereferentie operator overal gebruiken waar een variabele gebruikt mag worden, bijvoorbeeld bij een optelling of in een test.

```
1 int i = 2, *pint;
2 ...
3 pint = &i;
4 ...
5 i = *pint + 1;    /* add 1 to variable i */
6 ...
7 if (*pint > 5) {
8     printf("Variabele_is_%d\n", *pint);
9 }
```

Listing 6.4: Gebruik van een pointer bij het afdrukken van een variabele.

Overigens kan tijdens declaratie ook gelijk de initialisatie van een pointer plaatsvinden. Deze declaratie kan verwarrend zijn. In onderstaande listing wordt de pointer `pint` gedeclareerd en geïnitieerd met het adres van variabele `i`. Het betreft hier dus *geen* dereferentie.

```

1 int i = 2;
2 int *pint = &i;      /* declare and initialize pint */

```

Listing 6.5: Declaratie en initialisatie van een pointer.

6.2 De NULL-pointer

In principe wijst een pointer naar een variabele (of beter: naar een geheugenadres). Om aan te geven dat een pointer niet naar een variabele wijst, kunnen we de *NULL-pointer* gebruiken. Let erop dat de NULL-pointer niet hetzelfde is als een niet-geïnitieerde pointer. Een NULL-pointer is een pointer waarin alle bits 0 zijn². Een niet-geïnitieerde pointer heeft een willekeurige waarde³. In C is een preprocessor-macro genaamd NULL te gebruiken om een pointer als NULL-pointer te initialiseren. De C-standaard schrijft voor dat NULL wordt gedefinieerd in `locale.h`, `stddef.h`, `stdio.h`, `stdlib.h`, `string.h`, `time.h`, en `wchar.h`. Slechts een van deze header-bestanden is noodzakelijk om NULL te definiëren.

```

1 #include <stdio.h>
2
3 int *p = NULL;    /* p is initialized as NULL-pointer */

```

Listing 6.6: Declaratie en initialisatie van een NULL-pointer.

NULL-pointers kunnen *niet* gebruikt worden bij dereferentie. Dat veroorzaakt over het algemeen dat de executie van een programma wordt afgebroken.

```

1 #include <stdio.h>
2
3 int *p = NULL;    /* p is initialized as NULL-pointer */
4
5 *p = *p + 1;      /* Oops, dereference of NULL-pointer! */

```

Listing 6.7: Dereferentie van een NULL-pointer.

In een vergelijking zullen twee NULL-pointers altijd `true` opleveren.

```

1 int *p = NULL, *q = NULL;
2 ...
3 if (p == q) { /* true */ }

```

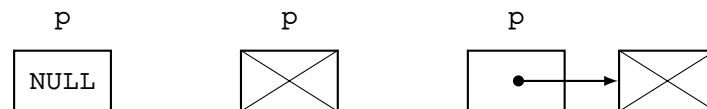
Listing 6.8: Vergelijken van twee NULL-pointers.

² In het algemeen is de inhoud van een NULL-pointer het getal 0, maar dat is niet in alle gevallen zo. De C-standaard definieert de NULL-pointer als een pointer die niet naar een bekende variabele wijst.

³ In geval een pointer als globale variabele wordt gedeclareerd, zorgt de compiler ervoor dat de inhoud op 0 gezet wordt.

De NULL-pointer wordt door diverse standaard functies gebruikt om aan te geven dat er een fout is geconstateerd. Zo geeft de functie `fopen` de waarde `NULL` terug als het niet gelukt is om een bestand te openen. De functie `malloc` geeft de waarde `NULL` terug als het niet gelukt is om een stuk geheugen te alloceren.

Bij het voorstellen van NULL-pointers zijn diverse mogelijkheden die gebruikt worden. Bij de linker voorstelling wordt het woord `NULL` in een vakje gezet, bij de middelste voorstelling wordt een kruis in het vakje gezet en bij de rechter voorstelling wordt een pijl getrokken naar een vakje met een kruis erin.



Figuur 6.4: Drie voorstellingen van *NULL*-pointers.

6.3 Pointer naar `void`

De void-pointer, ook wel *generieke pointer* genoemd, is een speciaal type pointer die naar elk type variabele kan wijzen. Een void-pointer wordt net als een gewone pointer gedeclareerd middels het keyword `void`. Toekenning aan een void-pointer gebeurt met de adres-operator `&`.

```
1 int i;  
2 char c;  
3 double d;  
4  
5 void *p; /* void-pointer */  
6  
7 p = &i; /* valid */  
8 p = &c; /* valid */  
9 p = &d; /* valid */
```

Listing 6.9: Declaratie en initialisatie van een void-pointer.

Omdat het type van een void-pointer niet bekend is, kan een void-pointer niet zonder meer in een dereferentie gebruikt worden. De void-pointer moet expliciet gecast worden naar het correcte type. In de onderstaande listing gebruiken we pointer `p` om naar een `int` te wijzen. De type cast `(int *)` zorgt ervoor dat pointer `p` naar een `int` wijst. Door gebruik te maken van de dereferentie operator `*` kunnen we bij de inhoud van variabele `i`. De constructie `*(int *)` is dus een expliciete dereferentie naar een integer.

6.4 Afdrukken van pointers

Het afdrukken van de waarde van een pointer kan met de `printf`-functie en de format specifier `%p`. Merk op dat de pointer van het type `void` moet zijn, maar veel compilers accepteren pointers naar een datatype.

```

1 int i = 2;
2 void *p = &i;                                /* void-pointer */
3
4 ...
5 *(int *) p = *(int *) p + 1;    /* explicit type cast */
6
7 ...
8 printf("De_waarde_is_%d\n", *(int *) p);

```

Listing 6.10: Declaratie, initialisatie en deference van een void-pointer.

```

1 #include <stdio.h>
2
3 int main() {
4
5     int i = 2, *p = &i;
6
7     printf("Pointer:_%p\n", (void *) p);
8
9     return 0;
10 }

```

Listing 6.11: Afdrukken van een pointer.

Noot: Bij 32-bits compilers is de grootte van een pointer 32 bits (4 bytes). Zo'n pointer kan maximaal 4 GB adresseren. Bij 64-bits compilers is de grootte van een pointer 64 bits (8 bytes). Zo'n pointer kan maximaal 16 EB (exa-bytes) adresseren.

6.5 Pointers naar array's

We kunnen een pointer ook laten wijzen naar het eerste element van een array. Dit is te zien in listing 6.12. De array bestaat uit negen elementen van het type `int`. De pointer `p` laten we wijzen naar het eerste element van de array. We gebruiken hiervoor de adres-operator `&` en elementnummer 0.

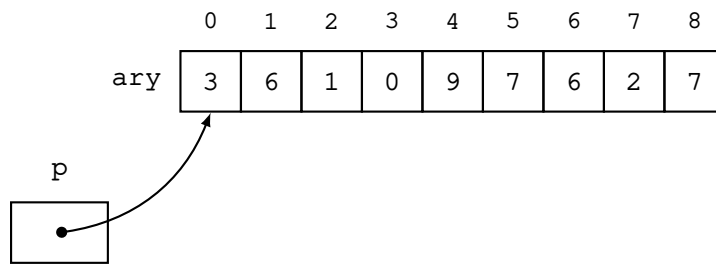
```

1 int ary[] = {3,6,1,0,9,7,6,2,7};
2
3 int *p = &ary[0]; /* p points to first element for array */

```

Listing 6.12: Een pointer naar het eerste element van een array.

De uitbeelding hiervan is te zien in figuur 6.5. Omdat deze toekenning zeer vaak in een C-programma voorkomt, is er een verkorte notatie mogelijk. We kunnen in plaats van `&ary[0]` ook de naam van de array gebruiken: *de naam van een array is een synoniem voor een adres van het eerste element van de array*. Zie listing 6.13.



Figuur 6.5: Uitbeelding van een pointer naar het eerste element van een array.

```

1 int ary[] = {3,6,1,0,9,7,6,2,7};
2
3 int *p = ary; /* p points to first element of array */

```

Listing 6.13: Een pointer naar het eerste element van een array.

Omdat de naam van een array een synoniem is, mag het dus niet gebruikt worden aan de linkerkant van een toekenning.

```

1 int *p, ary[] = {3,6,1,0,9,7,6,2,7};
2
3 p = ary;      /* correct use of pointer and array name */
4 ary[2] = *p; /* correct use of pointer and array element */
5
6 ary = p;      /* ERROR: array name cannot be used in this context */

```

Listing 6.14: Een pointer naar het eerste element van een array.

Het is ook mogelijk om een pointer naar een ander element van een array te laten wijzen. De compiler test niet of de toekenning binnen de array-grenzen ligt.

```

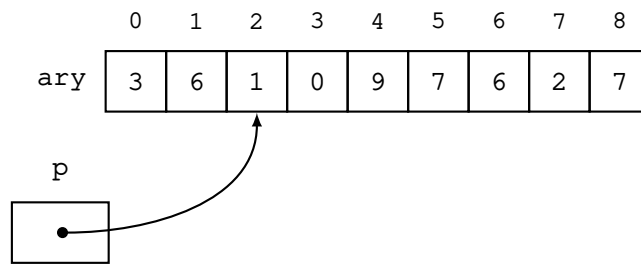
1 int ary[] = {3,6,1,0,9,7,6,2,7};
2
3 int *p = &ary[2]; /* p points to third element of array */

```

Listing 6.15: Een pointer naar het derde element van een array.

Een uitbeelding is te zien in figuur 6.6. In de figuur wijst p naar het derde element van ary. We kunnen nu de inhoud van dit element opvragen door ary[2] en door *p. Het is zelfs mogelijk om p als de naam van een array te beschouwen. Zie paragraaf 6.8.

We kunnen de lengte van ary bereken met de sizeof-operator. Dat kan alleen via ary omdat de compiler de lengte kan uitrekenen. Het kan *niet* via pointer p want dat is een pointer naar een int; pointer p “weet” niet dat er naar een array gewezen wordt.



Figuur 6.6: Uitbeelding van een pointer naar het derde element van een array.

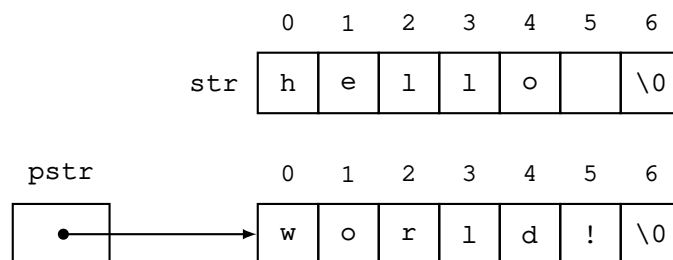
6.6 Strings

Een string in C is niets anders dan een array van karakters, afgesloten met een nul-karakter ('`\0`'). Er is dus altijd één geheugenplaats meer nodig dan het aantal karakters in de string. Een nul-karakter is niet hetzelfde als een NULL-pointer. Een nul-karakter is een byte met de inhoud 0 (alle bits zijn 0), een NULL-pointer is een pointer met de inhoud 0. in listing 6.16 zijn twee declaraties met strings te zien, een echte array met en string als inhoud en een pointer naar een string in het geheugen.

```
1 char str[] = "Hello_";
2 char *pstr = "world!";
```

Listing 6.16: Declaratie en initialisatie van twee C-strings.

Merk op dat `str` niet aangepast mag worden, want dit is de naam van een array. Pointer `pstr` mag wel aangepast worden want `pstr` is een pointer naar het eerste element van de array. Een voorstelling van beide strings is te zien in figuur 6.7.



Figuur 6.7: Uitbeelding van twee C-strings.

Ook hier merken we op dat we de lengte van `str` kunnen berekenen met de `sizeof`-operator. De compiler heeft genoeg informatie beschikbaar. We kunnen de lengte van de tweede string *niet* door de compiler laten uitrekenen, want `pstr` is een pointer naar een `char`. Pointer `pstr` “weet” dus niet dat er naar een string gewezen wordt.

Toch is het mogelijk om tijdens het draaien van een programma de lengte van de string te vinden. We kunnen namelijk uitgaan van het feit dat een string in ‘C’ wordt afgesloten met een nul-karakter. Dit wordt uitgelegd in de volgende paragraaf.

6.7 Rekenen met pointers

Pointers kunnen rekenkundig worden aangepast dat vooral nuttig is bij het gebruik van array's. In het onderstaande programma wijst pointer `p` in eerste instantie naar het begin van de array `ary` (dus `ary[0]`). Daarna wordt `p` twee maal met 1 verhoogd en daarna met 3 verhoogd. Bij rekenkundige operaties op pointers wordt rekening gehouden met de grootte van de datatypes. De grootte van een `int` is in de regel vier bytes. Door de pointer met 1 te verhogen wordt dus naar de volgende `int` gewezen.

```
1 int ary[] = {3,6,1,0,9,7,6,2,7};
2 int *p = ary; /* p pointe to ary[0] */
3
4 p = p + 1;     /* p points to ary[1] */
5 ...
6 p = p + 1;     /* p points to ary[2] */
7 ...
8 p = p + 3;     /* p points to ary[5] */
```

Listing 6.17: Rekenen met pointers.

Een mooi voorbeeld van het rekenen met pointers is het bepalen van de lengte van een C-string. In listing 6.18 wordt pointer `str` gedeclareerd en wijst naar het begin van de string. Pointer `begin` wijst ook naar het begin van de string. Daarna verhogen we pointer `str` totdat het einde van de string is bereikt. Daarna drukken we het verschil van de twee pointers af.

```
1 #include <stdio.h>
2
3 int main() {
4
5     char *str = "Hallo_wereld!";
6     char *begin = str;
7
8     while (*str != '\0') { /* while not end of string ... */
9         str = str + 1;     /* point to the next character */
10    }
11
12    printf("Lengte_is_%d\n", str-begin);
13 }
```

Listing 6.18: Berekenen van de lengte van een string met behulp van pointers.

Let erop dat de twee pointers naar elementen in dezelfde array moeten wijzen (of één na het laatste element). Alleen dan levert de aftrekking `str-begin` een gedefinieerd resultaat. De aftrekking is van het type `ptrdiff_t` (dat meestal gelijk is aan een `int`) en levert het verschil in elementen. Het onderstaande programmafragment geeft als uitvoer de waarde 3.

Vergelijken van twee pointers kan ook. Zo kunnen pointers op gelijkheid worden vergeleken, maar ongelijkheid kan ook. We zouden de `printf`-regel van listing 6.18 kunnen vervan-

```

1 int ary[] = {3,6,1,0,9,7,6,2,7};
2 int *p = &ary[2];
3 int *q = &ary[5];
4
5 printf("Verschil_is_%d\n", q-p);

```

Listing 6.19: *Het berekenen van het verschil van twee pointers.*

gen door de onderstaande programmafragment. Uiteraard moeten de twee pointers naar hetzelfde datatype wijzen en heeft de vergelijking alleen zin als de pointers naar elementen binnen dezelfde array wijzen.

```

1     if (str>begin) {
2         printf("Lengte_is_%d\n", str-begin);
3     } else {
4         printf("De_string_is_leeg\n");
5     }

```

Listing 6.20: *Vergelijken van twee pointers.*

6.8 Relatie tussen pointers en array's

De relatie tussen pointers en array's zijn zo sterk in 'C' verankerd, dat we er een aparte paragraaf aan wijden. In listing 6.21 zijn de declaratie en initialisatie van een array en een pointer te zien. De pointer `p` wijst na initialisatie naar het eerste element van de array.

```

1 int ary[] = {3,6,1,0,9,7,6,2,7};
2 int *p = ary;

```

Listing 6.21: *Declaratie en initialisatie van een array en een pointer.*

Om toegang te krijgen tot eerste element uit de array kunnen natuurlijk `ary[0]` gebruiken. Maar we kunnen via `p` ook bij het eerste element komen. Hiervoor gebruiken we `*p`. We mogen echt `p` ook lezen als de naam van een array. Om het eerste element te komen, mogen we dus ook `p[0]` gebruiken. Om alle elementen van de array bij elkaar op te tellen, kunnen we dus schrijven:

```

1 int ary[] = {3,6,1,0,9,7,6,2,7};
2 int *p = ary;
3
4 int sum = p[0]+p[1]+p[2]+p[3]+p[4]+p[5]+p[6]+p[7]+p[8];

```

Listing 6.22: *Bepalen sum van elementen in een array.*

Aan de andere kant mogen we de naam van de array ook lezen als een pointer naar het eerste element. We kunnen de naam gebruiken in een dereference. Dat betekent dat `*ary` identiek is aan `ary[0]` en dat `*(ary+2)` identiek is aan `ary[2]`. We hebben echter wel haken nodig bij `*(ary+2)` omdat de dereferentie-operator voorgaat op de optelling. Om de som van de array te bepalen mogen we dus schrijven:

```
1 int ary[] = {3,6,1,0,9,7,6,2,7};
2 int *p = ary;
3
4 int sum = *ary + *(ary+1) + *(ary+2) + *(ary+3) + ... ;
```

Listing 6.23: *Bepalen sum van elementen in een array.*

Het is mogelijk om een pointer naar een willekeurig element van de array te laten wijzen door de naam van de array als pointer te beschouwen. Als we `p` willen laten wijzen naar het derde element gebruiken we gewoon de toekenning `p = ary+2`. Na deze toekenning kunnen het derde element afdrukken door `p` als pointer of als array te beschouwen. Zie listing 6.24.

```
1 int ary[] = {3,6,1,0,9,7,6,2,7};
2
3 int *p = ary+2;                /* p points to third element */
4
5 printf("Contents_is_%d\n", *p); /* prints third element */
6 printf("Contents_is_%d\n", p[0]); /* prints third element */
```

Listing 6.24: *Pointer die naar een element in een array wijst.*

Let erop dat we in het bovenstaande programmafragment `p[0]` hebben gebruikt. De pointer wijst naar het derde element dus `p[0]` betekent dat de inhoud van het derde element wordt afgedrukt.

Als we zeker weten dat we een correct element gebruiken, mogen we ook negatieve waarden voor het elementnummer gebruiken. In listing 6.25 wordt het adres van het derde element uit de array toegekend aan pointer `p`. We mogen dan `p[-1]` gebruiken omdat dit het tweede element uit de array betreft. We kunnen echter niet `ary[-1]` gebruiken want dit leidt tot het gebruik van een element buiten de array. Let erop dat C-compilers over het algemeen niet testen of een adressering binnen de array-grenzen ligt.

6.9 Pointers als functie-argumenten

Net als “gewone” variabelen, kunnen ook pointers als argumenten bij het aanroepen van een functie gebruikt worden. Let erop dat een kopie van de pointers worden meegegeven. Via die kopie kunnen we bij de variabelen komen waar de pointers naartoe wijzen. We kunnen dus niet de pointers zelf aanpassen.

In listing 6.26 is te zien hoe we een functie `swap` definiëren die de inhoud van twee variabelen verwisseld. Bij het aanroepen van de functie geven we de adressen mee van

```

1 int ary[] = {3,6,1,0,9,7,6,2,7};
2
3 int *p = ary+2;    /* p points to third element */
4
5 int a = p[-1];     /* legal: points to second element */
6 int b = ary[-1];   /* ILLEGAL: points outside array */

```

Listing 6.25: Pointer die naar een element in een array wijst.

de te verwisselen variabelen. In de functie gebruiken we pointers om de inhoud te verwisselen.

```

1 void swap(int *pa, int *pb) {
2     int temp;
3
4     temp = *pa;
5     *pa = *pb;
6     *pb = temp;
7 }

```

Listing 6.26: Het verwisselen van twee variabelen met behulp van pointers.

Bij het aanroepen van de functie geven we de adressen van de variabelen mee. Dit is te zien in listing 6.27. Deze manier van argumentenoverdracht wordt *Call by Reference* genoemd.

```

1 void swap(int *pa, int *pb);    /* prototype of function swap */
2
3 int main(void) {
4
5     int a=2, b=3;               /* declare variables */
6
7     swap(&a, &b);               /* swap variables */
8
9     return 0;
10 }

```

Listing 6.27: Aanroep van de functie.

Een typisch voorbeeld is een functie die een string kopieert naar een andere string. De functie krijgt twee pointers naar strings als argumenten mee. Bij het aanroepen van de functie worden de namen van de twee strings als argumenten meegegeven. De naam van een string is immers een pointer naar het eerste karakter van de string. Het programma is te zien in listing 6.28.

Merk op dat de geheugenruimte voor de kopie groot genoeg moet zijn om de kopie op te slaan en dat de twee string elkaar in het geheugen niet mogen overlappen. De standaard

```

1 void string_copy(char *to, char *from) {
2
3     if (from == NULL) {          /* sanity check */
4         return;
5     }
6
7     while (*from != '\0') {      /* while not end of string ... */
8         *to = *from;             /* copy character */
9         to = to + 1;             /* point to the next character */
10        from = from + 1;
11    }
12    *to = '\0';                  /* terminate string */
13 }
14
15 int main() {
16
17     char stra[] = "Hello_world!";
18     char strb[100];
19
20     string_copy(strb, stra);     /* copy stra to strb */
21
22     return 0;
23 }
24

```

Listing 6.28: Functie voor het kopiëren van een string.

C-bibliotheek heeft een functie `strcpy` die een efficiënte implementatie is van het kopiëren van strings.

Noot: een array kan alleen maar via een pointer als argument aan een functie worden doorgegeven. Dit is veel efficiënter dan de hele array mee te geven. In listing 6.28 wordt dus *niet* de hele array meegegeven, maar alleen de pointers naar de eerste elementen. We mogen daarom de pointers `to` en `from` ook als namen van array's beschouwen. Zie listing 6.29. Merk op dat we dus de lengte van de meegegeven array *niet* kunnen uitrekenen met de `sizeof`-operator. Er wordt immers een pointer meegegeven. Dat we toch het einde van een string kunnen bepalen, komt doordat een string wordt afgesloten met een nul-byte.

Bij het overdragen van een array aan een functie moet expliciet de grootte worden opgegeven. Het is niet mogelijk om *in de functie* de grootte van de array uit te rekenen, er wordt immers een pointer meegegeven. Er is dus een extra parameter nodig waarmee we de grootte opgeven. Bij het aanroepen van de functie rekenen we de grootte uit en geven dit mee. Dit is te zien in listing 6.30. Let erop dat in `main` wél de grootte van de array kan worden uitgerekend, want daar wordt de array gedeclareerd. We gebruiken twee keer de `sizeof`-operator, want `sizeof` geeft de grootte van een object in bytes. We moeten de grootte van de array in bytes delen door de grootte van één element in bytes.⁵

```

1 void string_copy(char to[], char from[]) {
2
3     int i=0;
4
5     while (from[i] != '\0') { /* while not end of string */
6         to[i] = from[i];      /* ... copy character */
7         i = i + 1;           /* point to next character */
8     }
9     to[i] = '\0';            /* .. and terminate string */
10 }

```

Listing 6.29: Functie voor het kopiëren van een string.

```

1 #include <stdio.h>
2 #include <stdlib.h>
3
4 void printarray(int ary[], int len) {
5
6     int i;
7
8     for (i=0; i<len; i++) {
9         printf("%d_", ary[i]);
10    }
11 }
12
13 int main(void) {
14
15     int list[] = {1,2,3,4,5,6,7,8,9};
16     int length = sizeof(list)/sizeof(list[0]);
17
18     printarray(list, length);
19
20     return 0;
21 }

```

Listing 6.30: Meegeven van de grootte van een array.

6.10 Pointer als return-waarde

Een pointer kan ook als return-waarde dienen. In het onderstaande voorbeeld wordt in een string gezocht naar een bepaalde karakter in een string. Als het karakter gevonden is, wordt een pointer naar het karakter teruggegeven. Als het karakter niet wordt gevonden wordt NULL teruggegeven. Na het uitvoeren van de functie moet hier op getest worden. Tevens wordt in het begin getest of de pointer naar de string wel een geldige waarde heeft. Geldig wil zeggen dat de pointer niet NULL is. Het is *good practice* om altijd te testen of een pointer NULL is.

Let goed op de definitie van de functie `find_token`. Voor de functienaam is de dereferentie-

```

1 char *find_token(char *str, char ch) {
2
3     if (str == NULL) {          /* sanity check */
4         return NULL;
5     }
6
7     while (*str != '\0') {      /* while not end of string */
8         if (*str == ch) {       /* if character found ... */
9             return str;         /* return pointer */
10        }
11        str++;
12    }
13
14    return NULL;                 /* character not found */
15 }

```

Listing 6.31: Pointer als return-waarde.

operator geplaatst. Dit betekent dat de functie een pointer naar een karakter teruggeeft. Deze vorm wordt vaak verward met pointers naar functies. Zie paragraaf 6.15.

6.11 Pointers naar pointers

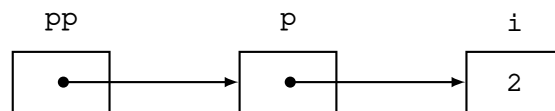
Een pointer kan ook gebruikt worden om naar een andere pointer te wijzen. In figuur 6.8 is te zien dat pointer *p* wijst naar variabele *i*. Met behulp van de dereferentie **p* kunnen we bij de inhoud van variabele *i* komen. De pointer *pp* wijst naar *p*. We hebben nu een *dubbele dereferentie* de nodig om via pointer *pp* bij de inhoud van variabele *i* te komen.

Voor de dubbele dereferentie gebruiken we twee keer de dereferentie-operator ***. Om

CALL BY REFERENCE...

In de programmeerwereld is het gebruikelijk om onderscheid te maken tussen twee manieren van argumentenoverdracht: *Call by Value* en *Call by Reference*. Bij *Call by Value* wordt een kopie van de waarde van een variabele aan de functie meegegeven. Alleen de kopie kan veranderd worden door de functie. De variabele waarvan de kopie is gemaakt kan dus niet op deze manier veranderd worden. Bij *Call by Reference* wordt het adres van de variabele aan de functie meegegeven. Via dit adres is het dus wel mogelijk om de originele variabele te veranderen.

Sommige programmeurs beweren dat *Call by Reference* eigenlijk niet bestaat. En daar is wat voor te zeggen. Er wordt immers een waarde aan de functie meegegeven en dat is het adres van een variabele. Dit adres kan in de functie veranderd worden maar het originele adres waar de variabele in het geheugen staat wordt niet aangepast. Toch maken programmeurs onderscheid tussen deze twee manieren van argumentenoverdracht.



Figuur 6.8: *Uitbeelding van een pointer naar een pointer naar een int.*

toegang te krijgen tot de inhoud van variabele `i` via pointer `pp` gebruiken we dus `**pp`. Dit is te zien in de onderstaande listing.

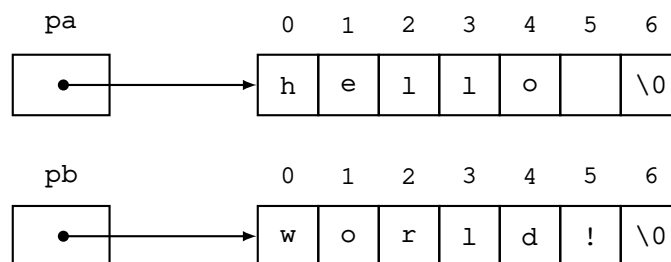
```

1 int i = 2;           /* the integer */
2 int *p = &i;        /* p points to i */
3 int **pp = &p;      /* pp points to p */
4
5 printf("De_waarde_is_%d\n", **pp);

```

Listing 6.32: *Voorbeeld van een pointer naar een pointer.*

Met behulp van pointers naar pointers kunnen de inhoud van twee pointers verwisselen. In figuur 6.9 is te zien dat de pointers `pa` en `pb` wijzen naar twee strings in het geheugen. Als we nu de strings willen “verwisselen”, hoeven we alleen maar de pointers er naartoe te verwisselen.



Figuur 6.9: *Uitbeelding van pointers naar strings.*

We kunnen dat doen met het onderstaande programmafragment. We declareren drie pointers en laten `pa` en `pb` wijzen naar strings. We gebruiken de pointer `temp` om de verwisseling tot stand te brengen.

```

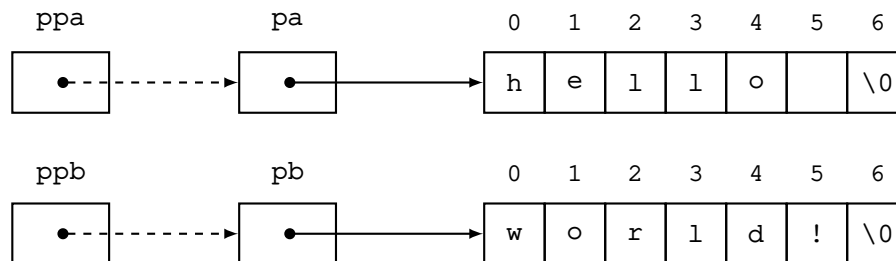
1 char *pa = "hello_";
2 char *pb = "world!";
3 char *temp;
4
5 temp = pa;           /* swap pa and pb */
6 pa = pb;
7 pb = temp;

```

Listing 6.33: *Verwisselen van twee pointers.*

Maar stel dat we zulke verwisselingen vaker in een programma moeten uitvoeren. Dan is het handig om een functie te gebruiken die dat voor ons doet. We geven aan de functie

de adressen van de pointers mee zodat de functie ze kan verwisselen. Hoe dit eruit ziet, is te zien in figuur 6.10. De pointers `pa` en `pb` wijzen naar de strings. In de functie zijn twee pointers gedefinieerd die wijzen naar pointers naar strings. Dus `ppa` wijst naar pointer `pa` en `ppb` wijst naar pointer `pb`. We kunnen nu in de functie de inhoud van `pa` en `pb` verwisselen.



Figuur 6.10: Uitbeelding van pointers naar pointers naar strings.

In listing 6.34 is de functie te zien voor het verwisselen van twee pointers. De functie heeft twee parameters `ppa` en `ppb` die een pointer zijn naar een pointer naar een string. In de functie declareren we een pointer `temp` die een pointer is naar string (of eigenlijk: een karakter). Met behulp van de dereferentie `*ppa` kopiëren we de inhoud van pointer `pa` naar `temp`. Daarna kopiëren we `pb` naar `pa` en als laatste kopiëren we `temp` naar `pb`.

```

1  #include <stdio.h>
2  #include <stdlib.h>
3
4  void swapstr(char **ppa, char **ppb) {
5
6      char *temp;
7
8      temp = *ppa;      /* copy pa into temp */
9      *ppa = *ppb;      /* copy pb into pa */
10     *ppb = temp;      /* copy temp into pb */
11 }
12
13 int main()
14 {
15     char *pa = "hello_";
16     char *pb = "world!";
17
18     printf("%s%s\n", pa, pb);
19     swapstr(&pa, &pb);
20     printf("%s%s\n", pa, pb);
21     return 0;
22 }

```

Listing 6.34: Functie voor het verwisselen van twee pointers.

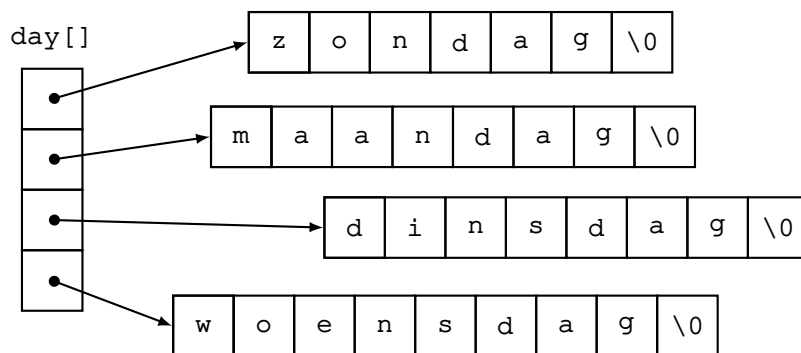
De uitvoer van dit programma is te zien in de onderstaande figuur.

C:\ Command Prompt

```
hello world!  
world!hello
```

6.12 Array van pointers

Uiteraard kunnen we ook een array van pointers maken. We demonstreren dat aan de hand van een array van pointers naar karakters. Omdat het pointers zijn, kan een pointer ook wijzen naar het begin van een array van karakters, oftewel strings. Dit is te zien in figuur 6.11. Merk op dat de vier pointers naar karakters wijzen. Dat daar toevallig vier strings aan gekoppeld zijn, is vanuit het perspectief van de pointers niet belangrijk.



Figuur 6.11: Voorstelling van een array van pointers naar strings.

We declareren een array van vier pointers naar karakters. Daarna laten we de pointers naar strings wijzen. Zie listing 6.35.

```
1 char *day[4];           /* array of four pointers to char */
2
3 day[0] = "zondag";      /* points to the 'z' */
4 day[1] = "maandag";     /* points to the 'm' */
5 day[2] = "dinsdag";     /* points to the 'd' */
6 day[3] = "woensdag";    /* points to the 'w' */
```

Listing 6.35: Een array van pointers.

Omdat dit soort toekenningen veel voorkomen, mogen de strings ook bij declaratie aan de pointers worden toegekend. Dit is te zien in listing 6.36.

```
1 char *day[4] = {"zondag", "maandag", "dinsdag", "woensdag"};
```

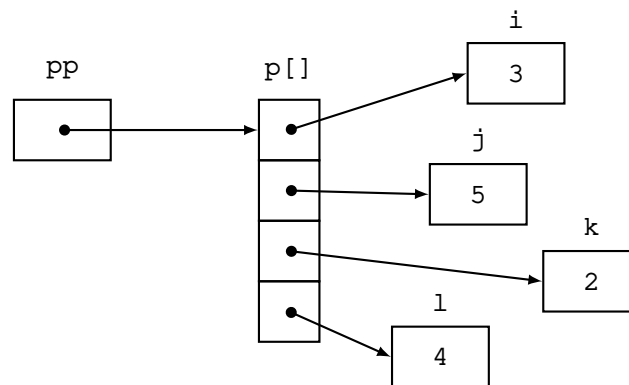
Listing 6.36: Een array van pointers naar strings met initialisatie.

Opmerking: de C++-standaard verbiedt het gebruik van dit soort declaraties en initialisaties. Veel compilers geven echter een waarschuwing en gaan gewoon verder. Omdat de strings in

dit soort constructies meestal niet veranderen, kan het keyword `const` voor de declaratie gezet worden, waarmee wordt aangegeven dat de strings niet veranderen. C++-compilers accepteren deze constructie.

6.13 Pointers naar een array van pointers

We bekijken nu een wat complexer voorbeeld van het gebruik van pointers. We declareren een array `p` van vier pointers naar integers. We vullen de array met de adressen van de integers `i`, `j`, `k` en `l`. Daarna declareren we een pointer `pp` die wijst naar (het eerste element van) de array. Een voorstelling van de variabelen is te zien in figuur 6.12.



Figuur 6.12: Voorstelling van een pointer naar een array van pointers naar integers.

De array `p` wordt gedeclareerd als:

```
int *p[4];
```

De rechte haken hebben een hogere prioriteit dan de dereferentie-operator. We lezen de declaratie dus als: `p` is een array van vier elementen en elk element is een pointer die wijst naar een integer. De pointer `pp` wordt gedeclareerd als:

```
int **pp;
```

Let goed op wat hier staat: `pp` is een pointer naar een pointer naar een integer. Vanuit pointer `pp` is niet af te leiden dat `pp` wijst naar een array, alleen maar dat er twee dereferenties nodig zijn om bij een integer te komen. We moeten dat in het C-programma zelf scherp in de gaten houden.

We gebruiken de pointers zoals te zien is in listing 6.37. In regel 6 worden de vier integers gedeclareerd. In regel 7 declareren we array `p` en initialiseren de array met de adressen van de integers. We geven geen array-grootte op want de C-compiler kan dat zelf uitrekenen.

In regel 8 declareren we de pointer `pp` (let op het gebruik van de dubbele dereferentie-operator) en initialiseren `pp` met het adres van `p`. In regel 10 drukken alle adressen van de pointers af. In de regels 12 t/m 15 drukken we de adressen van de integers en de inhoud van de array-elementen af. Om variabele `i` af te drukken hebben we drie mogelijkheden: `i` (de variabele), `*p[0]` (waar `p[0]` heen wijst) en `**pp`. Die wijst dus via dubbele dereferentie ook naar `i`.


```

1 #include <stdio.h>
2 #include <stdlib.h>
3
4 int main(void) {
5
6     int i=3, j=5, k=2, l=4;
7     int *p[] = {&i, &j, &k, &l};
8     int **pp = p;
9
10    printf("&p:_%p,_%pp:_%p,_%p[0]:_%p,_%p[1]:_%p,_%p[2]:_%p,_%p[3]:_%p\n\n", &p, &pp, &p[0], &p[1], &p[2], &p[3]);
11
12    printf("&i:_%p,_%p[0]:_%p\n", &i, p[0]);
13    printf("&j:_%p,_%p[1]:_%p\n", &j, p[1]);
14    printf("&k:_%p,_%p[2]:_%p\n", &k, p[2]);
15    printf("&l:_%p,_%p[3]:_%p\n", &l, p[3]);
16
17    printf("\ni:_%d,_%p[0]:_%d,_%pp:_%d\n", i, *p[0], **pp);
18
19    return 0;
20 }

```

Listing 6.37: Voorbeeld van het gebruik van pointers.

Een mogelijke uitvoer is te zien in de onderstaande figuur. We zeggen hierbij mogelijk omdat het draaien van het programma op een computer andere waarden (adressen) kan opleveren. In de onderstaande figuur is te zien dat pointer met acht hexadecimale cijfers worden afgedrukt. Dat betekent dat de pointers met 32 bits worden opgeslagen. We hebben gebruik gemaakt van een 32-bits C-compiler.

 Command Prompt

```

&p: 0061FDF0, &pp: 0061FDE8, &p[0]: 0061FDF0, &p[1]: 0061FDF8,
&p[2]: 0061FE00, &p[3]: 0061FE08

&i: 0061FE1C, p[0]: 0061FE1C
&j: 0061FE18, p[1]: 0061FE18
&k: 0061FE14, p[2]: 0061FE14
&l: 0061FE10, p[3]: 0061FE10

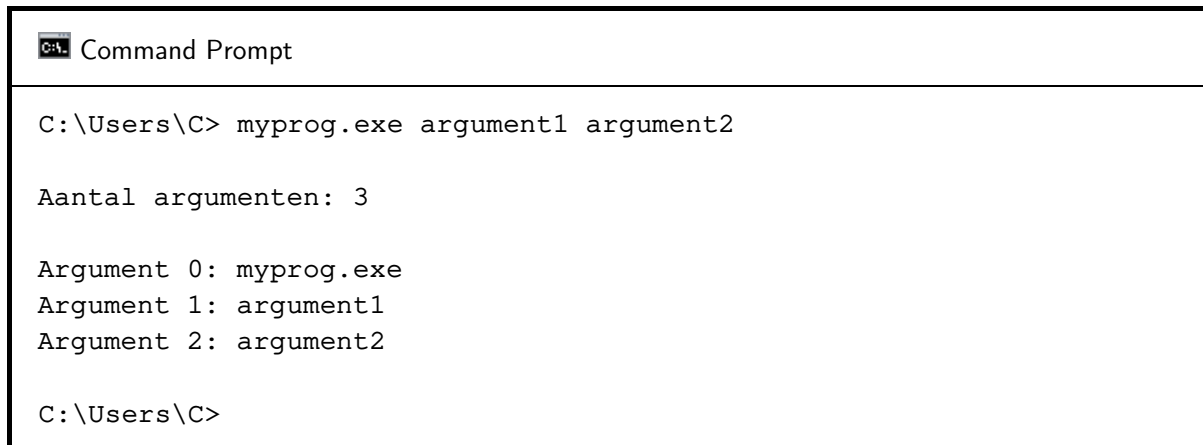
i: 3, *p[0]: 3, **pp: 3

```

6.14 Argumenten meegeven aan een C-programma

In besturingssystemen die C ondersteunen zoals Windows, Linux en Mac OS-X, is het mogelijk om een C-programma *command line argumenten* mee te geven. Bij het starten van een programma kunnen we (optioneel) gegevens invoeren en overdragen aan een gecompileerd

C-programma. Als voorbeeld starten we het programma `myprog.exe` met de argumenten `argument1` en `argument2`. Het programma drukt eenvoudigweg alle argumenten op het beeldscherm af. Te zien is dat ook de programmaam als argument wordt meegegeven.



```
C:\Users\C> myprog.exe argument1 argument2

Aantal argumenten: 3

Argument 0: myprog.exe
Argument 1: argument1
Argument 2: argument2

C:\Users\C>
```

Elk C-programma krijgt per definitie de twee parameters `argc` en `argv` mee, die aan `main` worden meegegeven. Dit is te zien in listing 6.38.

```
1 int main(int argc, char *argv[]) {
2
3     /* rest of the code */
4
5     return 0;
6 }
```

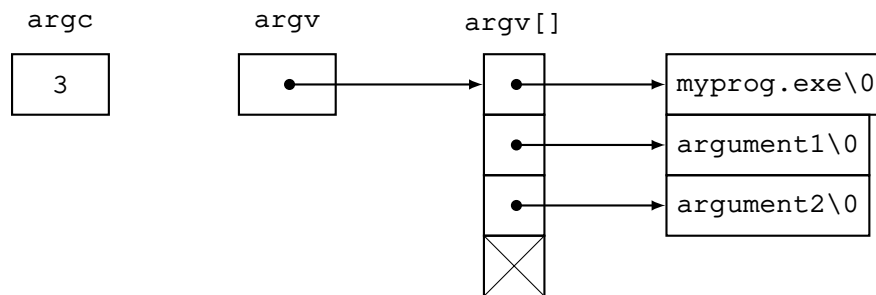
Listing 6.38: Declaratie van de command line parameters.

De integer `argc` (**argument count**) geeft aan hoeveel parameters aan het C-programma zijn meegegeven. De pointer `argv` (**argument vector**) is een pointer naar een lijst van pointers naar strings, gedeclareerd als `*argv[]`. Elke string bevat een argument. Per definitie wijst `argv[0]` naar een string waarin de programmaam vermeld staat. Dat houdt in dat `argc` dus minstens 1 is. Er zijn dan geen optionele argumenten meegegeven.

In het voorbeeldprogramma is `argc` dus 3 en zijn `argv[0]`, `argv[1]` en `argv[2]` pointers naar respectievelijk `myprog.exe`, `argument1` en `argument2`. In figuur 6.13 is een uitbeelding van de variabelen `argc` en `argv` te zien. De strings worden, zoals gebruikelijk in C, afgesloten met een nul-karakter. De C-standaard schrijft voor dat de lijst van pointers naar strings wordt afgesloten met een `NULL`-pointer.

Het programma `myprog.exe` is te zien in listing 6.39. Het programma drukt eerst de variabele `argc` af. Met behulp van een `for`-lus worden de argumenten één voor één afgedrukt. Merk op dat `argv[i]` een pointer is naar het i^e argument. We kunnen `argv[i]` dus direct gebruiken voor het afdrukken van de bijbehorende string.

Merk op dat `argv` een echte pointers is en niet de naam van een array. We mogen `argv` dus aanpassen. Dit is te zien in listing 6.40.



Figuur 6.13: Voorstelling van de variabelen *argc* en *argv*.

```

1 #include <stdio.h>
2 #include <stdlib.h>
3
4 int main(int argc, char *argv[]) {
5
6     int i;
7
8     printf("\nAantal_argumenten:_%d\n\n", argc);
9     for (i = 0; i<argc; i++) {
10         printf("Argument_%d:_%s\n", i, argv[i]);
11     }
12     return 0;
13 }

```

Listing 6.39: Het programma *myprog.exe*.

```

1 #include <stdio.h>
2 #include <stdlib.h>
3
4 int main(int argc, char **argv) {
5
6     printf("\nArguments:_%");
7     while (argc>0) {
8         printf("%s_", *argv);
9         argv = argv + 1;
10        argc = argc - 1;
11    }
12    return 0;
13 }

```

Listing 6.40: Afdrukken van argumenten.

6.15 Pointers naar functies

Functies zijn stukken programma die ergens in het geheugen liggen opgeslagen. De naam van een functie is het adres van de eerste instructie van de functie. Het is dus mogelijk om de naam van een functie te gebruiken als een pointer. Dit worden *functie-pointers* genoemd.

In listing 6.41 is te zien hoe een functie-pointer wordt gedeclareerd. In de eerste regel wordt een functie gedefinieerd, een zogenoemde prototype, die een `int` als parameter meekrijgt en een `int` teruggeeft. In de tweede regel wordt `pf` gedeclareerd als een pointer naar een functie die een `int` meekrijgt en een `int` teruggeeft.

```
1 int func(int a);    /* func is a function returning an int */
2 int (*pf)(int a);  /* pf is a pointer to a function
3                      returning an int */
4
5 pf = func;          /* assign pf */
```

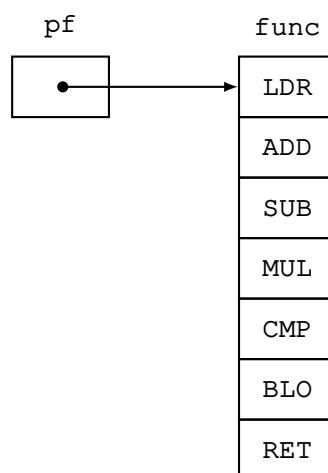
Listing 6.41: Een functie en een pointer naar een functie.

Let op het gebruik van de haakjes. Die zijn nodig om prioriteiten vast te leggen. Zonder de haken staat er `int *pf(int a)` en dan is `pf` een functie die een `int` als parameters meekrijgt en een pointer naar een `int` teruggeeft. Zie listing 6.42.

```
1 int (*pf)(int a);  /* pf: pointer to a function returning an int */
2 int *pf(int a);    /* pf: function returning a pointer to an int */
```

Listing 6.42: Een functie en een pointer naar een functie.

In figuur 6.14 is te zien hoe `pf` wijst naar de functie `func`. We kunnen nu `pf` gebruiken om de functie aan te roepen. In de figuur is een aantal instructies gezet.



Figuur 6.14: Voorstelling van een pointer naar een functie.

Het gebruik van functie-pointers komt niet zo vaak voor in C-programma's. Het is meer in gebruik bij het schrijven van besturingssystemen. We willen toch even twee functies de revue laten passeren waarbij functie-pointers gebruikt worden.

Quicksort

Quicksort is een sorteeralgoritme ontworpen door C.A.R. Hoare in 1962. Het is een van de efficientste sorteeralgoritmes voor algemeen gebruik. De exacte werking zullen we niet bespreken, er zijn genoeg boeken die dit beschrijven. De standard library bevat een implementatie onder de naam `qsort`. Het prototype van `qsort` is:

```
void qsort(void *base, size_t nitems, size_t size,
           int (*compar)(const void *, const void*));
```

Hierin is `base` een pointer naar het eerste element van de array, `nitems` het aantal elementen in de array en `size` de grootte (in bytes) van één element. De functie `compar` behoeft wat speciale aandacht. Dit is een functie (door de programmeur zelf te schrijven) die twee pointers naar twee elementen in de array meekrijgt. De pointers zijn van het type `void *` want we weten niet wat het datatypes van de elementen van de array zijn. De functie *moet* een getal geven kleiner dan 0 als het eerste argument kleiner is dan het tweede element, 0 als de twee elementen gelijk zijn en een getal groter dan 0 als het eerste element groter is dan het tweede element.

```
1 #include <stdio.h>
2 #include <stdlib.h>
3
4 int values[] = { 88, 56, 100, 2, 25 };
5
6 int cmpint (const void * a, const void * b) {
7     /* explicit type case to pointers to integers */
8     return ( *(int*)a - *(int*)b );
9 }
10
11 int main () {
12     int n;
13
14     printf("De_array_voor_sorteren:_\n");
15     for( n = 0 ; n < 5; n++ ) {
16         printf("%d_", values[n]);
17     }
18
19     qsort(values, 5, sizeof(int), cmpint);
20
21     printf("\nDe_array_na_sorteren:_\n");
22     for( n = 0 ; n < 5; n++ ) {
23         printf("%d_", values[n]);
24     }
25
26     return(0);
27 }
```

Listing 6.43: Sorteren van een array met *quicksort*.

In listing 6.43 is een programma te zien dat een array van integers sorteert. De functie `cmpint` vergelijkt twee integers uit de array. Let op de constructie om bij de integers te komen. De pointers zijn van het type `void *` dus er is een expliciete type case nodig naar een pointer naar een integer. Dat wordt gerealiseerd door `(int *)`. Daarna wordt de pointer gebruikt om bij de integer te komen. Om parameter `a` te verkrijgen is dus `*(int *) a` nodig. Op deze wijze wordt ook parameter `b` gevonden en de functie geeft eenvoudigweg het verschil tussen de twee parameters terug.

6.16 Pointers naar vaste adressen

C is erg coulant bij het toekennen van adressen aan pointers. Zo is het mogelijk om een pointer naar een vast adres te laten wijzen. De toekenning en initialisatie

```
int *p = (int *) 0x40fe;
```

zorgt ervoor dat `p` een pointer is naar een integer op adres $40FE_{16}$ (hexadecimale notatie). Er is een expliciete type cast nodig om de integer `0x40fe` om te zetten naar een adres. Nu zullen dit soort toekenningen niet voorkomen op systemen waar een operating system op draait. Het gebruik van de pointer zal hoogst waarschijnlijk een crash van het programma veroorzaken. Maar op kleine computersystemen zonder operating system, de zogenoemde *bare metal*-systemen, is het vaak de enige manier om informatie naar binnen en naar buiten te krijgen. We geven hieronder een voorbeeld van het gebruik van dit soort pointers op een ATmega-microcontroller van Atmel.

```
1 /*      Moet ik nog invullen */
```

Listing 6.44: Het gebruik van pointers op een ATmega-microcontroller.

6.17 Dynamische geheugenallocatie

Bij het schrijven van een programma declareren we de variabelen die we nodig hebben. Deze geheugenplaatsen blijven in principe bezet. Wel is het zo dat lokale variabelen worden aangemaakt en afgebroken aan het begin respectievelijk het einde van een functie (tenzij ze als `static` gekwalificeerd zijn), maar er komt geen extra geheugen erbij. Via het besturingssysteem is het mogelijk om extra geheugenruimte aan te vragen als dat nodig is. Is de aangevraagde geheugenruimte niet meer nodig dan geven we het terug aan het besturingssysteem. We noemen dit *dynamische geheugenallocatie*.

De standard library kent een aantal functie op dit gebied. Om deze functies te gebruiken moet de header-bestand `malloc.h` geladen worden. De functie `malloc` vraagt aan het besturingssysteem extra geheugenruimte *in bytes*. Als dat lukt, dan wordt een void-pointer naar het begin van de geheugenruimte teruggegeven. De void-pointer moet dus gecast worden naar het juiste type. Het geheugen wordt niet geïnitieerd, bijvoorbeeld met nul-bytes. Als het niet lukt, wordt een NULL-pointer teruggegeven. Hier moet uiteraard op getest worden. De functie `free` geeft eerder gealloceerd geheugen weer vrij. Als argument wordt een pointer naar het gealloceerde geheugenruimte meegegeven. Let erop dat de geheugenruimte niet eerder is vrijgegeven, want dan stopt het programma met een crash.

Een van de meest voorkomende problemen is dat gealloceerd geheugen niet wordt vrijgegeven. Dit wordt een *memory leak* genoemd. Op zich is dat niet zo erg, want als het programma stopt, ruim het besturingssysteem alle gealloceerde geheugenruimtes op. Maar als een programma alsmäär geheugen vraagt en niets vrijgeeft, kan het zijn dat uiteindelijk geen geheugen meer beschikbaar is.

Een eenvoudig gebruik van `malloc` en `free` is te zien in listing 6.45. De regel 6 definiëren we een *macro* met de grootte van 1000 bytes. Regel 13 zorgt ervoor dat Visual Studio de functie `strcpy` accepteert. In regel 17 declareren we een karakterpointer en initialiseren we die met `NULL`. Regel 19 roept de functie `malloc` aan. De functie geeft een void-pointer terug, dus we moeten een type cast naar `char *` doen. We testen in de regels 21 t/m 24 of de allocatie gelukt is en zo niet, dan drukken we een foutmelding af en stoppen we het programma.

```
1 #include <stdio.h>
2 #include <malloc.h>
3 #include <string.h>
4
5 /* Define 1000 bytes of space*/
6 #define SIZE 1000
7 /* Define 10 GB of space */
8 /* #define SIZE 100000000000 */
9 /* Define 100 GB of space */
10 /* #define SIZE 1000000000000 */
11
12 /* Keep Visual Studio happy */
13 #define strcpy(A,B) strcpy_s(A,SIZE,B);
14
15 int main(void) {
16
17     char *pstr = NULL;
18
19     pstr = (char *) malloc(SIZE);
20
21     if (pstr == NULL) {
22         printf("Kan_geheugen_niet_alloceren!");
23         return 0;
24     }
25
26     strcpy(pstr, "Vul_maar_wat_in");
27
28     printf("%s", pstr);
29
30     free(pstr);
31
32     return 0;
33 }
```

Listing 6.45: Gebruik van `malloc` en `free`.

In regel 26 t/m 28 kopiëren we een string naar het geheugen en drukken het af. In regel 30 geven we het geheugen weer vrij. We hebben twee macro's toegevoegd die we naar believen kunnen activeren: één voor 10 GB en één voor 100 GB. De lezer wordt gevraagd om deze te testen.

We vermelden nog even twee andere functies: `realloc` en `calloc`. De functie `realloc` heralloceert een stuk geheugen. Met de aanroep

```
pnew = realloc(porig, size);
```

wordt het stuk geheugen waar `porig` naar wijst uitgebreid of ingekrompen, `pnew` wijst naar de nieuwe geheugenruimte. Er zijn een paar zaken waar we op moeten letten: 1) als het geheugen wordt ingekrompen, blijft de originele inhoud ongewijzigd; 2) als het geheugen wordt uitgebreid blijft de originele inhoud ongewijzigd en het nieuwe, extra geheugen wordt niet geïnitieerd, 3) als het niet lukt om het geheugen op de originele plaats uit te breiden, wordt een compleet nieuw geheugenblok gealloceerd en wordt de inhoud van het originele inhoud gekopieerd, de rest wordt niet geïnitieerd, het oude geheugenblok wordt vrijgegeven; 4) als er geen geheugenruimte meer vrij is wordt een NULL-pointer teruggegeven en blijven de originele pointer en geheugen intact.

De functie `calloc` alloceert een aantal veelvoud van een elementair datatype en vult de geheugenruimte met nul-bytes. Dus

```
int *pint = calloc(1000, sizeof(int));
```

alloceert 1000 integers en vult ze met nul-bytes. Als het niet lukt, wordt een NULL-pointer teruggegeven.

6.18 Subtiele verschillen en complexe declaraties

We zullen in de praktijk nauwelijks complexere situaties tegenkomen dan dat we tot nu toe zijn tegengekomen. Toch willen we een bloemlezing geven van enkele bekende en onbekende, complexe declaraties.

<code>int *p</code>	<code>p</code> : pointer to int
<code>int **pp</code>	<code>pp</code> : pointer to pointer to int
<code>int ***ppp</code>	<code>ppp</code> : pointer to pointer to pointer to int
<code>int **pp[3]</code>	<code>pp</code> : array[3] of pointer to pointer to int
<code>char **argv</code>	<code>argv</code> : pointer to pointer to char
<code>char *argv[]</code>	<code>argv</code> : array[] of pointer to char
<code>int *list[5]</code>	<code>list</code> : array[5] of pointer to int
<code>int (*list)[5]</code>	<code>list</code> : pointer to array[5] of int
<code>int *(*list)[5]</code>	<code>list</code> : pointer to array[5] of pointer to int
<code>int *pf()</code>	<code>pf</code> : function returning a pointer to int
<code>int (*pf)()</code>	<code>pf</code> : pointer to function returning an int
<code>int *(*pf)()</code>	<code>pf</code> : pointer to function returning a pointer to int
<code>char ((*x())[])()</code>	<code>x</code> : function returning pointer to array[] of pointer to function returning char.
<code>char ((*x[3])())[5]</code>	<code>x</code> : array[3] of pointer to function returning pointer to array[5] of char.



Structures

8

Invoer en uitvoer

8.1 Inlezen van een getal met `scanf`

De functie `scanf` is een veelzijdige functie voor het inlezen van variabelen. Zo kan je een getal, een letter of een string inlezen. Maar hoe weet `scanf` nu wat er ingelezen moet worden? Dat geven we op met een format string. Een enkel geheel getal wordt inlezen met:

```
scanf("%d", &getal);
```

De format specification `%d` vertelt `scanf` dat een geheel getal moet worden ingelezen. De variabele waarin het getal moet komen te staan, wordt voorzien van de adres-operator `&` zodat `scanf` de variabele kan vinden. Het is ook mogelijk om meerdere getallen in te lezen:

```
scanf("%d %d %d", &getala, &getalb, &getalc);
```

Bij het invoeren moeten de getallen gescheiden worden door één of meerdere spaties. Het aantal spaties maakt dus niet uit; deze worden overgeslagen.

We kunnen ook andere datatypes inlezen zoals een `float` of een `char`

```
scanf("%f %c", &fl, &ch);
```

8.1.1 Problemen met `scanf`

Als we in een C-programma een getal inlezen met `scanf` dan gebeuren er vreemde dingen als we in plaats van een getal letters intypen. Bijvoorbeeld:


```

1 #include <stdio.h>
2
3 int main(void) {
4     int getal = -12345;
5
6     printf("Geef_een_geheel_getal:_");
7     scanf("%d", &getal);
8
9     printf("Het_getal_is_%d.\n", getal);
10
11     return 0;
12 }

```

Listing 8.1: *Getal inlezen met scanf.*

Als we dit programma uitvoeren en als invoer het woord `Hallo` intypen dan verschijnt de volgende uitvoer:

 Command Prompt
<pre> Geef een geheel getal: Hallo Het getal is -12345. </pre>

Het blijkt dat de waarde van variabele `getal` niet is veranderd. Het eerste teken dat `scanf` tegenkomt is de `H` en dat is geen cijfer, dus wordt er gestopt met inlezen van het toetsenbord. Het wordt echt problematisch als er een getal ingelezen moet worden dat aan een voorwaarde moet voldoen:

```

1 int main(void) {
2     int getal = -12345;
3
4     do {
5         printf("Geef_een_geheel_getal_groter_dan_0:_");
6         scanf("%d", &getal);
7     } while (getal < 1);
8
9     printf("Het_getal_is_%d.\n", getal);
10    return 0;
11 }

```

Listing 8.2: *Het inlezen van een geheel getal dat groter is dan 0.*

De uitvoer is nu:

Het programma gaat nu als een razende te werk en de enige manier om het te stoppen is het programma af te sluiten. Maar hoe komt dat nou?

Command Prompt

```
Geef een geheel getal groter dan 0: Hallo
Geef een geheel getal groter dan 0: Geef een geheel getal groter dan
0: Geef een geheel getal groter dan 0: Geef een geheel getal groter
dan 0: Geef een geheel getal groter dan 0: Geef een geheel getal gro
ter dan 0: Geef een geheel getal groter dan 0: Geef een geheel getal
enzovoorts
```

Als `scanf` voor de eerste keer iets inleest dan zijn er nog geen karakters vanaf het toetsenbord ingevoerd. Daarom vraagt `scanf` aan het besturingssysteem (Windows, Linux, OS-X) om een karakter. Het besturingssysteem weet dat er geen karakters beschikbaar zijn en gaat via interne routines karakters opvragen. Er wordt echter niet één karakter opgevraagd maar een hele reeks die afgesloten moet worden met een enter-toets. We moeten als gebruiker dus de invoer altijd afsluiten met een enter-toets, ook als we via `scanf` maar één karakter inlezen. De ingelezen karakters worden ergens in het geheugen opgeslagen. Dit wordt *buffering* genoemd en de geheugenruimte wordt *invoerbuffer* genoemd. Er zijn nu karakters beschikbaar en het eerste ingevoerde karakter wordt aan `scanf` gegeven.

Het eerste karakter is de letter H en dat is geen cijfers. Dus stopt `scanf` direct met het inlezen van karakters (die dus alleen cijfers mogen zijn). Het karakter H blijft hierbij in de invoerbuffer staan. De H wordt dus niet verwijderd.

Omdat `scanf` geen cijfers heeft kunnen inlezen, wordt de opgegeven variabele niet veranderd en blijft zijn originele waarde behouden. Daarom drukt het programma in listing ?? het getal `-12345` af. Het programma in listing ?? blijft in een `do-while`-lus steeds een getal inlezen als het getal kleiner is dan `-1`. Aangezien `scanf` geen cijfers inleest en de opgegeven variabele niet aanpast, blijft de waarde `-12345` behouden en dat is kleiner dan `-1`. Dus wordt de lus nog een keer uitgevoerd. De H staat nog steeds in de invoerbuffer en lukt het `scanf` niet om een getal in te lezen.

Dit kun je vrij eenvoudig voorkomen door de returnwaarde van de functie `scanf` te testen. Deze functie geeft het aantal variabelen terug dat succesvol is geconverteerd en ingelezen (dat kan dus ook 0 zijn), anders geeft de functie de integer waarde `EOF` terug. Je kan dat als volgt doen:

```
1 #include <stdio.h>
2
3 int main(void) {
4     int getal = -12345;
5     int ret;
6
7     printf("Geef een geheel getal groter dan 0: ");
8     ret = scanf("%d", &getal);
9
10    if (ret == 1) {
11        printf("Het getal is %d.\n", getal);
12    } else {
```

```

13 printf("Geen_getal_ingevoerd!\n");
14 }
15
16 printf("Druk_op_de_Enter_toets_om_dit_window_te_sluiten.");
17 getchar();
18 return 0;
19 }

```

De uitvoer wordt dan:

```

1 Geef een geheel getal groter dan 0: Hallo
2 Geen getal ingevoerd!
3 Druk op de Enter toets om dit window te sluiten.

```

Er is nog één probleem. De karakters van `Hallo` staan nog steeds in de invoerbuffer. Die zullen we er een voor een uit moeten halen. Dat kan met het onderstaande programma:

```

1 #include <stdio.h>
2
3 int main(void) {
4     int getal, ret;
5     do {
6         printf("Geef een geheel getal: ");
7         ret=scanf("%d", &getal);
8         if (ret == 0) {
9             printf("Dat_was_geen_getal!\n");
10            printf("Maar_het_karakter_%c.\n", getchar());
11        }
12        else if (ret == EOF) {
13            printf("Er_is_een_fout_opgetreden_bij_het_lezen!\n");
14        }
15    } while (ret != 1);
16    printf("Het_getal_is_%d.\n", getal);
17    printf("Druk_op_de_Enter_toets_om_dit_window_te_sluiten.");
18    getchar();
19    return 0;
20 }

```

De functie `getchar` leest één karakter uit de invoerbuffer. Daarna proberen we `scanf` opnieuw. Dat mislukt telkens als er een letter gevonden wordt. Als je nu `Hallo` invoert, krijg je de volgende uitvoer:

```

1 Geef een geheel getal: Hallo
2 Dat was geen getal!

```

```

3 Maar het karakter H.
4 Geef een geheel getal: Dat was geen getal!
5 Maar het karakter a.
6 Geef een geheel getal: Dat was geen getal!
7 Maar het karakter l.
8 Geef een geheel getal: Dat was geen getal!
9 Maar het karakter l.
10 Geef een geheel getal: Dat was geen getal!
11 Maar het karakter o.
12 Geef een geheel getal:

```

Natuurlijk willen niet steeds dat bij elk karakter een melding op het scherm wordt afgedrukt. We kunnen een aantal `printf`-regels verwijderen maar niet de regel waarin de gebruiker wordt gevraagd om een geheel getal in te voeren. Anders weet de gebruiker niet wat hij/zij moet doen. Helaas wordt deze regel herhaaldelijk afgedrukt. We zullen op een iets andere manier de karakters moeten inlezen.

Gelukkig zorgt het operating system ervoor dat ook het end-of-line-karakter in de invoerbuffer terecht komt. We kunnen dus hierop testen. Hoe dat moet, is te zien in de onderstaande code:

```

1 do { ch = getchar(); } while (ch != '\n' && ch != EOF);

```

We lezen een karakter van de invoerbuffer en dat doen we zolang dat karakter ongelijk is aan `\n` (end-of-line-karakter) en ongelijk is aan `EOF` (end-of-file).

Omdat we vaker gebruik willen maken van het legen van de invoerbuffer plaatsen we de code een functie. We kunnen nu aan de gebruiker vragen om een getal in te voeren:

```

1 #include <stdio.h>
2
3 void purge_stdin(void) {
4 int ch;
5 do {
6 ch = getchar();
7 } while (ch != '\n' && ch != EOF);
8 }
9
10 int main(void) {
11 int getal, ret;
12 do {
13 printf("Geef een geheel getal:_");
14 ret=scanf("%d", &getal);
15 if (ret == 0) {
16 purge_stdin();
17 }

```

```

18 } while (ret != 1);
19 printf("Het_getal_is_%d.\n", getal);
20 printf("Druk_op_de_Enter_toets_om_dit_window_te_sluiten.");
21 purge_stdin();
22 getchar();
23 return 0;
24 }

```

Er is nog een andere manier om de invoerbuffer te legen. We kunnen de invoerbuffer legen met de functie `fflush` (file flush):

```

1 fflush(stdin);

```

Als parameter van `fflush` wordt `stdin` opgegeven. Dat staat voor *standard input* en daar wordt in de regel het toetsenbord mee bedoeld¹. Daarnaast kennen we nog `stdout` (*standard output*, het beeldscherm) en `stderr` (*standard error*, meestal ook het beeldscherm).

Deze manier werkt echter niet met alle C-compilers en operating systems. Dat heeft te maken met de definitie van de functie `fflush`. De functie `fflush` is bedoeld om uitvoerbuffers te legen. Als de uitvoer naar het beeldscherm is, wordt de buffer naar het beeldscherm geschreven. Als de uitvoer naar een bestand is, wordt de buffer naar het bestand geschreven. De C-standaard schrijft echter alleen voor dat *flushen* van een uitvoerbuffer gedefinieerd is, niet van een invoerbuffer. Niet zo gek eigenlijk, want wat wordt er nou bedoeld met flushen van de invoerbuffer. Flushen van het toetsenbord is het nog te begrijpen maar flushen van een invoerbestand niet. Moeten we dan helemaal tot einde van het bestand flushen? Of alleen maar de bijbehorende buffer? Dat levert een onvoorspelbaar programma op want we weten immers niet hoeveel karakters in de buffer staan.

Toch zijn er wel C-implementaties die het flushen van een invoerbuffer uitvoeren, bijvoorbeeld de GNU C-compiler op Linux en MinGW op Windows (zit o.a. in Code::Blocks).

Hieronder is de code te vinden van het flushen van de invoerbuffer van het toetsenbord m.b.v. `fflush`:

¹ Op de bekende operating systems is het mogelijk om de inhoud van bestanden door te geven aan de *standard input*. Het programma krijgt dan data uit een bestand i.p.v. het toetsenbord. Dat wordt *redirection* genoemd.

```

1  /* Please note: might not works on all Operating Systems. */
2  #include <stdio.h>
3
4  int main(void) {
5      int getal, ret;
6      do {
7          printf("Geef_een_geheel_getal:_");
8          ret=scanf("%d", &getal);
9          if (ret == 0) {
10             fflush(stdin);
11         }
12     } while (ret != 1);
13     printf("Het_getal_is_%d.\n", getal);
14     printf("Druk_op_de_Enter_toets_om_dit_window_te_sluiten.");
15     fflush(stdin);
16     getchar();
17     return 0;
18 }

```


Bibliografie

Veel materiaal is tegenwoordig (alleen) via Internet beschikbaar. Voorbeelden hiervan zijn de datasheets van ic's die alleen nog maar via de website van de fabrikant beschikbaar worden gesteld. Dat is veel sneller toegankelijk dan boeken en tijdschriften. De keerzijde is dat websites van tijd tot tijd veranderen of verdwijnen. De geciteerde weblinks werken dan niet meer. Helaas is daar niet veel aan te doen. Er is geen garantie te geven dat een weblink in de toekomst beschikbaar blijft.

- [1] B.W. Kernighan en D.M. Ritchie. *The C Programming Language*. Prentice-Hall software series. Prentice Hall, 1988. ISBN: 9780131103627 (blz. iii).
- [2] LaTeX Project Team. *LaTeX – A document preparation system*. URL: <https://www.latex-project.org/> (bezocht op 30-12-2018) (blz. iii).
- [3] Tex User Groups. *TeX Live Website*. URL: <https://www.tug.org/texlive/> (bezocht op 30-12-2018) (blz. iii).
- [4] Benito van der Zander. *TeXstudio, LaTeX made comfortable*. URL: <https://www.texstudio.org/> (bezocht op 04-05-2019) (blz. iii).
- [5] Bitstream Inc. *Charter Fonts*. URL: <https://www.ctan.org/pkg/charter> (bezocht op 30-12-2018) (blz. iii).
- [6] Artifex. *Nimbus 15 Mono*. Okt 2015. URL: <https://www.ctan.org/tex-archive/fonts/nimbus15> (bezocht op 30-12-2018) (blz. iii).
- [7] J. Hoffman. *listings – Typeset source code listings using LaTeX*. URL: <https://www.ctan.org/pkg/listings> (bezocht op 30-12-2018) (blz. iii).
- [8] T. Tantau. *pgf – Create PostScript and PDF graphics in TeX*. Aug 2015. URL: <https://www.ctan.org/pkg/pgf> (bezocht op 30-12-2018) (blz. iii).
- [9] HBO Engineering. *Body of Knowledge and Skills Elektrotechniek*. Nov 2016. URL: http://www.hbo-engineering.nl/_asset/_public/competenties/BOKS_elektrotechniek_met_specialisaties_def_11nov2016.pdf (bezocht op 01-03-2018) (blz. iv).
- [10] D.E. Knuth. *Structured Programming with go to Statements*. 1974. URL: https://web.archive.org/web/20130731202547/http://pplab.snu.ac.kr/courses/adv_pl05/papers/p261-knuth.pdf (bezocht op 10-04-2020) (blz. 25).
- [11] V. Eijkhout. *An ASCII wall chart*. Nov 2010. URL: <https://ctan.org/pkg/ascii-chart> (bezocht op 29-07-2018) (blz. 84).
- [12] American Standards Association. *ASA standard X3.4-1963*. URL: <http://worldpowersystems.com/J/codes/X3.4-1963/> (bezocht op 26-07-2018) (blz. 83).
- [13] A.K. Maini. *Digital Electronics: Principles, Devices and Applications*. Wiley, 2007, p. 28. ISBN: 978-0-470-03214-5 (blz. 83).



De ASCII-tabel

We hebben gezien dat binaire coderingen worden gebruikt om numerieke gegevens (of informatie) weer te geven. Maar niet alle gegevens zijn numeriek. Informatie kan ook bestaan uit letters en leestekens. Om er voor te zorgen dat computers informatie kunnen uitwisselen is de ASCII-code bedacht.

De naam ASCII betekent *American Standard Code for Information Interchange* en dat geeft al goed aan waarvoor de code bedoeld is: op een gestandaardiseerde wijze informatie uitwisselen. De code is in 1963 voor het eerst gepubliceerd [12] en in die tijd was er nog geen noodzaak om andere tekens te gebruiken dan de bekende westerse letters, cijfers en leestekens. Vandaar dat het aantal tekens beperkt is.

De ASCII-code bestaat uit 128 7-bits tekens zoals te zien is in tabel A.1. De tekens zijn verdeeld in leesbare tekens, zoals letters, cijfers en leestekens en zogenoemde *besturingstekens*. De besturingstekens zijn nodig om informatie-overdracht af te bakenen en om een bepaalde *handshake* (uitwisselingsprotocol) te regelen. Zo zijn er codes voor de *carriage return* (CR, code 13_{10}) en de *backspace* (BS, code 8_{10}). De eerste 32 tekens zijn besturingstekens waarvan de meeste tegenwoordig niet meer gebruikt worden [13].

De codes zijn niet willekeurig toegekend dat we goed kunnen zien bij de cijfers en letters. De tabel is zo opgesteld dat de cijfers elkaar opvolgen. Dat is handig bij het afdrukken van een (decimaal) getal. Hetzelfde geldt voor de letters, ook die volgen elkaar op. De makers hebben ook nagedacht over de positie van hoofd- en kleine letters. Deze verschillen in de tabel in slechts één bit (bit b_6 in de tabel). Bij het gebruik van de Caps Lock-toets of Shift-toets hoeft dus maar één bit (in combinatie met een letter) gewijzigd te worden.

Een aantal besturingstekens is ook in C direct te gebruiken. Een besturingsteken begint altijd met een *backslash* (`\`) gevolgd door een letter, cijfer of teken. Zo is het teken voor een horizontale tab `\t` en voor de *line feed* is het teken `\n`. Met een line feed gaat de cursor naar het begin van de volgende regel.

Tabel A.1: De ASCII-code [11].

ASCII CONTROL CODE CHART

b7 b6 b5 BITS b4 b3 b2 b1	0	0	0	0	1	1	1	1
	0	0	1	1	0	0	1	1
	CONTROL		SYMBOLS NUMBERS		UPPER CASE		LOWER CASE	
0 0 0 0	0 NUL	16 DLE	32 SP	48 0	64 @	80 P	96 ' p	112
0 0 0 1	1 SOH	17 DC1	33 !	49 1	65 A	81 Q	97 a q	113
0 0 1 0	2 STX	18 DC2	34 "	50 2	66 B	82 R	98 b r	114
0 0 1 1	3 ETX	19 DC3	35 #	51 3	67 C	83 S	99 c s	115
0 1 0 0	4 EOT	20 DC4	36 \$	52 4	68 D	84 T	100 d t	116
0 1 0 1	5 ENQ	21 NAK	37 %	53 5	69 E	85 U	101 e u	117
0 1 1 0	6 ACK	22 SYN	38 &	54 6	70 F	86 V	102 f v	118
0 1 1 1	7 BEL	23 ETB	39 ' 7	55	71 G	87 W	103 g w	119
1 0 0 0	8 BS	24 CAN	40 (8	56	72 H	88 X	104 h x	120
1 0 0 1	9 HT	25 EM	41) 9	57	73 I	89 Y	105 i y	121
1 0 1 0	10 LF	26 SUB	42 * :	58	74 J	90 Z	106 j z	122
1 0 1 1	11 VT	27 ESC	43 + ;	59	75 K	91 [107 k {	123
1 1 0 0	12 FF	28 FS	44 , <	60	76 L	92 \	108 l	124
1 1 0 1	13 CR	29 GS	45 - =	61	77 M	93]	109 m }	125
1 1 1 0	14 SO	30 RS	46 . >	62	78 N	94 ^	110 n ~	126
1 1 1 1	15 SI	31 US	47 / ?	63	79 O	95 -	111 o DEL	127
	F 17	1F 37	2F 57	3F 77	4F 117	5F 137	6F 157	7F 177

LEGEND:

dec	CHAR
hex	oct

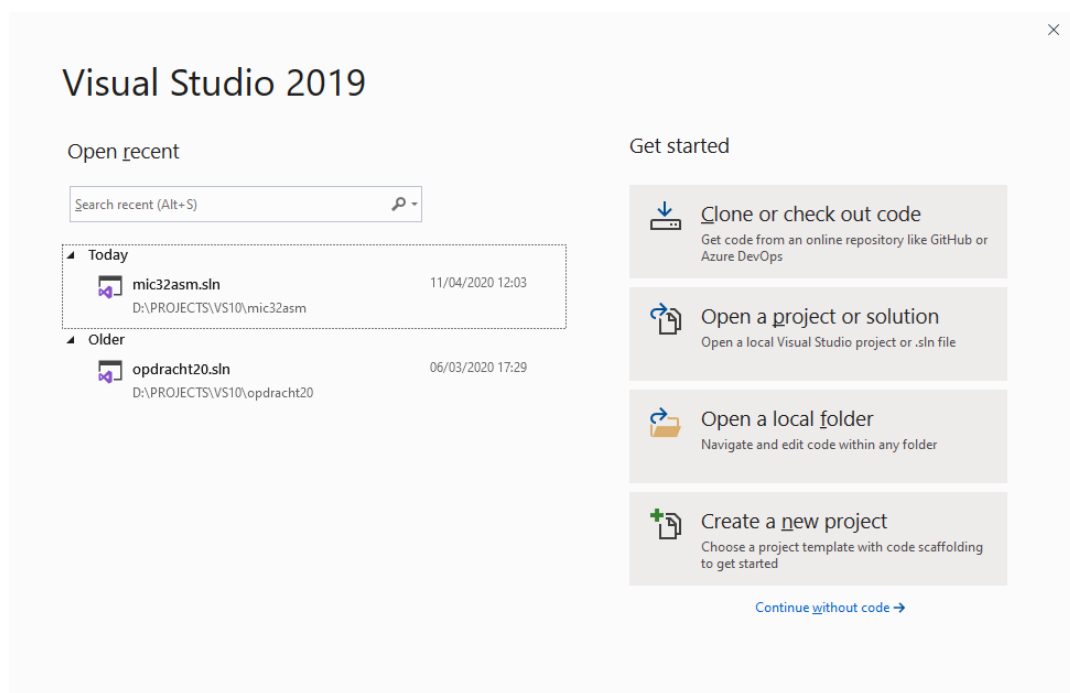
Victor Eijkhout
TACC
Austin, Texas, USA

B

Tutorial Visual Studio

We bespreken in deze bijlage het opzetten van een project in Visual Studio 2019. Visual Studio 2019 kan gedownload worden van de website van Microsoft. De precieze installatie van de software verschilt van computer tot computer, afhankelijk van de al eerder geïnstalleerde software, met name de *runtime libraries*.

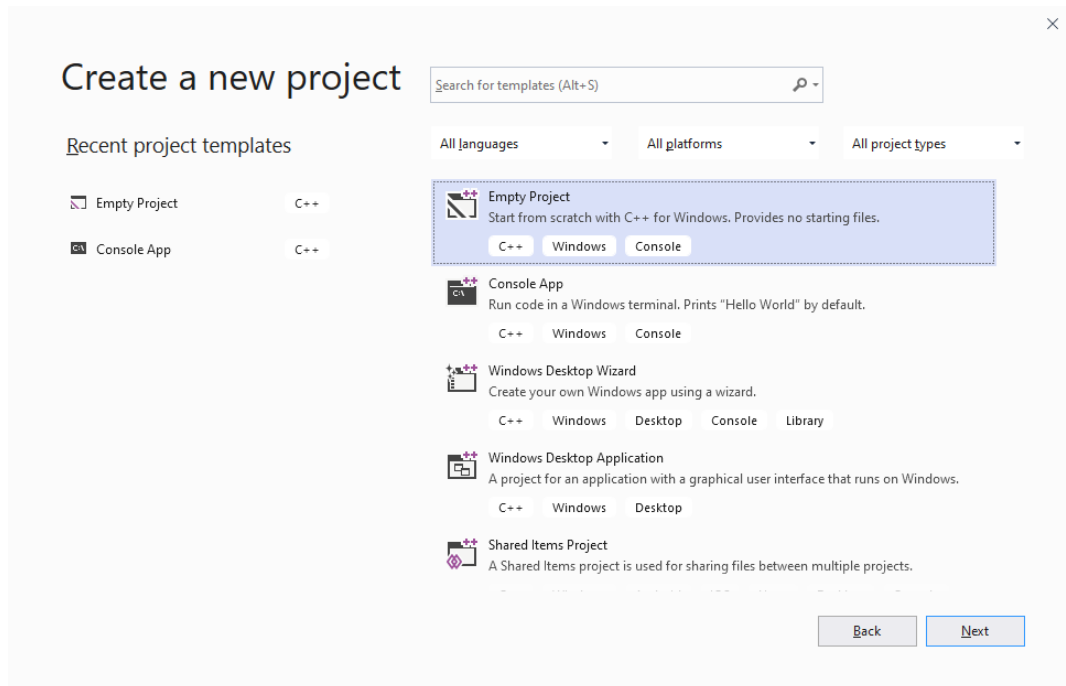
Start Visual Studio door op het icoon te klikken. Visual Studio opent een beginscherm waarin een nieuw project kan worden aangemaakt. Dit is te zien in figuur B.1. Klik op het kader `Create a new project`.



Figuur B.1: Aanmaken van een nieuw project.

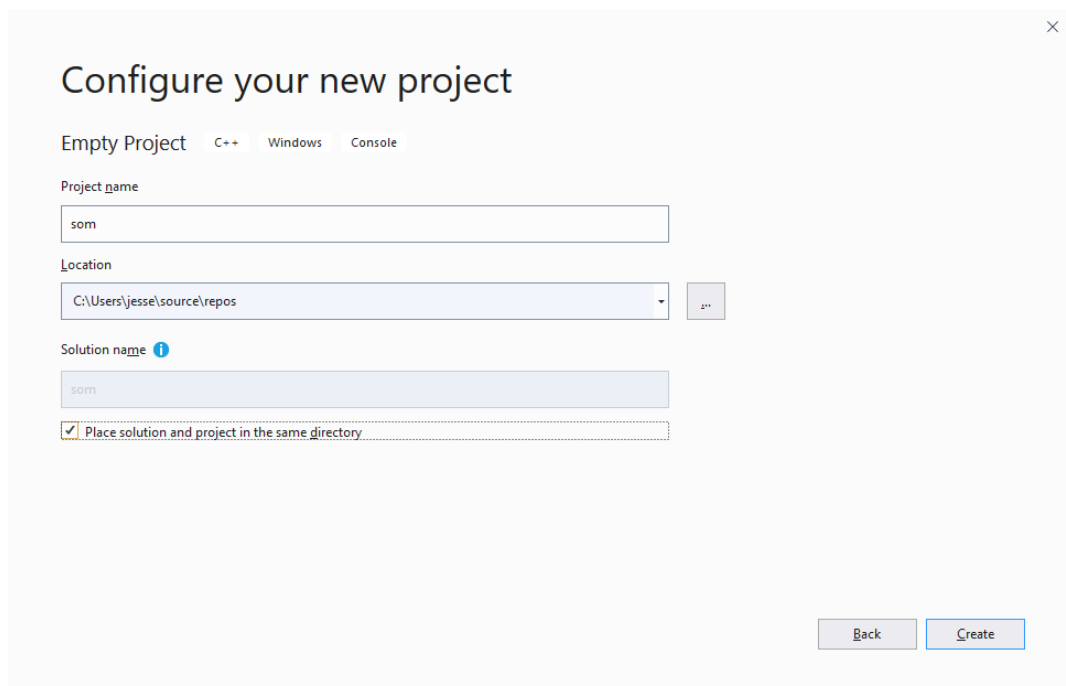
Er wordt een nieuw scherm geopend, zie figuur B.2. Klik daarin op het kader `Empty`

Project. **Klik niet op Console App.**



Figuur B.2: Aanmaken van een leeg project.

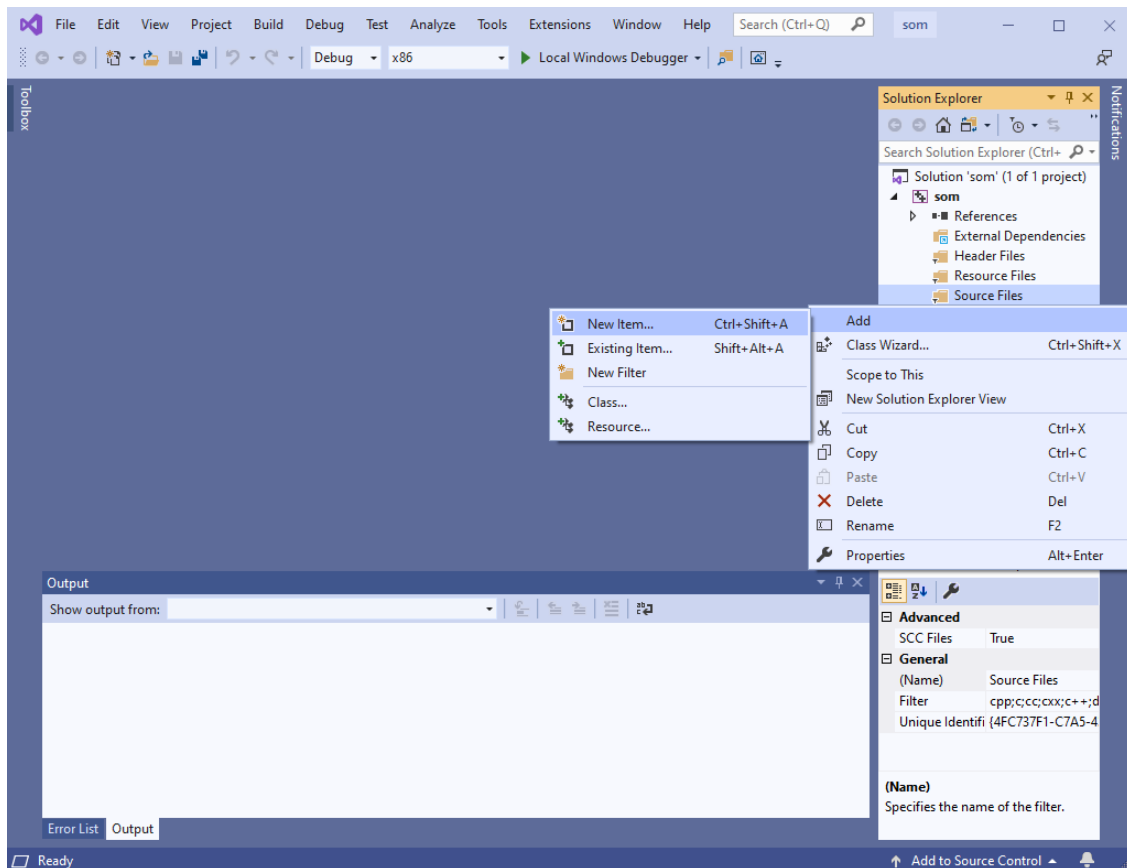
Daarna moeten wat gegevens worden ingevuld. Vul de projectnaam in en de map waarin het project terecht moet komen. Vink de checkbox onderaan aan en klik op de knop Create. Zie figuur B.3.



Figuur B.3: Gegevens van het project invoeren.

Visual Studio komt nu met het hoofdscherm waarin een aantal vensters (Engels: pane) te zien zijn. Er is nog geen C-bestand aangemaakt, dat moeten we zelf doen. In de *Solution Explorer* aan de rechterkant is een map *Source Files* te zien. Ga met de muispointer daar op staan en klik op de **rechter** muisknop.

Selecteer daarna de optie *Add* en daarna *New item...* . . . Zie figuur B.4.



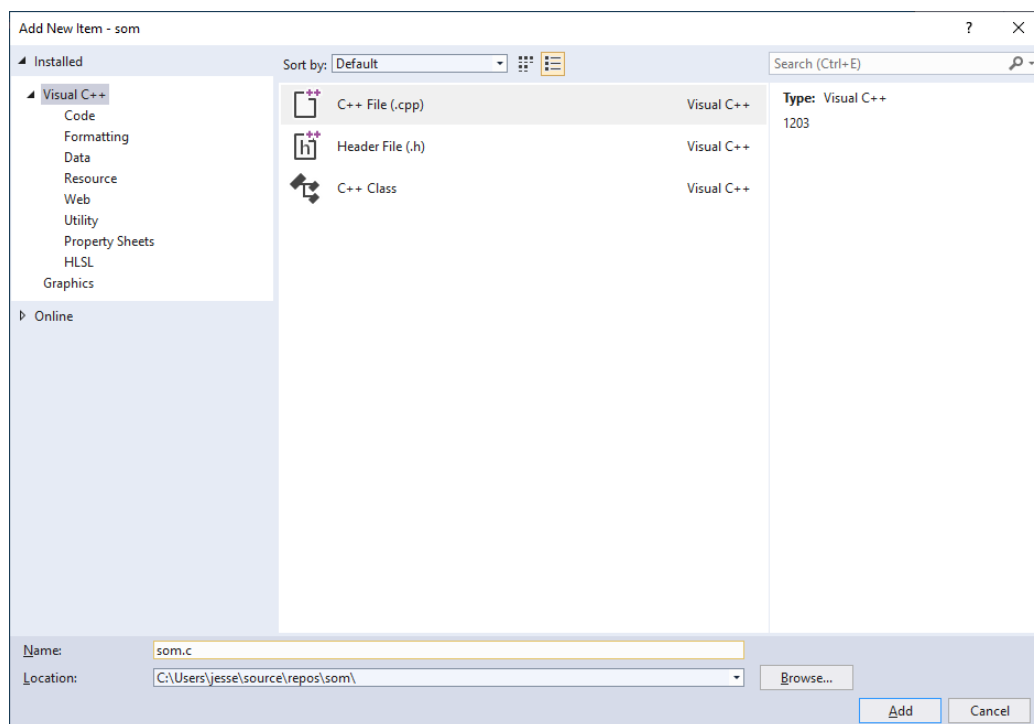
Figuur B.4: Een nieuw bestand aanmaken.

In het volgende scherm moet een bestandstype en een naam worden opgegeven. Klik op *C++ File (.cpp)*. Vul onderin bij *Name* de naam van het bestand in **en zorg ervoor dat de naam eindigt met .c**, anders wordt een C++-bestand aangemaakt. Klik daarna op de knop *Add*. Zie figuur B.5.

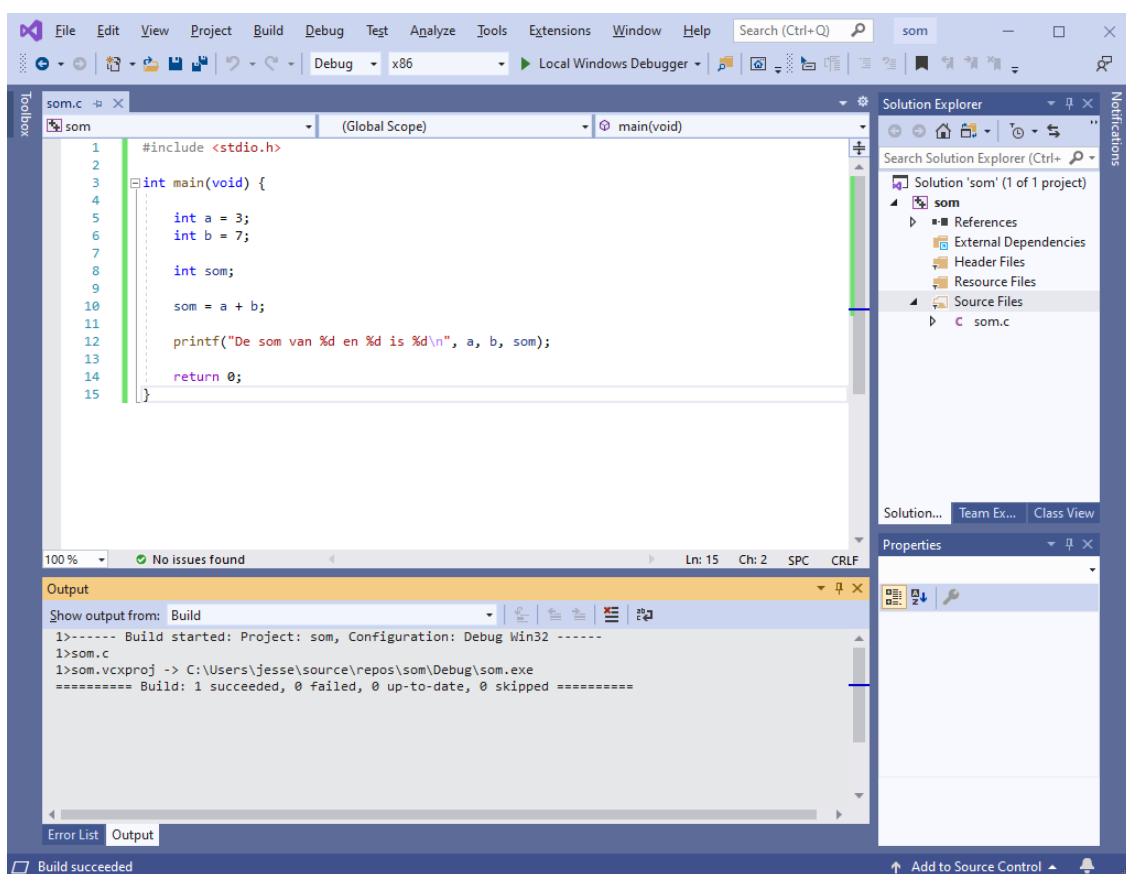
Voer het programma in zoals te zien is in figuur B.6. Klik daarna op de knop *Local Windows Debugger*. Het programma wordt nu gecompileerd en als er geen fouten zijn gevonden, wordt het programma uitgevoerd. Herstel eventuele fouten die door de compiler gevonden worden en herstart de compilatie.

Het programma drukt de regel *De som van 3 en 7 is 10* af. Dit wordt gedaan in een zogenoemde *console*. Dit is te zien in figuur B.7.

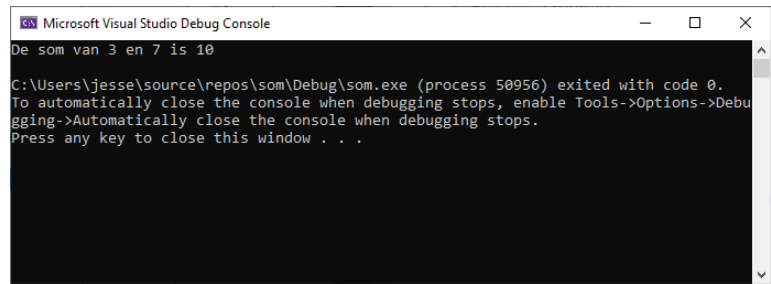
De tutorial is hiermee ten einde.



Figuur B.5: Gegevens van het C-bestand invullen.



Figuur B.6: Compileren en starten van de executable.



```
Microsoft Visual Studio Debug Console
De som van 3 en 7 is 10
C:\Users\jesse\source\repos\som\Debug\som.exe (process 50956) exited with code 0.
To automatically close the console when debugging stops, enable Tools->Options->Debugging->Automatically close the console when debugging stops.
Press any key to close this window . . .
```

Figuur B.7: *Uitvoer van het programma in een console.*



Vorrangsregels van operatoren

In deze tabel zijn alle operatoren in C opgesomd. De voorrang of prioriteit is in aflopende volgorde van hoog naar laag. Twee opmerkelijke operatoren zijn (*type cast*) en *sizeof*. Bij (*type cast*) wordt het casten van een enkelvoudig datatype bedoeld, *sizeof* berekent de grootte in bytes van een datatype of variabele tijdens *compile time*.

Tabel C.1: Voorrangsregels van de alle operatoren.

Operator	Associativiteit
() [] -> .	links naar rechts
! ~ + - ++ -- (type cast) sizeof * &	rechts naar links
* / %	links naar rechts
+ -	links naar rechts
<< >>	links naar rechts
< <= > >=	links naar rechts
== !=	links naar rechts
&	links naar rechts
^	links naar rechts
	links naar rechts
&&	links naar rechts
	links naar rechts
?:	rechts naar links
= += -= *= /= %= &= ^= = <<= >>=	rechts naar links
,	links naar rechts

Index

Operatoren en constanten

- !, logische negatie, 17
- , 15
- *, dereferentie, 45
- *, vermenigvuldigen, 15
- +, 15
- /, delen, 15
- <<, links schuiven, 17
- >>, rechts schuiven, 17
- [], operator, 32
- %, 15
- &, adres, 44
- &, bitsgewijze AND, 17
- &, operator, 73
- &&, logische AND, 16
- ^, bitsgewijze EXOR, 17
- ~, bitsgewijze inverse, 17
- \0, nul-karakter, 13, 40, 50
- \\, backslash karakter, 13
- \n, newline karakter, 13
- \r, carriage return karakter, 13
- \t, horizontal tab karakter, 13
- |, bitsgewijze OR, 17
- ||, logische OR, 16

A

- aantal elementen bepalen, 34
- adres, 12
- alfabetische ordening, 41
- argument, 6
- argumenten
 - aan een C-programma, 62
- array, 31
 - als argument, 37
 - als parameter, 37
 - declaratie van, 31
 - van pointers, 60
 - vergelijken van, 39
- array element, 31
- assignment, 9
- associativiteit, 15
- ATmega, 19

B

- backslash, 13
- backslash karakter, 13
- bare metal, 3
- bepalen aantal elementen, 34
- bestand, 2
- bestandssysteem, 2
- besturingssyteem, 2
- besturingsteken, 83
- bibliotheek, 4
- bit, 20
- bitsgewijze operatoren, 17
- buffer overflow, 41
- buffering, 75

C

- Call by Reference, 57
- calloc, functie, 69
- carriage return, 13
- char, keyword, 11
- chip, 3
- command line argumenten, 62
- compile time, 34, 91
- compile-time, 3
- compiler, 3
- compileren, 3
- computer, 2
- concatenation, 41
- conditionele expressie, 23, 37
- const, keyword, 14
- constante, 12
- constante array, 33
- cross compiler, 3

D

- datatype, 10
- datatypeconversie, 19
- debuggen, 4
- declaratie, 6, 9
- double, 12
- dubbele dereferentie, 57
- dynamische geheugenallocatie, 67

E

eendimensionaal, 33
element, 31
embedded system, 3
enkelvoudige datatypes, 12
escape sequence, 13
executable, 3
expressie, 15

F

float, 12
floating point, 12
fractie, 16
free, functie, 67
functie, 4

G

geheugen, 2
geheugenallocatie, 67
generieke pointer, 47
getchar, 24

H

header-bestand, 5, 20
horizontal tab, 13

I

I/O, 2
I/O-adressen, 20
initialisatie, 6, 10
instructie, 2
int, keyword, 11
integer, 9
inverteren, 18
invoerbuffer, 75
ISA, 3

K

karakterconstante, 13
kast-ongevoelig, 10
keyword, 6
kolomnummer, 35

L

leap year, 17
library, 4
linken, 4
logische operatoren, 16
long, keyword, 11
long double, 12
long long, keyword, 11

M

machinecode, 3
macro, 68
main, 6
malloc, functie, 67
malloc.h, header-bestand, 67

memory leak, 68
microcontroller, 3
modulus operator, 15

N

negatieoperator, 17
newline, 13, 24, 37
niet-nul, 17
nul-bytes, 67, 69
nul-karakter, 13, 40, 50
NULL-pointer, 46, 67

O

one's complement, 18
operating system, 2
overflow, 20

P

pointer, 12, 43
 als functie-argument, 53
 als return-waarde, 56
 array van, 60
 naar array, 48
 naar array van pointers, 61
 naar functie, 64
 naar pointer, 57
 naar vast adres, 67
 naar void, 47
 NULL, 46
 rekenen met, 51
postfix, 22
prefix, 13, 22
processor, 2
programma, 2
programmeren, 3
promotie, 19

Q

qualifier, 11
quicksort, 66

R

RAM, 2
realloc, functie, 69
relatieve operatoren, 16
rijnummer, 35
ROM, 2
runtime, 3

S

samengesteld datatype, 31
schrikkeljaar, 17
schuifoperator, 18
secundair geheugen, 2
short, keyword, 11
shortcut evaluation, 17
sizeof, operator, 21, 34
software, 2

- spatie, 37, 73
- standard library, 4
- statement, 5
- strcat, functie, 41
- strcmp, functie, 41
- strcpy, functie, 41
- string, 4, 6, 13, 31, 40, 50
- string array, 13
- string constante, 13
- string.h, header-bestand, 41
- strlen, functie, 41

T

- tekenbit, 18
- toekennen, 9
- toekenning, 6

- toolchain, 3
- tweedimensionaal, 35
- type cast, 19

U

- uitvoerbaar bestand, 3
- unaire operator, 15
- underscore, 10
- unsigned, 11

V

- variabele, 5, 6, 9
- void-pointer, 47, 67
- voorrangsregels, 24

W

- waarheidstabel, 17