ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Физический факультет

К. Ф. ЛЫСАКОВ

ОСНОВЫ ПРОГРАММИРОВАНИЯ

Учебное пособие

Новосибирск
2010

УДК
ББК
Л 886

Лысаков К. Ф. Основы программирования: Учеб. пособие / Новосиб. гос. ун-т. Новосибирск, 2010. 104 с.

В пособии рассмотрены основы структурного программирования. При этом основной упор делается на алгоритмическое решение задачи, а программная реализация рассматривается лишь как завершающий этап.

В качестве языка программирования используется смешение языков C и C++, что позволяет облегчить восприятие языка программирования.

Рецензент
доц., зам. зав. кафедрой АФТИ ФФ НГУ М. Ю. Шадрин

Учебное пособие подготовлено в рамках реализации программы развития НИУ-НГУ на 2009-2018 гг.

ОГЛАВЛЕНИЕ

1. ВВЕДЕНИЕ

1.1. Понятие программирования

Программирование — это процесс создания (разработки) программы, который может быть представлен последовательностью следующих шагов:

1. Спецификация (определение, формулирование требований к программе).
2. Разработка алгоритма.
3. Кодирование (запись алгоритма на языке программирования).
4. Отладка.
5. Тестирование.

В большинстве своем, различные самоучители и другие печатные издания, во главу угла ставят изучение какого-либо языка программирования, на примере которого решают задачи. При этом, для людей впервые встречающихся с программированием, наибольшую сложность представляют именно два первых пункта.

В данном учебном пособии, основной упор делается именно на постановку задачи и на ее алгоритмическое решение. Если алгоритмическое решение составлено верно, то кодирование может происходить на любом языке программирования.

В качестве языка программирования, в пособии рассмотрен симбиоз языков C и C++, которые позволяет описывать программные реализации значительно проще чем классический язык C, но при этом не использованы основы объектно-ориентированного программирования (классы, полиморфизм, перегрузка и наследование).

1.2. Описание алгоритмов решения

Для решения любой задачи, необходимо выполнить следующие этапы:

1. Четко определить условия задачи, входные данные и какой результат должен быть получено после решения задачи.
2. Какие дополнительные данные необходимы для решения задачи.
3. Составить блок-схему решения задачи и записать ее в виду удобного описания.
4. Анализ всех возможных проблем и усовершенствование алгоритма.

Рассмотрим на примере, каким образом выполняются перечисленные этапы при решении задачи вычисления корней квадратного уравнения.

1.2.1. Этап первый: постановка задачи

Для определенности будем решать уравнение следующего вида:

$$ax^2 + bx + c = 0$$

Таким образом, входными данными для нашей задачи являются три коэффициента: a, b и c.

Решение задачи предполагает вычисление возможных корней уравнения. При этом, так как количество корней возможно от 0 до 2, то необходимо в качестве решения указать количество корней и собственно перечислить их.

1.2.2. Этап второй: дополнительные данные

При решении квадратного уравнения необходимо вычислить значение дискриминанта. В нашей задаче, дискриминант является промежуточным результатом, необходимым для решения задачи.

1.2.3. Этап третий: составление блок-схемы решения

Описание блок-схемы может производиться с помощью различных средств и обозначений. Основной принцип заключается в наглядности шагов исполнения и однозначности переходов при ветвлении.

На рисунке (рис. 1) приведена основная блок-схема, к которой чаще всего приходят студенты при решении квадратного уравнения.

Рис. 1. Блок-схема решения квадратного уравнения

1.2.4. Этап четвертый: корректность и усовершенствования

Данный этап предполагает анализ созданной блок-схемы решения задачи. При этом необходимо ответить на два вопроса:

1. Будет ли схема корректно работать во всем диапазоне входных параметров?
2. Возможны ли оптимизации, которые позволят ускорить процесс выполнения задачи?

При анализе совершаемых действий, очевидно, что при значении коэффициента **a = 0**, происходит деление на **0!** В качестве решения данной ситуации можно предложить два основных метода:

- Добавить проверку корректности введенных пользователем значений, и при вводе **a=0**, выдавать сообщение об ошибке: «Уравнение не является квадратным!».
- Допустить возможность решения линейных уравнений, по сути, расширив диапазон применения вашей реализации. Для этого необходимо добавить такую проверку до вычисления дискриминанта, и идти описанным путем лишь при коэффициенте **a** отличном от 0, иначе добавить еще одну ветвь исполнения.

Что касается оптимизации, то основной ее смысл в удалении лишний действий, а также в недопустимости совершения дублирующий действий. На приведенной выше схеме, дублированным действием является извлечение квадратного корня из дискриминанта.

1.3. Отладка программной реализации

После создания программной реализации, нередки случаи когда ее функционирование отличается от запланированного. Для выяснения причин некорректного поведения существуют два основных метода отладки:

- метод отладочной печати,
- метод пошаговой отладки.

1.3.1. Метод отладочной печати

Одним из основных средств отладки является отладочная печать, позволяющая получить данные о ходе и состоянии процесса вычислений. Обычно разрабатываются специальные отладочные методы, вызываемые в критических точках программы - на входе и выходе программных модулей, на входе и выходе циклов и так далее. Искусство отладки в том и состоит, чтобы получить нужную информацию о прячущихся ошибках, проявляющихся, возможно, только в редких ситуациях.

Иногда пошаговая отладка невозможна. Например, программа может быть связана с внешним процессом, который не будет ждать, пока разработчик проверяет значения переменных. Другой весьма вероятный вариант — программа скомпилирована без отладочной информации с оптимизацией. В таких случаях единственным способом узнать, что происходит во время исполнения, остаётся вывод сообщений, по которым можно судить о состоянии программы.

Суть метода заключается в том, что программист вставляет в код программы вывод определенных сообщений, по которым впоследствии можно проанализировать какая именно ветвь программы исполняется, какие получаются промежуточные результаты и сравнить это с предполагаемым ходом выполнения и вычислениями.

1.3.2. Метод пошаговой отладки

Метод пошаговой отладки, заключается в том, что программист заставляет компьютер выполнять программу по шагам, инструкцию за инструкцией, а сам следит за состоянием программы на каждом шаге. При поиске ошибки очень удобно видеть, какая строка программы сейчас

выполняется, и выполнять программу строка за строкой в соответствии со строками исходного кода.

Команды управления отладкой собраны в меню Debug. Для удобства разработчика на самые нужные команды назначены клавиши.

Программа запускается в отладочном режиме командой Debug (при стандартных настройках клавиша F5). При этом за работой программы следит отладчик, переводящий исполнение в пошаговый режим в двух случаях: при возникновении ошибки, которая в обычном режиме привела бы к аварийному завершению программы, или по достижении точки останова.

Точка останова (англ. ***Breakpoint***), может быть установлена в любой строке программы, содержащей выполняемую инструкцию. Это может быть начало цикла, вызов функции, выражение или даже фигурная скобка, закрывающая блок. При стандартных настройках точка останова устанавливается или снимается нажатием клавиши F9.

После того как отладчик перевёл исполнение в пошаговый режим, можно заняться непосредственно отладкой.

Для того чтобы увидеть значение переменной, существующей в текущем контексте, достаточно навести указатель мыши на её идентификатор в исходном тексте. Этой же цели служит окно ***Watch***: впишите в левую колонку интересующие вас идентификаторы и их значения будут показаны всё время, пока идентификаторы имеют смысл.

Две команды пошагового исполнения позволят вам проконтролировать работу интересующего вас фрагмента кода:

- ***Step Over*** выполняет очередную строку и останавливает программу на следующей строке. Если текущая строка содержит вызов функции, то данная команда выполнит этот вызов в обычном режиме, не показывая по шагам подробности внутренней работы вызываемой функции.

- ***Step Into*** заставляет отладчик войти в функцию, вызов которой находится в очередной строке. Когда интересующий вас фрагмент пройден, можно выйти из пошагового режима командой ***Debug*** (***F5***). Исполнение программы продолжится до следующей точки останова или критической ошибки.

1.4. Работа с проектами в Microsoft Visual C++

В данном пособии в качестве среды разработке рассмотрена MSVS (Microsoft Visual Studio), как наиболее дружелюбная системы для разработки программ на языках C и C++.

1.4.1. Общие сведения

Для того, чтобы писать программу, вам *необходимы* проект (*project*) и рабочее пространство (*solution*).

Проект — это специальный файл, имя которого имеет расширение *vcproj*. В проекте содержится информация о том, из каких исходных файлов строится программа. Всё, что перечислено в проекте, будет скомпоновано в один целевой модуль (обычно исполняемый файл — **.exe*, но есть и другие варианты); если вам нужно получить два исполняемых файла, придётся сделать несколько проектов.

Среда разработки следит за изменениями файлов, перечисленных в файле проекта, и при сборке проекта компилирует заново только файлы, изменённые после предыдущей компиляции.

Компилятор обрабатывает только те файлы, которые перечислены в файле проекта. При этом совершенно всё равно, как называется папка в проекте — *Source files* или *Include files*: инструмент для обработки файла выбирается, исходя из расширения имени файла, а папки вы можете создавать для собственного удобства в любом количестве.

Рабочее пространство — это специальный файл, имеющий расширение *sln*, который описывает зависимости между отдельными

проектами, входящими в рабочее пространство: разрабатывая большую программу, состоящую из ряда модулей, удобно собрать несколько проектов в одно пространство и получать последнюю версию всех модулей нажатием одной клавиши!

Visual Studio устроена так, что проект обязательно должен содержаться внутри рабочего пространства. Если у вас уже есть рабочее пространство, то открывать в *Visual Studio* нужно его, а не проект или файл с исходным текстом!

Файл рабочего пространства создаётся автоматически, когда вы создаёт новый проект.

1.4.2. Создание проекта

Visual Studio предоставляет большое количество шаблонов проектов для разных целей. Полный список шаблонов, сгруппированных по категориям, можно увидеть, выбрав в меню:

File->New->Project... (рис. 2).

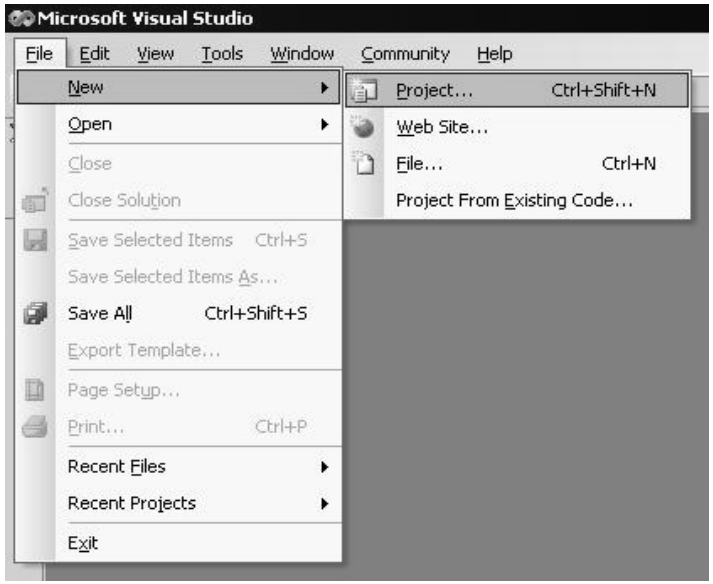


Рис. 2. Создание нового проекта

В зависимости от выбранного шаблона, среда добавляет во вновь создаваемый проект те или иные файлы, выполняя за разработчика часть рутинной работы по первоначальной настройке рабочего окружения.

Для большинства учебных заданий идеальным стартом является пустой проект консольного приложения. Для того чтобы создать такой проект, нужно выбрать подкатегорию *General* в категории проектов *Visual C++*. В окне выбора шаблона выбрать пустой проект *Empty project* (рис. 3).

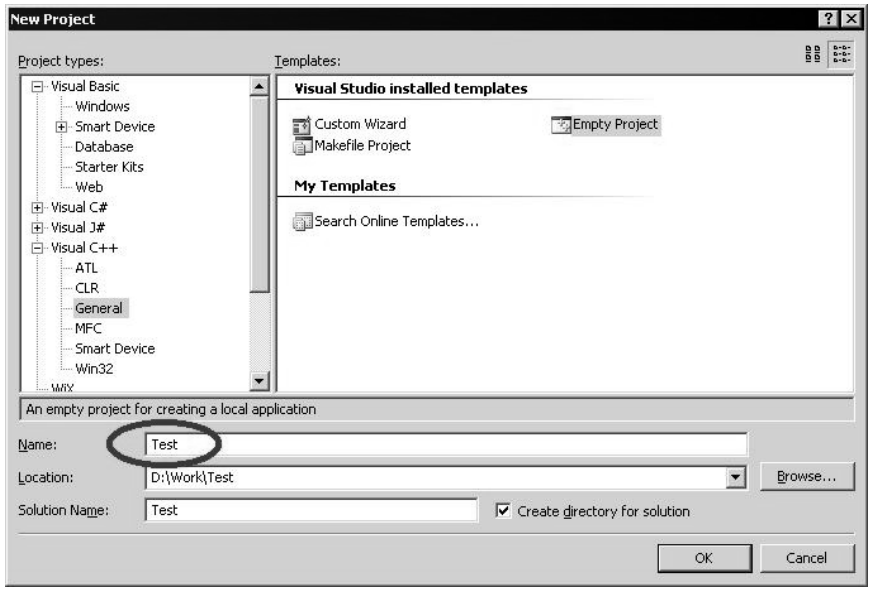


Рис. 3. Выбор пустого проекта

Новому проекту нужно дать имя. Правила для имён проектов такие же, как для имён файлов, но с русскими буквами или пробелами могут возникнуть проблемы. Имя проекта вводится тут же в поле Name (обведено на рисунке). После этого можно нажать OK, и Visual Studio создаст файлы рабочего пространства и проекта. Рабочее пространство получит то же имя, что и проект.

Ниже (рис. 4) показано, как выглядит новое рабочее пространство после добавления пустого проекта. Окно, отображающее структуру рабочего пространства, называется Solution Explorer, его можно открыть через меню View->Solution Explorer.

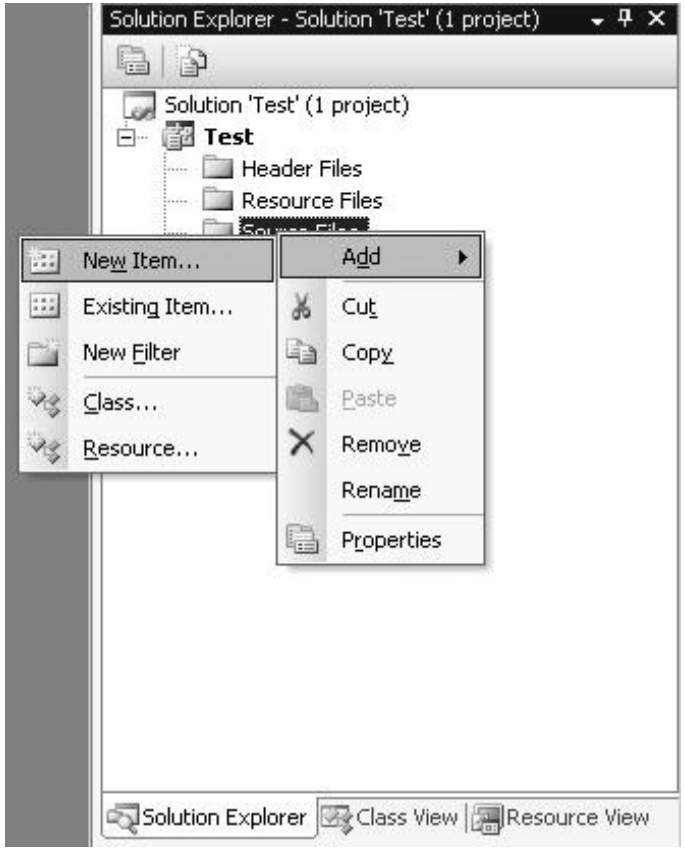


Рис. 4. Добавление файлов в проект

Как видно, проект *Test* не содержит никаких исходных текстов; для того чтобы начать программировать, нужно добавить в проект файл исходного кода. Откройте панель структуры проекта (*Solution Explorer*),

щёлкните правой кнопкой мыши на папке *Source Files* и выберите в меню *Add->New Item...*, затем выберите категорию *Code*, в ней выберите *C++ File* и впишите имя нового файла (рис. 5).

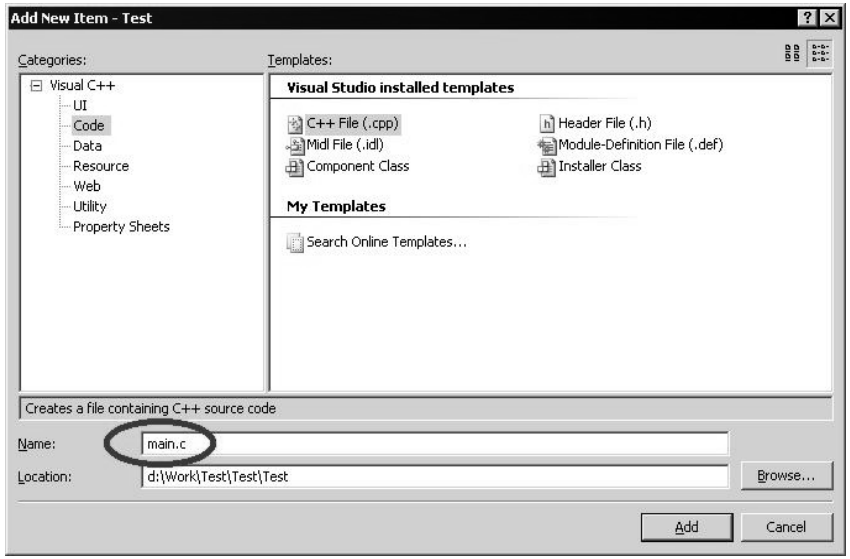


Рис. 5. Добавление файла программы на языке C

Важно явно указать расширение имени: «*.c*» для программы на *C* и «*.cpp*» для программы на *C++* (обведено на рис. 4). Если расширение не указано, *Visual Studio* автоматически добавит расширение «*.cpp*»; такой файл будет обрабатываться компилятором *C++*, что может помешать, если вы пишете программу на *C*.

Введите имя нового файла, нажмите *Add*, и вы увидите, что в списке файлов проекта появился новый элемент. Файл автоматически откроется в редакторе — можно начинать программировать (рис. 6).

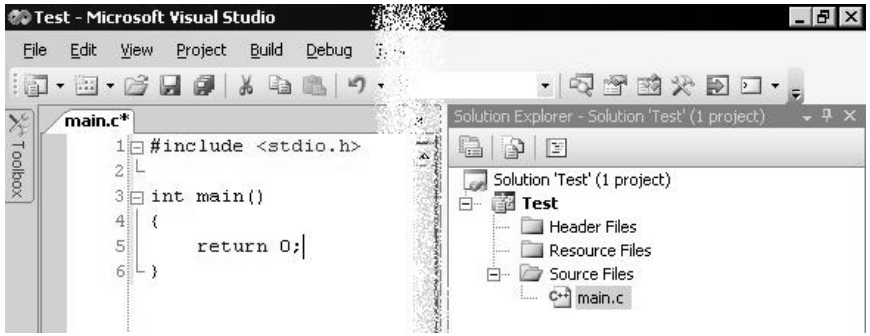


Рис. 6. Вид рабочего проекта

1.4.3. Сборка проекта

После того как программист написал некоторое количество исходного кода, у него обычно возникает желание посмотреть, как этот код работает. Для этого нужно скомпилировать и скомпоновать программу.

Компилятор и компоновщик запускаются из меню *Build->Build Solution* (или нажатием соответствующей «горячей клавиши»). Кроме того, можно скомпилировать только тот файл, который вы в данный момент редактируете (*Build->Compile*), однако в этом случае запустить исполнение кода не получится: для этого необходимо скомпилировать и скомпоновать весь проект.

Если всё получилось (в коде нет ошибок), то результатом работы компилятора и компоновщика будет исполняемый файл (с расширением *exe*), который можно будет запустить. Запускать можно прямо из среды разработки: выбор *Debug->Start Without Debugging* запускает программу в виде самостоятельного процесса, *Debug->Start Debugging* запускает программу в режиме отладки.

1.4.4. Конфигурации проектов

В реальной жизни разработчику часто бывает необходимо иметь несколько вариантов одной и той же программы: одну — с отладочной информацией и без оптимизации, для рабочей отладки, другую — без

отладочной информации, но удобную для тестирования (например, предварительно настроенную на тестовые данные), третью — оптимизированную, готовую для демонстрации заказчику и т. д.

Очевидно, все эти версии программы должны строиться из одних и тех же исходных файлов, иначе работа программиста превратится в кошмар. Возможность настроить различные варианты сборки программы дают *конфигурации*.

Конфигурация — это набор настроек, которому дано имя. Конфигурация определяется в рабочем пространстве, но распространяет влияние также и на все проекты, входящие в пространство. На уровне рабочего пространства конфигурация определяет, во-первых, какие проекты необходимо собирать, а во-вторых, куда помещать основной вывод (*Primary Output*) этих проектов. На уровне проекта можно для каждой возможной конфигурации задать полный набор настроек.

По умолчанию новое рабочее пространство имеет две конфигурации: *Debug* и *Release*. Настройки этих конфигураций соответствуют названиям: *Debug* предназначена для сборки отладочной версии, *Release* — для рабочей. Соответственно отладочная версия содержит полную отладочную информацию, код не оптимизируется компилятором, объявлена директива препроцессора *_DEBUG*. В рабочей версии всё наоборот: код оптимизируется, отладочная информация не генерируется, объявлена директива препроцессора *_RELEASE*.

1.4.5. Файловая структура рабочего пространства

Когда вы создаёте новый проект с именем, например, *sample*, *Visual Studio* создаёт следующие директории и файлы:

- [sample] — корневая директория рабочего пространства. Всё, что относится к рабочему пространству и вложенным в него проектам, создаётся и содержится в этой директории, если ничего не менять в настройках. Все директории и файлы, перечисленные ниже, создаются в корневой директории;

- [sample \ debug] — директория, в которой создаются результаты сборки проектов, входящих в рабочее пространство *sample*, при выборе конфигурации *debug*. Как правило, это основной вывод (*Primary Output*) проектов — исполняемые файлы (**.exe)*, базы данных для отладки (**.pdb*) и некоторые другие файлы. После того как вы закрыли среду разработки, можно удалить эту директорию со всем содержимым, потому что она будет создана заново при следующей сборке проектов, входящих в рабочее пространство;
- [sample \ release] — аналогично [sample \ debug], но для конфигурации *release*;
- Sample \ sample.sln — рабочее пространство *sample*;
- Sample \ sample.ncb, sample \ sample.suo — служебные файлы Visual Studio, относящиеся к рабочему пространству *sample*. Эти два файла генерируются средой разработки, и их тоже можно удалить, если свободного места на диске остро не хватает, закрыв предварительно соответствующее рабочее пространство;
- [sample \ sample] — корневая директория проекта *sample*, входящего в рабочее пространство *sample*. Все файлы, относящиеся к проекту *sample*, по умолчанию создаются в этой директории;
  - [sample \ sample\debug] — директория, в которой создаются все промежуточные файлы при сборке проекта *sample* в конфигурации *debug*. После того как вы закрыли среду разработки, можно удалить эту директорию со всем содержимым, потому что она будет создана заново при следующей сборке проекта;
  - [*sample* \ *sample* \ *release*] — аналогично предыдущему, только для конфигурации *release*;
  - [*Sample* \ *sample* \ *sample.vcproj*] — файл проекта *sample*, входящего в рабочее пространство *sample*.

Все файлы с исходными текстами (**.c*, **.cpp*) и заголовочные файлы (**.h*) среда разработки тоже предлагает создавать в корневой директории соответствующего проекта.

1.5. Использование справочной системы

Справочная система — без преувеличения главный помощник программиста. В ней описаны стандартные функции, классы, интерфейсы и библиотеки. Кроме того, справочная система содержит весьма подробное описание языка программирования, включающее сведения о синтаксисе, типах данных, приоритетах операций и т. п.

Справку можно вызвать двумя способами: через меню *Help->Index* или непосредственно из текстового редактора: по любому слову в программе, будь то идентификатор стандартной функции, ключевое слово языка или директива препроцессора, установив курсор в середине слова и нажав *F1*.

Для ускорения загрузки рекомендуется выбирать локальные файлы справки.

Обычно справочная система предлагает несколько статей. Многие стандартные функции имеют различные реализации для разных платформ, поэтому, если в настройках не указаны предпочтения, вам предложат выбор доступных вариантов. Для функций стандартной библиотеки *C* обычно следует выбирать вариант *C Runtime Library*.

В большинстве статей, описывающих стандартные функции, приводится пример использования (обычно в конце статьи).

2. БАЗОВЫЕ КОНСТРУКЦИИ ЯЗЫКОВ С И С++

Данная глава посвящена описанию основных принципов построения программ, включая не только функциональное описание программной реализации, но и стиль написания.

2.1. Потоки ввода / вывода

Поток — это абстракция источника (поток ввода) или назначения (поток вывода) при последовательной передаче данных. Посредством различных реализаций потоков программа может взаимодействовать с файлами на жёстком диске, клавиатурой, экраном, внешними устройствами и даже памятью, используя унифицированный набор операций.

Для того чтобы пользоваться стандартными потоками, необходимо указать библиотеку, в которой находится реализация этих потоков. Такая библиотека имеет название **iostream** (сокрашщение от InputOutputStream – потоки ввода-вывода) и включается в текст программы следующим образом:

```
#include <iostream>
```

Здесь ***#include*** – служебная команда, означающая, что необходимо к тексту описываемой ниже программы включить указанную в угловых скобках библиотеку.

После включения таким образом библиотеки, пользователь может использовать стандартный ввод-вывод следующим образом:

```
std::cout << "Hello"
```

- **cout** – поток вывода. Обычно поток cout связан с экраном, и все что передается в поток печатается в текстовом виде на экране консоли**.**

- **std::** – префикс, означающий, что используется реализация из стандартной библиотеки.
- **<<** – в данном контексте, это операция берущая данные, указанные слева, преобразующая их в текстовый вид, и отсылающая, в поток, указанный слева**.**

Для упрощения операций ввода-вывода существует возможность включить в программу библиотеку **iostream** и сразу указать, что все команды будут по умолчанию использоваться из стандартной библиотеки. Для этого необходимо написать следующим образом:

```
#include <iostream>
using namespace std;
```

В этом случае, префикс std:: можно опустить, и команда вывода строк на экран будет выглядеть следующим образом:

```
cout << "Hello"
```

2.2. Переменные

Переменная — это именованная область памяти, в которую могут быть записаны различные значения. Также из этой области памяти может быть извлечено значение переменной, используя ее имя. В каждый момент времени переменная может иметь только одно значение.

Значения, которые может хранить переменная, определяется ее **типом**.

В простейшем виде переменную можно определять следующим образом:

Тип список_имен_переменных

Тип переменной определяет значения, которые может принимать переменная.

Типы могут быть следующими:

char — целое значение, 8 бит (диапазон от –128 до 127);

int — целое значение, обычно 4 байта, зависит от платформы;

float — вещественные числа;

double— вещественные числа удвоенной точности.

Каждый из целочисленных типов может быть определен как знаковый **signed** либо как беззнаковый **unsigned** (по умолчанию **signed**).

Пример определения переменных:

```
char my_symbol
int val, val2
double Result
```

После того как переменная создана, ей можно присваивать значения. Это можно сделать как при определении переменной, так и после этого:

Тип имя_переменной = начальное_значение;

имя_переменной = начальное_значение;

Например:

```
int val = 5
val = 5;
```

Необходимо помнить, что задавать значения переменной можно только после того, как эта переменная создана. А если значение явно не определено при создании, то по умолчанию оно ничему не присваивается. То есть попытка вывести значения неинициализированной переменной на экран может привести либо к выводу «мусора», либо даже к ошибке выполнения программы (в зависимости от среды выполнения).

2.3. Структура программы

Классической первой программой, которую обычно пишут, является вывод на экран «Hello World!». Ниже приведен код программы, которая это делает.

```
1 #include <iostream>
2 using namespace std;
3
4 void main()
5 {
6    cout << "Hello World!";
7 }
```

Разберем подробно структуру программы.

В строках 1 и 2 происходит подключение стандартной библиотеки для использования потоками ввода-вывода

Строчка 3 остается пустой для большей наглядности программы

В строке 4 описывается функция void main(). Подробное изучение функций происходи ниже, поэтому пока такое описание необходимо принять за аксиому.

В 5 строке открывается фигурная скобка, которая обозначает начало функции main, а в 7 строке фигурная скобка закрывается, обозначая что функция закончена. Внутри этих скобок собственно и происходит описание функции.

Функция main является главной функцией программы, так как только она исполняется в программе. То есть при запуске программы, происходит выполнение тех команд, которые написаны в этой функции. В описанном примере это одна команда – вывод на экран текстового сообщения.

Внутри функции main (а также всех других функций), каждая строка должна заканчиваться символом **;**.

Написанные код, как видно, имеет удобное форматирование, удобное для чтения. Форматирование производится командой табуляции (кнопка **Tab** на клавиатуре). Но при этом, если написание программы происходит построчно, то при переходе на следующую строчку (кнопка **Enter** на клавиатуре), текст форматируется автоматически.

2.4. Ввод и вывод переменных

При работе программы, зачастую необходимо чтобы пользователь сообщал программе различные параметры, а программа, по ходу выполнения, выдавала результат вычислению пользователю.

В языке программирования С, для вывода значений переменных на консоль, необходимо дополнительно указывать тип данных. Поэтому вывод переменных выглядит следующим образом:

```
int a = 4;
char ch = 'Q';
double pi = 3.141592;

printf("%d", a);
printf("%c", ch);
printf("%f", pi);
```

Помимо этого, при вводе значений переменных с консоли, программист должен априорно быть знакомым с понятиями ссылок и указателей на объкты, а также помнить, что для дробных чисел ввод и вывод осуществляются по-разному.

```
double pi;
scanf("%lg", &pi);
```

Язык программирования С++, позволяет значительно упростить операции ввода-вывода значений переменных, за счет использования потоков.


```
int a = 4;
char ch = 'Q';
double pi = 3.141592;

cout << a;
cout << ch;
cout << pi;
```

Ввод значений переменных также осуществляется без дополнительной информации о переменной.

```
double pi;
cin >> pi;
```

- **cin** – поток ввода. Обычно поток cin связан с клавиатурой, и весь набираемый в консоли текст, хранится в потоке.
- **>>** – в данном контексте, это операция извлекающая данные из потока справа, преобразующая их из текстового вида, и складывающая их в переменную, указанную слева.

При операциях вывода значений переменных возможны комбинации переменных и текстовых выражений. При операциях ввода, такие средства не предусмотрены.

Ниже приведен фрагмент кода, запрашивающий у пользователя значение переменной, и выводящей ее квадрат.

```
int a;
int Res;

cout << "Input value a = ";
cin >> a;

Res = a*a;

cout << "a*a = " << Res << endl;
```

Обратите внимание, что символы >> и << показывают направление потока – в переменную, или из нее на консоль.

Специальная команда **endl** означает перевод строки.

2.5. Арифметические операции и их использование

Для вычислений выражений в языках C и C++ могут использоваться различные арифметические операции. При этом каждая операция имеет свой ранг, определяющий приоритет ее выполнения. Чем ниже ранг операции, тем больший приоритет имеет операция. Операции одного ранга исполняются согласно правилам ассоциативности либо слева направо (–>), либо справа налево (<–). В таблице разобраны типовые операции по рангам и тип их ассоциативности:

Ранг	Операции	Ассоциативность
1	() [] –> .	→
2	! ~ + – ++ -- & * (тип) sizeof()	←
3	* / % (мультипликативные бинарные)	→
4	+ – (аддитивные бинарные)	→
5	<< >> (поразрядного сдвига)	→
6	< <= >= > (отношения)	→
7	== != (отношения)	→
8	& (поразрядная конъюнкция «И»)	→
9	^ (поразрядное исключающее «ИЛИ»)	→
10	| (поразрядное дизъюнкция «ИЛИ»)	→
11	&& (конъюнкция «И»)	→
12	|| (дизъюнкция «ИЛИ»)	→
13	?: (условная операция)	←
14	= *= /= %= += –= &= ^= |= <<= >>=	←
15	, (операция «запятая» – перечисление)	→

2.5.1. Выражения и приведение арифметических типов

Каждое выражение состоит из одного или нескольких операндов, символов операций и ограничителей, в качестве которых чаще всего выступают круглые скобки ( ). Назначение любого выражения — формирование некоторого значения. Тип результата определяется типом выражений. Если значениями выражения являются целые и вещественные типы, то говорят об арифметических выражениях, в которых допустимы следующие операции:

+ — сложение;

– — вычитание;

* — умножение;

/ — деление;

% — получение остатка от целочисленного деления.

Примеры выражений с двумя операндами:

```
A + b
12.3 - x
3.14169 * Z
e / 8
8 % i
```

Также существуют специфичные унарные операции **++** и **--** для изменения операнда на 1. При этом эти операнды могут применяться как после операнда (a++), так и до него (--x). При этом результат будет различный:

- Если изначально a было равно 6. То после выполнения z = a++, результатом будет: z = 6; a = 7;
- Если изначально a было равно 6. То после выполнения z = --a, результатом будет: z = 5; a = 5.

2.5.2. Отношения и логические выражения

Отношение определяется как пара выражений, между которыми стоит знак операции отношения. Допустимы следующие операции:

== — равно (два знака «равно»);

!= — не равно;

> — больше;

< — меньше;

>= — больше или равно;

<= — меньше или равно.

Примеры отношений:

```
a - b > 6.8
(x - 8) * 3 == 15
5 >= 1
```

Логический тип (истина или ложь) в языке С отсутствует. Поэтому принято, что отношение имеет ненулевое значение (обычно 1), если оно истинно, и равно 0, если оно ложно.

Логических операций в языке С три:

!	— отрицание	— логическое НЕ
&&	— конъюнкция	— логическое И
||	— дизъюнкция	— логическое ИЛИ

Так как значением отношения является целое, то ничего не противоречит применению логических операций к целочисленным значениям. При этом любое положительное ненулевое значение будет восприниматься как истинное. Иначе говоря значением !0 будет 1, а !54 будет 0.

2.5.3. Приведение типов

Особенность языков С и С++ заключается в том, что при вычислении значений выражений, результат всегда приводится к типу первого операнда. То есть при делении двух целых значений результат будет также иметь тип целого числа, поскольку каждый из операндов представлен целым типом.

```
int a = 5;
int b = 2;
double res;

res = a / b;     // Результатом будет 2
```

Для получения вещественного типа необходимо, чтобы хотя бы один операнд имел тип вещественного числа. Для этого можно использовать приведение типов. В коде программ достаточно часто приводится явное приведение типов, что не влечет изменение типа самой переменной. Рекомендуется в случае подобных вычислений приводить к желаемому типу первый операнд, что упрощает чтение и последующую отладку программ.

```
int a = 5;
int b = 2;
double res;

res = (double)a / b;
```

2.5.4. Выражения с поразрядными операциями

Такие операции позволяют конструировать выражения, в которых обработка операндов выполняется побитно (в двоичном представлении числа). Возможны следующие операции:

~ — инвертирование битов; // ~170 равно 85

>> — сдвиг последовательности битов вправо; // 100 >> 2 равно 25

<< — сдвиг последовательности битов влево // 5 << 2 равно 20

^ — поразрядное исключающее ИЛИ;

| — поразрядное ИЛИ;

& — поразрядное И;

Операции побитового сдвига часто применяются в программировании, поскольку позволяют значительно ускорять операции умножения или целочисленного деления на 2^n.

```
int a;

cout << "Input value a = ";
cin >> a;

cout << "a*16 = " << (a << 4) << endl;
cout << "a/8 = " << (a >> 3) << endl;
```

2.6. Операторы ветвления

В языках C и C++ существует два оператора ветвления:

- условный оператор if.
- переключатель switch.

2.6.1. Оператор if

Оператор **if** может использоваться либо для условного выполнения определенного действия, либо для разветвления программы на небольшое количество ветвей.

В случае условного исполнения, оператор имеет следующий вид:

```
if ( условие )
    оператор;
```


При этом, если необходимо исполнить несколько операторов, то их необходимо выделить в ТЕЛО, которое обозначается фигурными скобками.

```
if ( условие )
{
    оператор_1;
    оператор_2;
    оператор_3;
}
```

В случае ветвления исполнения, оператор имеет следующий вид:

```
if ( условие )
    оператор_1;
else
    оператор_2;
```

При этом возможно ветвление и на большее количество ветвей:

```
if ( условие_1 )
    оператор_1;
else if(условие_2 )
    оператор_2;
else
    оператор_3;
```

В качестве условия могут использоваться арифметические выражения, отношения и логические выражения (x > 10 && x < 87).

Приведем фрагмент программы для определения корней уравнения вида

$$ax^2 + bx + c = 0$$

```
D = b*b-4*a*c;

if ( D < 0)
   cout << "No roots" << endl;

else if(D == 0)
{
   x = (double)-b / (2*a);
   cout << "1 root: x = " << x << endl;
}

else
{
   x1 = ( - b + sqrt(D)) / (2*a);
   x2 = ( - b - sqrt(D)) / (2*a);
   cout << "2 root: x1 = " << x1 << " x2 = " << x2;
}
```

2.6.2. Переключатель switch

Данный оператор используется для организации мультиветвления следующим и выглядит следующим образом:

switch (выражение)

{

case Константа_1: операторы_1;

case Константа_2: операторы_2;

….

default: операторы;

}

При первом совпадении значения выражения с константой происходит выполнение операторов, помеченных данной меткой. Если после их выполнения не предусмотрено никаких операторов перехода, то выполняются также все последующие операторы. То есть по сути, CASE

является меткой, обозначающей место выполнения программы после SWITCH.

Оператором перехода может быть **break**, который осуществляет выход из тела (фрагмент кода, обозначенный фигурными скобками).

Переключатели чаще всего используются, когда количество ветвей больше 3, либо необходимо перебрать заданные константы.

Для иллюстрации приведен фрагмент программы, которая выводит на экран название введенной цифры от 0 до 9:

```
int a;
cout << "Input a = ";
cin >> a;

switch (a)
{
  case 1: cout << a << " - One" << endl;
    break;
  case 2: cout << a << " - Two" << endl;
    break;
  case 3: cout << a << " - Three" << endl;
    break;
  case 4: cout << a << " - Four" << endl;
    break;
  case 5: cout << a << " - Five" << endl;
    break;
  case 6: cout << a << " - Six" << endl;
    break;
  case 7: cout << a << " - Seven" << endl;
    break;
  case 8: cout << a << " - Eight" << endl;
    break;
  case 9: cout << a << " - Nine" << endl;
    break;
  case 0: cout << a << " - Zero" << endl;
    break;
  default: cout << "The value is not from 0 to 9";
}
```

Еще раз отметим, что если бы в программе отсутствовали операторы break, то при вводе например цифры 6, на экран бы распечаталось следующее:

```
6 - Six
7 - Seven
8 - Eight
9 - Nine
0 - Zero
The value is not from 0 to 9
```

2.7. Операторы циклов

В языках C и C++ существуют три разных способа оргиназации циклов. Ниже мы их разберем с примерами использования.

2.7.1. Цикл for

Оператор **for** является, пожалуй базовым оператором для организации циклов. Цикл **for** — это параметрический цикл, то есть имеется возможность задавать параметры для выполнения цикла: начальные значение и условия. Форма записи выглядит следующим образом:

for (выражение_1; условие_цикла; выражение_2)
Тело цикла

выражение_1 – определяет действия, выполняемые до начала цикла.

условие_цикла – обычно логическое или арифметическое условие. Пока это условие истинно, выполняется цикл. Проверка происходит перед началом очередной итерации тела цикла.

выражение_2 — это действие, выполняемые в конце каждой итерации цикла.

.

В качестве примера рассмотрим программу, распечатывающую все числа, которые делятся на 2 до введенного пользователем значения.

```
int i, Max;

cin >> Max;

for(i = 0; i <= Max; i++)
{
   if(i%2 == 0)
      cout << i << endl;
}
```

Выражения, указывающиеся в скобках оператора **for** не являются обязательными. Ниже приведен полностью идентичный код, выполняющий точно такой же цикл.

Помните, что даже если программист не указывает внутри оператора for какие-либо выражения, все 3 поля должны быть оставлены пустыми!

```
int i, Max;
cin >> Max;
i = 0;
for(;;)
{
   if(i <= Max)
   {
      if(i%2 == 0)
      cout << i << endl;

      i++;
   }
   else
      break;
}
```

2.7.2. Цикл while

Цикл **while** — это цикл с предусловием, то есть перед каждой итерацией проверяется условие, и только если оно истинно, то происходит выполнение тела цикла. Он имеет следующий вид:

while (выражение условия цикла)

Тело цикла

Приведем фрагмент кода, который также распечатывает все числа, которые делятся на 2 до введенного пользователем значения

```
i = 0;
while(i <= Max)
{
   if(i%2 == 0)
      cout << i << endl;

   i++;
}
```

Как видно из примера, условие цикла записывается в более наглядном виде, чем при использовании **for**. Но при таком описании довольно часто возникают ошибки связанные с тем, что необходимо явно описывать инкрементацию счетчика цикла.

2.7.3. Цикл do-while

Цикл **do** — это цикл с постусловием, то есть проверка условия цикла происходит после каждой итерации. Это обеспечивает, по крайней мере, одну итерацию выполнения вне зависимости от условий. Он имеет следующий вид:

do

Тело цикла

while (выражение условия цикла);

2.8. Массивы данных

В случае, если программы обрабатывает множество однотипных данных, именовать каждую переменную не только неудобно, но порой и просто невозможно из-за их большого количества. Для обработки однотипных данных применяются ***массивы***, которые позволяют компоновать переменные.

Наиболее наглядными примерами применения массивов являются задачи статистической обработки данных. Также можно упомянуть обработку векторов различной размерности.

Для определения массива необходимо указать, какой тип будут иметь его элементы, и количество этих элементов:

```
int mas[10];
```

Таким образом, определен одномерный массив, состоящий из целых чисел, имеющий 10 элементов. Элементы в массиве нумеруются начиная с 0, поэтому в приведенном примере массив содержит элементы с номерами от 0 до 9. Обращаться к его элементам следует следующим образом:

mas[i]

Ниже приведен фрагмент кода, в котором пользователь задает 10 значений, а программа выводит их сумму.

```
int mas[10];
int i;

for(i=0; i < 10; i++)
{
   cout << "a[" << i << "] = ";
   cin >> mas[i];
}

int Sum = 0;
for(i=0; i < 10; i++)
```

```
        Sum += mas[i];

        cout << "Summa = " << Sum << endl;
```

Из примера видно, что работа с элементами массивов мало чем отличается от работы с любыми другими переменными.

2.9. Функции

Теперь некоторые пояснения к самой программе. Любая программа на языке С, состоит из одной или более "***функций***", указывающих фактические операции компьютера, которые должны быть выполнены. По своей сути, в функции перечисляется порядок действий, а также заводятся переменные для хранения данных.

Функция с именем **main** – особенная. Именно выполнение этой функции и происходит при запуске любой программы. Это означает, что каждая программа должна в каком-то месте содержать функцию с именем **main**. Для выполнения определенных действий функция **main** обычно обращается к другим функциям, часть из которых находится в той же самой программе, а часть - в библиотеках, содержащих ранее написанные функции.

Действия, выполняемые при обращении к функции, задает ее тело, которое выделяется фигурными скобками **{ }**. Структура определения функции имеет вид:

Тип_результата Имя функции (список_параметров);

Реализация функции происходит следующим образом:

Тип_результата Имя_Функции (список_параметров)
{
  ...
}

Необходимо помнить, что в отличие от объявления переменных, в списке параметров функции недопустимо перечислять несколько имен переменных одного типа – для каждой переменной должен быть явно указан тип.

Реализация функции всегда должна происходить после ее объявления. В ряде случаев (например когда вся программа состоит только из одного файла), допускается не делать отдельных объявлений функций.

Помните, что объявление (или реализация, если объявления не существут) должно производится до того места, где функция вызывается! То есть если создается пользовательская функция, которая будет вызываться из функции **main**, то она должна быть описана до функции **main**.

Если не указан тип возвращаемого значения, по умолчанию считается, что функция возвращает тип ***int***. Для возврата значения используется команда **return** (она может быть выполнена в любом месте кода тела, но выполнение функции после этого заканчивается). Если функция не возвращает никаких значений, то используется специальный тип **void**.

В качестве примера рассмотрим программу, в которой описана пользовательская функция для вычисления дискриминанта квадратного уравнения, вида

$$ax^2 + bx + c = 0:$$

```
#include <iostream>
using namespace std;

double Discr_Calc(int a, int b, int c)
{
   double D;
   D = b*b - 4*a*c;
   return D;
}

void main()
{
```

```
    int a, b, c;
    double D;
    a = 1;
    b = 5;
    c = 3;

    D = Discr_Calc(a, b, c);

    cout << "D = " << D << endl;
}
```

Функция может возвращать только одно значение.

Обычно функции используются для написания более простых для чтения программ. Так программа, которая содержит только функцию **main()**, состоящую из 13 окон теста, очень трудна для понимания, тем более сторонним человеком. Поэтому часто используется негласное соглашение, что текст любой функции не должен превышать по размеру одного экрана — размера, которые человек может охватить взглядом. При выборе имен функций программист также должен руководствоваться тем принципом, чтобы постороннему человеку стало интуитивно понятно, что делает функция.

2.10. Локальные и глобальные переменные

Все создаваемые в программе переменные, имеют определенную область видимости. Для переменных выделяют две основных области видимости: **локальные** и **глобальные**.

2.10.1. Глобальные переменные

Из названия этого класса переменных становится понятно, что они доступны всем. Под всеми подразумеваются все функции проекта.

Для того чтобы переменная была глобальная, она должна быть объявлена вне функций. Ниже приведен текст программы, в которой

переменная с именем D является глобальной. При этом к ней обращаются две различные функции, для которых эта переменная является общей.

```
#include <iostream>
using namespace std;

double D = 0;

void Discr_Calc(int a, int b, int c)
{
  D = b*b - 4*a*c;
}

void main()
{
  int a, b, c;
  a = 1; b = 5; c = 3;

  Discr_Calc(a, b, c);
  cout << "D = " << D << endl;
}
```

2.10.2. Локальные переменные

Локальные переменные обладают ограничением области видимости – они видны и доступны только внутри того блока, в котором они созданы. Важно помнить, что локальные переменные перестают существовать при выходе из блока, в котором они были объявлены.

Блоками, ограничивающими видимость, могут являться функции, либо просто набор действий, выделенных в отдельное тело.

Таким образом, если бы в тесте предыдущей программы описать функцию вычисления дискриминантом следующим образом:

```
void Discr_Calc(int a, int b, int c)
{
  double D;
  D = b*b - 4*a*c;
}
```

То в ней существует собственная локальная переменная D, которая перестает существовать, как только исполнение программы выходит из этой функции. Другими словами, локальные переменные перекрывают глобальные, в результате чего, выполнение описанный программы с такими изменениями приведет к тому, что на экран всегда будет выводиться НОЛЬ!

Ниже приведен пример еще одной неправильно написанной программы, которая не сможет быть даже скомпилирована.

```
#include <iostream>
using namespace std;

void Discr_Calc(int a, int b, int c)
{
   D = b*b - 4*a*c;
}

void main()
{
     int a, b, c;
     double D;
     a = 1;
     b = 5;
     c = 3;

     Discr_Calc(a, b, c);

     cout << "D = " << D << endl;
}
```

В приведенной программе ошибка заключается в том, что переменная с именем **D** определена в функции ***main()***, и не существует за ее пределами. Таким образом, в функции ***Discr_Calc()*** происходит обращение к несуществующей переменной **D**.

Переменные, перечисленные в списке параметров функции, также являются локальными для этой функции. Их отличие от описанных в теле функции только в том, что они будут гарантированно инициализированы к моменту вызова функции.

3. Работа с динамической памятью

Работа с динамической памятью позволяет программисту решать задачи с неизвестным количеством исходных данных, позволяя добавлять новые возможности в функционал программ.

Основными понятиями для работы с динамической памятью являются указатели и ссылки. Именно с помощью переменных этих типов осуществляется управление памятью, ее выделение и освобождение.

3.1. Указатели и работа с ними

Каждая переменная представляет собой именованный участок памяти ПК, в котором хранится ее значение. При этом, чтобы получить доступ к значению переменной, необходимо обратиться к ней по ее имени.

```
int a;
a = 1;
```

Для того чтобы получить значение адреса в памяти, по которому находится значение конкретной переменной, необходимо произвести **унарную операцию разыменования &.** Выражение **&a** позволяет получить адрес участка памяти, выделенного для переменной a.

```
int a;
a = 5;
cout << &a;
```

Имея возможность определять адрес переменной или другого объекта программы, нужно уметь его сохранять, преобразовывать и передавать. Для этих целей введены переменные типа «**указатель**».

Указатель — это переменная, значение которой является адресом объекта конкретного типа. Для обозначения значения указателя, который никуда не указывает, используется специальная константа **NULL**.

При определении указателя необходимо обозначать, на какой тип данных указывает эта переменная. Здесь и далее при объявлении указателей будем использовать символ «**p**» в названиях указателей. Это позволит легко отличать в тексте программы указатели, от собственно переменных. Примеры определения указателя и присваивания ему значения:

```
int Ch = 14;
int* pCh = NULL;
pCh = &Ch;
```

Чтобы получить значение, которое находится по указанному адресу, применяется операция разыменования – «*»:

```
int Ch = 14;
int* pCh = &Ch;
cout << *pCh;
```

3.2. Арифметика указателей и массивы

В языке С допустимы только две арифметические операции над указателями: суммирование и вычитание. При этом данные операции обеспечивают передвижение по памяти не по байтам и битам, а на размер типа данных, который указан при создании переменной указателя.

Указатели непосредственно связаны с массивами данных, поскольку при создании массива все его элементы создаются в памяти последовательно, т. е. массив представляет собой единую и неразрывную область памяти. При этом если обратиться к массиву через его имя, то результатом будет адрес первого элемента массива (имеющего индекс 0).

```
int mas[3] = {1, 2, 3};
int* pVal;
pVal = mas; // тождественно pVal = &mas[0]
cout << *pVal;
```

Для указателей определены арифметические операции. Таким образом, к указателям можно добавлять или отнимать целые значения. Так операция «++», примененная к указателю, изменяет адрес, хранящийся в указатели на число байт, соответствующее размеру одной переменной типа, соответствующую типа указателя. Иначе говоря, если у нас имеется указатель на тип ***int***, занимающий 4 байта, то операции ++ сместит указатель в памяти на 4 байта.

Поэтому для адресации внутри массива бывает удобно использовать указатели. В качестве примера приведем фрагмент кода для инициализации массива:

```
int mas[10], i;
int* pMas;

for(i = 0, pMas = mas; i < 10; i++)
   cin >> *(pMas+i);
```

2.3.1. Динамическая память. Массивы

До сих пор все массивы данных у нас были строго определенного размера. В некоторых задачах это бывает неудобно. Например, требуется, чтобы пользователь задал некий массив данных, при этом неизвестен его размер. Можно, конечно, заранее выделить под массив 100 МБайт памяти и считать, что этого всегда хватит. Но возможны два варианта:

- массив все-таки окажется недостаточным, поскольку данные занимают 101 Майт.
- Для конкретной задачи достаточно 10 элементов по 4 КБайт.

В обоих этих случаях необходимый объем памяти становится известен только на момент использования программы. Поэтому было бы намного удобнее, если бы программист имел возможность выделять необходимый объем памяти по требованию пользователя.

Такая возможность имеется и реализуется с помощью указателей и средств для динамического выделения памяти, из которых основными являются функции **new** и **delete**. Эти функции используются для выделения и освобождения памяти.

Приведем фрагмент кода, в котором происходит выделение необходимого количества памяти для хранения данных, количество которых задает пользователь. А также происходит инициализация значений, поиск максимального элемента и среднего арифметического значения.

```
int* pMas;
int Count;

cout << "Input elements count: ";
cin >> Count;

pMas = new int [Count];

int i;
for(i = 0; i < Count; i++)
   cin >> *(pMas+i);

int Sum = 0;
for(i = 0; i < Count; i++)
   Sum += *(pMas+i);

double Avg;
Avg = (double)Sum/Count;

cout << "Summa = " << Sum << endl;
cout << "Avg = " << Avg << endl;
```

В языке программирования C++ оператор **delete** возвращает память, выделенную оператором **new**. Вызов **delete** должен происходить для каждого вызова **new**, дабы избежать утечки памяти. После вызова **delete**

объект, указывающий на этот участок памяти, становится некорректным и не должен больше использоваться. При этом важно помнить, что если память была выделена при помощи **new**, то она освобождается при помощи **delete**, а если **new[]**, то должен быть вызван **delete[]**.

```
int* pA;
int* pMas;

pA = new int;
pMas = new int[10];

delete pA;
delete[] pMas;
```

3.3. Передача переменных по ссылке

Еще раз напомним про различие локальных и глобальных переменных.

```
void MyFunc(int a, int b, int c)
{
  a = 11;
  b = 12;
  c = 13;
}

void main()
{
  int a, b, c;
  a = 1; b = 2; c = 3;

  MyFunc(a, b, c);

  cout << a << b << c;
}
```

В результате исполнения приведенного кода, на экран будет выведено **1 2 3.** Другими словами, переменные в функции ***main()*** не изменят своего значения. В функции ***MyFunc()*** определены собственные локальные

переменные с такими же именами, но которые не имеют отношения к переменным в функции ***main()***.

При написании программ часто возникает желание произвести внутри функции манипуляции с переменными таким образом, чтобы в вызываемой функции они также изменили значения. Например, в программах часто применяется метод инициализации всех переменный внутри отдельного блока (функции). **Одним из способов** решения такой задачи является заведение глобальных переменных следующим образом.

```
int a, b, c;

void MyFunc()
{
  a = 11;
  b = 12;
  c = 13;
}

void main()
{

  a = 1; b = 2; c = 3;

  MyFunc();

  cout << a << " " << b << " "  << c;
}
```

В этом случае никаких переменных в функцию передавать не надо, так как она имеет непосредственный доступ ко всем глобальным переменным.

Но помните, если вы все-таки напишите следующим образом:

```
int a, b, c;

void MyFunc(int a, int b, int c)
{
   a = 11;
   b = 12;
   c = 13;
}

void main()
{

   a = 1; b = 2; c = 3;
   MyFunc(a, b, c);
   cout << a << " " << b << " "  << c;
}
```

то в функции будут использоваться уже не глобальные, а локальные переменные. Это вновь приведет к тому, что значение глобальных переменных не изменится после выполнения функции.

Второй способ заключается в передаче переменных по ссылке.

Оператор **&** — это унарный оператор, возвращающий адрес своего операнда. (Напомним, что унарный оператор имеет один операнд). Например, если написать **&count**, то результатом будет адрес переменной **count**. Оператор **&** можно представить себе как оператор, возвращающий адрес объекта.

Для хранения адресов используются указатели, а для присвоения значения переменной через указатель на нее, унарный оператор разыменования *****. Таким образом, корректной является запись

```
int* p;
int count;
p = &count;

*p = 5;
cout << count;
```


Таким образом, приведем фрагмент кода, в котором переменные передаются по ссылке в функции, которая изменяет их значения.

```
void MyFunc(int* a, int* b, int* c)
{
   *a = 11;
   *b = 12;
   *c = 13;
}

void main()
{
   int a, b, c;

   MyFunc(&a, &b, &c);

   cout << a << " " << b << " "  << c;
}
```

4. СТРУКТУРНОЕ ПРОГРАММИРОВАНИЕ

Методология структурного программирования появилась как следствие возрастания сложности решаемых на компьютерах задач, и соответственного усложнения программного обеспечения. В 70-е годы XX века объёмы и сложность программ достигли такого уровня, что «интуитивная» (неструктурированная, или «рефлекторная») разработка программ, которая была нормой в более раннее время, перестала удовлетворять потребностям практики. Программы становились слишком сложными, чтобы их можно было нормально сопровождать, поэтому потребовалась какая-то систематизация процесса разработки и структуры программ.

Наиболее сильной критике со стороны разработчиков структурного подхода к программированию подвергся оператор **GOTO** (оператор безусловного перехода), имевшийся тогда почти во всех языках программирования. Неправильное и необдуманное использование произвольных переходов в тексте программы приводит к получению запутанных, плохо структурированных программ (т.н. спагетти-кода), по тексту которых практически невозможно понять порядок исполнения и взаимозависимость фрагментов.

Следование принципам структурного программирования сделало тексты программ, даже довольно крупных, нормально читаемыми. Серьёзно облегчилось понимание программ, появилась возможность разработки программ в нормальном промышленном режиме, когда программу может без особых затруднений понять не только её автор, но и другие программисты. Это позволило разрабатывать достаточно крупные для того времени программные комплексы силами коллективов разработчиков, и сопровождать эти комплексы в течение многих лет, даже в условиях неизбежных изменений в составе персонала.

Перечислим некоторые достоинства структурного программирования:

1. Структурное программирование позволяет значительно сократить число вариантов построения программы по одной и той же спецификации, что значительно снижает сложность программы и, что ещё важнее, облегчает понимание её другими разработчиками.
2. В структурированных программах логически связанные операторы находятся визуально ближе, а слабо связанные — дальше, что позволяет обходиться без блок-схем и других графических форм изображения алгоритмов (по сути, сама программа является собственной блок-схемой).
3. Сильно упрощается процесс тестирования и отладки структурированных программ.
4.

4.1. Методология

Структурное программирование — методология разработки программного обеспечения, в основе которой лежит представление программы в виде иерархической структуры блоков, предложенная в 70-х годах XX века Э. Дейкстрой, а разработана и дополнена Н. Виртом.

В соответствии с данной методологией

1. Любая программа представляет собой структуру, построенную из трёх типов базовых конструкций:
    a. ***последовательное исполнение*** — однократное выполнение операций в том порядке, в котором они записаны в тексте программы;
    b. ***ветвление*** — однократное выполнение одной из двух или более операций, в зависимости от выполнения некоторого заданного условия;
    c. ***цикл*** — многократное исполнение одной и той же операции до тех пор, пока выполняется некоторое заданное условие (условие продолжения цикла).

2. Повторяющиеся фрагменты программы, либо представляющие собой логически целостные вычислительные блоки, оформляются в виде функций (подпрограмм).
3. Разработка программы ведётся пошагово, методом «сверху вниз».

4.1.1. Создание программ

При решении любой задачи сначала пишется текст основной программы, в котором, вместо каждого связного логического фрагмента текста, вставляется вызов подпрограммы, которая будет выполнять этот фрагмент. Вместо настоящих, работающих подпрограмм, в программу вставляются «заглушки», которые ничего не делают. Полученная программа проверяется и отлаживается. После того, как программист убедится, что подпрограммы вызываются в правильной последовательности (то есть общая структура программы верна), подпрограммы-заглушки последовательно заменяются на реально работающие, причём разработка каждой подпрограммы ведётся тем же методом, что и основной программы. Разработка заканчивается тогда, когда не останется ни одной «затычки», которая не была бы удалена.

Такая последовательность гарантирует, что на каждом этапе разработки программист одновременно имеет дело с обозримым и понятным ему множеством фрагментов, и может быть уверен, что общая структура всех более высоких уровней программы верна. При сопровождении и внесении изменений в программу выясняется, в какие именно процедуры нужно внести изменения, и они вносятся, не затрагивая части программы, непосредственно не связанные с ними. Это позволяет гарантировать, что при внесении изменений и исправлении ошибок не выйдет из строя какая-то часть программы, находящаяся в данный момент вне зоны внимания программиста.

4.2. Программа «бродилка»

Суть программы заключается в том, чтобы пользователю было представлено игровое поле, на котором были случайным образом разбросаны препятствия. При этом пользователь мог бы управлять игроком внутри поля и перемещать его внутри лабиринта.

4.2.1. Подключение внешней библиотеки

Для работы с текстовой графикой предлагается использовать библиотеку текстовой графики Conlib.

Библиотека предоставляется в виде двух файлов:

- **Conlib.h** – описание реализованных в библиотеке функций и описание их вызовов.
- **Conlib.lib** – реализация функций.

Для работы необходимо скопировать оба эти файла в директорию проекта. После этого прописать добавление деклараций функций библиотеки:

```
#include "conlib.h"
```

Если работа происходит в **MSVS2008**, то отдельно прописывать **Conlib.lib** нигде не нужно.

Если же работа происходит в **MSVS2005**, то необходимо явно указать в настройках проекта файл **Conlib.lib**. Это делается следующим образом (рис. 7):

Настройки проекта –> Configuration options -> Linker ->
-> Input -> Additional Dependencies -> … прописать название

После этого библиотека будет включаться в проект и будут доступны все ее функции.

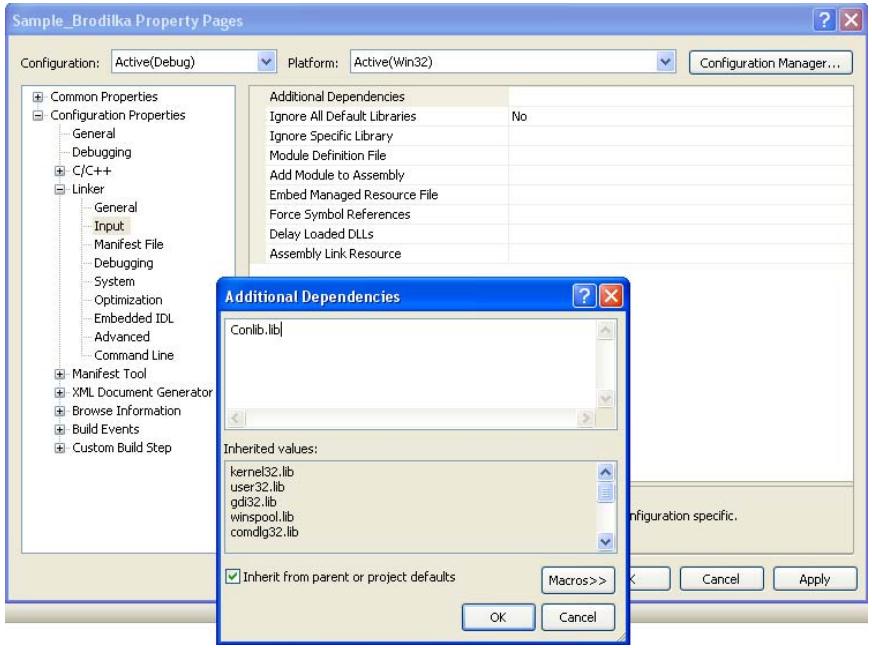


Рис. 7. Добавление внешней библиотеки в проект

4.2.2. Основные функции библиотеки Conlib

Основными функциями данной библиотеки для программы бродилка являются функции позиционирования на экране и обработка нажатий клавиш на клавиатуре для реализации интерактивности и перемещения игрока по экрану.

- ***int GotoXY(int x, int y);***
  функция перемещает курсор в заданную координату на экране.

- ***int MaxXY(int *px, int *py);***
  функция определяет максимальные размеры консоли

- ***int ClearConsole();***
  производит полную очистку консоли от всего содержимого

- ***int KeyPressed();***
  возвращает не 0, если была нажата какая-либо клавиша на клавиатуре

- ***int GetKey();***
  возвращает код нажатой клавиши

- ***int SetColor(short color);***
  позволяет задавать цвет шрифта и фона

4.2.3. Создание каркаса программы

Согласно сформулированным в первой главе описанием алгоритмов решения задачи, необходимо во-первых определиться с входными данными и результатом программы.

Итак, ***входными данными являются***:

- размер игрового поля;
- количество препятствий внутри лабиринта (либо процент заполнения лабиринта препятствиями).

Результатом является представление игрового поля в следующем виде (рис. 8):

Рис. 8. Проект «бродилка»

Далее необходимо составить алгоритм решения задачи. Если у человека не возникает больших трудностей с формулировками, то решение может быть сразу представлено в виде кода программы, в которой будут использованы функции в качестве логических блоков.

```
int main()
{

   int MaxX, MaxY;
   int BarrierCount;

   Init(&MaxX, &MaxY, &BarrierCount);

   int* pPole;
   pPole = new int[MaxX * MaxY];

   FillPole(pPole, MaxX, MaxY, BarrierCount);

   PrintPole(pPole, MaxX, MaxY);

  int PlayerX, PlayerY;

   PlayerX = MaxX / 2;
```



```
   PlayerY = MaxY / 2;

   Play(pPole, PlayerX, PlayerY, MaxX, MaxY);

   ClearConsole();
   GotoXY(10, 5);
   cout << "Press any key to continue...";
}
```

Разберем более подробно данный код:

- функция **Init()** определяет входные параметры для программы, в которые входят размеры игрового поля и количество препятствий;
- функция **FillPole()** производит заполнение игрового поля содержимым: границами и препятствиями
- функция **PrintPole()** распечатывает игровое поле на экран.
- После этого определяются начальные координаты игрока и вызывается функция **Play()**, которая отвечает за обработку клавиатуры и собственно перемещение игрока по полю.

4.2.4. Инициализация переменных программы

Функция **Init()** получает в качестве аргументов адреса переменных, куда производит запись данных. Таким образом, вызывая одну функцию, мы получаем все необходимые значения переменных.

```
void Init(int* pMaxX, int* pMaxY, int* pBarrierCount)
{
    MaxXY(pMaxX, pMaxY);

  int Pr;
  cout << "Input Barriers % = ";
  cin >> Pr;

  int PoleSize = (*pMaxX) * (*pMaxY);
```

```
   *pBarrierCount = PoleSize * Pr / 100;
}
```

4.2.5. Заполнение игрового поля

Разумным решением является разделить этот этап на несколько частей:

- Прописать все значения в поле нулем. Это необходимо, поскольку при выделении памяти этого автоматически не происходит.
- Расставить границы поля.
- Расставить препятствия внутри поля.

Помимо этого нам потребуются некоторые вспомогательные функции, которые мы должны реализовать сами.

Получение случайного числа

Стандартная функция rand() позволяет получить псевдослучайное число в диапазоне от 0 до RAND_MAX.

Описанная ниже функция позволяет получить псевдослучайное значение в заданном диапазоне.

```
int GetRandomValue(int Min, int Max)
{
   int Val = rand();

   int Range = Max - Min;

   double q = ((double)Val / RAND_MAX) * Range;

   return (q + Min);
}
```

Правило отображения игрового поля

Сложность состоит в том, что игровое поле является двумерным, а выделенная память – линейной. Разумным является расположить все

строки последовательно в памяти. В этом случае функция преобразования двумерных координат в индекс массива будет выглядеть следующим образом:

```
int GetPoleIndex(int x, int y, int MaxX)
{
   return (y*MaxX + x);
}
```

Определение границ игрового поля

Функция последовательно расставляет верхнюю, нижнбб, левую и правую границы. Особенностями является то, что горизонтальные границы закодированы единицей, а вертикальные – двойкой.

Это необходимо, поскольку в постановке задачи границы имеют различное отображение.

```
void CreatePoleBorders(int* pPole,
                       int MaxX, int MaxY)
{
   int x, y;

   y = 0;
   for(x= 0; x < MaxX; x++)
      *(pPole + GetPoleIndex(x, y, MaxX)) = 1;

   y = MaxY-1;
   for(x= 0; x < MaxX; x++)
      *(pPole + GetPoleIndex(x, y, MaxX)) = 1;

   x = 0;
   for(y= 0; y < MaxY; y++)
      *(pPole + GetPoleIndex(x, y, MaxX)) = 2;

   x = MaxX-1;
   for(y= 0; y < MaxY; y++)
      *(pPole + GetPoleIndex(x, y, MaxX)) = 2;
}
```

Результирующая функция заполнения игрового поля

С использование всех перечисленных вспомогательных функций, требуемый блок по заполнению поля препятствиями будет иметь следующий вид:

```
void FillPole(int* pPole, int MaxX,
              int MaxY, int BarrierCount)

{
  int x, y;

  for(y = 0; y < MaxY; y++)
    for(x= 0; x < MaxX; x++)
      *(pPole + GetPoleIndex(x, y, MaxX)) =
0;

  CreatePoleBorders(pPole, MaxX, MaxY);

  int i;

  for(i = 0; i < BarrierCount; i++)
  {
    x = GetRandomValue(1, MaxX-1);
    y = GetRandomValue(1, MaxY-1);

    *(pPole + GetPoleIndex(x, y, MaxX)) = 3;
  }

}
```

4.2.6. Отображение игрового поля

Отображение поля использует созданные выше вспомогательные функции отображения.

При отображении необходимо учесть возможность сопоставлять различным кодам – различные символы на экране. Удобнее всего использовать для этого переключатель **switch**.

```
void PrintPole(int* pPole, int MaxX, int MaxY)
{

  int x, y;
  GotoXY(0, 0);
  for(y = 0; y < MaxY; y++)
  {
    for(x= 0; x < MaxX; x++)
      switch(*(pPole + GetPoleIndex(x, y, MaxX)))
      {
      case 1: cout << '-'; break;
      case 2: cout << '|'; break;
      case 3: cout << '*'; break;
      default: cout << ' ';break;
      }
    cout << endl;
  }
}
```

4.2.7. Обработка клавиатуры и перемещение игрока

```
void Play(int *pPole, int CurX, int CurY,
          int MaxX, int MaxY)
{
  GotoXY(CurX, CurY);
  cout << 'X';

  int key;
  int bExit = 0;
  while(!bExit)
    {
     while(!KeyPressed());
     key = GetKey();

     GotoXY(CurX, CurY);
     cout << ' ';
```

```
   if (key == KEY_ESC)
      bExit = 1;
   else if (key == KEY_UP)
      CurY--;
   else if (key == KEY_DOWN)
      CurY++;
  else if (key == KEY_LEFT)
      CurX--;
  else if (key == KEY_RIGHT)
      CurX++;

   GotoXY(CurX, CurY);
   cout << 'X';
   }
}
```


5. Алгоритмы сортировки данных

Сортировка — это упорядочивание элементов по выбранному признаку. В случае чисел критерием может быть признак «больше» или «меньше», в результате чего мы получим последовательность, отсортированную по возрастанию (первый элемент будет самым маленьким) либо по убыванию (первый элемент будет иметь самое большое значение).

Единственного эффективнейшего алгоритма сортировки не существует ввиду множества параметров оценки эффективности:

1. **Время** — основной параметр, характеризующий быстродействие алгоритма. Называется также вычислительной сложностью. Для упорядочения важны худшее, среднее и лучшее поведение алгоритма в терминах размера списка (n). Для типичного алгоритма хорошее поведение — это O(n log n) и плохое поведение — это O(n2). Идеальное поведение для упорядочения — O(n). Алгоритмы сортировки, использующие только абстрактную операцию сравнения ключей, всегда нуждаются по меньшей мере в O(n log n) сравнениях в среднем.
2. **Память** — ряд алгоритмов требует выделения дополнительной памяти под временное хранение данных. При оценке используемой памяти не будет учитываться место, которое занимают исходный массив и не зависящие от входной последовательности затраты, например, на хранение кода программы.
3. **Устойчивость** (stability) — устойчивая сортировка не меняет взаимного расположения равных элементов.
4. **Естественность поведения** — эффективность метода при обработке уже упорядоченных или частично упорядоченных данных. Алгоритм ведёт себя естественно, если учитывает эту характеристику входной последовательности и работает лучше.

5.1. Сортировка пузырьком

Расположим массив сверху вниз, от нулевого элемента к последнему.

Идея метода: шаг сортировки состоит в проходе снизу вверх по массиву. По пути просматриваются пары соседних элементов. Если элементы некоторой пары находятся в неправильном порядке, то меняем их местами (рис. 9).

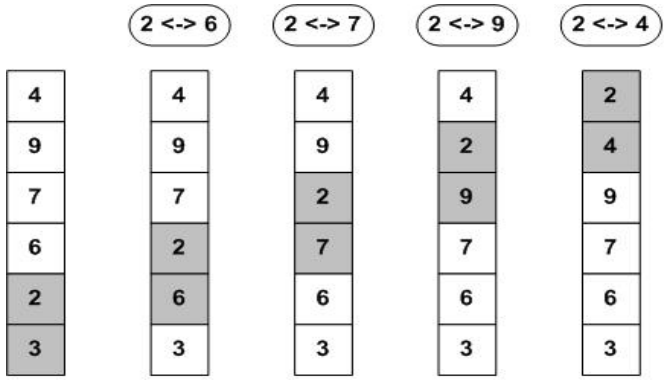


Рис. 9. Нулевой проход

После нулевого прохода по массиву «вверху» оказывается самый «легкий» элемент — отсюда аналогия с пузырьком. Следующий проход делается до второго сверху элемента, таким образом, второй по величине элемент поднимается на правильную позицию.

Делаем проходы по все уменьшающейся нижней части массива до тех пор, пока в ней не останется только один элемент (рис. 10). На этом сортировка заканчивается, так как последовательность упорядочена по возрастанию.

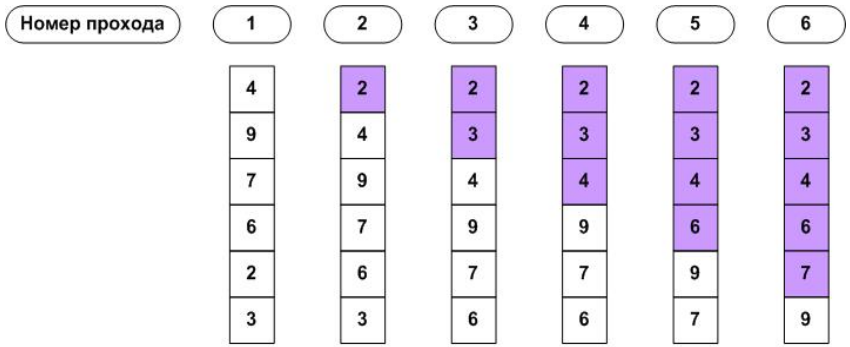


Рис. 10. Сортировка массива за шесть проходов

Среднее число сравнений и обменов имеют квадратичный порядок роста: $O(n^2)$, отсюда можно заключить, что алгоритм пузырька очень медлителен и малоэффективен.

Тем не менее, у него есть громадное преимущество в том, что он прост в реализации, и его можно улучшать различными способами.

5.2. Пирамидальная сортировка

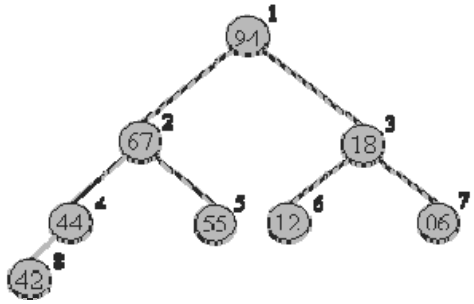
Пирамидальная сортировка является методом, быстродействие которого оценивается как O(n log n).

- Назовем пирамидой (Heap) бинарное дерево высоты k, в котором все узлы имеют глубину k или (k – 1) — дерево сбалансированное;
- при этом уровень (k – 1) полностью заполнен, а уровень k заполнен слева направо, т. е форма пирамиды имеет приблизительно такой вид:


;

- выполняется «свойство пирамиды»: каждый элемент меньше либо равен родителю.

Как хранить пирамиду? Наименее хлопотно поместить ее в массив.



Соответствие между геометрической структурой пирамиды как дерева и массивом устанавливается по следующей схеме:

- в a[0] хранится корень дерева;
- левый и правый сыновья элемента a[i] хранятся соответственно в a[2i + 1] и a[2i + 2].
-
- Поэтому для массива, хранящего в себе пирамиду, выполняется следующее свойство: a[i] >= a[2i + 1] и a[i] >= a[2i + 2].

Плюсы такого хранения пирамиды очевидны:

- никаких дополнительных переменных, нужно лишь понимать схему;
- узлы хранятся от вершины и далее вниз, уровень за уровнем;
- узлы одного уровня хранятся в массиве слева направо.

Запишем в виде массива пирамиду, изображенную выше. Слева направо, сверху вниз:

94 67 18 44 55 12 06 42.

На рисунке место элемента пирамиды в массиве обозначено цифрой справа и вверху от него.

Восстановить пирамиду из массива как геометрический объект легко: достаточно вспомнить схему хранения и нарисовать, начиная от корня.

Фаза 1 сортировки: построение пирамиды

Начать построение пирамиды можно с a[k]...a[n], k = [size / 2]. Эта часть массива удовлетворяет свойству пирамиды, так как не существует индексов i,j: i = 2i + 1 (или j = 2i + 2)... Просто потому, что такие i и j находятся за границей массива.

Следует заметить, что неправильно говорить о том, что a[k]..a[n] является пирамидой как самостоятельный массив. Это, вообще говоря, не верно: его элементы могут быть любыми. Свойство пирамиды сохраняется лишь в рамках исходного, основного массива a[0]...a[n].

Далее будем расширять часть массива, обладающую столь полезным свойством, добавляя по одному элементу за шаг. Следующий элемент на каждом шаге добавления — тот, который стоит перед уже готовой частью.

Чтобы при добавлении элемента сохранялась пирамидальность, будем использовать следующую процедуру расширения пирамиды a[i + 1]..a[n] на элемент a[i] влево:

1. Смотрим на сыновей слева и справа — в массиве это a[2i + 1] и a[2i + 2] и выбираем наибольшего из них.
2. Если этот элемент больше a[i] — меняем его с a[i] местами и идем к шагу 2, имея в виду новое положение a[i] в массиве. Иначе конец процедуры.

Новый элемент «просеивается» сквозь пирамиду.

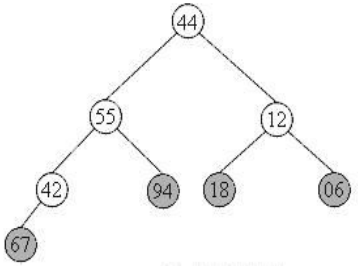
Ниже дана иллюстрация процесса для пирамиды из 8-и элементов:

```
   44    55    12    42  //  94    18    06    67
   44    55    12  //  67    94    18    06    42
   44    55  //  18    67    94    12    06    42
   44  //  94    18    67    55    12    06    42
 // 94    67    18    44    55    12    06    42
```

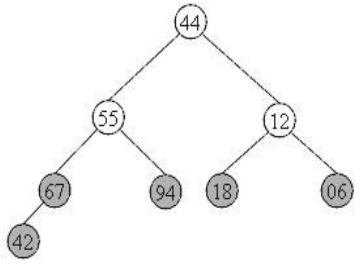
Справа — часть массива, удовлетворяющая свойству пирамиды. Остальные элементы добавляются один за другим, справа налево.

В геометрической интерпретации ключи из начального отрезка a[size / 2] ... a[n] является листьями в бинарном дереве, как изображено ниже. Один за другим остальные элементы продвигаются на свои места, и так будет продолжаться, пока не будет построена вся пирамида.

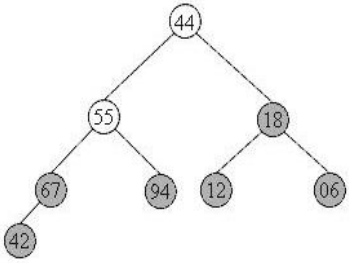
На рисунках ниже изображен процесс построения. Неготовая часть пирамиды (начало массива) окрашена в белый цвет, удовлетворяющий свойству пирамиды конец массива — в темный.



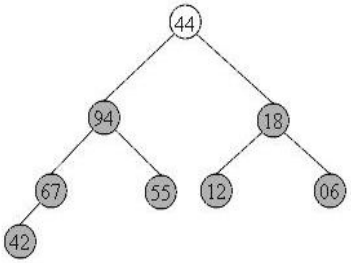
Исходная картина: 94, 18, 06, 67 – листья



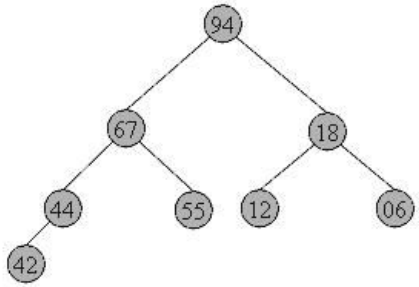
42 сравнили с 67 и поменяли их местами



12 сравнили с max (18, 6) = 18



55 сравнили с max (67, 94) = 94

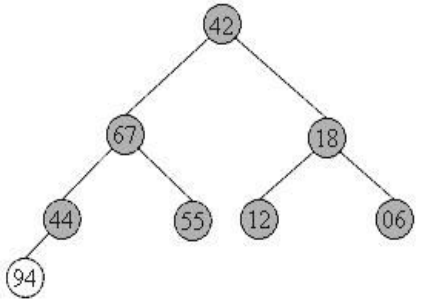


44 просеяли сквозь 94 и 67, остановились на 42

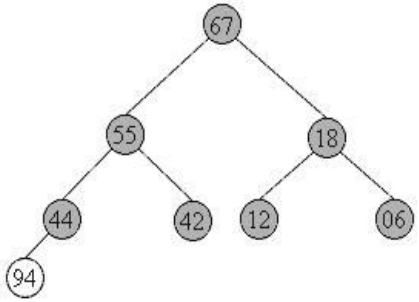
Фаза 2: собственно сортировка

Итак, задача построения пирамиды из массива успешно решена. Как видно из свойств пирамиды, в корне всегда находится максимальный элемент. Отсюда вытекает алгоритм фазы 2:

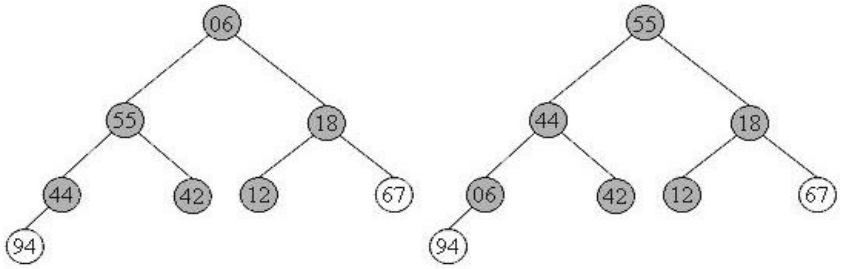
1. Берем верхний элемент пирамиды a[0]...a[n] (первый в массиве) и меняем его с последним элементом местами. Теперь «забываем» об этом элементе и далее рассматриваем массив a[0]...a[n–1]. Для превращения его в пирамиду достаточно просеять лишь новый первый элемент.
2. Повторяем шаг 1, пока обрабатываемая часть массива не уменьшится до одного элемента.



Обменяли 94 и 42, «забыли» о 94



Просеяли 42 сквозь 67 и 55



Обменяли 06 и 67, «забыли» о 67 Просеяли 06 сквозь 55 и 44

Очевидно, в конец массива каждый раз попадает максимальный элемент из текущей пирамиды, поэтому в правой части постепенно возникает упорядоченная последовательность.

Ниже приведена иллюстрация второй фазы сортировки во внутреннем представлении пирамиды.

```
94   67   18   44   55   12   06   42 //
67   55   44   06   42   18   12 // 94
55   42   44   06   12   18 // 67   94
44   42   18   06   12 // 55   67   94
42   12   18   06 // 44   55   67   94
18   12   06 // 42   44   55   67   94
12   06 // 18   42   44   55   67   94
06 // 12   18   42   44   55   67   94
```

- Пирамидальная сортировка не использует дополнительной памяти.
- Метод не является устойчивым: по ходу работы массив так «перетряхивается», что исходный порядок элементов может измениться случайным образом.
- Поведение неестественно: частичная упорядоченность массива никак не учитывается.

6. Производные типы данных. Списки данных

Из базовых типов языка С можно формировать производные типы данных, к которым относятся массивы, структуры и объединения. Также можно отнести сюда и строки, которых до сих пор не хватало, чтобы достаточно простым образом оперировать с различными именами и т. п.

6.1. Строки

Строка определяется как частный случай одномерного массива (поскольку состоит их набора символов). При этом последний элемент обязан быть равен **'/0'**, что обозначает конец строки.

Приведем пример определения строки и считывания имени пользователя, и затем его распечатки:

```
void main()
{
  char name[80];
  char zapros[80] = "Input your name: ";

  cout << zapros;

  cin >> name;
  cout << "Hello " << name << "!" << endl;
}
```

Если необходимо определить длину введенной строки, то можно использовать следующий способ:

```
int n;
for(n=0;;n++)
{
  if(name[n] == '\0')
    break;
}
cout << "Your name contain " << n << " letters";
```

6.2. Структурный тип

Этот тип задает внутреннее строение определяемых с помощью него структурных переменных. Часто применяется просто название **«структура»**.

Структура — это объединенное в единое целое множество поименованных элементов данных. В отличие от массивов данных, всегда состоящих из данных одного типа, структуры могут содержать различные типа данных. Например, может быть введена структура, описывающая студента университета со следующими компонентами (**полями**):

- имя;
- фамилия;
- номер группы;
- год рождения.

Для данного примера структура создается следующим образом:

```
struct Student
{
char Name[80];
char Surname[80];
int Group;
int Birth_year;
};
```

Таким образом, для того чтобы ввести в программу студентов определенного факультета, можно, например, создать массив, состоящий из таких структур данных. Это позволит упростить написание подобных программ.

Обратите внимание, что поскольку структурный тип данных является пользовательским, то он должен быть определен вне функций для обеспечения максимальной области видимости.

Переменная структурного типа создается таким же образом, как и любые другие переменные. Также можно создавать и указатели на

структуры. При этом объем необходимой памяти определяется с помощью функции **sizeof()**, которой в качестве параметра можно передать имя или тип структуры.

Доступ к полям структур осуществляется следующим образом:

Имя_Объекта . поле

либо ***Указатель_на_Объект -> поле***

Приведем в качестве примера фрагмент кода, в котором создается переменная структурного типа, осуществляется инициализация всех полей и распечатка через указатель на структуру.

```
#include <iostream>
using namespace std;

struct Student
{
  char Name[80];
  char Surname[80];
  int Group;
  int Birth_year;
};
void main()
{
  Student Vasya;

  cin >> Vasya.Name;
  cin >> Vasya.Surname;
  cin >> Vasya.Group;
  cin >> Vasya.Birth_year;

  Student* pVasya;
  pVasya = &Vasya;

  cout << pVasya->Name;
  cout << pVasya->Surname;
  cout << pVasya->Group;
  cout << pVasya->Birth_year;
}
```

6.3. Списки данных

Списки данных являются в некотором роде альтернативой динамическим массивам. В некоторых задачах, например база данных о студентах университета или телефонная книга, пользователь сам зачастую не знает, какое количество записей у него будет. Таким образом, нужен иной инструмент, позволяющий по ходу программы изменять количество элементов в некоей базе данных. Таким инструментом и являются списки данных.

Список — это упорядоченная последовательность связанных данных, связанных между собой. При этом списки бывают двух основных видов, в зависимости от характера связей элементов: односвязные и двусвязные.

Для хранения списка, и организации работы с ним достаточно лишь одного указателя, который будет хранить адрес первого элемента списка. Зная первый элемент, вы всегда сможете последовательно пройти по списку и извлечь любой элемент.

Помните, что если вы измените значение указателя начала списка (например перейдете к другому элементу в односвязном списке), то вы можете навсегда потерять первый элемент. Поэтому на практике, никогда не работают с указателем на начало списка, а создают временные указатели.

6.3.1. Односвязный список данных

Односвязный список характеризуется наличием одной связи у соседних элементов. Другими словами, каждый элемент знает об одном соседе (рис. 11). По такому списку можно передвигаться только последовательно в одном направлении: от начала к концу.

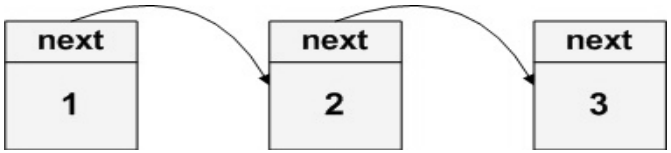


Рис. 11. Односвязный список

Разберем, каким образом будет выглядеть элемент односвязного списка, содержащий данные о студенте университета (описанные выше) и указатель на следующий элемент списка.

```
struct Student
{
   char Name[80];
   char Surname[80];
   int Group;
   int Birth_year;

   Student* pNext;
};
```

Заметим, что если программа компилируется с использованием C-компилятора, в типе указателя необходимо указывать не просто имея структуры, а именно с указанием того что это структура. Это связано с тем что определение структурного типа еще не закончено.

```
struct Student* pNext;
```

6.3.2. Двусвязный список данных

Двусвязный список отличается наличием двух связей у каждого элемента. При этом организуется двунаправленность списка (рис. 12). Другими словами, зная любой элемент списка, можно получить информацию как о следующем элементе списка, так и о предыдущем.

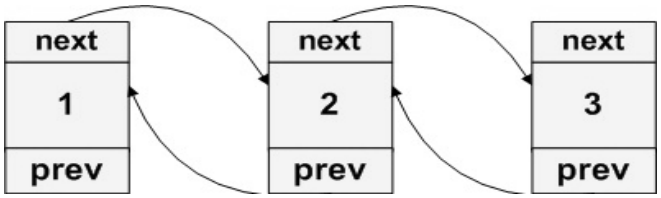


Рис.12. Двусвязный список

Для управления списком необходимо знать только первый его элемент — начало списка. Последний элемент определяется по признаку того, что следующего за ним не существует, т. е. указатель равен NULL (рис. 13).

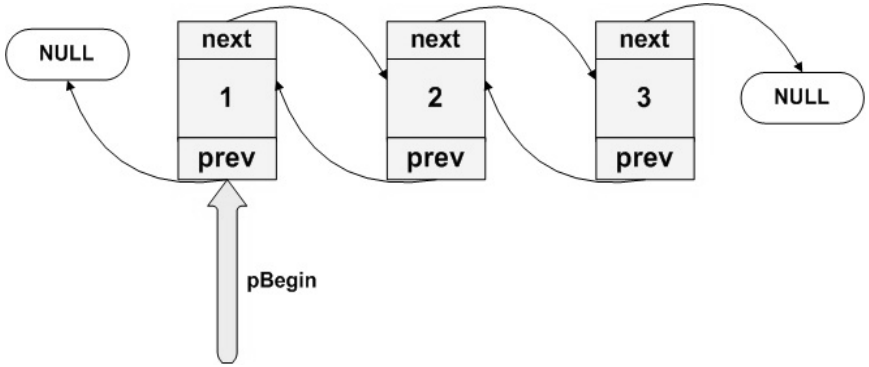


Рис.13. Указатель на начало двусвязного списка

Если следующего элемента не существует, то указатель pnext необходимо задать равным NULL.

6.3.3. Распечатка элементов списка

Функция для распечатки существующего списка может выглядеть следующим образом (в качестве параметра функция получает указатель на первый элемент списка):


```
void PrintAll(Student* pBegin)
{
   if (pBegin == NULL)
   {
      printf("No elements\n");
      return;
   }

   Student* pCur = pBegin;

   do
   {
      cout << pCur->Name;
      cout << pCur->Surname;
      cout << pCur->Group;
      cout << pCur->Birth_year;

      pCur = pCur->pNext;
   }
   while(pCur != NULL);
}
```

6.3.4. Добавление элемента в существующий список

Разберем на примере односвязного списка студентов процесс добавления элемента и реализуем для этого функцию.

В качестве входных параметров функция должна получить указатель на существующий список, собственно указатель на добавляемый элемент и номер, которым он должен стать в списке.

При добавлении элемента в список возможны три различных ситуации:

- список пуст, соответственно независимо от номера, элемент становится первым;
- элемент необходимо вставить в конец списка;
- элемент вставляется в середину списка.

Помните, что если элемент оказывается последним, то необходимо в поле указателя на следующий элемент прописать **NULL**.

```
void AddElement(Student* pBegin,
Student* pAdd, int number)
{
   // Если список был пуст - на первое место
   if(pBegin == NULL)
   {
      pBegin = pAdd;
      pAdd->pNext = NULL;
      return;
   }

   // Ищем в списке элемент
   // за которым должны вставить
   int n = 0;
   Student* pCur = pBegin;

   for(n = 0; n < number; n++)
      pCur = pCur->pNext;

   // Если нужно вставить последним
   if(pCur->pNext == NULL)
   {
      pCur->pNext = pAdd;
      pAdd->pNext = NULL;
      return;
   }

   // Определяем следующий элемент
   Student* pNext = pCur->pNext;

   // Вставляем элемент в список
   pCur->pNext = pAdd;
   pAdd->pNext = pNext;
}
```

Обратите внимание, что в приведенном примере не производится обработка ситуаций с некорректным номером элемента.

7. Работа с файлами

В данном пособии рассматривается подход к файловым операциям, основанный на языке С..

Первоначально язык С был реализован в операционной системе UNIX. Как таковые, ранние версии С (да и многие нынешние) поддерживают набор функций ввода/вывода, совместимый с UNIX. Этот набор иногда называют UNIX-подобной системой ввода/вывода или небуферизованной системой ввода/вывода. Однако когда С был стандартизован, то UNIX-подобные функции в него не вошли — в основном из-за того, что оказались лишними. Кроме того, UNIX-подобная система может оказаться неподходящей для некоторых сред, которые могут поддерживать язык С, но не эту систему ввода/вывода.

В этой главе описана работа с файловой системой в языке С. В языке С система ввода/вывода реализуется с помощью библиотечных функций, а не ключевых слов. Благодаря этому система ввода/вывода является очень мощной и гибкой. Например, во время работы с файлами данные могут передаваться или в своем внутреннем двоичном представлении или в текстовом формате, то есть в более удобочитаемом виде. Это облегчает задачу создания файлов в нужном формате.

Библиотека С поддерживает три уровня ввода-вывода: потоковый ввод-вывод, ввод-вывод нижнего уровня и ввод-вывод для консоли и портов.

7.1. Потоки и файлы

Перед тем как начать изучение файловой системы языка С, необходимо уяснить, в чем разница между потоками и файлами. В системе ввода/вывода С для программ поддерживается единый интерфейс, не зависящий от того, к какому конкретному устройству осуществляется доступ. То есть в этой системе между программой и устройством находится нечто более общее, чем само устройство. Такое обобщенное устройство ввода или вывода (устройство более высокого уровня

абстракции) называется потоком, в то время как конкретное устройство называется файлом. (Впрочем, файл — тоже понятие абстрактное.) Очень важно понимать, каким образом происходит взаимодействие потоков и файлов.

7.1.1. Потоки

Файловая система языка С предназначена для работы с самыми разными устройствами, в том числе терминалами, дисководами и накопителями на магнитной ленте. Даже если какое-то устройство сильно отличается от других, буферизованная файловая система все равно представит его в виде логического устройства, которое называется потоком. Все потоки ведут себя похожим образом. И так как они в основном не зависят от физических устройств, то та же функция, которая выполняет запись в дисковый файл, может ту же операцию выполнять и на другом устройстве, например, на консоли. Потоки бывают двух видов: текстовые и двоичные.

7.1.2. Текстовые потоки

Текстовый поток — это последовательность символов. В стандарте С считается, что текстовый поток организован в виде строк, каждая из которых заканчивается символом новой строки. Однако в конце последней строки этот символ не является обязательным. В текстовом потоке по требованию базовой среды могут происходить определенные преобразования символов. Например, символ новой строки может быть заменен парой символов — возврата каретки и перевода строки. Поэтому может и не быть однозначного соответствия между символами, которые пишутся (читаются), и теми, которые хранятся во внешнем устройстве. Кроме того, количество тех символов, которые пишутся (читаются), и тех, которые хранятся во внешнем устройстве, может также не совпадать из-за возможных преобразований.

7.1.3. Двоичные потоки

Двоичный поток — это последовательность байтов, которая взаимно однозначно соответствует байтам на внешнем устройстве, причем никакого преобразования символов не происходит. Кроме того, количество тех байтов, которые пишутся (читаются), и тех, которые хранятся на внешнем устройстве, одинаково. Однако в конце двоичного потока может добавляться определяемое приложением количество нулевых байтов. Такие нулевые байты, например, могут использоваться для заполнения свободного места в блоке памяти незначащей информацией, чтобы она в точности заполнила сектор на диске.

7.1.4. Файлы

В языке С файлом может быть все что угодно, начиная с дискового файла и заканчивая терминалом или принтером. Поток связывают с определенным файлом, выполняя операцию открытия. Как только файл открыт, можно проводить обмен информацией между ним и программой.

Но не у всех файлов одинаковые возможности. Например, к дисковому файлу прямой доступ возможен, в то время как к некоторым принтерам — нет. Таким образом, мы пришли к одному важному принципу, относящемуся к системе ввода/вывода языка С: все потоки одинаковы, а файлы — нет.

Если файл может поддерживать запросы на местоположение (указатель текущей позиции), то при открытии такого файла указатель текущей позиции в файле устанавливается в начало. При чтении из файла (или записи в него) каждого символа указатель текущей позиции увеличивается, обеспечивая тем самым продвижение по файлу.

Файл отсоединяется от определенного потока (т.е. разрывается связь между файлом и потоком) с помощью операции закрытия. При закрытии файла, открытого с целью вывода, содержимое (если оно есть) связанного

с ним потока записывается на внешнее устройство. Этот процесс, который обычно называют дозаписью потока, гарантирует, что никакая информация случайно не останется в буфере диска. Если программа завершает работу нормально, т.е. либо main() возвращает управление операционной системе, либо вызывается exit(), то все файлы закрываются автоматически. В случае аварийного завершения работы программы, например, в случае краха или завершения путем вызова abort(), файлы не закрываются.

У каждого потока, связанного с файлом, имеется управляющая структура, содержащая информацию о файле; она имеет тип FILE. В этом блоке управления файлом никогда ничего не меняйте.

Если вы новичок в программировании, то разграничение потоков и файлов может показаться излишним или даже "заумным". Однако надо помнить, что основная цель такого разграничения — это обеспечить единый интерфейс. Для выполнения всех операций ввода/вывода следует использовать только понятия потоков и применять всего лишь одну файловую систему. Ввод или вывод от каждого устройства автоматически преобразуется системой ввода/вывода в легко управляемый поток.

Функции библиотеки ввода-вывода языка С, поддерживающие обмен данными с файлами на уровне потока, позволяют обрабатывать данные различных размеров и форматов, обеспечивая при этом буферизованный ввод и вывод. Таким образом, поток — это файл вместе с предоставляемыми средствами буферизации.

7.2. Основы файловой системы

Файловая система языка С состоит из нескольких взаимосвязанных функций. Самые распространенные из них показаны в ниже таблице. Для их работы требуется заголовок <stdio.h>.

Часто используемые функции файловой системы С

fopen()	Открывает файл
fclose()	Закрывает файл
putc()	Записывает символ в файл
fputc()	То же, что и putc()
getc()	Читает символ из файла
fgetc()	То же, что и getc()
fgets()	Читает строку из файла
fputs()	Записывает строку в файл
fseek()	Устанавливает указатель позиции на определенный байт
ftell()	Возвращает текущее значение указателя в файле
fprintf()	Для файла то же, что printf() для консоли
fscanf()	Для файла то же, что scanf() для консоли
feof()	Возвращает значение true, если достигнут конец файла
ferror()	Возвращает значение true, если произошла ошибка
rewind()	Устанавливает указатель позиции в начало файла
remove()	Стирает файл
fflush()	Дозапись потока в файл

Заголовок <stdio.h> предоставляет прототипы функций ввода/вывода и определяет следующие три типа: size_t, fpos_t и FILE. size_t и fpos_t представляют собой определенные разновидности такого типа, как целое без знака. А о третьем типе, FILE, рассказывается в следующем разделе.

Кроме того, в <stdio.h> определяется несколько макросов. Из них к материалу этой главы относятся NULL, EOF, FOPEN_MAX, SEEK_SET,

SEEK_CUR и SEEK_END. Макрос NULL определяет пустой (null) указатель. Макрос EOF, часто определяемый как -1, является значением, возвращаемым тогда, когда функция ввода пытается выполнить чтение после конца файла. FOPEN_MAX определяет целое значение, равное максимальному числу одновременно открытых файлов. Другие макросы используются вместе с fseek() — функцией, выполняющей операции прямого доступа к файлу.

7.3. Указатель файла

Указатель файла — это то, что соединяет в единое целое всю систему ввода/вывода языка С. Указатель файла — это указатель на структуру типа FILE. Он указывает на структуру, содержащую различные сведения о файле, например, его имя, статус и указатель текущей позиции в начало файла. В сущности, указатель файла определяет конкретный файл и используется соответствующим потоком при выполнении функций ввода/вывода. Чтобы выполнять в файлах операции чтения и записи, программы должны использовать указатели соответствующих файлов. Чтобы объявить переменную-указатель файла, используйте такого рода оператор:

```
FILE *fp;
```

7.4. Открытие файла

Функция fopen() открывает поток и связывает с этим потоком определенный файл. Затем она возвращает указатель этого файла. Чаще всего (а также в оставшейся части этой главы) под файлом подразумевается дисковый файл. Прототип функции fopen() следующий:

FILE *fopen(const char *имя_файла, const char *режим);

где имя_файла — это указатель на строку символов, представляющую собой допустимое имя файла, в которое также может входить спецификация пути к этому файлу. Строка, на которую указывает режим, определяет, каким образом файл будет открыт. Ниже в таблице показано, какие значения строки режим являются допустимыми. Строки, подобные "r+b" могут быть представлены и в виде "rb+".

Допустимые значения режим	
r	Открыть текстовый файл для чтения
w	Создать текстовый файл для записи
a	Добавить в конец текстового файла
rb	Открыть двоичный файл для чтения
wb	Создать двоичный файл для записи
ab	Добавить в конец двоичного файла
r+	Открыть текстовый файл для чтения/записи
w+	Создать текстовый файл для чтения/записи
a+	Добавить в конец текстового файла или создать текстовый файл для чтения/записи
r+b	Открыть двоичный файл для чтения/записи
w+b	Создать двоичный файл для чтения/записи
a+b	Добавить в конец двоичного файла или создать двоичный файл для чтения/записи

Как уже упоминалось, функция fopen() возвращает указатель файла. Никогда не следует изменять значение этого указателя в программе. Если при открытии файла происходит ошибка, то fopen() возвращает пустой (null) указатель.

В следующем коде функция fopen() используется для открытия файла по имени TEST для записи.

```
FILE *fp;
fp = fopen("test", "w");
```

Хотя предыдущий код технически правильный, но его обычно пишут немного по-другому:

```
FILE *fp;

if ((fp = fopen("test","w"))==NULL)
  printf("Ошибка при открытии файла.\n");
```

Этот метод помогает при открытии файла обнаружить любую ошибку, например, защиту от записи или полный диск, причем обнаружить еще до того, как программа попытается в этот файл что-либо записать. Вообще говоря, всегда нужно вначале получить подтверждение, что функция - fopen() выполнилась успешно, и лишь затем выполнять с файлом другие операции.

Хотя название большинства файловых режимов объясняет их смысл, однако не помешает сделать некоторые дополнения. Если попытаться открыть файл только для чтения, а он не существует, то работа fopen() завершится отказом. А если попытаться открыть файл в режиме дозаписи, а сам этот файл не существует, то он просто будет создан. Более того, если файл открыт в режиме дозаписи, то все новые данные, которые записываются в него, будут добавляться в конец файла. Содержимое, которое хранилось в нем до открытия (если только оно было), изменено не будет. Далее, если файл открывают для записи, но выясняется, что он не существует, то он будет создан. А если он существует, то содержимое, которое хранилось в нем до открытия, будет утеряно, причем будет создан новый файл. Разница между режимами r+ и w+ состоит в том, что если файл не существует, то в режиме открытия r+ он создан не будет, а в

режиме w+ все произойдет наоборот: файл будет создан! Более того, если файл уже существует, то открытие его в режиме w+ приведет к утрате его содержимого, а в режиме r+ оно останется нетронутым.

Из таблицы видно, что файл можно открыть либо в одном из текстовых, либо в одном из двоичных режимов. В большинстве реализаций в текстовых режимах каждая комбинация кодов возврата каретки (ASCII 13) и конца строки (ASCII 10) преобразуется при вводе в символ новой строки. При выводе же происходит обратный процесс: символы новой строки преобразуются в комбинацию кодов возврата каретки (ASCII 13) и конца строки (ASCII 10). В двоичных режимах такие преобразования не выполняются.

Максимальное число одновременно открытых файлов определяется FOPEN_MAX. Это значение не меньше 8, но чему оно точно равняется – это должно быть написано в документации по компилятору.

7.5. Закрытие файла

Функция fclose() закрывает поток, который был открыт с помощью вызова fopen().Функция fclose() записывает в файл все данные, которые еще оставались в дисковом буфере, и проводит, так сказать, официальное закрытие файла на уровне операционной системы. Отказ при закрытии потока влечет всевозможные неприятности, включая потерю данных, испорченные файлы и возможные периодические ошибки в программе. Функция fclose() также освобождает блок управления файлом, связанный с этим потоком, давая возможность использовать этот блок снова. Так как количество одновременно открытых файлов ограничено, то, возможно, придется закрывать один файл, прежде чем открывать другой. Прототип функции fclose() такой:

```
int fclose(FILE * файл);
```

Возвращение нуля означает успешную операцию закрытия. В случае же ошибки возвращается EOF. Чтобы точно узнать, в чем причина этой ошибки, можно использовать стандартную функцию ferror() (о которой вскоре пойдет речь). Обычно отказ при выполнении fclose() происходит только тогда, когда диск был преждевременно удален (стерт) с дисковода или на диске не осталось свободного места.

7.6. Запись символа

В системе ввода/вывода языка С определяются две эквивалентные функции, предназначенные для вывода символов: putc() и fputc(). (На самом деле putc() обычно реализуется в виде макроса.) Две идентичные функции имеются просто потому, чтобы сохранять совместимость со старыми версиями С. В этой книге используется putc(), но применение fputc() также вполне возможно.

Функция putc() записывает символы в файл, который с помощью fopen() уже открыт в режиме записи. Прототип этой функции следующий:

```
int putc(int ch, FILE * файл);
```

Указатель файла сообщает putc(), в какой именно файл следует записывать символ. Хотя ch и определяется как int, однако записывается только младший байт.

Если функция putc() выполнилась успешно, то возвращается записанный символ. В противном же случае возвращается EOF.

7.7. Чтение символа

Для ввода символа также имеются две эквивалентные функции: getc() и fgetc(). Обе определяются для сохранения совместимости со старыми

версиями С. В этой книге используется getc() (которая обычно реализуется в виде макроса), но если хотите, применяйте fgetc().

Функция getc() записывает символы в файл, который с помощью fopen() уже открыт в режиме для чтения. Прототип этой функции следующий:

```
int getc(FILE * файл);
```

Функция getc() возвращает целое значение, но символ находится в младшем байте. Если не произошла ошибка, то старший байт (байты) будет обнулен.

Если достигнут конец файла, то функция getc() возвращает EOF. Поэтому, чтобы прочитать символы до конца текстового файла, можно использовать следующий код:

```
do
{
  ch = getc(fp);
}
while(ch!=EOF);
```

Однако getc() возвращает EOF и в случае ошибки. Для определения того, что же на самом деле произошло, можно использовать ferror().

7.8. Использование fopen(), getc(), putc(), и fclose()

Функции fopen(), getc(), putc() и fclose() — это минимальный набор функций для операций с файлами. Следующая программа, представляет собой простой пример, в котором используются только функции putc(), fopen() и fclose(). В этой программе символы считываются с клавиатуры и записываются в дисковый файл до тех пор, пока пользователь не введет знак доллара. Имя файла определено в программа «test.txt».

```
void main()
{
  FILE *fp;  char ch;

  if((fp=fopen("test.txt", "w"))==NULL)
  {
    printf("ERROR\n");
    return;
  }

  do {
     ch = getchar();
     putc(ch, fp);
  }
  while (ch != '$');

  fclose(fp);
}
```

Следующая программа, читает любой текстовый файл и выводит его содержимое на экран.

```
void main()
{
  FILE *fp;  char ch;

  if((fp=fopen("test.txt", "r")) == NULL)
  {
    printf("ERROR.\n");
    return;
  }

  ch = getc(fp);

  while (ch!=EOF)
  {
    putchar(ch);
    ch = getc(fp);
  }

  fclose(fp);
}
```


7.9. Использование feof()

Как уже говорилось, если достигнут конец файла, то getc() возвращает EOF. Однако проверка значения, возвращенного getc(), возможно, не является наилучшим способом узнать, достигнут ли конец файла. Во-первых, файловая система языка С может работать как с текстовыми, так и с двоичными файлами. Когда файл открывается для двоичного ввода, то может быть прочитано целое значение, которое, как выяснится при проверке, равняется EOF. В таком случае программа ввода сообщит о том, что достигнут конец файла, чего на самом деле может и не быть. Во-вторых, функция getc() возвращает EOF и в случае отказа, а не только тогда, когда достигнут конец файла. Если использовать только возвращаемое значение getc(), то невозможно определить, что же на самом деле произошло. Для решения этой проблемы в С имеется функция feof(), которая определяет, достигнут ли конец файла. Прототип функции feof() такой:

```
int feof(FILE * файл);
```

Если достигнут конец файла, то feof() возвращает true (истина); в противном же случае эта функция возвращает нуль. Поэтому следующий код будет читать двоичный файл до тех пор, пока не будет достигнут конец файла:

```
while(!feof(fp)) ch = getc(fp);
```

Ясно, что этот метод можно применять как к двоичным, так и к текстовым файлам.

В следующей программе, которая копирует текстовые или двоичные файлы, имеется пример применения feof(). Файлы открываются в двоичном режиме, а затем feof() проверяет, не достигнут ли конец файла.

```
void main()
{
   FILE *in, *out;  char ch;

   if((in=fopen("in.txt", "rb"))==NULL)
   {
      printf("ERROR\n");
      return;
   }

   if((out=fopen("out.txt", "wb")) == NULL)
   {
      printf("ERROR\n");
      return;
   }

   while(!feof(in))
   {
      ch = getc(in);
      if(!feof(in)) putc(ch, out);
   }

   fclose(in);
   fclose(out);
}
```

7.10. Ввод / вывод строк: fputs() и fgets()

Кроме getc() и putc(), в языке С также поддерживаются родственные им функции fgets() и fputs(). Первая из них читает строки символов из файла на диске, а вторая записывает строки такого же типа в файл, тоже находящийся на диске. Эти функции работают почти как putc() и getc(), но читают и записывают не один символ, а целую строку. Прототипы функций fgets() и fputs() следующие:

```
int fputs(const char *cmp, FILE * файл);
char *fgets(char *cmp, int длина, FILE * файл);
```

Функция fputs() пишет в определенный поток строку, на которую указывает cmp. В случае ошибки эта функция возвращает EOF.

Функция fgets() читает из определенного потока строку, и делает это до тех пор, пока не будет прочитан символ новой строки или количество прочитанных символов не станет равным длина-1. Если был прочитан разделитель строк, он записывается в строку, чем функция fgets() отличается от функции gets(). Полученная в результате строка будет оканчиваться символом конца строки (***'0'***). При успешном завершении работы функция возвращает cmp, а в случае ошибки — пустой указатель (null).

В следующей программе показано использование функции fputs(). Она читает строки с клавиатуры и записывает их в файл, который называется TEST. Чтобы завершить выполнение программы, введите пустую строку. Так как функция gets() не записывает разделитель строк, то его приходится специально вставлять перед каждой строкой, записываемой в файл; это делается для того, чтобы файл было легче читать:

```
#include <stdio.h>
#include <stdlib.h>
#include <string.h>

void main(void)
{
   char str[80]; FILE *fp;

   if((fp = fopen("test.txt", "w"))==NULL)
   {  printf("Error\n");
      return;
   }

   do {
      printf("Input string\n");
      gets(str);
      strcat(str, "\n");
      fputs(str, fp);
   } while(*str!='\n');

   fclose(fp);
}
```

7.11. Функция rewind()

Функция rewind() устанавливает указатель текущей позиции в файле на начало файла, указанного в качестве аргумента этой функции. Иными словами, функция rewind() выполняет "перемотку" (rewind) файла. Вот ее прототип:

```
void rewind(FILE * файл);
```

Чтобы познакомиться с rewind(), изменим программу из предыдущего раздела таким образом, чтобы она отображала содержимое файла сразу после его создания. Чтобы выполнить отображение, программа после завершения ввода "перематывает" файл, а затем с помощью fback() читает его с самого начала. Обратите внимание, что сейчас файл необходимо открыть в режиме чтения/записи, используя в качестве аргумента, задающего режим, строку "w+".

```
#include <stdio.h>
#include <stdlib.h>
#include <string.h>

void main(void)
{
  char str[80]; FILE *fp;

  if((fp = fopen("test.txt", "w+"))==NULL)
  {
    printf("Ошибка при открытии файла.\n"); return;
  }

  do
  {
    printf("Введите строку (пустую - для
выхода):\n");
    gets(str);
    strcat(str, "\n");  /* ввод разделителя строк */
    fputs(str, fp);
```



```
  }
  while(*str!='\n');

  rewind(fp);

  while(!feof(fp))
  {
    fgets(str, 79, fp);
    printf(str);
  }

  fclose(fp);
}
```

7.12. Функция ferror()

Функция ferror() определяет, произошла ли ошибка во время выполнения операции с файлом. Прототип этой функции следующий:

```
int ferror(FILE *файл);
```

Функция возвращает значение true (истина), если при последней операции с файлом произошла ошибка; в противном же случае она возвращает false (ложь). Так как при любой операции с файлом устанавливается свое условие ошибки, то после каждой такой операции следует сразу вызывать ferror(), а иначе данные об ошибке могут быть потеряны.

7.13. Стирание файлов

Функция remove() стирает указанный файл. Вот ее прототип:

```
int remove(const char *имя_файла);
```

В случае успешного выполнения эта функция возвращает нуль, а в противном случае — ненулевое значение.

7.14. Функции fread() и fwrite()

Для чтения и записи данных, тип которых может занимать более 1 байта, в файловой системе языка С имеется две функции: fread() и fwrite(). Эти функции позволяют читать и записывать блоки данных любого типа. Их прототипы следующие:

```
size_t fread(void *буфер, size_t колич_байт,
             size_t счетчик, FILE * файл);

size_t fwrite(const void *буфер, size_t колич_байт,
              size_t счетчик, FILE * файл);
```

Для fread() буфер — это указатель на область памяти, в которую будут прочитаны данные из файла. А для fwrite() буфер — это указатель на данные, которые будут записаны в файл. Значение счетчик определяет, сколько считывается или записывается элементов данных, причем длина каждого элемента в байтах равна колич_байт. (Вспомните, что тип size_t определяется как одна из разновидностей целого типа без знака.) И, наконец, уф — это указатель файла, то есть на уже открытый поток.

Функция fread() возвращает количество прочитанных элементов. Если достигнут конец файла или произошла ошибка, то возвращаемое значение может быть меньше, чем счетчик. А функция fwrite() возвращает количество записанных элементов. Если ошибка не произошла, то возвращаемый результат будет равен значению счетчик.

7.14.1. Использование fread() и fwrite()

Как только файл открыт для работы с двоичными данными, fread() и fwrite() соответственно могут читать и записывать информацию любого типа. Например, следующая программа записывает в дисковый файл данные типов double, int и long, а затем читает эти данные из того же

файла. Обратите внимание, как в этой программе при определении длины каждого типа данных используется функция sizeof().

```
#include <stdio.h>
#include <stdlib.h>

void main(void)
{
  FILE *fp;
  double d = 12.23;
  int i = 101;
  long l = 123023L;

  if((fp=fopen("test", "wb+"))==NULL) {
    printf("Error");
    return;
  }

  fwrite(&d, sizeof(double), 1, fp);
  fwrite(&i, sizeof(int), 1, fp);
  fwrite(&l, sizeof(long), 1, fp);

  rewind(fp);

  fread(&d, sizeof(double), 1, fp);
  fread(&i, sizeof(int), 1, fp);
  fread(&l, sizeof(long), 1, fp);

  printf("%f %d %ld", d, i, l);

  fclose(fp);
}
```

Как видно из этой программы, в качестве буфера можно использовать (и часто именно так и делают) просто память, в которой размещена переменная. В этой простой программе значения, возвращаемые функциями fread() и fwrite(), игнорируются. Однако на практике эти значения необходимо проверять, чтобы обнаружить ошибки.

Одним из самых полезных применений функций fread() и fwrite() является чтение и запись данных пользовательских типов, особенно структур. Например, если определена структура:

```
struct struct_type
{
   float balance;
   char name[80];
} cust;
```

то следующий оператор записывает содержимое cust в файл, на который указывает fp:

```
fwrite(&cust, sizeof(struct struct_type), 1, fp);
```

7.15. Функции fprinf() и fscanf()

Кроме основных функций ввода/вывода, о которых шла речь, в системе ввода/вывода языка С также имеются функции fprintf() и fscanf(). Эти две функции, за исключением того, что предназначены для работы с файлами, ведут себя точно так же, как и printf() и scanf(). Прототипы функций fprintf() и fscanf() следующие:

```
int fprintf(FILE *файл, const char *строка, ...);
int fscanf(FILE *файл, const char *строка, ...);
```

Операции ввода/вывода функции fprintf() и fscanf() выполняют с тем файлом, указанным в параметрах функции.

В качестве примера предлагается рассмотреть следующую программу, которая читает с клавиатуры строку и целое значение, а затем записывает их в файл на диске; имя этого файла — **test**. После этого программа читает

этот файл и выводит информацию на экран. После запуска программы проверьте, каким получится файл **test**. Как вы и увидите, в нем будет вполне удобочитаемый текст.

```
#include <stdio.h>
#include <io.h>
#include <stdlib.h>

void main(void)
{
  FILE *fp;
  char s[80];
  int t;

  if((fp=fopen("test", "w")) == NULL)
  {
    printf("Error\n");
    return;
  }

  printf("Введите строку и число: ");

  /* читать с клавиатуры */
  fscanf(stdin, "%s%d", s, &t);

   /* писать в файл */
  fprintf(fp, "%s %d", s, t);
  fclose(fp);

  if((fp=fopen("test","r")) == NULL)
  {
    printf("Error\n");
    return;
  }

  fscanf(fp, "%s%d", s, &t);
  fprintf(stdout, "%s %d", s, t);
}
```

Маленькое предупреждение. Хотя читать разносортные данные из файлов на дисках и писать их в файлы, расположенные также на дисках, часто легче всего именно с помощью функций fprintf() и fscanf(), но это не всегда самый эффективный способ выполнения операций чтения и записи. Так как данные в формате ASCII записываются так, как они должны появиться на экране (а не в двоичном виде), то каждый вызов этих функций сопряжен с определенными накладными расходами. Поэтому, если надо заботиться о размере файла или скорости, то, скорее всего, придется использовать fread() и fwrite().

Учебное издание

Лысаков Константин Федорович

СТРУКТУРНЫЙ И ОБЪЕКТНО-БАЗИРОВАННЫЙ ПОДХОДЫ В ПРОГРАММИРОВАНИИ НА ПРИМЕРЕ ЯЗЫКОВ С И С++

УЧЕБНОЕ ПОСОБИЕ