

Информатика: семестр 3

Программирование на C++

Конспекты лекций

Лектор: Хирьянов Т.Ф.

3 октября 2017 г.

Оглавление

1	Введение в C++	3
1.1	Этапы компиляции	3
1.2	Синтаксис	4
1.3	Арифметические операции и сравнение	4
1.4	Циклы, условия и конструкции	5
1.5	Тернарный оператор	7
1.6	Оператор switch	7
2	Типы в C++	8
2.1	Строгая статическая типизация в C++	8
2.2	Представление целых чисел в C++	8
2.3	Хранение чисел с плавающей точкой	9
2.4	Явное и неявное преобразование типов	9
2.5	Типы целых чисел	10
2.6	Логические операции	11
2.7	Битовые операции C++	11
2.8	Автоматические типы переменных	11
2.9	Простейшие массивы	12
3	Адреса и указатели	13
3.1	Доступ к памяти	13
3.2	Модификатор static	15
3.3	Указатель	16
3.4	Несколько слов о массивах	17
3.5	Динамическая память	18
3.6	Передача параметров по указателю	18
4	Динамическая память. Продолжение.	20
4.1	Динамические массивы	20
4.2	Структуры в C++	21
4.3	Массивы и функции	23
4.4	Двумерный массив и передача его функции	24

4.5	Статические массивы	25
4.6	Динамические массивы	25
5	Строки и файлы в C++	27
5.1	Хранение строк	27
5.2	Операции со строками	27
5.3	C++ – строки	28
5.4	Считывание строки	29
5.5	Файловые потоки	29
5.6	Строковые потоки	30
5.7	Односвязный список	30

Лекция 1

Введение в C++

1.1 Этапы компиляции

C++ — компилируемый язык. Перевод с человеческого языка на машинный код происходит не во время исполнения, а заранее. Сборка программы при этом проходит в несколько этапов:

1. Препроцессинг: `my.cpp` → `my.cpp'`

Здесь происходит подстановка всех макросов, а также удаление комментариев:

Код 1.1: Пример макроса

```
1 #include "mylib"
2 #define A 123456
```

2. Компиляция: `my.cpp'` → `main.o`

Он содержит переведенную на машинный код инструкции, но все еще содержит сигнатуру функций. Это означает, что объектный файл может содержать названия функций, их входные данные, но не саму реализацию.

3. Компоновка (линковка): `my.o` → `my.exe`

Здесь уже все файлы связываются и функции, не реализованные в п.3, получают реализацию. Программа может быть разделена на модули так, что разные файлы используют функции других файлов. При этом, если один файл ссылается на ресурсы других файлов, эти ресурсы необходимо объявить в отдельном файле: header-файле (хидер на жаргоне). Хидер содержит лишь сигнатуры функций и переменных, однако не входит в не конечный результат компиляции. Во время компоновки эти хидеры будут показывать, где искать функции.

Подобное деление лежит в основе принципа раздельной компиляции: программы делятся на компоненты, которые собираются по отдельности. Это уменьшает время компиляции.

1.2 Синтаксис

Код 1.2: Демонстрация синтаксиса

```
1 #include<iostream>
2 int main()
3 {
4     int x, y;
5     std::cin >> x;
6     std::cout << "x=" << x << "\n"; // "\n" === std::endl
7     << "y=" << y << std::endl;
8     return 0;
9 }
```

В языке C++ есть особенности, отличающие его от Python. Так, при создании переменной необходимо явно указать ее тип (строка 4).

В первой строке — подключение библиотеки для работы ввода/вывода.

В строке 5 стоит оператор ввода из стандартного ввода (это консоль) и спец. оператор `>>`. Результат считывания уходит в переменную `x`. `std` — это пространство имен, о них будет сказано позже.

В строке 6 — вывод, по аналогии с вводом. При этом строки можно выводить конвейером (как видно `'cout << "x=" << x'` — эти строки будут сшиты при выводе).

В C++ табуляция не имеет значения, перевод строки после (6) и дальнейшие команды показывают это.

Строка 8 содержит код возврата — он будет возвращен по окончании программы (0 означает успех).

1.3 Арифметические операции и сравнение

`x = y` — арифметическая операция (не присваивание, как в питоне). Можно писать такой код

Код 1.3: Неясное поведение

```
1 y=(x=3)+(z=5)
```

В коде выше все работает, однако такое поведение программы будет неочевидным. В C++ все еще работают `x += 1`. Помимо этого имеются два оператора: `x++` \approx `++x` \approx `x+=1`. Разница следующая: `y = (x++)` запишет в `y` **старый** `x`, в то время как в `y=(++x)` будет записан **новый** `x`. Операции `**` (возведение в степень) *нет*. Примеры некоторых арифметических действий:

1. `x*y` — при применении к целым числам результат будет целым числом.

2. x/y — при применении к целым числам будет применено целочисленное деление. Проблему можно решить, если добавить в конце ".". Так, $2/3$ станет $2/3.$ — точка превратит число в вещественное.
3. $x > y$, $x \geq y$, $x == y$. Это тоже арифметические операции, возвращают, правда, значение типа *bool*: **true** или **false**.

1.4 Циклы, условия и конструкции

В C++ три вида циклов, их синтаксис отличается от Python:

Код 1.4: Виды циклов

```
1 // 1 type
2 while(condition)
3 {
4     some code;
5 }
```

Условие надо обязательно кидать в скобки, иначе C++ не сможет отличить его от тела.

Код 1.5: Виды циклов

```
1 // 2 type
2 while(condition)
3 {
4     if(condition)
5     {
6         std::cout << "Exiting .. \n";
7         break;
8     }
```

Здесь мы выходим из цикла по команде **break**.

Код 1.6: Виды циклов

```
1 // 3 type
2 do {
3     ...
4 } while(condition);
```

Как обычный **while**, но сначала выполняет тело цикла, а затем проверяет условие.

Код 1.7: Виды циклов

```
1 // 4 type
2 for (op1; condition; op3)
3 {
4     some code;
5 }
6 // EQUAL:
7 op1;
8 while (condition)
9 {
10    some code;
11    op3;
12 }
13 // Example:
14 for (i=0; i < 10; i++)
15 {
16     some code;
17 }
```

Это очень похоже на то, что в питоне было под видом

Код 1.8: Аналог цикла for в Python

```
1 for i in range(0, 10, 1):
2     some code
```

Переменные существуют локально, поэтому можно переменную *i* сделать только для цикла. Более того, переменная "живет" до тех пор, пока не будет завершен соответствующий блок:

Код 1.9: Виды циклов

```
1 for (int i=0; i < 10; i++) // "i" won't be seen outside
2 {
3     some code;
4 }
5
6 // else == elif
7 if (condition)
8 {
9     ..
10 } else if (condition2)
11 {
12     ..
13 }
14 else {
15     ..
16 }
```

1.5 Тернарный оператор

Тернарный оператор заменяет условия (иногда полезно):

Код 1.10: Пояснение к тернарному оператору

```
1 z = (a > 0) ? x : y;  
2 // EQUALS TO:  
3 if (a > 0)  
4 {  
5   z = x;  
6 }  
7 else  
8 {  
9   z = y;  
10 }
```

Такой подход здорово уменьшает код, однако при неправильном использовании может сделать его нечитаемым.

1.6 Оператор switch

Код 1.11: Оператор switch

```
1 switch (x)  
2 {  
3   case 0: op1; // if x=0  
4     op2;  
5     break;  
6   case 1: op1; // if x=1  
7     op2;  
8     break;  
9   default: op1; // By default  
10 }
```

break — обязательно в конце, так как **case 0** — это метка (такое поведение берет корни из Assembler) для перехода. Как только значение *x* будет подходящим, программа прыгнет в соответствующую метку, но что будет после выполнения основного кода? Дальше она просто пойдет по остальным, ведь метка не накладывает никаких условий для входа (она просто обозначает кусок программы). Именно поэтому необходим **break** — он завершает выполнение **switch**.

Лекция 2

Типы в C++

2.1 Строгая статическая типизация в C++

Проверка типов происходит в момент компиляции.

Код 2.1: Неправильное присваивание

```
1 int x;  
2 x = "Hello"; // won't work
```

Помимо этого, присваивание является арифметической операцией (делает дело и возвращает число).

2.2 Представление целых чисел в C++

Существует три способа хранить целые отрицательные числа в памяти компьютера:

1. Прямой ход. Сделаем первый бит знаковым. 0 — знак минус, 1 — знак плюс. Недостатком будет то, что данное число можно интерпретировать как беззнаковое. Пример: $10010011_2 = -19_{10}$. Есть еще одна проблема. 0 трактуется дважды: $-0 = 0$
2. Обратный ход. Мы инвертируем исходное число: $10010011 \rightarrow 11101100$. Просто поменяли 0 на 1 и 1 на 0 (это называется побитовое отрицание).
3. Дополнительный код. Обратный код, к которому прибавлена единица. В чем profit: $11101101 = -19$. Попробуем к нему прибавить число 31:

+31		00011111
-19		11101101
<hr/>		
12		00001100

Обратите внимание: здесь ушла единица, которая возникла слева при сложении

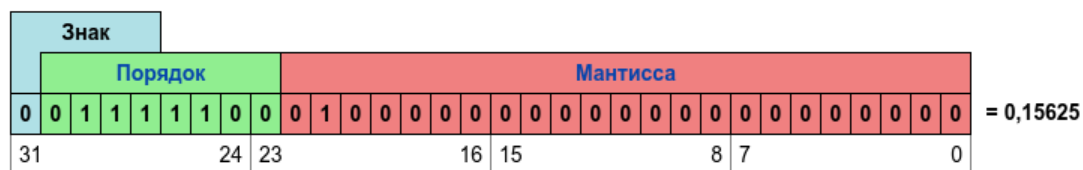
(от увеличения старшего разряда). В этом и заключается вся прелесть дополнительного кода: можно использовать тот же сумматор для двух чисел.

Помимо этого есть *тупоконечное* и *остроконечное* представление. Когда число сохраняется в память, его биты уходят в память в определенном порядке. В тупоконечном порядке сначала в памяти идут старшие разряды, т.е., если мы будем "читать" содержимое ОЗУ слева направо сверху вниз, то все будет идти в правильном порядке. В остроконечном порядке - наоборот. В компьютерах Intel, например, используется остроконечная система.

2.3 Хранение чисел с плавающей точкой

Мы можем условиться: первые биты брать как целую часть, дальше - дробную. Поскольку любое число можно представить бесконечной периодической двоичной дробью, мы сможем представить все числа.

IEEE 754



Стандарт для хранения чисел с плавающей точкой. Можно почитать здесь: [вики](#).

Числа одинарной точности с плавающей запятой обеспечивают относительную точность 7-8 десятичных цифр в диапазоне от 10^{-38} до примерно 10^{38} . Для вычисления показателя степени из восьмиразрядного поля порядка вычитается смещение порядка равно $127_{10} = 7F_{16} = 01111111_2$ (то есть, $01111110_2 - 01111111_2 = 124_{10} - 127_{10} = -3_{10}$). Так как в нормализованной двоичной мантиссе целая часть всегда равна единице, то в поле мантиссы записывается только её дробная часть. Для вычисления мантиссы к единице добавляется дробная часть мантиссы из 23-х разрядного поля дробной части мантиссы $1,01000000000000000000000_2$. Число равно произведению мантиссы со знаком на двойку в степени порядка $= 1,01_2 * 2_{10}^{-3_{10}} = 101_2 * 2_{10}^{-5_{10}} = 5_{10} * 2_{10} - 5_{10} = 0,15625_{10}$.

(Взято из [другой статьи](#))

2.4 Явное и неявное преобразование типов

Код 2.2: Неявное преобразование

```
1 double x;
2 int a = 2, b = 5;
```

```
3 x = b/a; // NOT DOUBLE without casting
4
5 x = double(b)/double(a); // explicit casting
6
7 x = b; // auto convert b to double, then will write to x
```

В коде выше используется неявное преобразование: дает возможность писать код гибче, однако требует осторожнее (так, в строке 3 используется *целочисленное* деление).

Код 2.3: Не всегда удастся преобразовать типы

```
1 std::string s;
2 s = std::string(x); //won't work
```

2.5 Типы целых чисел

В C++ есть несколько типов целых чисел:

1. **char** - код символа (по умолчанию 1 байт)
2. **int** - целое число (сколько байт - не известно).

Не стоит быть уверенным, что **char** - 1 байт, на разных процессорах по-разному. В C++ так устроено: язык заточен под процессор, на котором он будет выполняться. Для обоих типов можно указать **signed** и **unsigned** - знаковый тип или нет.

Код 2.4: Пример доп. указателей к типам

```
1 signed int x;
2 unsigned int y;
3 long int z;
4 // size
5 sizeof(int); // about 4
6 long int // 4 bytes, too
7 unsigned long long int // 8 bytes now
8 short int // 2 bytes, wow
9
10 int8_t x8;
11 int16_t x16;
12 int32_t x32;
13 int64_t x64;
14 // They all has fixed size, dude
```

. Стоит быть внимательным: сравнение для знаковых и беззнаковых может пойти не так, как хотелось бы. Пример: $-1 > 50$ - верное равенство, если -1 имеет тип **signed**, а 50 - **unsigned**. Проблема заключается в том, что -1 соответствует максимальному числу из **unsigned**.

2.6 Логические операции

Работают над переменными типа **bool**.

не	!	A
и	&&	A B
или		A B

Здесь почти то же самое, что в Python, лишь синтаксис другой.

2.7 Битовые операции C++

Как можно поменять две переменные без третьей? Для этого могут помочь логические операции. В C++ их несколько:

1. Инверсия – x – работает над *любым* числом, не обязательно **bool**. Заменяет все ноли и единички в представлении числа на противоположные.
2. Битовое "и" – $x \& y$ – умножает биты в двоичном представлении двух чисел. Иногда битовое "и" называют *наложением маски*.
3. Битовое "или" – $x | y$ – складывает биты в двоичном представлении двух чисел.
4. хог - $x \wedge y$. С его помощью можно как раз осуществить обмен переменных:
 $x = x \wedge y$
 $y = x \wedge y$
 $x = x \wedge y$

2.8 Автоматические типы переменных

Код 2.5: Автоматический тип переменной

```

1 #include<iostream>
2
3 int get_x(); // 'auto x' will be int
4
5 double get_x(); // now 'auto x' will be double
6 // and so on...
7
8 int main()
9 {
10     auto x = get_x();
11     return 0;
12 }
```

Код будет присваивать переменной `x` различные типы, однако этот трюк будет работать, только если компилятору сказать использовать современный стандарт (например, C++11)!

2.9 Простейшие массивы

В C++, как и во многих языках, существуют массивы.

Код 2.6: Пример объявления простейшего массива

```
1  const int N_max = 1000;
2
3  int main()
4  {
5      int A[N_max];
6      int B[N_max] = {1, 2, 3};
7
8      // Let's iterate!
9      for (int i = 0; i < N_max; i++)
10         A[i] = i*i;
11
12 }
```

При объявлении массива надо указать его размер. По массиву можно итерироваться через обращение к элементу (точно так же, как в Питоне): `A[1]` – второй элемент и так далее; нумерация идет с *нуля*.

Подробнее о массивах будет рассказано на следующей лекции.

Лекция 3

Адреса и указатели

3.1 Доступ к памяти

Память, выделяемая программе, неоднородна – она делится на **сегменты**. Первый сегмент — коды программы (режим "только чтение"). Второй – сегмент DATA – содержит данные программы. Более того, элементы DATA (к примеру, начальные значения переменных) записывается в .exe файл. Эти данные уходят в оперативку при запуске программы. Переменные, записанные в DATA, называются *статическими*. Помимо этого, для функционирования программы необходима область *стека* — такой сегмент памяти, который заполняется снизу вверх и изначально пуст. Когда ОС только загрузилась, в стек выделяется некоторое количество автоматической памяти – для *локальных переменных*.

Память программе выделяется на время ее работы; после ее завершения память высвобождается.

Код 3.1: Пример многократных вызовов функции

```
1 #include<iostream>
2
3 using namespace std;
4
5 void count(int n)
6 {
7     int k = 1;
8     while(k <= n)
9     {
10         cout << k << " ";
11         k++;
12     }
13     cout << "\n";
14 }
15
16 int main()
```

```
17 {  
18     count(5);  
19     count(10);  
20 }
```

Рассмотрим, как происходит работа с памятью в примере 5.1. Перед вызовом функции в стек программы будет выделено место для хранения "5" (оно в программе будет интерпретировано как *n*), *stack pointer* поднимается. *stack pointer* - это указатель, показывающий на текущее положение стека. Затем, после вызова функции, ее адрес возврата будет записан в стек - *stack pointer* двинется выше. Потом в стек уйдет "1" (мы же объявили ее).

Попробуем написать функцию, считающую количество рекурсивных вызовов.

Код 3.2: Функция для подсчета количества собственных вызовов

```
1 #include <iostream>  
2  
3 using namespace std;  
4  
5 int k2 = 0;  
6  
7 void count2()  
8 {  
9     k2++;  
10    cout << k2 << " ";  
11 }  
12  
13 void reducer()  
14 {  
15     k2--;  
16 }  
17  
18 int main()  
19 {  
20     count2();  
21     reducer();  
22     // any your staff ..  
23 }
```

Здесь *k2* — глобальная переменная. Ее присваивание происходит в *момент компиляции* программы (т.е. она запишется в *.exe* и будет переноситься в память в момент загрузки). Функция **reducer** уменьшает значение *k2*.

Однако глобальные переменные могут быть опасны! Считается, что нормальная функция не заглядывает в глобальные переменные (в таких случаях функцию называют "чистой" в противном - "грязной").

3.2 Модификатор static

Код 3.3: Та же программа через static

```
1 #include<iostream>
2
3 using namespace std;
4
5 //int k2 = 0;
6
7 void count2()
8 {
9     static int k3 = 0;
10    k3 ++ ;
11    cout << k3 << endl;
12 }
13
14 void reducer()
15 {
16     k3 --; // won't work, k3 is not defined in this block
17 }
18
19 int main()
20 {
21     count2();
22     reducer();
23     // any your staff ..
24 }
25
26 // Recursive func sample
27 void count(int n)
28 {
29     int k = 1;
30     if (n > 0)
31         count(n-1);
32 }
```

Мы переделали программу 5.2, теперь переменная k3 не видна в других блоках, но по-прежнему она записывается в сам .exe . Если бы функция была рекурсивной, то была бы такая ситуация, когда переменных k было бы много - на каждом слое рекурсии она создается. Они просто будут записаны в разные ячейки памяти и уйдут после того, как рекурсивный вызов завершится.

Код 3.4: Подсчет рекурсии

```
1 #include<iostream>
2
3 using namespace std;
```



```
4
5 //int k2 = 0;
6
7 void count2 ()
8 {
9     static int k3 = 0;
10    k3++ ;
11    cout << k3 << endl;
12    if (n > 0)
13        count2 (n-1)
14    k3 --;
15 }
16
17 int main()
18 {
19     count2 ();
20     // any your staff ..
21 }
```

В коде выше как раз **static** делает `k` одним для всех вызовов рекурсии — оно при каждом вызове не будет присваиваться в ноль. Вместо этого переменная будет увеличиваться на 1, таким образом, мы сможем увидеть глубину рекурсии.

3.3 Указатель

Указатель - переменная адресного типа.

Код 3.5: Адрес `x` можно указать

```
1 int x;
2 &x; // pointer to x
3     // it has type 'int*'
4 int* p = &x;
5 double* pd = &x; // won't work
```

Указатели тоже работают только со своими типами: в коде выше нельзя в **double*** записать число **int**. Дело в том, что возникнет ложная интерпретация – ссылка на `double` превратится в ссылку на `int`, а поскольку в памяти `double` и `int` имеют разные размеры, то можно даже вылезти за пределы своей памяти. Компилятор это запретит, однако все можно обойти:

Код 3.6: Пример обхода запрета

```
1 int x;
2 double* pd = (double*) &x;
```

```
3  
4 int main()  
5 {  
6     return 0;  
7 }
```

У указателей могут тоже быть указатели:

Код 3.7: Указатель на указатель

```
1 int x;  
2 int* p = &x;  
3 int** pp = &p;  
4 int*** ppp = &pp;  
5 // ...
```

Обратиться к переменной можно несколькими способами:

Код 3.8: Разница

```
1 int x;  
2 int* p = &x;  
3 cout << x; // print x  
4 cout << p; // print ADDRESS of x  
5 cout << *p; // print x
```

В последней строке происходит *разыменование* адреса - грубо говоря, считывание данных из того адреса, куда указывает указатель. Разыменовывать надо последовательно:

Код 3.9: Двойное разыменование

```
1 cout << x;  
2 cout << *p;  
3 cout << **pp;  
4 cout << ***ppp;
```

3.4 Несколько слов о массивах

Массивов в C++ нет. Массив - это константный указатель на область памяти, представляющую содержимое!

Код 3.10: Работа массива

```
1 int A[5] = {0, 1, 2, 3, 4};
2 int* p = A; // works!
3 *p = 7
4 p + 1 // magic, read in text
5
6 // Furthermore:
7 A[i] == *(A+i)
```

Команда объявления массива по факту означает следующее: создать n переменных типа `int`.

С указателями можно прийти к *адресной арифметике*: так, указатель $p + 1$ выше означает, что мы сдвигаемся по памяти от места, куда указывает p , на `sizeof(1)`. Это иллюстрирует строка 7: `A[i]` абсолютно равносильна указателю `*(A+i)`. Проверки на *i нет*, то есть, в теории, мы можем убежать вообще в произвольную область памяти.

3.5 Динамическая память

Код 3.11: Пример выделения из дин. памяти

```
1 int get ()
2 {
3     int* p = new int[1000]
4     return p;
5 }
6
7 int main()
8 {
9     int* q;
10    q = get ();
11    // ...
12    // after your code
13    q = get (); // takes another amount of mem
14    delete [] q;
15 }
```

Динамическая память не является автоматической, т.е. ей надо явным образом указывать, когда ее выделять, а когда удалять. Если без конца писать вызов `get()`, то можно постоянно выбивать у системы память и это все приведет к *memory leak* - утечке памяти.

3.6 Передача параметров по указателю

Передача по параметрам бывает полезна, когда хочется поменять значение переменной внутри самой функции. Функции передается указатель и она, имея ссылку непо-

средственно на область памяти, может ее менять:

Код 3.12: Передача указателей функции

```
1 void changer(int* x)
2 {
3     (*x)++;
4 }
5
6 // You can use another construction in C++
7
8 void changer1(int& y)
9 {
10     y++;
11 }
12
13
14 int main()
15 {
16     int x;
17     changer(&x);
18     changer1(x); // no need for &
19     cout << x;
20 }
```

Подробнее – в следующей лекции.

Лекция 4

Динамическая память. Продолжение.

4.1 Динамические массивы

Как было сказано, массивов в C++ нет. Повторим объявление указателей:

Код 4.1: Пример объявления указателей

```
1 double A[100]; // has type 'const double*'
2 double* p = A;
3 //or
4 double* p1 = &A[5]
5 //equals to
6 double* p1 = A + 5
7
8 double* p = new double(5.);
```

В строке 4 мы получаем указатель на **A[5]**. В строке 8 мы создаем число, да и еще присваиваем ему значение.

Код 4.2: Объявление массива через **new**

```
1 double* p = new double[50];
2
3 for(int k = 0; k < 50; k++)
4     p[k] = k*k;
5
6 delete[] p;
```

При объявлении массива через **new** нужно обязательно его удалять через **delete[]** – оператор удаления.

4.2 Структуры в C++

В C++ существует такая вещь, как структура – набор переменных под одной крышей.

Код 4.3: Простейшая структура

```
1 #include <iostream>
2
3 struct Student {
4     int8_t age;
5     int16_t group;
6     std::string name;
7 };
8
9 void student_print(Student x) {
10     std::cout << x.name << " " << x.age
11     << " " << x.group << "\n";
12 }
13
14 int main()
15 {
16     Student vasya = {16, 649, "Vasilii Ivanov"};
17     Student masha = {18, 649, "Maria Vasilyeva"};
18
19     Student *p = nullptr;
20
21     p = &vasya;
22
23     std::cout << (*p).name << "\n";
24     // This code is equal to
25     std::cout << p -> name << "\n";
26     student_print(masha);
27
28     return 0;
29 }
30
31 void studentPrint1(Student *x) {
32     std::cout << x->name << " " << x->age
33     << " " << x->group << "\n";
34 }
35
36 void studentPrint2(const Student &x) {
37     std::cout << x.name << " " << x.age
38     << " " << x.group << "\n";
39 }
```

Обратите внимание на строки **23** и **25**: такое сокращение еще часто будет встречаться (для указателей вместо "." используется "—>"). Метод для распечатки мы вынесли

отдельно для удобства.

Подход к распечатке в самой первой функции имеет минус: при вызове функции мы создаем локальную копию `Student`. Метод **studentPrint1** лишен этого – там передается по ссылке. Но возникает проблема, ведь мы внутри функции случайно можем изменить переменную `x` и этот результат будет виден извне. Решение дает передача по константной ссылке, что сделано в **studentPrint2**.

Существует еще одно решение задачи, в котором мы сначала создаем объект `Student` в памяти, получаем на него указатель, а затем через него меняем свойства объекта:

Код 4.4: Еще одно решение со структурой

```
1 #include <iostream>
2
3 struct Student {
4     int8_t age;
5     int16_t group;
6     std::string name;
7 };
8
9 void studentPrint(const Student &x);
10
11 int main()
12 {
13     Student *p_student = new Student;
14     p_student->age = 17;
15     p_student->group = 649;
16     p_student->name = "Vasya Ivanov";
17
18     studentPrint(*p_student);
19     return 0;
20 }
21
22 void studentPrint(const Student &x) {
23     std::cout << x.name << " " << x.age
24     << " " << x.group << "\n";
25 }
```

Из структур можно составлять массивы, что неудивительно, учитывая, что массив – это просто несколько переменных. В качестве примера можно привести следующий код:

Код 4.5: Массив из структур

```
1 #include <iostream>
2
3 struct Student {
4     int8_t age;
5     int16_t group;
6     std::string name;
7 };
8
9 void studentPrint(const Student &x);
10
11 int group_max_size = 10;
12 int main()
13 {
14     Student *my_group = new Student[group_max_size];
15     for(int k = 0; k < group_max_size; k++)
16     {
17         my_group[k].age = 17 + k;
18         my_group[k].group = 600 + k;
19         my_group[k].name = "Vasya Ivanov";
20     }
21
22     for(int k = 0; k < group_max_size; k++)
23     {
24         studentPrint(my_group[k]);
25     }
26     return 0;
27 }
28
29 void studentPrint(const Student &x) {
30     std::cout << x.name << " " << x.age
31     << " " << x.group << "\n";
32 }
```

4.3 Массивы и функции

В функциях можно менять массивы, но их нельзя возвращать. Код внизу – пример изменения массива внутри функции.

Код 4.6: Изменение массива внутри функции

```
1 #include<iostream>
2
3 void squarer(int* B, const int& B_size)
4 {
5     for(int k = 0; k < B_size; k++) {
6         B[k] = B[k] * B[k];
7     }
8 }
9
10 int main()
11 {
12     int A[5];
13     squarer(A, 5);
14     return 0;
15 }
```

4.4 Двумерный массив и передача его функции

Двумерные массивы работают так же, как в Питоне. Тем не менее, у них есть некоторые особенности, связанные с их способом представления в памяти (помним, что массив по факту лишь набор переменных). Одна из этих особенностей видна в коде ниже, где мы передаем двумерный массив в функцию:

Код 4.7: Передача массива в функцию

```
1 #include<iostream>
2
3 //void squarer(A);
4
5 void printer(double A[][4])
6 {
7     for(int k = 0; k < 3; k++)
8     {
9         for(int j = 0; j < 4; j++)
10             std::cout << A[k][j] << 't';
11         std::cout << 'n';
12     }
13 }
14
15 int main()
16 {
17     double A[3][4];
18     for(int k = 0; k < 3; k++)
19         for(int j = 0; j < 4; j++)
20             A[k][j] = 1 + k*4 + j;
```

```
21  
22     printer(A);  
23  
24     return 0;  
25 }
```

Обратите внимание: в объявлении функции **printer** мы обязательно указываем длину (вторая квадратная скобка).

4.5 Статические массивы

Код 4.8: Указание на элемент массива

```
1 A[i][j] == *(A+i*y+j)
```

В коде выше видно, что для прохода по элементам массива в функции необходимо заранее знать его длину. Если посмотреть на выражение 4.8, то станет видно, что эта длина уходит в инструкции прохода по массиву еще на этапе компиляции, поэтому ее и нужно было передавать заранее, как это сделано в 4.7. Но можно ли сделать функцию **printer** без явного указания длины массива? Ответ дают динамические массивы.

4.6 Динамические массивы

Динамические массивы — это массивы, хранящиеся в так называемой динамической памяти (*heap*). Динамические массивы объявлять несколько сложнее: так, для двумерного случая необходимо делать указатель на область памяти с указателями, каждый указатель в которой, в свою очередь, будет указывать на область памяти с нашими объектами (к примеру, **int**). Звучит не очень, однако пример 4.10 может немного прояснить этот момент.

Код 4.9: Объявление динамического массива

```
1 #include <iostream>  
2  
3 //void squarer(A);  
4  
5 // TODO: func must take width and height, fix it  
6 void printer(double A[][4])  
7 {  
8     for(int k = 0; k < 3; k++)  
9     {  
10         for(int j = 0; j < 4; j++)  
11             std::cout << A[k][j] << '\t';
```

```
12         std::cout << '\n';
13     }
14 }
15
16 int main()
17 {
18     int width = 4, height = 3;
19     double **A = nullptr;
20     A = new double*[height];
21     for (int k = 0; k < height; k++)
22         A[k] = new double[width];
23
24     for (int k = 0; k < height; k++)
25         for (int j = 0; j < width; j++)
26             A[k][j] = 1 + k*width + j;
27
28     printer(A);
29
30     for (int k = 0; k < height; k++)
31         delete [] A[k];
32     delete A;
33
34     return 0;
35 }
```

A – это массив из указателей типа **double***. Каждый элемент этого массива - массив размера **width**. Таким образом, **A** есть двумерный массив. Обращение к его элементам происходит так же, как и при обычном объявлении массива, что видно в строке **26**. Мы сделали массив, распечатали его, а затем удалили, чтобы освободить память. При этом следует помнить такой трюк:

Код 4.10: Порядок разыменования

```
1 A[k][j] == (*(A + k) + j)
```

Лекция 5

Строки и файлы в C++

5.1 Хранение строк

Строки делятся на *C-строки* и *ANSI-строки*. В разных языках строки реализованы по-разному. Так, в Pascal они хранятся как *ANSI-строки*: в первой ячейке памяти записана длина строки, а затем последовательно идут символы строки. Главный минус такого подхода — ограниченность длины строки (255 символов). Этот недостаток отсутствует в *C-строках*.

В *C-строках* используется следующая схема хранения: в памяти просто идет последовательность символов и завершается **нулем** (типа **int**, не "0"). Так, строка "Hello" будет записана как последовательность 'H', 'e', 'l', 'l', 'o', 0x0 (строки такого типа называются *zero-ending string*). НО! Если мы сделаем обращение и изменим какой-либо символ на 0 (типа целочисленное), то строка поделится на две, ведь теперь у нее два символа окончания. Это — один из минусов такого подхода.

5.2 Операции со строками

Конкатенация строки можно делать с помощью стандартной библиотеки **cstring**.

Код 5.1: Работа со строками

```
1 #include <cstring>
2 #include <iostream>
3
4 using namespace std;
5
6 int main()
7 {
8     char s1[10000], s2[1000];
9     s1[0] = 0;
10    std::cin.getline(s2, 1000);
11    strcat(s1, s2); // concatenate
```

12 | }

В этом коде мы считываем строки, а затем их конкатенируем. Но Т.Ф. не любит эту библиотеку. Дело в том, что `strcat` сначала проходит и за линейное время находит ноль в конце строки, дабы узнать, где ее конец, а затем уже объединяет обе. Если такой код засунуть в цикл, то (из свойств арифметической прогрессии) мы подарим нашему коду квадратичную асимптотику.

В MacOS сделано несколько интереснее: используется как число, равное длине строки, в ее начале, так и символ окончания в конце (комбинированный подход).

При работе со строками следует быть осторожными: при операции `+=` можно уйти за аллоцированную (выделенную) память для данной строки и стереть ее содержимое. Кстати, в C++ строка — это объект. Для нас сейчас это значит, что операции, которые нам кажутся простыми (наподобие `+=`), на деле такими не являются — они реализованы по уму. К тому же, у строк есть свои методы и поля. К примеру, метод `s.size()` вернет длину строки за $O(1)$. У строк автоматическая работа с памяти — язык автоматически выделяет столько памяти, сколько нужно.

Т.Ф. рекомендует пользоваться C++ — строками вместо C-строк.

5.3 C++ — строки

В C++ строки входят в окружение `std`.

Код 5.2: Объявление строки в C++

```
1 #include <string>
2 #include <iostream>
3
4 int main()
5 {
6     std::string s("Hello");
7     for(int k=0; k < s.size(); k++)
8         std::cout << s[k] << std::endl;
9 }
```

В Python строки кодируются стандартом UTF-8. В таком подходе память используется в разных количествах для разных символов, поэтому при вставке символа в середину строки возникают проблемы переноса (одна из причин, почему в Python строки неизменяемы). C++ лишен такого недостатка и здесь строку можно менять посимвольно.

C++ не разделяет типы символа и *кода* символа. Символ = код символа. Вообще, подробности можно изучить [здесь](#).

5.4 Считывание строки

Существует несколько способов читать строки:

Код 5.3: Считывание строки

```
1 while (cin >> s) // (cin >> s) returns true or false
2 {
3     cont << s << '*';
4 }
5
6 // Another way
7 getline(cin, s)
```

Обратите внимание: `cin >> s` возвращает **true** или **false**, в зависимости от того, закончился ли ввод.

Другой способ заключался в методе **getline**, которому на вход поступает поток ввода (см. ниже) и строка `s`, куда будет записано. Различие следующее: **cin** читает до символа пробела, тогда как **getline** читает до символа переноса строки.

5.5 Файловые потоки

Стандартная библиотека для работы с потоками — это **fstream**. Она дает три потока для работы: **ifstream** для ввода и **ofstream** для вывода и **fstrim** для обоих вариантов.

Код 5.4: Пример работы с потоками

```
1 #include <fstream>
2 #include <iostream>
3
4 using namespace std;
5
6 void read_array(int *A, int A_size, string f_name)
7 {
8     ifstream fin(f_name); // input stream based on file
9     for(int k=0; k < A_size; k++)
10         fin >> A[k];
11 }
```

В коде выше было бы неплохо закрыть файл после окончания работы с ним. Более того, операционная система может дать отказ, если мы затребуем открытие слишком большого количества файлов. Однако C++ умный, он закроет файл после окончания существования объекта (в ходе выполнения деструктора, о них будет сказано в следующих лекциях).

Во время открытия потока можно указать для него *режим*: только для чтения и т.п.

Код 5.5: Режимы потока

```
1 fstream f("1.txt", "rwt")
```

Подробнее о режимах можно почитать в документации.

5.6 Строковые потоки

Работа со строковыми потоками тоже проходит через стандартную библиотеку:

Код 5.6: Строковые потоки

```
1 #include <sstream>
2
3 using namespace std;
4
5 int main()
6 {
7     std::stringstream t;
8     t << "x = " << 7 << "." << true << std::endl; // any type is compatible
9     string s;
10    s = sout.str(); // cast stream to string
11    return 0;
12 }
```

5.7 Односвязный список

Попробуем сделать односвязный список в C++ по аналогии с тем, что мы делали в Python.

Код 5.7: Реализация односвязного списка

```
1 #include <iostream>
2
3 using namespace std;
4
5 typedef std::string DataType;
6
7 struct tNode
8 {
9     DataType data;
10    tNode* next = nullptr;
11 };
12
13 tNode * insert_node(tNode *p_begin, DataType data);
```

```
14
15 int main()
16 {
17
18     tNode* p = new tNode;
19     p->data = "Hello";
20     // Sequently add nodes
21     p = insert_node(p, "World");
22     p = insert_node(p, "Reverse order");
23
24     while(p) {
25         std::cout << p->data << std::endl;
26         p = p->next;
27     }
28
29     return 0;
30 }
31
32 tNode * insert_node(tNode *p_begin, DataType data)
33 {
34     tNode *p = new tNode;
35
36     p->data = data;
37     p->next = p_begin;
38     return p;
39 }
```

1. **typedef** — это создание синонима типа **string**. Это сделано только для того, чтобы не писать каждый раз **std::string**, другого назначения у такой конструкции для нас нет.
2. **struct tNode** — объявление структуры ноды. В ней всего две переменные: свое значение и ссылка на следующую ноду.
3. **tNode * ...** — сигнатура функции *insert_node*. Ее мы опишем позже.
4. В методе **main** мы создаем ноду через **new**, затем сохраняем ссылку на созданный объект. После этого мы записываем туда какую-то информацию ("Hello" в примере), далее прикрепляем к ней другие ноды.
5. В цикле мы прогуливаемся по всем нодам и выводим их содержимое. Обратите внимание на переход к следующей ноды в последней строке цикла.
6. Функция **insert_node** — изюминка нашего кода. Ей передается указатель на нашу ноду. Внутри функции мы создаем новую ноду, заполняем в нее данные из **data**, а в качестве следующей ноды записываем ту, что мы передали функции.

Из этого получается, что мы добавляем ноды в обратном порядке, так как только что созданная нода встает впереди старой.