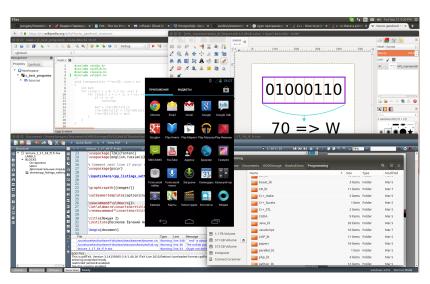
posevg@yandex.ru

12 февраля 2020

# До изучения языка программирования: взгляд снизу на типы данных

# Работа с информацией: что видим мы?



# Работа с информацией: что видит компьютер?

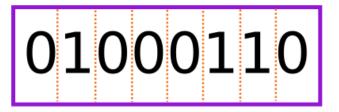
Современные процессоры работают с потоком **бит** - элементарных ячеек для хранения информации, принимающих, как правило, только два значения

# Работа с информацией: что видит компьютер?

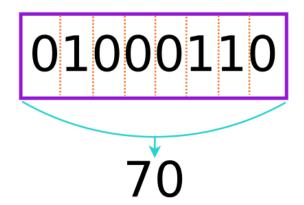
Поток бит делятся на блоки различной длины. Учитывая определение бита, каждый блок представляет собой некоторое число в двоичной системе исчисления

## Работа с информацией: что видит компьютер?

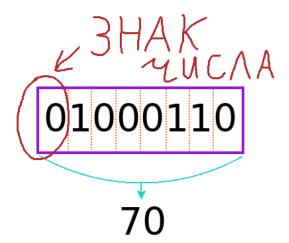
Минимальным блоком в современных ЭВМ является **байт**. На всех популярных ОС 1 байт состоит из 8 бит

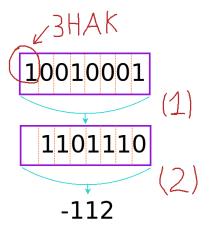


Целое число: восьми бит хватит для хранения 256 чисел в диапазоне [0; 255]



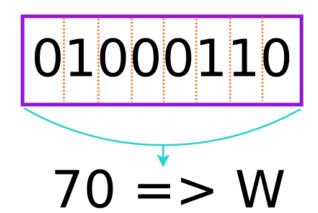
Целое число со знаком: первый бит отвечает за знак, диапазон теперь - [-128; 127]





Когда бит знака равен единице - число отрицательно. (1) - инвертирование значащих битов, (2) - получение отрицательного числа в десятичной системе исчисления.

Некоторый символ: вводится таблица соответствия между числом и набором текстовых символов



# Работа с информацией: формализация "типа данных"

#### Определение

Тип данных в языках программирования представляет собой:

- заданное хранилище (storage): кратная числу байт область в памяти;
- конечное множесто допустимых значений и их интерпретацию;
- заданные операции для работы с допустимыми значениями.

**Переменная** конкретного типа данных при её использовании в программе  $\mathcal{B}cez\partial a$  занимает ограниченный блок байт в оперативной памяти ЭВМ.

# Переходим к конкретному языку программирования

## Где получить информацию о С++?

#### Литература

- Алекс Эллайн, "С++. От ламера до программера 2015 г.,
   Питер ("Jumping into C++")
- Татьяна Павловская, "С/С++. Процедурное и объектно-ориентированное программирование 2014 г., Питер
- Бьярне Страуструп "Программирование. Принципы и практика с использованием С++ 2016 г., Вильямс
- Питер Готшлинг, "Современный С++ для программистов, инженеров и учённых 2016 г., Вильямс

#### Альтернатива онлайн

- github.com/posgen/OmsuMaterials
- en.cppreference.com/w/cpp ( ru.cppreference.com/w/ )
- www.cplusplus.com/reference/
- https://www.cprogramming.com/tutorial/c++-tutorial.html

## Язык программирования С++

#### Является

- императивным: для написания программы используются операторы и управляющие конструкции языка программирования;
- компилируемым:



**О СТАТИЧЕСКИ ТИПИЗИРОВАННЫМ:** ТИПЫ ВСЕХ ПЕРЕМЕННЫХ ПРОГРАММЫ ИЗВЕСТНЫ В МОМЕНТ КОМПИЛЯЦИИ.

## Язык программирования С++

- Позволяет писать программы в процедурном, объектно-ориентированном и обобщённом стилях.
- На C++ созданы: Chrome, Firefox, Opera, Safari, IE Edge
- Используется для прикладных программ: создания/редактирование изображений, работа с видео, создание графических интерфейсов
- Обработка больших объёмов данных
- Машинное обучение: фреймворк tensorflow

## Язык программирования С++: историческая справка

- Компания Bell Labs, Бьярне Страуструп (Bjarne Straustrup)
- 1985, первый коммерческий релиз языка С++



- 1998 первый стандарт ISO C++98
- На данный момент выпущены C++03, C++11, C++14, C++17

# Какие особенности определяют язык программирования?

- Фундаментальные типы данных и операторы для работы с переменными этих типов.
- Управление ходом выполнения программы (циклы, условные и безусловные переходы).
- Способ обособления блоков кода, для неоднократного использования (функции).
- Специальные типы данных и способы работы с ними.
- **5** Конструкции языка для создания *пользовательских* типов данных.

#### Пример использования С++

#### Математика:

$$force(x, y) = 2x + 4ysin(x)$$
  
 $a = 5.5$   
 $b = -8.34$   
 $res = force(a, b)$ 

#### Пример использования С++

#### Математика:

```
force(x,y) = 2x + 4ysin(x)

a = 5.5

b = -8.34

res = force(a,b)
```

#### Определение

**Идентификатором** в языке программирования называется непрерывная последовательность символов, которые используются для именования переменных, функций, пользовательских типов данных.

В языке C++ в состав идентификатора могут входить только **буквы, цифры и символ нижнего подчёркивания** "\_". При этом начинаться каждый идентификатор должен только с **буквы или символа подчёркивания**.

Кроме того, в языках программирования существует определённый набор слов (в широком смысле - символьных последовательностей), которые не могут быть использованы в качестве идентификаторов. Такие слова называются ключевыми, вот они:

alignas alignof and and eq asm auto bitand bitor bool break case catch char char16 t char32 t class compl const constexpr const cast continue decitype default delete do double dynamic cast else enum explicit export extern false float for friend goto if inline int long mutable namespace new noexcept not not eq nullptr operator or or eq private protected public register reinterpret cast return short signed sizeof static static assert static cast struct switch template this thread local throw true try typedef typeid typename union unsigned using virtual void volatile wchar t while xor xor eq union unsigned using virtual void volatile wchar t while xor xor eq

**Коментарии** - произвольный текст в файле с исходным кодом, который игнорируется компилятором и никак не влияет на ход программы.

```
    // Это однострочный комментарий
    /*
    Пример
    многострочного
    комментария
    */
```

**Предупреждение!** Многострочные комментарии не могут быть вложенными

Общий вид создания переменной в С++:

Эта форма называется объявлением переменной.

Первый тип данных для знакомства: int

- предназначен для хранения целых чисел со знаком;
- диапазон значений (на большинстве 64-битных ОС тип int занимает 4 байта): [-2\_147\_483\_648; 2\_147\_483\_647]
- При объявлении переменной под неё выделяется место в оперативной памяти
- Однако С++ не даёт никаких гарантий относительно того, какое значение (целое число) окажется в объявленной переменной (фактически, не гарантируется, что в выделенном ОС блоке памяти сохранено конкретное число)

Рассматриваемый тип данных: **int** Пример объявляния переменных:

```
1 // Переменные конкретного типа можно объявлять по одной
2 int counter;
3 int velocity;
4
5 // А можно и несколько одновременно, через запятую
6 int width, height, area;
```

Рассматриваемый тип данных: **int** Пример объявляния переменных:

```
1 // Переменные конкретного типа можно объявлять по одной
2 int counter;
3 int velocity;
4
5 // А можно и несколько одновременно, через запятую
6 int width, height, area;
```

Изменение значений, хранящихся в переменных, происходит с помощью **операторов**.

#### Определение

**Оператором** в языках программирования называется символьная конструкция, которая производит действия над одним или более объектом.

#### Операторы в С++ делятся на:

- **•• унарные** требуется один объект для совершения действия, ставится после оператора;
- бинарные два объекта, по одному до и после оператора.

#### +1

Также в языке программирования присутствует единственный **тернарный** оператор - он состоит из двух символов и требует три объекта для своей работы. Будет расмотрен в разделе о ходе выполнения программы.

Тип int: присвоение значений с помощью оператора «=»

int acceleration\_rate;

// Присвоение переменной начального значения

acceleration\_rate = -7;

// Присвоение можно (и нужно!) делать

при объявлениии переменной.

\*/

int good\_rate = 16, bad\_rate = -10, karma;

karma = good rate + bad rate;

#### Определение

Присвоение значения переменной в момент её объявления называется её **инициализацией**.

```
Тип int: арифметические операции - «+, -, *, /, %»
1 int balance = 10, rate, total;
2
3 rate = balance - 8;
4 total = 2 * balance * (6 - rate);
5
6 // Целочисленное деление — дробная часть \leftarrow
      отсекается
7 rate = 7 / 2; // rate равен 3
8
9 // Взятие остатка от деления
10 rate = 7 % 2; // rate равен 1
11
12 // Ошибка времени выполнения:
13 // rate = 11 / 0;
```

Обратите внимание: при работе с целыми числами нельзя допускать деления на нуль, это ошибка, вызывающая немедленную остановку выполнения программы.

Все возможные виды объявления и их инициализации переменной:

```
(1) [...] <тип> <идентификатор> [= <значение>];
(2) [...] <тип> <идентификатор> [{ <значение> }];
(3) [...] <тип> <идентификатор> [( <значение> )];
```

, где треугольные скобки означают обязательные части, квадратные - опциональные, троеточие - дополнительные характеристики переменной или указания компилятору.

**Предупреждение!** (3) форма не рекомендуется к использованию с фундаментальными типами данных.

#### Правило хорошего тона

По возможности, **каждая** переменная программы должна быть проинициализированна разумным начальным значением.

Примеры форм (1) и (2) инициализации переменных

```
1 int count = 8; // (1)
2 int rate {101}; // (2)
```

#### Правило хорошего тона

Для фундаментальных типов данных предпочтительной является форма (1).

Целочисленный тип	Размер на 64-битных ОС
short <del>int</del>	2 байта
unsigned short int	
int	4 байта
unsigned <del>int</del>	
long <del>int</del>	8 байт
unsigned long int	
long long <del>int</del>	8 байт
unsigned long long int	
size t	8 байт, беззнаковый тип

- Перечёркнутый int может быть пропущен
- Стандарт языка не определяет конкретного размера каждого типа, только соотношение размеров между ними: sizeof(short) <= sizeof(int) <= sizeof(long) <= ...</li>
- size\_t размер определяется максимальным числом оперативной памяти на данной архитектуре ЭВМ

В примерах с присвоением значений переменным были использованы конкретные числа (-7, 16, -10). Поскольку язык С++ является типизированным, то данные числа в момент компиляции должны также получить тип данных. Когда в программе встречается целое число, то компилятор пробует сначало поместить его в тип int, затем в long, потом в long long. А если не получилось - выдаст ошибку компиляции.

#### Определение

Значения конкретных типов данных, записанные в программе в явном виде, называются **литералами**.

Целочисленные типы: расширенные операторы присваивания, инкремент/декремент

```
1 int balance = 5, rate = 10;
2
3 // Оператор присваивания сначало вычисляет
4 // выражение справа
5 balance = balance + 1;
6 // Сокращённая запись
7 balance += 1;
8 /* Также опре∂елены: -=, *=, /=, %= */
10 // Инкремент/Декремент — yВеличение/yменьшение \leftarrow
      значения переменной на единицу
11 rate++; // rate стал равным 11
12 ++rate; // -//- 12
13 rate--; // снова 11
14 --rate; // теперь 10
```

#### Целочисленные типы: тонкости инкремента/декремента

```
1 // Разница пре— и пост— инкремента
2 int balance = 5, total = 5;
3
4 printf("balance = %d\n", balance++);
\mathbf{5} // напечатает \mathcal{B} консоли "balance = \mathbf{5}"
6 // значение balance равно 6
8 printf("total = %d\n", ++total);
9 // Haneyamaem "total = 6"
10 // значение total равно 6
11
12 // Пример унарных операторов
13 int number = -6;
14 +number;
15 -number;
```

Целочисленные типы: побитовые операции (только для положительных целочисленных значений)

```
1 unsigned number = 4, next_number;
2
3 // Побитовый сдвиг вправо: number * (2^2)
4 next number = number << 2;</pre>
5
6 // Побитовый сдвиг влево: number / (2^3)
7 next number = number >> 3;
8
9 // Побитовое "И"
10 next number = number & 2;
11
12 // Побитовое "ИЛИ"
13 next number = number | 3;
14
15 // Побитовое исключащее "ИЛИ" (xor)
16 next number = number ^ 3;
17 // Поддерживаются >>=, |=, \hat{}= и аналоги
```

Целочисленные типы: оператор sizeof

```
1 // Двоичный литерал, С++14
2 int i num = 0b01100010;
3 long 1_num = 0345; // Восьмиричный
4 size t sz num = 0xff11c;// Шестнадцатиричный
5
6 printf("Size of int: %lu\n", sizeof(int));
7 printf("Size of long: %lu\n",
8 sizeof(l num));
9 printf("Size of size t: %lu\n",
10
        sizeof(size_t));
```

Возможный вывод в консоли:

```
Size of int: 4
Size of long: 8
Size of size_t: 8
```

#### Определение

**Адресом перменной** называют номер первого байта связанного с ней блока памяти (хранилища) в памяти.

Унарный оператор взятия адреса - амперсанд &

```
1 int degree = 120;
2 int rank = 8;
3
4 printf("Addr of \"degree\": %p\n", &degree);
5 printf("Addr of \"rank\": %p\n", &rank);
```

Возможный вывод в консоли:

```
Addr of "degree": 0x7ffff89f4200
Addr of "rank": 0x7ffff89f4204
```

bool - логический (булев) тип данных, принимающий только два значения - true или false

Операция	Оператор				
отрицание	!				
логическое И	&&, and				
логическое ИЛИ	, <b>or</b>				

```
1 bool truth = true, falsey = false, result;

2 3 // Ποευчесκοε "И"

4 result = truth && falsey; // result paßeн false

5 // Ποευчесκοε "ИЛИ"

6 result = truth || falsey; // result paßeн true

7 // ∂οηνςσωμου ω πακ

8 result = truth or falsey; // result paßeн true

9 

10 // Οπρυμαние:

11 result = !falsey; // result paßeн true
```

- 1 bool var1, var2;
- 2 // Переменным присваиваем значения...

Таблица истинности											
	var1	var2	Результат								
&&	true	true	true								
&&	true	false	false								
&&	false	true	false false								
&&	false	false									
	true	true	true								
	true	false	true								
	false	true	true								
	false	false	false								

bool to int (3): 1

```
Логический тип: совместимость с языком С
1 bool to be or not = 0;
2 // to be or not равна false
3 printf("bool to int (1): %d\n", to_be_or_not);
5 to_be_or_not = 1;
6 // to be or not равна true
7 printf("bool to int (2): %d\n", to_be_or_not);
8
9 to_be_or_not = 118;
10 printf("bool to int (3): %d\n", to be or not);
  Вывод в консоль:
  bool to int (1): 0
  bool to int (2): 1
```

Переменные типа **bool** могут учавствовать в арифметических выражениях (но не должны)

#### Операторы сравнения

Операция	Оператор				
равенство	==				
неравенство	! =				
больше	>				
меньше	<				
больше или равно	>=				
меньше или равно	<=				

Все операторы сравнения в C++ возращают значения типа **bool**.

```
1 int width = 5, height = 4;
2 bool status = width > height;
3 // status равен true
4
5 status = (width == height);
6 // status равен false
```

```
Числа с плавающей точкой
float - число с плавающей точкой одинарной точности (\approx 7-8
знаков после запятой). Максимальное значение - \approx 10^{38}.
минимальное - \approx 10^{-38}.
double - число с плавающей точкой двойной точности (pprox
15-16 знаков после запятой). Максимальное значение -
\approx 10^{308}, минимальное - \approx 10^{-308}.
long double - двойной точности (\approx 15-16 знаков после
запятой), расширенный диапазон: 10^{4932}.
Pasмep типов: sizeof(float) -> 32 бита, sizeof(double) -> 64
бита. Тип long double) занимает как минимум 80 бит.
```

```
1 double speed = 5.5, distance;
2 distance = speed / 3.0;
3
4 speed *= 4.0;
```

Числа с плавающей точкой: как представлены в памяти

	3н	ак														
		(11 би	т)	(52 бита)												
Порядок									Мантис	ca						
	63	56	55	48	47	40	39	32	31	24	23	16	15	8	7	0

Конкретное число вычисляется как:

$$(-1)^{ ext{3-HaK}} \left(1 + \sum_{i=1}^{52} b_{52-i} 2^{-i} 
ight) 2^{\Pi ext{орядок - } 1023}$$

, где *b<sub>i</sub> - i-*ый бит мантиссы.

Формат чисел с плавающей запятой одинарной и двойной точности стандартизирован: IEEE 754.

В точности числа не учитывается десятичная экспонента:  $1184 \rightarrow 1.184 \times 10^3$ 

Числа с плавающей точкой: на нуль делить разрешается

```
1 #include <cmath>
2
2
3 double super_rate = 5.5;
4 super_rate /= 0.0;
5
6 bool is_infinity = isinf(super_rate);
7 // is_infinity paßha true
```

Числа с плавающей точкой: понятие Not-A-Number (NaN)

```
1 #include <cmath>
2
3 // sqrt — вычисление квадратного корня
4 double super_rate = sqrt(-4.5);
5
6 bool is_nan = isnan(super_rate);
7 // is_nan paвнa true
```

#### Предупреждение

Сравнение чисел с плавающей точкой неоднозначно

```
1 double first_rate = 0.2 + 0.1 * 2;
2 double second_rate = 0.2 * 2;
3
4 bool is_equal = (first_rate == second_rate);
5 // что будет β is equal — непонятно
```

Числа с плавающей точкой: подход в сторону правильного сравнения

```
#include <cmath>

double rate1 = 0.4, rate2 = 0.8 / 2;

// Определяем для себя приемлемую точность

double eps = 1.5E-8;

// abs — вычисляет модуль аргумента

bool is_equal = abs(rate1 - rate2) < eps;

// есть уверенность в is_equal: равна false
```

- не универсальный способ;
- для примера: https://randomascii.wordpress.com/2012/02/25/comparingfloating-point-numbers-2012-edition/

Числа: взаимные неявные преобразования.

- При работе с числами язык C++ осуществляет неявные преобразования целых значений в действительные и наоборот.
- Любое действительное число преобразуется в целое путём отбрасывания всей дробной части
- Любое целое значение преобразуется в действительное путём добавления нулей в дробную часть
- Если в арифметическом **операторе** есть хотя бы одно действительное число, результат оператора будет преобразован к действительному типу
- Если в арифметическом **операторе** операндами являются безнаковое и знаковое целые значения, то число со знаком приводится к беззнаковому и результат тоже будет беззнаковый

Числа: использование в арифметических выражениях.

```
1 int five = 5, seven = 7;
2 double expr = five / seven + 4.25;
3 printf("expr = %f\n", expr);
```

Числа: использование в арифметических выражениях.

```
1 int five = 5, seven = 7;
2 double expr = five / seven + 4.25;
3 printf("expr = %f\n", expr);
```

#### Вывод на консоль:

```
expr = 4.250000
```

Числа: явные преобразования.

- использование типа + круглых скобок
- использование оператора static\_cast

#### Вывод на консоль:

```
expr2 = 4.964286
expr3 = 4.964286
```

Любая переменная может быть задана константной (неизменяемой) с помощью ключевого слова const.

```
1 const double LIGHT_SPEED = 2.99792458e8;
2 printf("Speed of light: %.4f"\n, LIGHT_SPEED);
3 // Значение поменять нельзя:
4 // LIGHT_SPEED = 3e8; // Ошибка компиляции
5
6 const size_t PARTS_COUNT = 10;
```

#### Совет

На название констант C++ не накладывает ограничений, то есть может быть использован любой уникальный идентификатор, но по общим практикам рекомендуется использовать *верхний регистр* для их именования, разделяя слова символом *нижнего подчёркивания*.

Далее - управляющие конструкции

if ( max\_score < min\_score ) {
 printf("He должно так быть\n");</pre>

printf("Мы этого и ждали\n");

2

**7** }

5 } else {

Условный переход: общий синтаксис

if ( <логическое выражение> ) {
 <набор инструкций>;
 } [else {
 <набор инструкций>;
 }]

int max\_score = 3, min\_score = 5;

Условный переход: отсутствие пары фигурных скобок

```
1 int max score = 8, min score = 2;
2 // Опасный синтаксис:
3 if ( max_score < min_score )</pre>
  printf("He должно так быть\n");
   printf("A не отношусь к условному переходу\n");
5
6
7 int current score = 4;
8 // Логическое выражение может быть составным
9 if ( (current_score >= min_score) and (←)
     current_score <= max_score) ) {</pre>
10 printf("Bcë oκ\n");
11 } else {
12 printf("Какая-то аномалия\n");
13 }
```

#### Условный переход: ещё пример

```
1 int current_score = 7, last_score = 4;
2
3 if ( last score == 4 ) {
  current score += 4;
5 } else if ( last_score == 5 ) {
6 current score += 1;
7 } else if ( last_score == 6 ) {
8 current score -= 10;
9 } else {
10    current_score = 0;
11 }
```

2

```
Тернарный оператор ?:
    <погическое выражение>
       <?> <выражение, когда истинно>
       <:> <выражение, когда ложно>;
1 int max score = 13, min score = 5;
3 int result;
4 result = ((max_score - min_score) < 5 ) ? ←
    min score * 2 : max score / 2;
```

### Конструкция switch

```
switch (<выражение>) {
    case <значение 1>:
        [инструкции]
        <break>;
    case <значение 2>:
        [инструкции]
        <break>;
    case <значение N>:
        [инструкции]
        <break>;
    [default:
        [инструкции]
```

```
1 int rate, score;
2 // Вычисляем rate
3
4 switch ( rate ) {
5
    case 2:
6
      score = 1;
7
       break;
8
   case 5:
9
      score = 2;
10
       break;
11
   case 8:
12
      printf("Выпала восьмёрка!\n");
13
      score = 3;
14
      break:
15 default:
16
      score = 5;
17 }
```

```
1 \text{ score} = 4;
2
  switch ( score ) {
    case 4:
      printf("Вполне неплохо.\n");
5
       // не ставим *break*
6
    case 5:
7
8
       printf("Совсем здорово!\n");
       break:
9
     case 6:
10
       printf("Недосягаемая высота!\n");
11
       break:
12
     default:
13
      printf("Недотянули.");
14
15 }
```

В примере, когда в переменной **score** будет четвёрка, то будут выполнены строчки **5** и **8** 

- Фактически switch производит сопоставления результата выражения с значениями в "ветках" case
- Его недостатком, во-первых является то, что результат выражения должен быть приводим к числу
- А во-вторых необходимость не забыть поставить break в каждой ветке case

#### Цикл

Под **циклом** в программировании понимается обособленный набор повторяющихся N раз инструкций. Причём,  $N \ge 0$ 

#### Дополнительные определения

- Тело цикла набор повторяющихся инструкций
- Итерация (или проход цикла) однократное выполнение всех инструкций из тела цикла

```
Циклы while и do ... while
while (<логическое выражение>) {
  [инструкции];
do {
  [инструкции];
} while (<логическое выражение>);
```

Цикл **do** ... **while**: печать квадратов чисел от 0 до 9 включительно, возрастающий порядок

```
1 int counter = 0;
2
3 do {
4   printf("%i\n", counter * counter);
5   counter++;
6 } while ( counter < 10 );</pre>
```

Цикл **while**: печать таблицы умножения двойки на числа от 1 до 10 включительно

### Бесконечный цикл на примере while

```
1 while( true ) {
2  printf("*");
3 }
```

- Инициализация: как правило, задание начальных значений переменных, которые будут использованы внутри цикла. Или определение новых.
- Условие выхода: логическое выражение, проверяемой до начала и после каждой итерации.
- Итеративное изменение: задание выражений (изменение счётчика цикла, как пример), которые выполняются после каждой итерации. Перечисляются через запятую.

#### Цикл for

#### Цикл for

```
1 // Бесконечный цикл
2 for (;;) {
3
    for (int i = 0; i <= 3; i++) {
4
       switch (i) {
5
      case 0:
6
           printf("\r-"); break;
7
      case 1:
8
           printf("\r\\"); break;
9
      case 2:
10
           printf("\r|"); break;
11
      case 3:
12
          printf("\r/"); break;
13
14
15 }
```

#### Управление циклами continue и break

```
1 for (int counter = 0; ; ++counter) {
    if ( (counter % 2) == 0 ) {
3
      printf("%i\n", counter);
      continue;
5
6
7
    if ( (counter > 15) ) {
      break:
10
    printf("Итерация закончена\n");
11
12 }
```