ОЦЕНКА ХАРАКТЕРИСТИК ПРОГРАММ НА ОСНОВЕ ЛЕКСИЧЕСКОГО АНАЛИЗА

Метрики Холстеда

# ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ

Особенности формирования словаря программы

Любая программа определяет последовательность действий над *операндами* с помощью *операторов*. Исходный текст программы, записываемой на том или ином языке программирования, представляет собой набор текстовых строк, которые записываются по специальным правилам, и в том числе имеет свои элементы. *Операнд –* это некоторый объект или величина, обрабатываемая в программе, а *оператор* представляет собой обозначение конкретного действия, выполняемого по отношению к операнду.

Если считать, что словарь любой программы состоит только из имен операторов и операндов, то тексты программ всегда удовлетворяют следующим условиям:

* маловероятно появление какого-либо имени оператора или операнда много раз подряд – языки программирования, как правило, позволяют создавать такие конструкции, в которых подобные фрагменты программы имеют минимальную длину;
* циклическая организация программ исключает многократное повторение какой-либо группы операторов и операндов – более компактные варианты текстов получаются при разумном использовании развитых возможностей языков программирования, причем многообразие языков предоставляет богатую палитру инструментов;
* блоки программ, требующие периодического повторения при ее исполнении, обычно оформляются как процедуры или функции, поэтому в текстах программ достаточно применения только их имен;
* имя каждого операнда должно появляться в тексте программы хотя бы один раз – многие среды

программирования обращают внимание программистов на неиспользуемые имена, которые следует удалять из текста программ, чтобы сократить объем памяти, используемой при объявлении переменных.

Измеряемые свойства программ

При разработке программы формируется ее текст на каком-либо языке программирования, реализующий алгоритм получения искомого результата на основании обработки заданной совокупности данных. В тексте программы можно идентифицировать все операнды, определенные как переменные или константы, используемые в данной реализации. Аналогичным образом идентифицируются все операторы, определенные как символы или комбинации символов, влияющие на способ обработки, изменение значения или порядок следования и преобразования операндов. Исходя из идентификации операторов и операндов, можно определить ряд измеримых категорий, обязательно присутствующих в версиях любого алгоритма. Они определяются метриками, с помощью которых могут быть получены основные характеристики качества программ.

В состав измеримых свойств любого представления алгоритма (или программы) могут быть включены следующие метрические характеристики:

* 𝜂1 – число простых (уникальных) операторов, появляющихся в данной реализации;
* 𝜂2 – число простых (уникальных) операндов, появляющихся в данной реализации;
* 𝑁1 – общее число всех операторов, появляющихся в данной реализации;
* 𝑁2 – общее число всех операндов, появляющихся в данной реализации;
* 𝑓1𝑗 – число появлений в программе 𝑗-го оператора, где 𝑗 = 1,2,3, … , 𝜂1;
* 𝑓1𝑗 – число появлений в программе 𝑗-го оператора, где 𝑗 =

1,2,3, … , 𝜂2.

Учитывая эти основные метрические характеристики для программы, в конкретной реализации текста программы можно определить:

 словарь 𝜂 = 𝜂1 + 𝜂2;

* длину реализации программы 𝑁 = 𝑁1 + 𝑁2;

 длину программы 𝑁̃ = (𝜂1 ∙ 𝑙𝑜𝑔2𝜂1)+ (𝜂2 ∙ 𝑙𝑜𝑔2𝜂2).

Следует отметить, что помимо своего прямого назначения

метрики длины программы и длины реализации можно использовать для выявления несовершенств программирования, которые являются следствием применения не самых удачных приемов программирования. Если расчетные значения длины программы и длины реализации отличаются более чем на 10 %, то это свидетельствует о возможном наличии в программе следующих шести классов несовершенств*:*

1. Наличие последовательности дополняющих друг друга операторов к одному и тому же операнду, например 𝐴 + 𝐶 −

𝐴. Понятно, что в подобном случае будет выполнено два совершенно ненужных действия, дополняющих переменную

𝐶 одной и той же величиной, взятой с противоположными знаками.

1. Наличие неоднозначных операндов, например 𝐴 = 𝐷 и A = C *.* При выполнении таких действий программа будет поставлена в затруднительное положение, поскольку присвоение осуществляется путем приравнивания значения операнда, указанного в левой части, значению, приведенному в правой части выражения. В лучшем случае произойдет ненужное присвоение нового значения уже имеющемуся.
2. Наличие синонимичных операндов, например А = В и С = В*.* Поскольку одно и то же значение должно быть присвоено разным переменным, то для данного примера переменная В вообще может не использоваться. Более лаконичным вариантом является простое приравнивание значений переменных А и С.
3. Наличие общих подвыражений, например:

(𝐴 + 𝐵) ∙ 𝐶 + 𝐷 ∙ (𝐴 + 𝐵).

Здесь применено совсем не обязательное повторение суммирования переменных А и В *,* что приводит к дополнительному времени выполнения программы.

1. Ненужное присваивание, например C = A + B *,* если переменная 𝐶 используется в программе только один раз. При однократном выполнении каких-либо операций над переменной нецелесообразно вводить дополнительный операнд, это ведет к увеличению объема памяти, резервируемой под переменные программы, и увеличивает размер словаря.
2. Наличие выражений, которые не представлены в свернутом виде как произведение множителей, например:

𝑋 ∙ 𝑋 + 2 ∙ 𝑋 ∙ 𝑌 + 𝑌 ∙ 𝑌

Данное преобразование можно представить как

(X + Y) ∙ (X + Y),

т. е. свернуть выражение до квадрата суммы переменных 𝑋 и 𝑌.

Такое представление окажется более лаконичным и сократит время, необходимое для выполнения программы.

В соответствии с приведенными определениями применяются следующие соотношения:

𝜂

𝑁 = ∑ 𝑓 ;

1 𝑗=1 1,𝑗

𝜂

𝑁 = ∑ 𝑓 ;

2 𝑗=1 2,𝑗

𝜂

2 ∑ 𝑓 .

𝑁 = ∑

𝑖=1 𝑗=1 𝑖,

Таким образом, длина реализации и объем программы определяются исключительно на основе анализа текста программы путем подсчета количества операндов и операторов, а также числа их вхождений в текст программы, т. е. на основе лексического анализа текста программы. Длина программы представляет собой математическое ожидание количества слов в тексте программы при фиксированном словаре.

Другой важной характеристикой программы является ее объем 𝑉. В отличие от длины программы 𝑁 объем измеряется не количеством слов, а числом двоичных разрядов. Если в словаре имеется 𝜂 слов, то для задания номера любого из них требуется минимум 𝑙𝑜𝑔2𝜂 бит.

Объем программы определяется следующим образом:

𝑉 = 𝑁 ∙ 𝑙𝑜𝑔2𝜂 = 𝜂 ∙ 𝑙𝑜𝑔2𝜂

2

Тогда с точностью до обозначений полученные соотношения

окажутся совершенно идентичными, хотя смысл их будет различным:

𝑁 ≈ 𝜂 ∙ 𝑙𝑜𝑔2𝜂*;*

2

𝑉 = 𝜂 ∙ 𝑙𝑜𝑔2 𝜂*.*

В первом случае зафиксирована взаимная связь между

длиной программы 𝑁 и размером словаря 𝜂, во втором – между величиной словаря и объемом программы 𝑉. Следовательно, по известному размеру словаря 𝜂 можно найти значения 𝑁 и 𝑉 . Идентичность этих выражений говорит о том, что соотношение между величиной словаря и длиной текста единственно и взаимно однозначно.

Выше было отмечено, что словарь программы состоит только из операторов и операндов. Учитывая принятые обозначения, соотношение Холстеда примет следующий вид:

𝑁 = 𝜂1 ∙ 𝑙𝑜𝑔2𝜂 + 𝜂2 ∙ 𝑙𝑜𝑔2𝜂 ≈ 𝜂1 ∙ 𝑙𝑜𝑔2𝜂1 + 𝜂2 ∙ 𝑙𝑜𝑔2𝜂2 = 𝑁1 + 𝑁2

Как правило, при проведении статистических исследований

текстов программ к словарю операторов относят следующие элементы:

* имена арифметических и логических операций;
* присваивания;
* условные и безусловные переходы;
* разделители;
* скобки (парные);
* имена процедур и функций;
* выражения типа BEGIN...END, IF...THEN...ELSE, DO...WHILE.

Выражения типа BEGIN...END, IF...THEN...ELSE,

DO...WHILE и им подобные, осуществляющие блочную группировку операторов, при этом рассматриваются как единые операторы (то же относится и к парам скобок).

Величины количества операторов и операндов 𝜂1 и 𝜂2 независимы и могут принимать произвольные значения. Однако этого нельзя сказать относительно 𝑁1 и 𝑁2, т. е. числа появления всех операторов 𝑁1 = 𝜂1 ∙ 𝑙𝑜𝑔2𝜂1 , и всех операндов 𝑁2 = 𝜂2 ∙

𝑙𝑜𝑔2𝜂2 в тексте программы: между ними можно установить

приблизительное соответствие, причем оно будет взаимно однозначным. В каждом конкретном случае каждый операнд не может позиционироваться в программе обособленно, он входит в текст, по крайней мере, хотя бы с одним оператором: например, с разделителем (точка с запятой), отделяющим его от других операторов, или другим набором символов, определяющим способ действия над этим операндом. В то же время применение нескольких операторов к одному операнду маловероятно. Поэтому можно утверждать, что 𝑁1 ≈ 𝑁2, хотя величины словарей

𝜂1 и 𝜂2 могут сильно отличаться друг от друга. Это позволяет прийти к весьма важному практическому выводу относительно

объема программы:

𝑁 ≈ 2 ∙ 𝑁2 = 2 ∙ 𝜂2 ∙ 𝑙𝑜𝑔2𝜂2.

Основной исходный параметр, на котором базируются все

расчеты метрических характеристик будущего ПС, – количество имен входных и выходных переменных 𝜂∗ , представленных в предельно краткой записи (с точки зрения алгоритмической сложности – сжатой). Например, для задания одномерного массива (т. е. строки), каково бы ни было число его элементов, требуется всего два имени:

2

* указатель адреса начала массива;
* количество элементов в нем.

Точно так же для задания двумерного массива достаточно иметь три параметра:

* указатель адреса первого элемента;
* число столбцов;
* число строк.

Если параметр 𝜂∗ , определенный таким образом,

2

рассматривать как размер генеральной совокупности имен входных и выходных переменных, то величина словаря

программы 𝜂2 в соответствии с соотношением Холстеда не будет

∗ ∗

превосходить 𝜂2 ∙ 𝑙𝑜𝑔2𝜂2. Таким образом, можно считать, что

∗ ∗

𝜂2 ≈ 𝜂2 ∙ 𝑙𝑜𝑔2𝜂2

Предположим, что существует некоторый язык

программирования, называемый *потенциальным*, в котором все программы (по крайней мере - для некоторой предметной области) уже написаны и представлены в виде заранее подготовленных

процедур или функций. Тогда для реализации любого алгоритма на таком языке потребуется всего два оператора (функция и присваивание) и 𝜂∗ имен входных и выходных переменных. Это объясняется тем, что в таком потенциальном языке программист должен будет только выбрать нужную процедуру или функцию и применить ее к нужной переменной. Поскольку в такой записи никакие слова не повторяются, то длина программы совпадает с ее объемом и равна:

2

∗ ∗

𝑉∗ = (𝜂2 + 2) ∙ 𝑙𝑜𝑔2(𝜂2 + 2)

Эта величина называется потенциальным объемом

(минимально возможным), соответствующим максимально компактному тексту программы, реализующей данный алгоритм. Это объясняется тем, что в потенциальном языке минимизировано число операторов, а все операнды сведены к перечню процедур или функций и списку входных и выходных переменных.

В таком случае можно определить *уровень реализации*

программы, который рассчитывается с помощью отношения

𝑉∗

𝐿 =

𝑉

Уровень реализации представляет собой метрический

показатель, который характеризует степень компактности программы, экономичность использования средств алгоритмического языка. Чем ближе значение 𝐿 к единице, тем более совершенна программа.

При переводе алгоритма с одного языка на другой его потенциальный объем 𝑉∗ не изменяется, но действительный объем 𝑉 может увеличиваться или уменьшаться в зависимости от развитости языков программирования.

Для потенциального языка справедливо равенство 𝑉 = 𝑉∗, для любого менее развитого языка следует учитывать

соотношение 𝑉 > 𝑉∗ Это обусловлено тем, что для

потенциального языка 𝑁∗ = 𝜂∗, в то время как для всех других языков применяется уравнение длины и учет соотношения 𝑁 > 𝜂.

Оценка уровня языков программирования

Если 𝑁 - длина программы, а 𝜂— словарь программы, то общее количество выборок необходимых элементов словаря (т. е.

фактически – *работа программирования*) в соответствии с законом Хика будет равна объему программы 𝑁 ∙ 𝑙𝑜𝑔𝜂.

В то же время необходимо еще учесть уровень реализации программы 𝐿: количество выборок при этом возрастет в 1⁄𝐿 раз. Обозначив работу программирования символом 𝐸 и учитывая

формулу для вычисления уровня реализации программы, окончательно получим:

𝐸 =

𝑁 ∙ 𝑙𝑜𝑔2𝜂 = 𝑉

𝐿 𝐿

𝑉2

= 𝑉∗

В начале 1980-х гг. Холстед ввел формальное определение

уровня языка программирования, определив этот уровень языка следующим образом:

𝜆 = 𝐿 ∙ 𝑉∗,

где 𝐿 - уровень реализации программы, 𝑉∗ - ее потенциальный объем.

Для любого алгоритма, который программируется с использованием разных языков, с увеличением объема уровень реализации уменьшается в той же пропорции. В результате произведение уровня 𝐿 на объем 𝑉 равняется потенциальному объему 𝑉∗ данного алгоритма. С другой стороны, если язык реализации остается одним и тем же, а разрешено менять сам алгоритм, имеется другое, но похожее соотношение. В этом случае с увеличением потенциального объема 𝑉∗ уровень программы 𝐿 уменьшится в том же отношении. Следовательно, произведение 𝐿 на 𝑉∗ остается неизменным для любого языка.

Определим работу программирования при использовании потенциального языка. В записи на потенциальном языке программа имеет минимально возможную длину, и так как слова (операнды и операторы) в ней не повторяются, то она совпадает с объемом:

∗ ∗

𝑉∗ = (𝜂2 + 2)log(𝜂2 + 2)

Работа программирования в потенциальном языке сводится к

выбору из конечного, но огромного по масштабам числа имен функций и процедур:

21 + 22 + 23 + ⋯ + 2𝑉−1 + 2𝑉 ≈ 2𝑉+1,

где 𝑉- объем программы.

Существует закон Хика, в соответствии с которым время реакции при выборе из некоего числа альтернативных сигналов зависит от их количества. Тогда по закону Хика работа выбора из библиотеки функций составит log 2𝑉+1 ≈ 𝑉 , и полная работа программирования при использовании потенциального языка будет определяться следующим образом:

𝐸пот = 𝑉∗ ∙ 𝑉.

В таком случае мы получим следующее соотношение:

𝜆 =

𝐸пот

𝐸

𝑉∗2

=

𝑉

Полученное значение 𝜆 как раз можно считать

количественной мерой уровня любого алгоритмического языка.

Из последнего выражения видно, что для постоянства 𝜆 , увеличение объема программы должно квадратично зависеть от увеличения объема информации по внешним связям. Поэтому алгоритмически сложные программы вычисления малого числа переменных будут давать значительно более низкое значение 𝜆, чем программы вычисления большого числа переменных по элементарным выражениям.

В связи с этим метрику уровня языка программирования 𝜆 для сравнения языков следует применять только для конкретной предметной области и близких типов задач. Ниже приведены данные об уровнях некоторых известных языков программирования (табл. 1).

Таблица 1

Уровни языков программирования

|  |  |  |
| --- | --- | --- |
| **Язык** | 𝜆 | **Отклонения** |
| Естественный язык | 2,16 | 0,74 |
| PL/1 | 1,53 | 0,92 |
| Алгол | 1,21 | 0,74 |
| Паскаль | 1,25 | 0,76 |
| Бейсик | 1,22 | 0,72 |
| Фортран | 1,14 | 0,81 |
| Ассемблер | 0,88 | 0,42 |

Поскольку известны соотношения для работы программирования 𝐸 и уровня языка 𝜆

𝐸 = 𝑉 и 𝜆 = 𝑉 ,

2

∗2

𝑉∗ 𝑉

то, исключая из выражений величину 𝑉, получим:

∗3

𝐸 = 𝑉 .

𝜆2

Тогда квалификационное время программирования будет

определяться следующим образом:

𝑇 = 𝐸 = 𝑉 ,

∗3

𝑆 𝑆𝜆2

где 𝑆 - число Страуда ( 5 < 𝑆 < 20 , среднее его значение

принято считать равным 18).

Метрика числа ошибок в программе

Значительная часть усилий и времени, которые затрачиваются на создание большинства программных продуктов, приходится на их отладку, т. е. выявление и устранение ошибок типа «лишних и недостающих элементов», внесенных в начальный период написания программы. Следовательно, любое обоснованное представление о количестве первоначальных ошибок, ожидаемых в данной программе, даст важную оценку для практики. Рассматриваемая метрика позволяет предсказать число первоначальных ошибок (до отладки и тестирования), однако не может служить свидетельством правильности (корректности) программы, даже если ее значение равно нулю.

Время, требуемое на разработку программы, характеризуется числом элементарных мысленных различений 𝐸. Следовательно, число моментов, в которые можно сделать ошибочное различение, также определяется значением 𝐸 или связанным с ним значением объема программы 𝑉.

Пусть каждый объект так же, как и результат, соответствует

∗

единице уникальных операндов в потенциальном языке, т. е. 𝜂2 = 6.

С помощью равенств

∗ ∗ ∗ ∗ ∗ ∗

𝜂1 = 2 и 𝑉 = (𝜂1 + 𝜂2)𝑜𝑔2(𝜂1 + 𝜂2)

получаем предельное значение потенциального объема:

∗ = (2 + 6) ∙ 𝑙𝑜𝑔

(2 + 6) = 24

𝑉крит 2

Далее из уравнения уровня языка программирования имеем:

𝐸 =

𝑉∗3

𝜆2

а из таблицы уровней языка известно, что для естественного

(английского) языка 𝜆 = 2,16.

Тогда для описания программы на уровне английского языка приходим к выводу:

𝐸 = 243 = 3000.

крит 2

2.16

Определим теперь 𝐸0 как среднее число элементарных различений между возможными ошибками в программировании, а

𝐵 - как число переданных ошибок в программе. Можно ожидать, что

𝐵 = 𝐸 ,

𝐸0

но при этом не будет учтено наличие какой-либо избыточности в

создаваемой программе, т. е. если в программе применены какие- либо фрагменты неоптимальной структуры, то этот аспект в данном расчете во внимание не принимается.

Однако уровень реализации программы 𝐿 , собственно, и является мерой такой избыточности. Заметим, что только в потенциальном языке, на котором любая программа может быть выражена в виде вызова процедуры, не повторяются ни операторы, ни операнды. Для потенциального языка уровень реализации 𝐿 = 1, для всех остальных языков 𝐿 уменьшается с увеличением избыточности, неминуемо присущей любому языку в связи с введением дополнительных условий записи функций и процедур.

Следовательно, вместо последнего полученного уравнения реальнее ожидать, что на количество ошибок в программе будет влиять и сам язык программирования, который мы можем учесть при помощи уровня реализации программы:

𝐸

𝐵 = 𝐿 ∙

𝐸0

Учитывая, что произведение 𝐿 ∙ 𝐸 можно заменить на

значение объема программы 𝑉, получим:

𝑉

𝐵 =

𝐸0

Если теперь приравнять 𝐸0 значению 𝐸крит , получим соотношение

𝐵 =

𝑉

3000

С другой стороны, подставив в это выражение значение для

𝑉 из формулы для определения уровня языка, получим:

𝑉∗2

𝐵 =

3000 ∙ 𝜆

Из этого выражения следует, что поскольку для определения

потенциального объема необходимо только знание числа независимых входных и выходных параметров программы, задаваемого в техническом задании на разработку программы, то после выбора языка программирования потенциальное количество ошибок можно оценить до начала написания программы, т. е. до начала проектирования.

Опытно установлено, что количество ошибок в текстах программ пропорционально работе программирования, которая, как показано выше, может быть вычислена. Кроме того, формирование текста программы человеком происходит не в виде цельного готового продукта, а некоторыми фрагментами ограниченного объема, причем размеры этих фрагментов зависят от очень многих факторов, включая настроение программиста. Предположив, что с каждым из таких фрагментов связана, по крайней мере, одна ошибка (это доказано на основе многочисленных статистических исследований), получим:

𝐸

𝐵0 = 𝐸

0

𝑉

=

3000

где 𝑉- расчетный объем программного средства.

# ПРИМЕР ОЦЕНКИ ХАРАКТЕРИСТИК

Задача «Расчет значений функции»

Разработать программу для вычисления значений заданной функции 𝐹:

𝐹 = {

sin(𝑥) + 𝑐𝑜𝑠2(𝑦), при 𝑥 < 𝑦; ln(𝑥) , при 𝑥 = 𝑦;

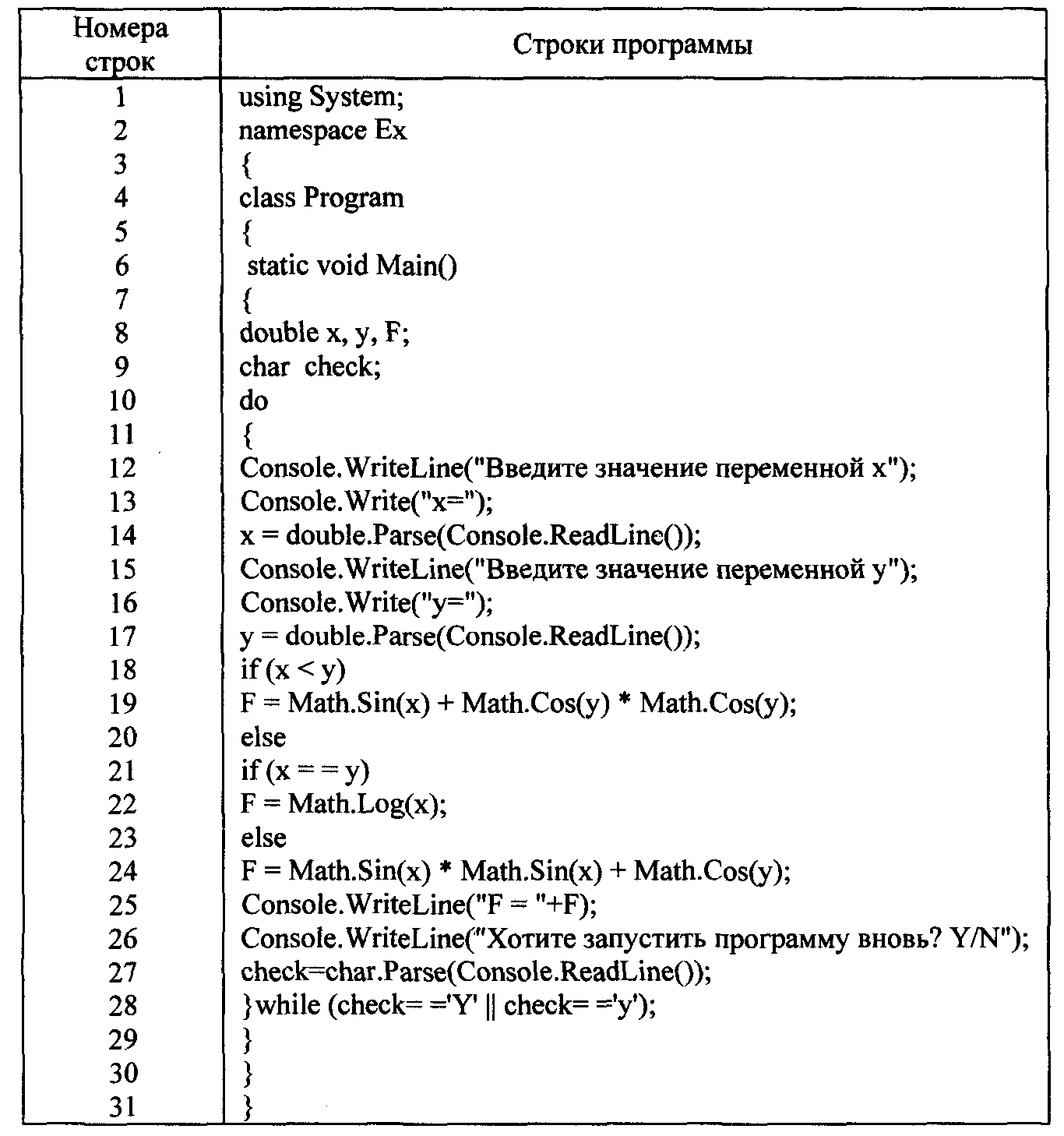
𝑠𝑖𝑛2(𝑥) + cos(𝑦) , при 𝑥 > 𝑦.

Значения аргументов функции ввести с клавиатуры. На экран монитора вывести значение функции.

*Определить значения метрик Холстеда, на основе которых дать оценку качества разработанного исходного текста программы.*

Реализация программы

Текст программы для реализации возможного решения поставленной задачи, разработанной с использованием языка программирования С#, приведен на [рис. 1](#_bookmark0).



*Рис. 1. Текст программы*

Словарь программы

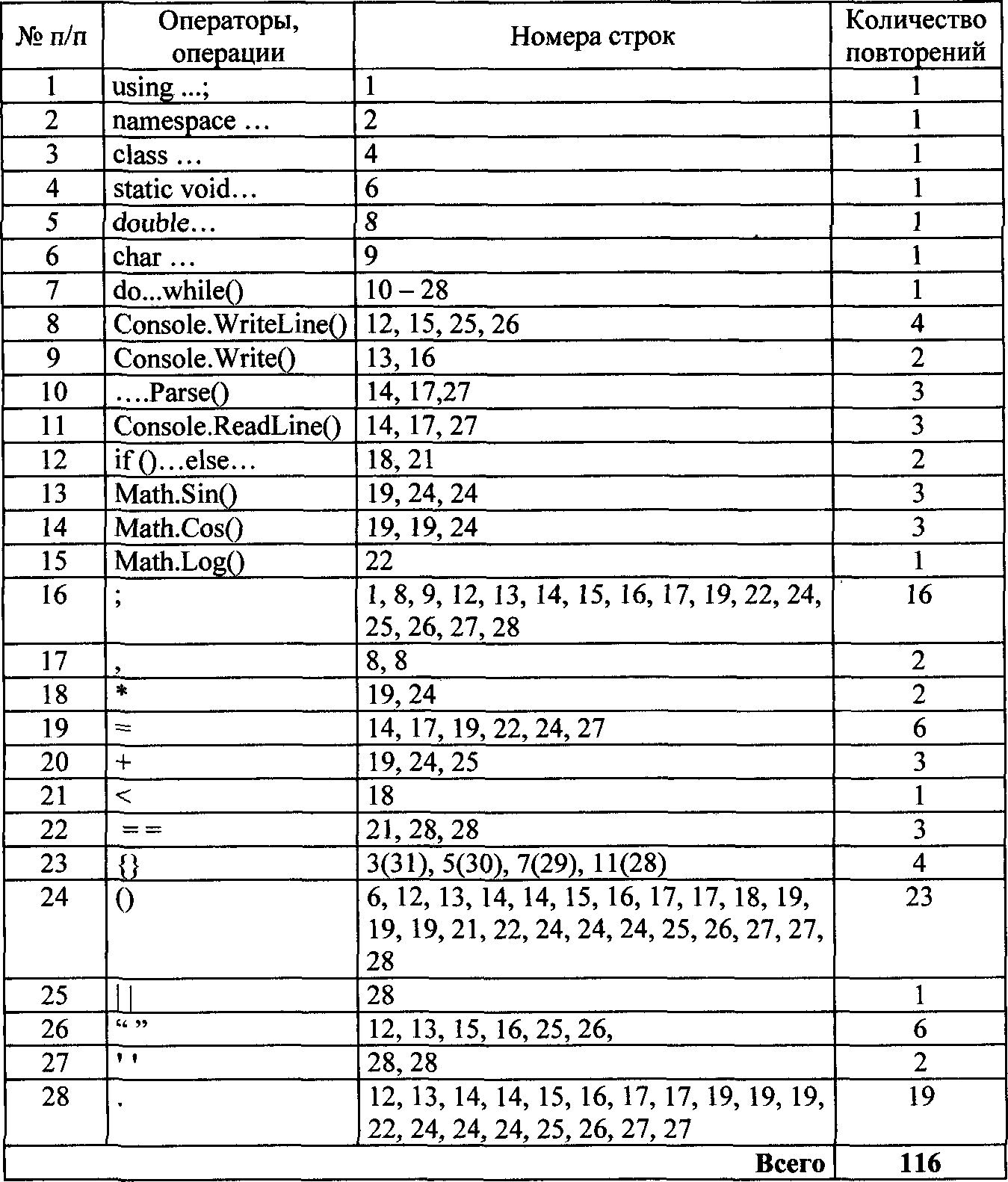
В табл. 2 приведены операторы и операции, используемые в программе (столбец 2). Номера строк исходной программы, где встречается каждый оператор или операция, указаны в третьем столбце. В четвертом столбце указано число повторений каждого оператора или операции в исходном тексте программы.

Таким образом, количество строк этой таблицы есть число уникальных операторов и операций, появляющихся в данном

тексте. Если вычислить сумму значений из четвертого столбца, то получим общее число всех операторов и операций, используемых в исходном тексте программы. Отметим, что для фигурных скобок, определяющих блок, приведены два номера строки. Первый определяет левую фигурную скобку, открывающую блок, а второй – закрывающую. Отметим, что такая пара в словаре учитывается только один раз.

Таблица 2.

Словарь операторов и операций программы



Словарь всех операторов в исходном тексте этой реализации программы сведен в столбце 2. В третьем столбце этой таблицы приведены номера строк исходной программы, где встречаются операторы. В последнем столбце приводится количество повторений (число вхождений) операторов в тексте программы.

Таким образом, количество строк этой таблицы есть число уникальных операторов программы. Сложив значения четвертого столбца, получим общее число вхождений всех операторов.

Проведем подробный анализ исходного текста программы в соответствии с полученной таблицей, начиная с первой позиции (первая строка программы (𝑢𝑠𝑖𝑛𝑔 𝑆𝑦𝑠𝑡𝑒𝑚;).

Ключевое слово 𝑢𝑠𝑖𝑛𝑔 представляет собой команду (инструкцию), обеспечивающую доступ к именам пространства имен 𝑆𝑦𝑠𝑡𝑒𝑚 . Следовательно, команду 𝑢𝑠𝑖𝑛𝑔 можно отнести к выполняемым операторам. Оператор 𝑢𝑠𝑖𝑛𝑔 встречается в программе всего один раз. Слово 𝑆𝑦𝑠𝑡𝑒𝑚 представляет собой имя, над которым осуществляется операция 𝑢𝑠𝑖𝑛𝑔. Таким образом, имя

𝑆𝑦𝑠𝑡𝑒𝑚 заносится в таблицу словаря операндов (табл. 3). Имя

𝑆𝑦𝑠𝑡𝑒𝑚 встречается в программе один раз.

Следующая строчка программы 𝑛𝑎𝑚𝑒𝑠𝑝𝑎𝑐𝑒 𝐸𝑥 состоит из оператора 𝑛𝑎𝑚𝑒𝑠𝑝𝑎𝑐𝑒 и операнда 𝐸𝑥 , которые также присутствуют в тексте программы в единственном экземпляре. Оператор занесен в таблицу операторов, а операнд 𝐸𝑥 - в таблицу.

Строки

 𝑐𝑙𝑎𝑠𝑠 𝑃𝑟𝑜𝑔𝑟𝑎𝑚;

 𝑠𝑡𝑎𝑡𝑖𝑐 𝑣𝑜𝑖𝑑 𝑀𝑎𝑖𝑛𝑄;

 𝑑𝑜𝑢𝑏𝑙𝑒 𝑥, 𝑦, 𝐹;

 𝑐ℎ𝑎𝑟 𝑐ℎ𝑒𝑐𝑘;

также представляют собой сочетание операторов и операндов, которые встречаются в тексте один раз, где ключевые слова 𝑐𝑙𝑎𝑠𝑠,

𝑠𝑡𝑎𝑡𝑖𝑐 𝑣𝑜𝑖 , 𝑑𝑜𝑢𝑏𝑙𝑒 и 𝑐ℎ𝑎𝑟 представляют соответственно операции, a 𝑃𝑟𝑜𝑔𝑟𝑎𝑚, 𝑀𝑎𝑖𝑛, 𝑥, 𝑦, 𝐹 и 𝑐ℎ𝑒𝑐𝑘 – имена (операнды). Все операции попадают в словарь операторов, а имена – в словарь операндов.

Следующий оператор 𝑑𝑜 . . . 𝑤ℎ𝑖𝑙𝑒() представляет собой инструкцию реализации циклического алгоритма, которая

используется в тексте программы один раз. Рассмотрим тело цикла (блок операторов, заключенных между ключевыми словами

𝑑𝑜 . . . 𝑤ℎ𝑖𝑙𝑒). Первой строчкой цикла является операция вызова функции вывода строк на экран монитора 𝐶𝑜𝑛𝑠𝑜𝑙𝑒. 𝑊𝑟𝑖𝑡𝑒𝐿𝑖𝑛𝑒().

Данная операция повторяется в тексте программы 4 раза. В каждом из этих случаев применения оператора вызова функции (метода) 𝐶𝑜𝑛𝑠𝑜𝑙𝑒. 𝑊𝑟𝑖𝑡𝑒𝐿𝑖𝑛𝑒() входным параметром функции является строка (строковая константа). Значение строковой константы в каждом случае применения оператора разное:

* + "Введите значение переменной 𝑥 ";
  + "Введите значение переменной у ";

 "𝐹 = ";

* + "Хотите запустить программу вновь? 𝑌/𝑁".

Все перечисленные константы являются операндами и заносятся в таблицу операндов. Каждый из перечисленных операндов используется один раз.

Следующим по ходу выполнения программы выполняется оператор 𝐶𝑜𝑛𝑠𝑜𝑙𝑒. 𝑊𝑟𝑖𝑡𝑒(); вызова функции вывода символов на экран.

Оператор используется 2 раза с разными операндами:

 "х = ";

 "у = ",

каждый из которых используется в программе однократно. Оператор включается в таблицу операторов, операнды - в таблицу операндов.

Следующая операция

х = 𝑑𝑜𝑢𝑏𝑙𝑒. 𝑃𝑎𝑟𝑠𝑒(𝐶𝑜𝑛𝑠𝑜𝑙𝑒. 𝑅𝑒𝑎𝑑𝐿𝑖𝑛𝑒());

включает три оператора:

 = – оператор присваивания используется в программе 6 раз;

* + . . . 𝑃𝑎𝑟𝑠𝑒() – оператор вызова функции преобразования строки в заданный тип используется в программе 3 раза;

 𝐶𝑜𝑛𝑠𝑜𝑙𝑒. 𝑅𝑒𝑎𝑑𝐿𝑖𝑛() – оператор вызова функции считывания строки с клавиатуры используется в программе 3 раза.

Следующий оператор 𝑖𝑓(). . . 𝑒𝑙𝑠𝑒. .. используется дважды в тексте программы для ветвления алгоритма.

Операции

 𝐹 = 𝑀𝑎𝑡ℎ. 𝑆𝑖𝑛(𝑥) + 𝑀𝑎𝑡ℎ. 𝐶𝑜𝑠(𝑦) ∗ 𝑀𝑎𝑡ℎ. 𝐶𝑜𝑠(𝑦);

 𝐹 = 𝑀𝑎𝑡ℎ. 𝐿𝑜𝑔(𝑥);

 𝐹 = 𝑀𝑎𝑡ℎ. 𝑆𝑖𝑛(𝑥) ∗ 𝑀𝑎𝑡ℎ. 𝑆𝑖𝑛(𝑥) + 𝑀𝑎𝑡ℎ. 𝐶𝑜𝑠 (𝑦);

включают следующие ранее не рассмотренные операторы:

𝑀𝑎𝑡ℎ. 𝑆𝑖(𝑥) – оператор вызова функции вычисления синуса, используется 3 раза;

𝑀𝑎𝑡ℎ. 𝐶𝑜𝑠 (у) – оператор вызова функции вычисления косинуса, применяется 3 раза;

𝑀𝑎𝑡ℎ. 𝐿𝑜(𝑥) – оператор вызова функции вычисления логарифма, используется один раз.

Имена 𝐹, 𝑥 и 𝑦 являются операндами: 𝐹 используется 5 раз,

𝑥 - 8 раз, у -7 раз.

Символы «;», «,», «∗» и «+», используемые в программе, обозначают следующие операции:

* + ; - операция определения завершения оператора, используется 16 раз;
  + , - операция отделения элементов списка, используется 2 раза;
  + ∗ - операция умножения, используется 2 раза;

 + - операция сложения (сцепления строк), используется 3 раза.

Символы «<» и «= =» используются для определения логических операций сравнения:

* + < - операция сравнения «меньше», используется один раз;

 = = - операция сравнения «равно», используется 3 раза. В позициях 23 и 24 табл. 2. представлены символы,

определяющие следующие операции:

* + {} - операция начала и завершения блока инструкций, используется 4 раза;
  + () - операция начала и завершения списка параметров или условия, используется 22 раза.

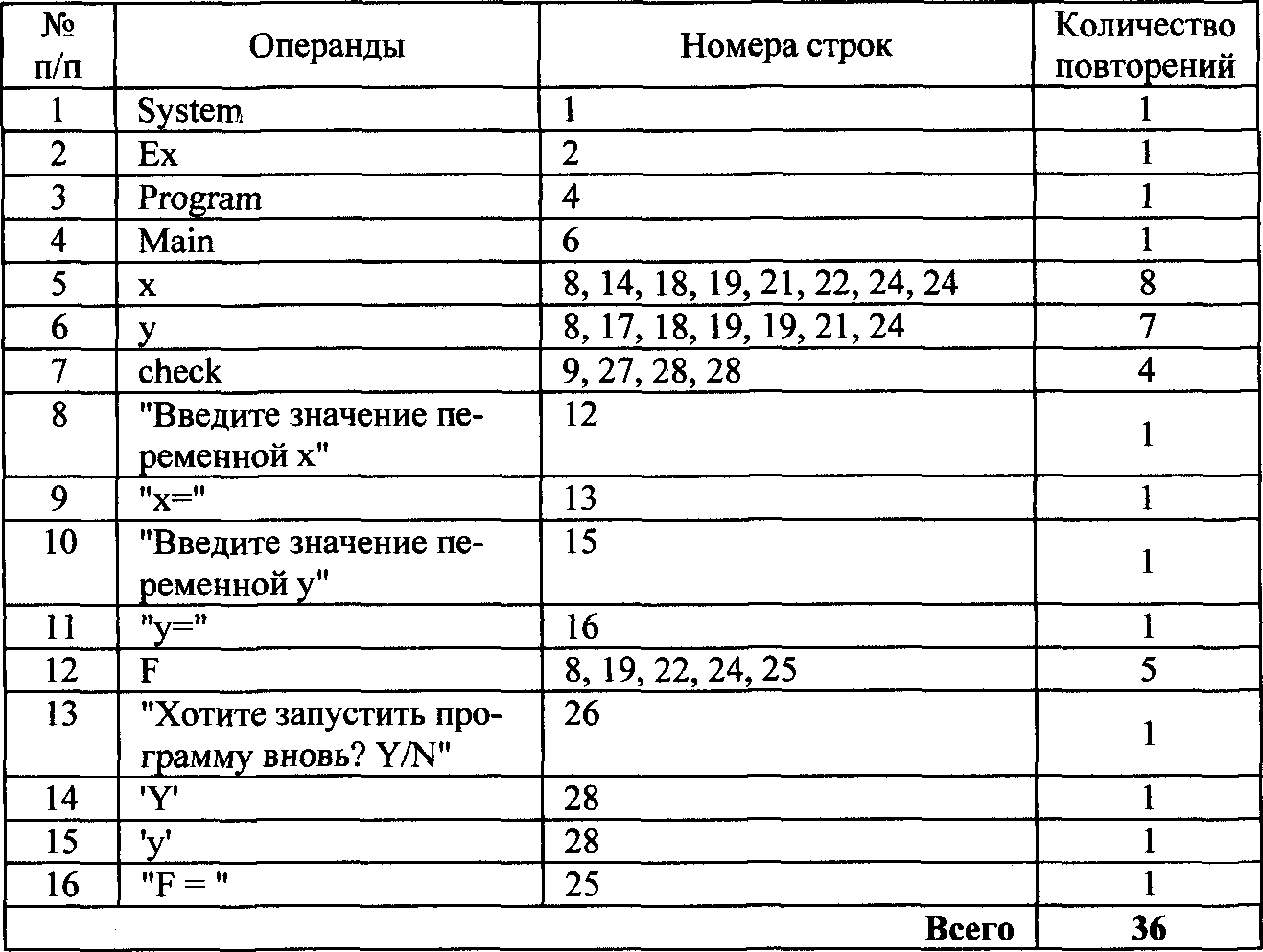
Оставшиеся четыре позиции табл. 2 содержат символы:

* + || - операция логического сложения (дизъюнкция), используется один раз;
  + " " - операция определения строковых констант, используется 6 раз;
  + ′ ′ - операция определения символьных констант, используется 2 раза;
  + . - операция связывания имен, используется 19 раз.

В табл. 3 приведены операнды рассматриваемой программы.

Таблица 3

Словарь операндов программы



Проанализируем содержание табл. 3. Позиции 1, 2, 3 и 4 содержат имена операндов 𝑆𝑦𝑠𝑡𝑒𝑚 , 𝐸𝑥 , 𝑃𝑟𝑜𝑔𝑟𝑎𝑚 , 𝑀𝑎𝑖𝑛() , которые используются в программе по одному разу. Строковые константы (позиции 8, 9,10, 11,13 и 16 табл. 3):

* + "Введите значение переменной х ";

 "х=";

* + Введите значение переменной у ;

 "у = ";

* + "Хотите запустить программу вновь? 𝑌/𝑁";

 "𝐹 = ";

используются в тексте программы однократно.

Символьные константы ′𝑌′ и ′у′ применяются также по одному разу. Имена переменных 𝑥, у, 𝑐ℎ𝑒𝑐𝑘 и 𝐹 повторяются в программе соответственно 8,7,4 и 5 раз.

Для рассматриваемой программы список входных и выходных параметров (табл. 4) не обладает большим разнообразием. Входными параметрами являются значения переменных:

 х = 𝑑𝑜𝑢𝑏𝑙𝑒. 𝑃𝑎𝑟𝑠𝑒(𝐶𝑜𝑛𝑠𝑜𝑙𝑒. 𝑅𝑒𝑎𝑑𝐿𝑖𝑛𝑒());

 у = 𝑑𝑜𝑢𝑏𝑙𝑒. 𝑃𝑎𝑟𝑠𝑒(𝐶𝑜𝑛𝑠𝑜𝑙𝑒. 𝑅𝑒𝑎𝑑𝐿𝑖𝑛𝑒());

 𝑐ℎ𝑒𝑐𝑘 = 𝑐ℎ𝑎𝑟. 𝑃𝑎𝑟𝑠𝑒(𝐶𝑜𝑛𝑠𝑜𝑙𝑒. 𝑅𝑒𝑎𝑑𝐿𝑖𝑛𝑒()).

Таблица 4.

Входные и выходные переменные программы

|  |  |
| --- | --- |
| **Входные переменные** | **Выходные переменные** |
| 𝑋 | "Введите значение переменной х" |
| У | "х = " |
| 𝑐ℎ𝑒𝑐𝑘 | "Введите значение переменной у" |
|  | "у = " |
|  | "𝐹 = " |
|  | "Хотите запустить программу вновь? 𝑌/𝑁" |
|  | 𝐹 |

Выходными значениями являются шесть констант, для которых имена совпадают со значениями, и одна переменная 𝐹:

 𝐶𝑜𝑛𝑠𝑜𝑙𝑒. 𝑊𝑟𝑖𝑡𝑒𝐿𝑖𝑛("Введите значение переменной х");

 𝐶𝑜𝑛𝑠𝑜𝑙𝑒. 𝑊𝑟𝑖𝑡𝑒("𝑥 = ");

 𝐶𝑜𝑛𝑠𝑜𝑙𝑒. 𝑊𝑟𝑖𝑡𝑒𝐿𝑖𝑛("Введите значение переменной у");

 𝐶𝑜𝑛𝑠𝑜𝑙𝑒. 𝑊𝑟𝑖𝑡𝑒("𝑦 = ");

 𝐶𝑜𝑛𝑠𝑜𝑙𝑒. 𝑊𝑟𝑖𝑡𝑒𝐿𝑖𝑛("𝐹 = " + 𝐹) − в этом случае два выходных параметра: строковая константа "𝐹 = " и переменная 𝐹;

 𝐶𝑜𝑛𝑠𝑜𝑙𝑒. 𝑊𝑟𝑖𝑡𝑒𝐿𝑖𝑛("Хотите запустить программу вновь? 𝑌/

𝑁").

Оценка характеристик программы

Используя сформированные таблицы с необходимыми параметрами для расчета и применяя соотношения Холстеда, вычислим характеристики рассматриваемой программы:

* словарь программы:

𝑛 = 𝑛1 + 𝑛2 = 28 + 16 = 44;

* длина реализации:

𝑁 = 𝑁1 + 𝑁2 = 116 + 36 = 152;

* длина программы:

𝑁̃ = 1 ∙ log2𝑛1 + 𝑛2 ∙ log2𝑛2 = 28 ∙ 𝑙𝑜𝑔228 + 16 ∙ 𝑙𝑜𝑔216

= 28 ∙ 4.81 + 16 ∙ 4 = 134.68 + 64 = 198.68;

* объем программы в битах:

𝑉 = (𝑁1 + 𝑁2) ∙ 𝑙𝑜𝑔2(𝑛1 + 𝑛2) = (116 + 36) ∙ 𝑙𝑜𝑔2(28 + 16)

= 152 ∙ 𝑙𝑜𝑔244 = 152 ∙ 5.46 = 829.92;

* потенциальный объем программы:

∗ ∗

𝑉∗ = (𝑛2 + 2) ∙ 𝑙𝑜𝑔2(𝑛2 + 2) = (10 + 2) ∙ 𝑙𝑜𝑔2(10 + 2)

= 12 ∙ 3.58 = 42.96;

* уровень программы:
* уровень языка:

𝐿 =

𝑉∗

𝑉

42.96

=

829.92

* 1. ;

𝜆 = 𝐿 ∙ 𝑉∗ = 0.052 ∙ 42.96 = 2.23;

* + - интеллектуальное содержание программы:

𝐼 = 𝐿 ∙ 𝑉 = 0.052 ∙ 829.92 = 43.16;

* + - работа по программированию:

𝑉

𝐸 =

𝐿

829.92

=

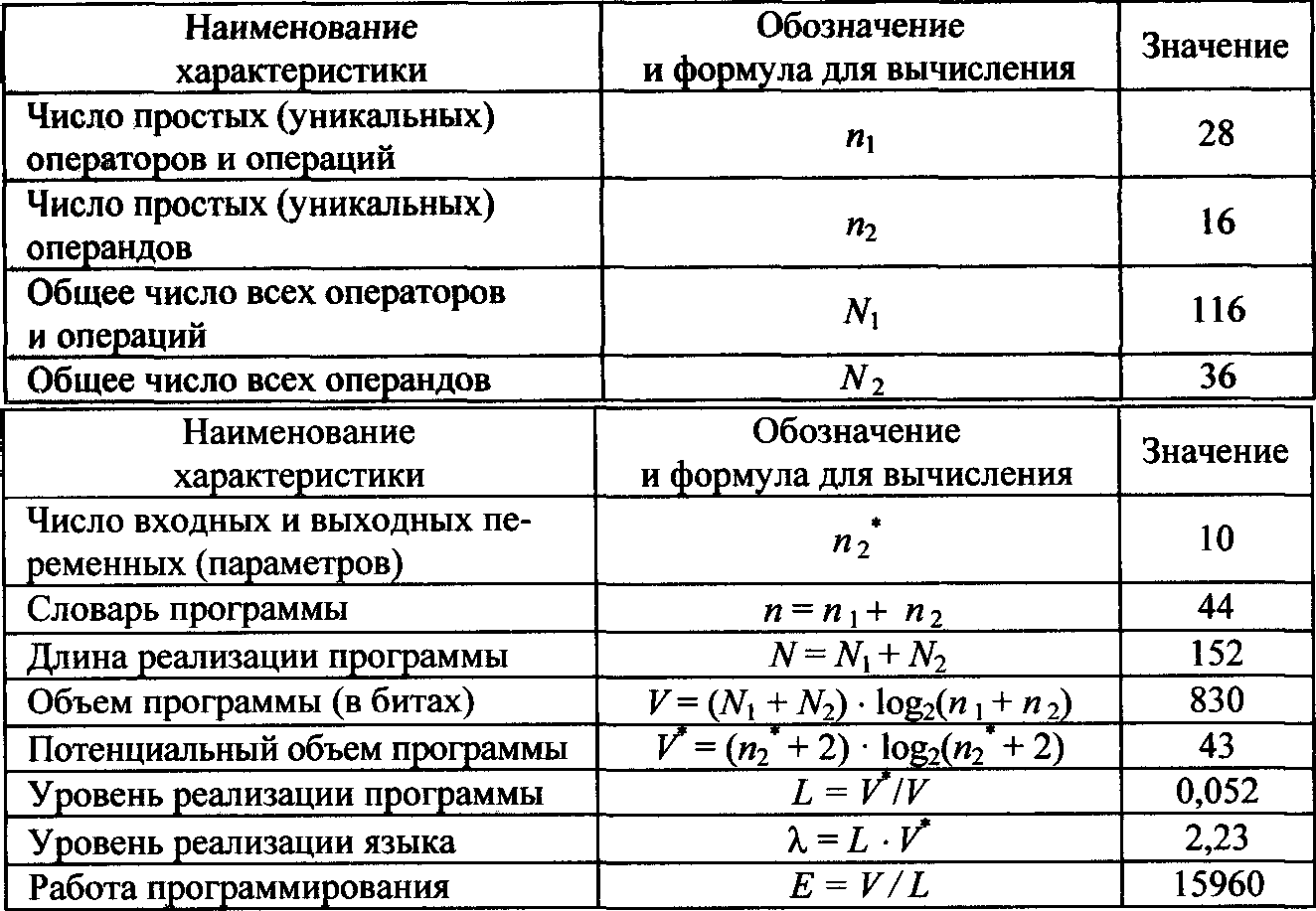
0.052

= 15960.

Сведем все результаты расчетов метрик Холстеда в табл. 5.

Таблица 5.

Значения метрик Холстеда для программы



Уровень исследуемой программы весьма низкий, так как

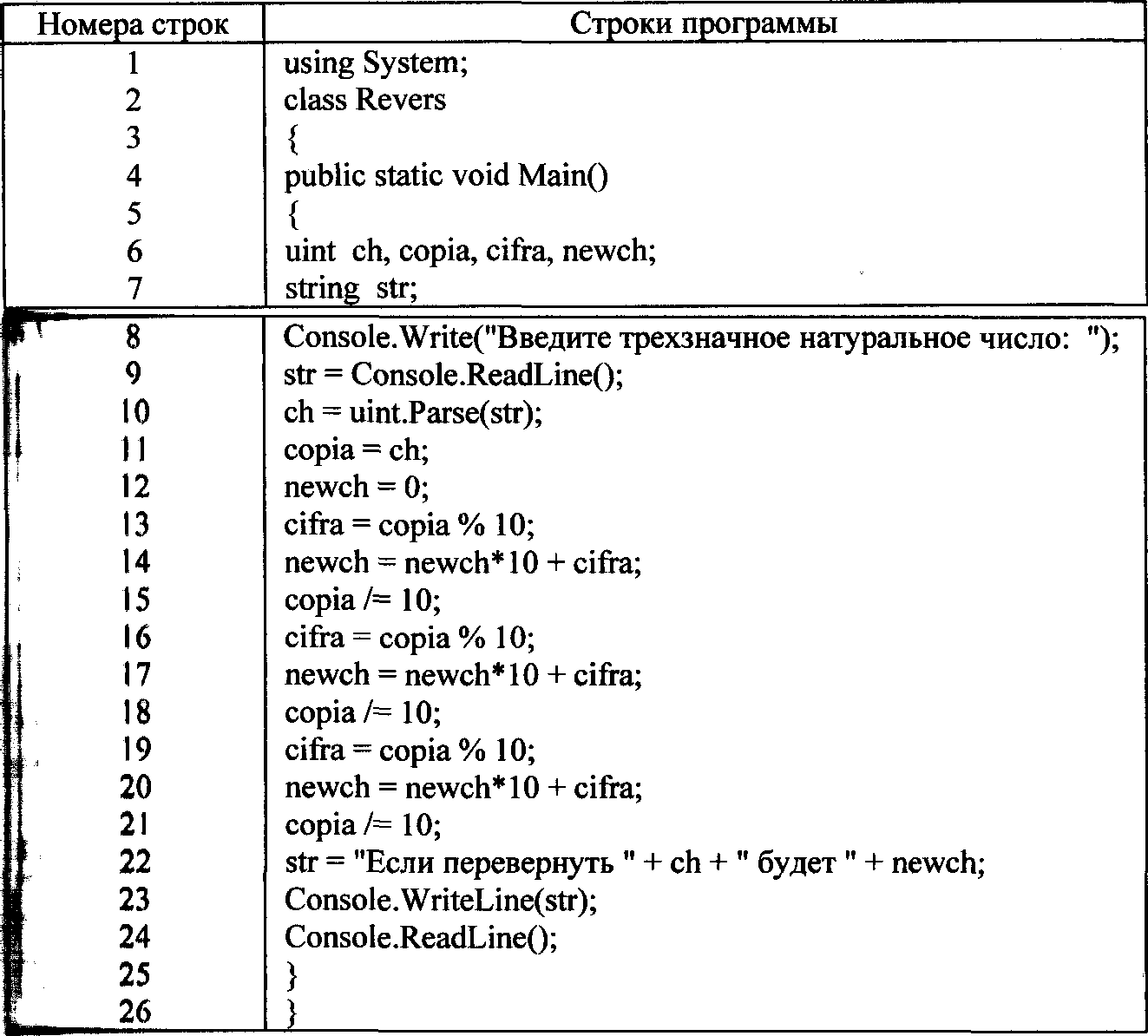
потенциальный объем программы в значительной степени меньше ее реального объема.

Задача «Зеркальное число»

Ввести с клавиатуры трехзначное натуральное число. Вычислить и вывести на экран зеркальное число, полученное путем изменения цифр числа на обратное по отношению к исходной позиции (например, для числа 253 должно быть получено 352). Разработать программу и определить значения метрик Холстеда, на основе которых дать оценку качества разработанного исходного текста программы.

Реализация программы

Текст программы для реализации возможного решения поставленной задачи, разработанной с использованием языка программирования С#, приведен на рис. 3.

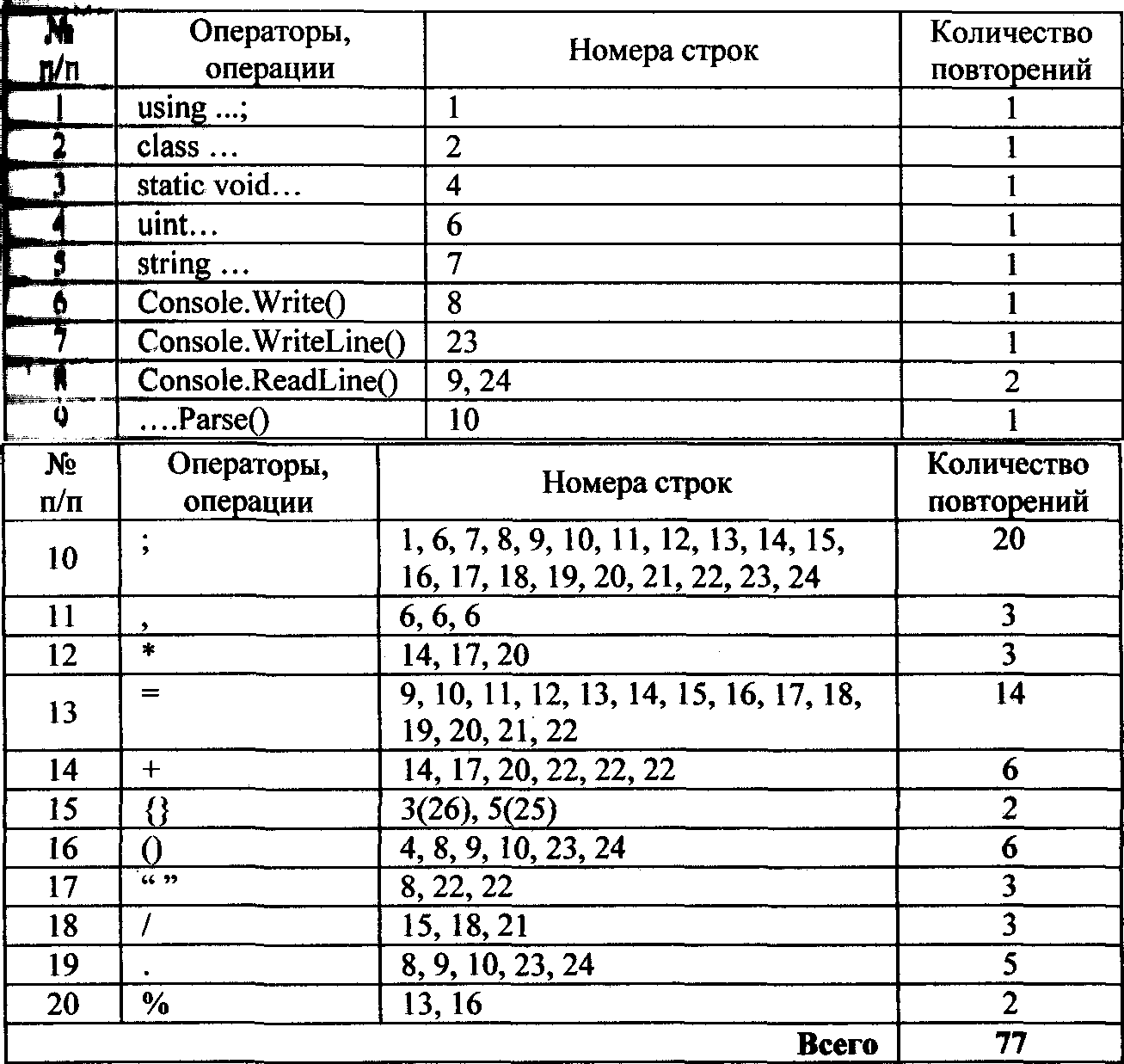


*Рис. 3. Текст программы задачи «Зеркальное число»*

Словарь программы

В табл. 6 приведены операторы и операции, используемые в программе (столбец 2). Номера строк исходной программы, где встречается каждый оператор или операция, указаны в третьем столбце. В четвертом столбце указано число повторений каждого оператора или операции в исходном тексте программы.

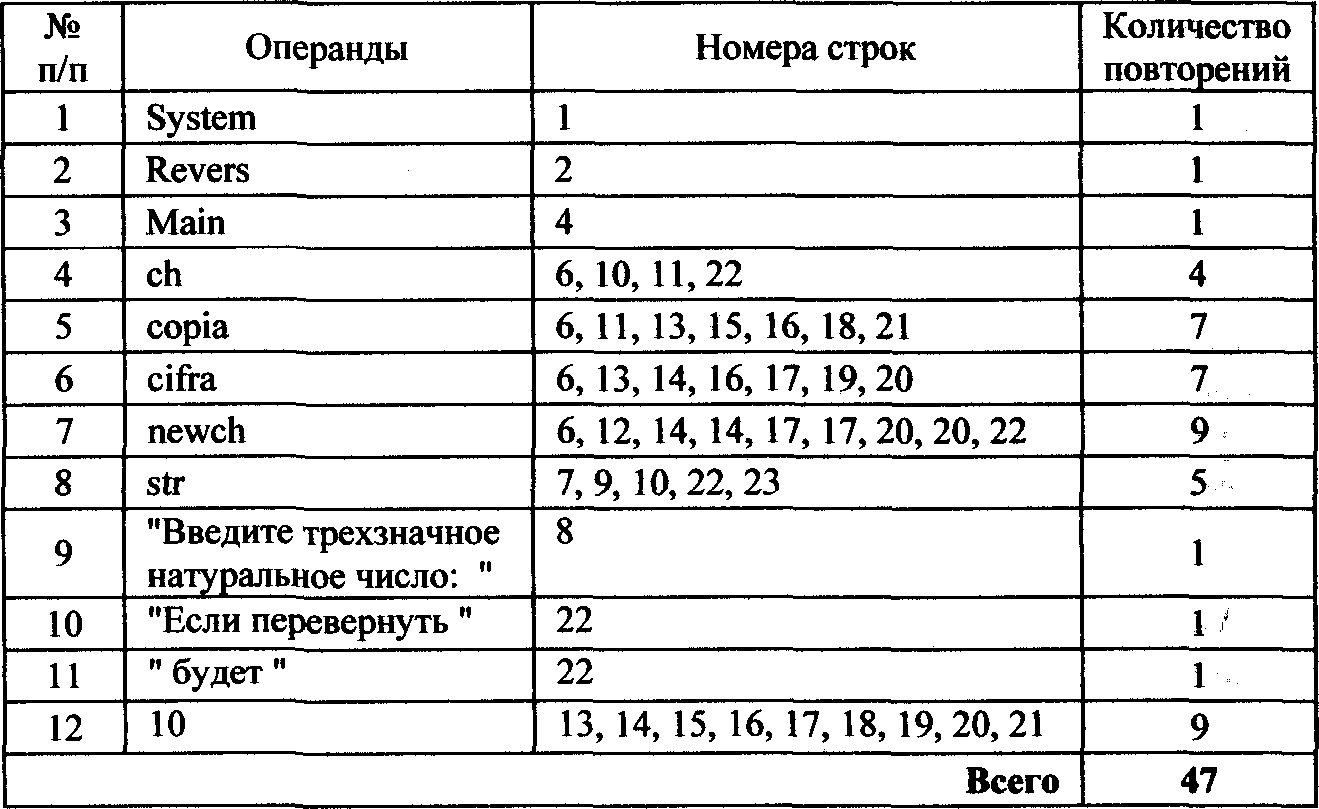
Таблица 6 Словарь операторов и операций программы



В таб. 7 приведены операнды рассматриваемой программы.

Таблица 7.

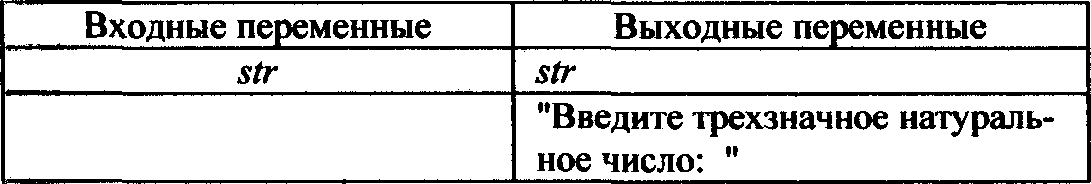
Словарь операндов программы



Для рассматриваемой программы список входных и выходных параметров (табл. 8) включает две переменные и одну константу.

Таблица 8

Входные и выходные переменные программы

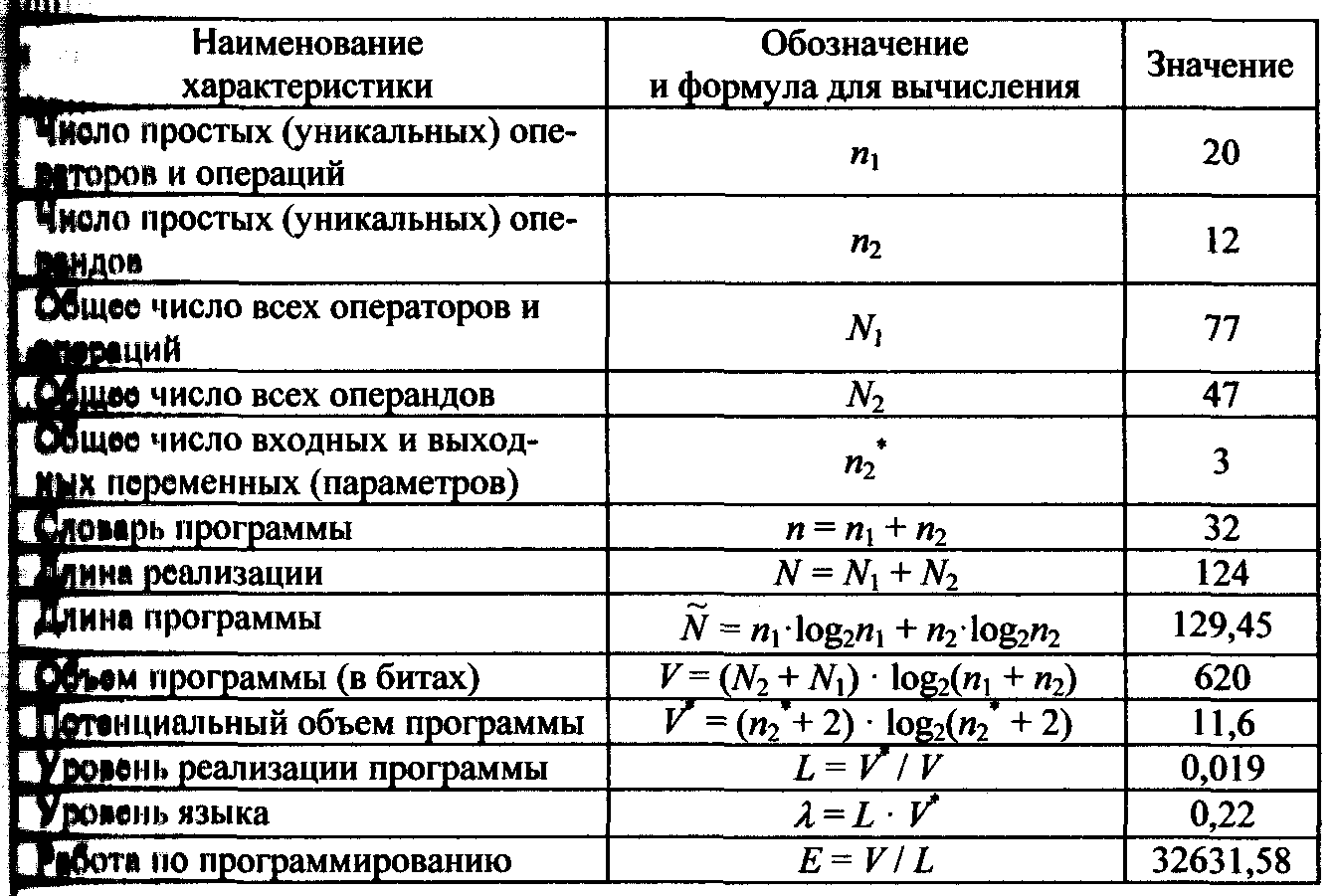


Оценка характеристик программы

Используя сформированные таблицы с необходимыми параметрами и для расчета и применяя соотношения Холстеда, вычислим характеристики рассматриваемой программы (табл. 9).

Таблица 9.

Значения метрик Холстеда для программы



Анализируя полученные показатели, можно сказать, что уровень исследуемой программы весьма низкий, так как потенциальный объем программы в значительной степени меньше ее реального объема, данном решении возможности языка программирования С# использованы на низком уровне. Длина реализации значительно меньше расчетной длины программы. Следовательно, несовершенства программирования в представленном решении задачи отсутствуют.

Задача «Заправка бака топливом

Определить класс «Бак», описывающий понятие

«Топливный бак». Данный класс должен иметь следующие поля:

* ширина, длина и высота в сантиметрах;
* вид топлива.

В классе должны присутствовать операции (методы):

* полная заправка бака заданным видом топлива;
* вычисление стоимости полной заправки бака.

Бак имеет вид параллелепипеда. Стандартным считается бак, имеющий вид куба (ширина, длина и высота совпадают).

Возможные разновидности баков:

* бак с одинаковой шириной и длиной и индивидуальной высотой;
* бак с индивидуальными размерами по ширине, длине и высоте.

Стоимость одного кубического сантиметра топлива определяется полями класса «Топливо». Стоимость топлива в рамках решения задачи неизменна. Выдача стоимости топлива должна быть реализована специальной операцией этого класса.

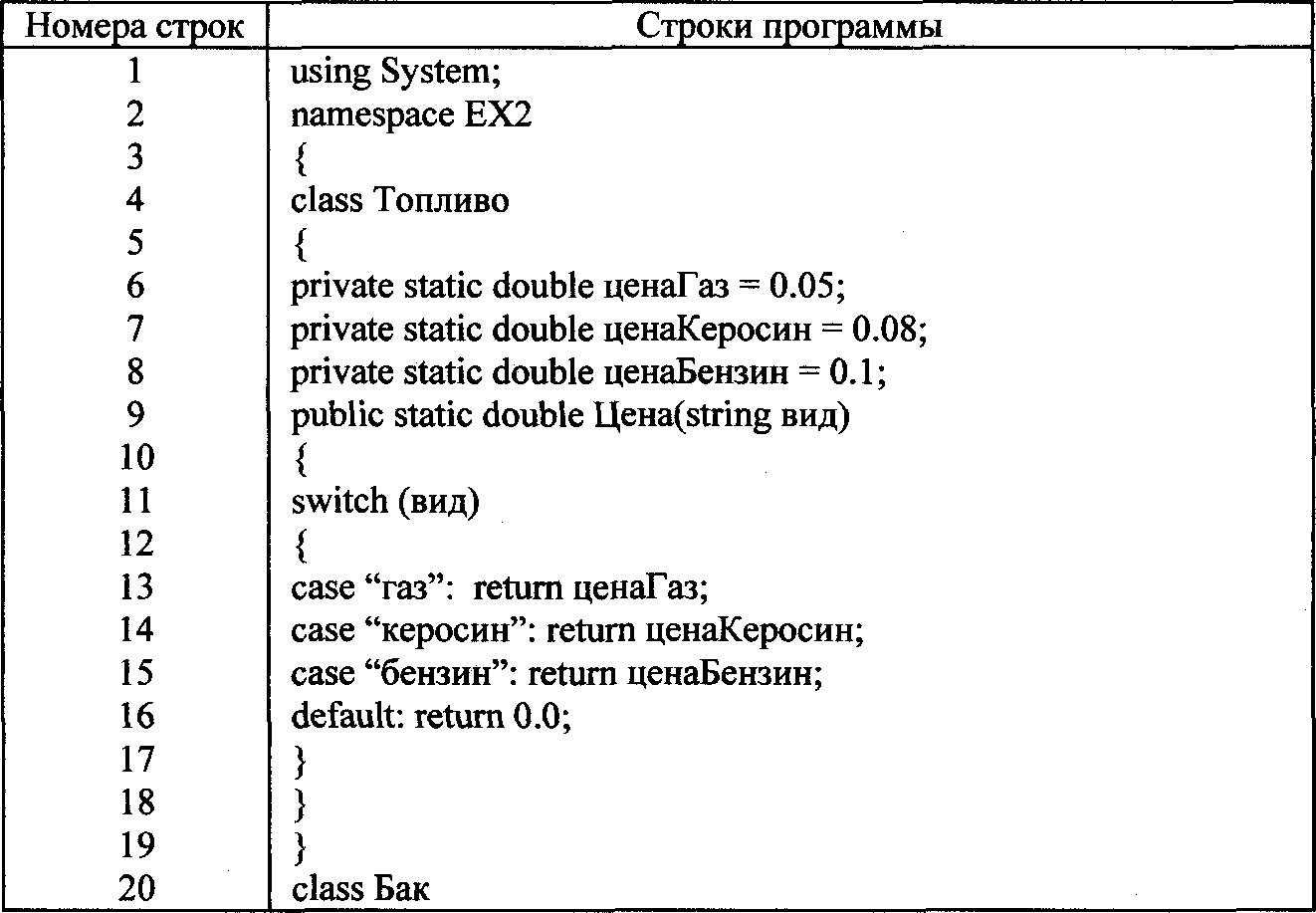
Заполнить и вывести на экран стоимость заправки четырех баков:

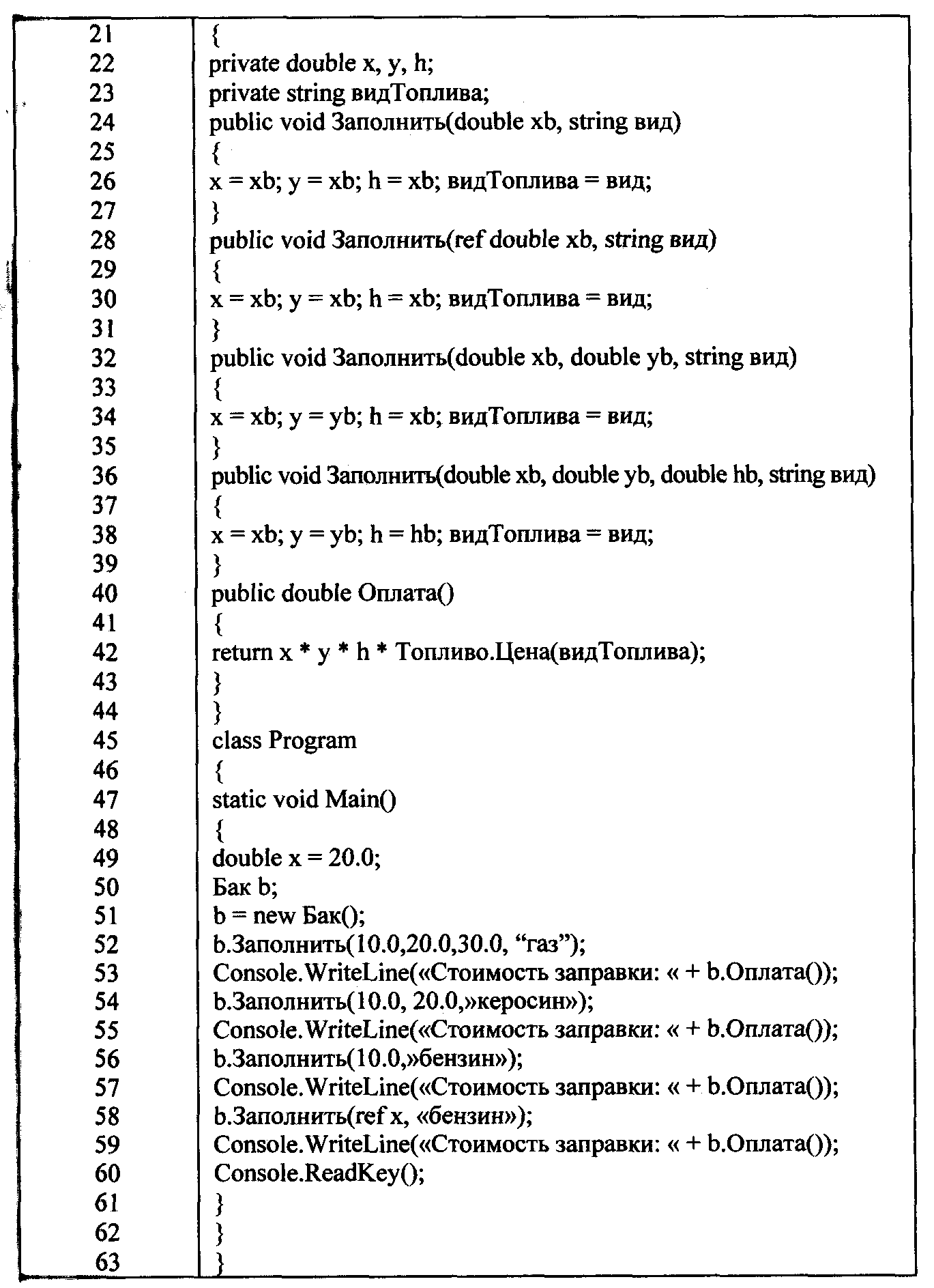
* 10x20x30 см, топливо - газ;
* 10х 10x20 см, топливо - керосин;
* 10x10x10 см, топливо - бензин;
* 20x20x20 см, топливо - бензин.

Разработать программу и определить значения метрик Холстеда, на основе которых дать оценку качества разработанного исходного текста программы.

Реализация программы

Текст программы приведен на рис. 4



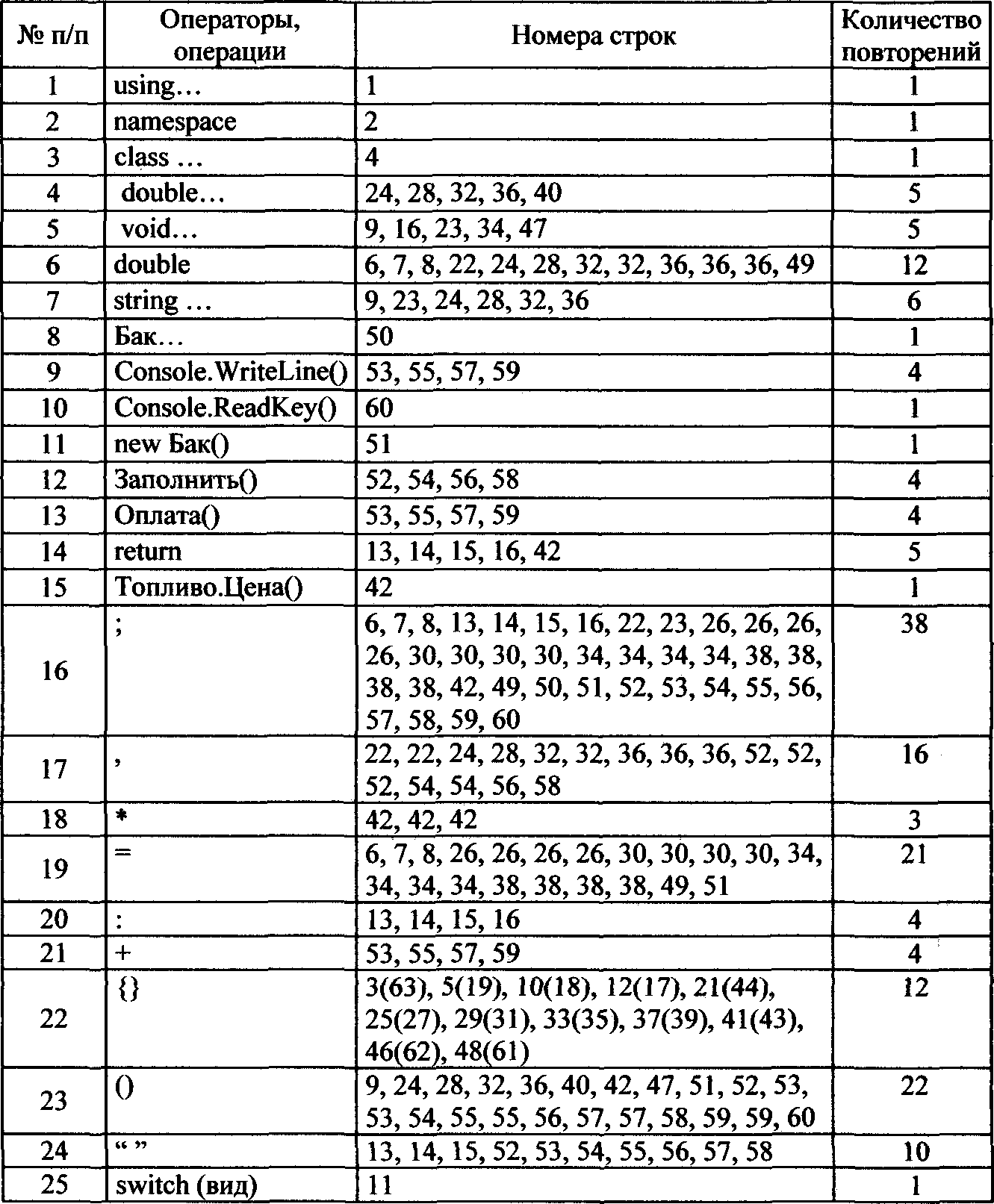


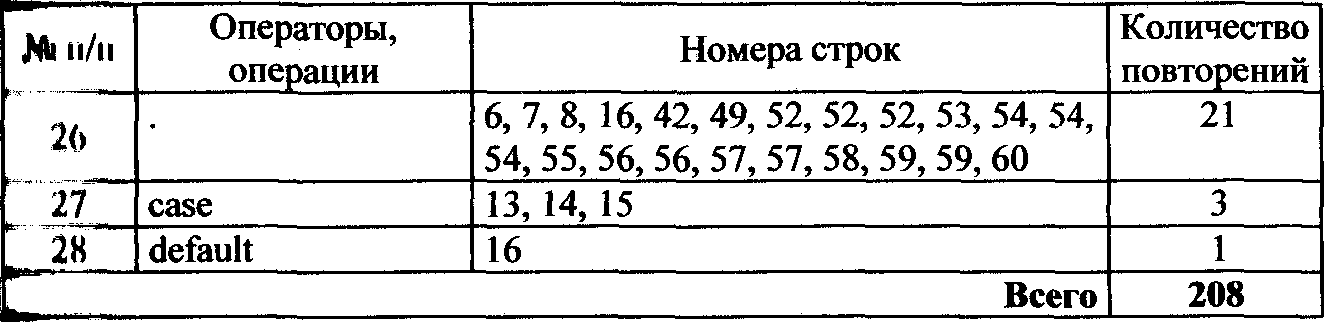
*Рис. 4. Текст программы задачи «заправка бака топливом»*

Словарь программы

В табл. 10 приведены операторы и операции, используемые в программе.

Таблица 10 Словарь операторов и операций программы



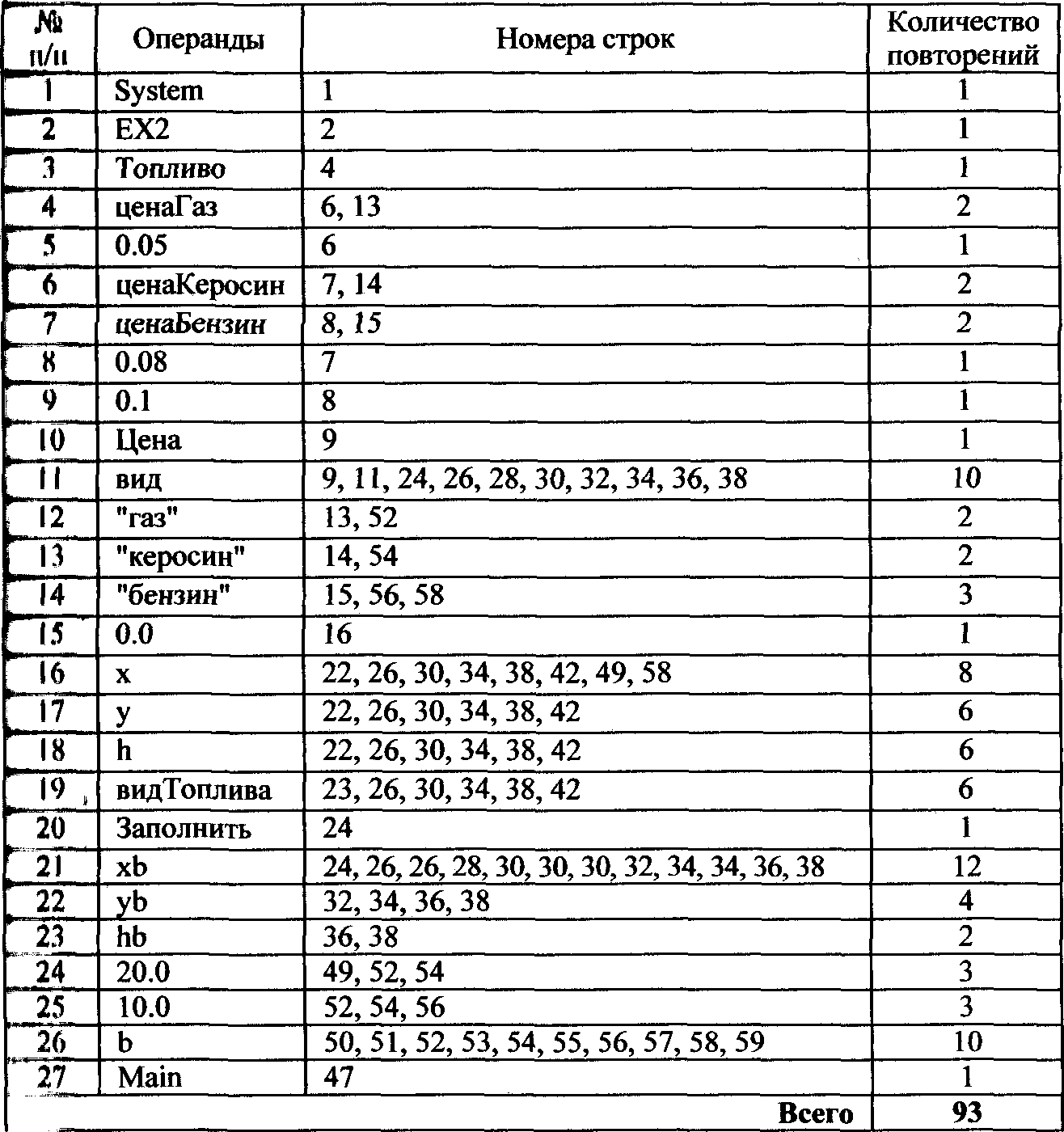


В табл. 11 приведены операнды рассматриваемой

программы.

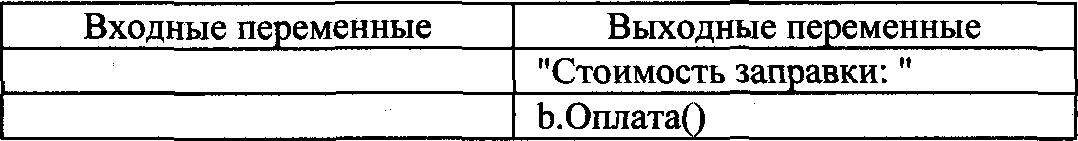
Таблица 11

Словарь операндов программы



Для рассматриваемой программы список входных и выходных параметров (табл. 12) включает одну строковую константу и возвращаемое значение метода Оплата().

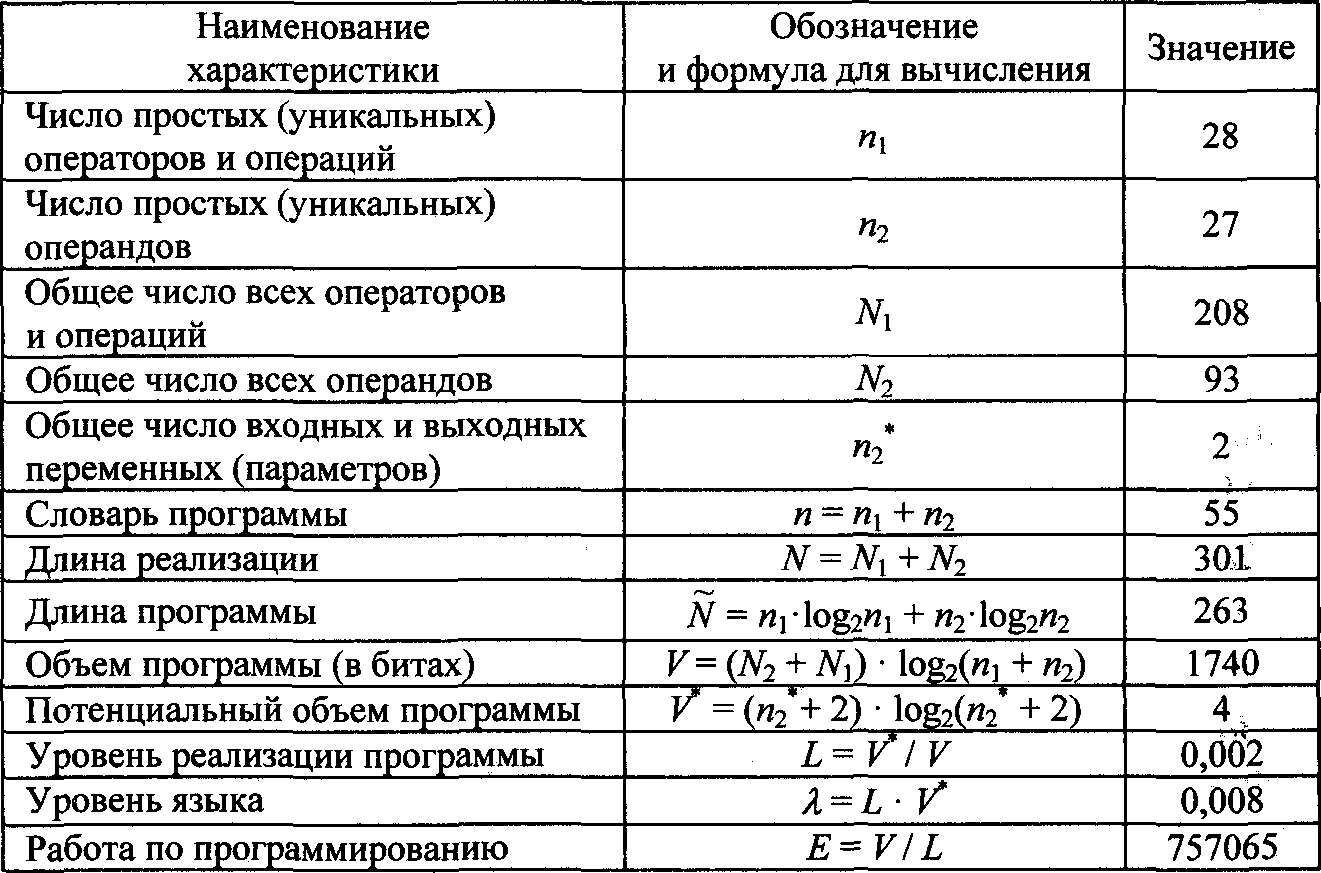
Таблица 20 Входные и выходные переменные программы



Оценка характеристик программы

Используя сформированные таблицы с необходимыми параметрами для расчета и применяя соотношения Холстеда, вычислим характеристики рассматриваемой программы (табл. 21).

Таблица 21 Значение метрик Холстеда для программы



Уровень исследуемой программы весьма низкий, так как потенциальный объем программы в значительной степени меньше ее реального объема. В данном решении возможности языка программирования С# использованы на низком уровне. Длина реализации превышает расчетную длину программы больше чем

на 10%. В программе присутствуют несовершенства. Проявлением несовершенства программы являются строки:

х = xb; у = xb; h = xb; х = xb; h = xb;.

Представленная последовательность присваиваний представляет собой набор операций с синонимичными операндами, что в значительной степени снижает качественные характеристики программы.

# ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

В задачах, предлагаемых для самостоятельного решения, необходимо выполнить следующее:

* разработать программу, реализующую заданный алгоритм (рекомендуется использовать язык программирования С#);
* сформировать словарь программы, охватывающий операнды, а также операторы и операции;
* словари оформить в виде таких же таблиц, как в рассмотрен- ных примерах;
* рассчитать метрики Холстеда, оформив результат в виде итоговой таблицы;
* провести анализ полученных результатов, сформировав содержательные выводы.

*Задача 1*. Написать и протестировать функцию, которая

«переворачивает» строку, передаваемую ей в качестве параметра, в зеркальное состояние.

*Задача 2*. Дано натуральное число 𝑁 . Вывести на экран число, которое получится после выписывания цифр числа 𝑁 в обратном порядке. Для получения нового числа составить функцию.

*Задача 3*. Написать и протестировать функцию, подсчитывающую количество минимальных элементов в каждой строке целочисленной матрицы.

*Задача 4*. Написать и протестировать функцию для вычисления числа сочетаний по формуле

𝑛!

𝐶

= =

𝑘

𝑛 𝑘! (𝑛 − 𝑘)!

(𝑛 − 1) … (𝑛 − 𝑘 + 1)

.

𝑘!

*Задача 5*. По заданным значениям 𝑋[20], 𝑌[20] вычислить

19

∑ 𝑥𝑖

2

15

; при ∑ 𝑥𝑖 ∙ 𝑦𝑖 > 0 ;

𝑢 =

𝑖=0 19

2

𝑖=1

15

∑ 𝑦𝑖 ; при ∑ 𝑥𝑖 ∙ 𝑦𝑖 ≤ 0.

{ 𝑖=10

𝑖=1

*Задача 6*. Написать и протестировать функцию, преобразующую строку восьмеричных цифр в эквивалентное ей целое десятичное число.

*Задача 7*. Описать функцию 𝑚𝑖𝑛𝑚𝑎𝑥(𝑥, 𝑦) , которая присваивает первому параметру большее, а второму – меньшее из значений 𝑥 и у . Используя эту функцию, перераспределить введенные значения переменных 𝐴, 𝐵, 𝐶 так, чтобы стало 𝐴 ≤ 𝐵 ≤

𝐶.

*Задача 8*. Даны две квадратные матрицы. Напечатать ту из них, которая имеет минимальный «след», т. е. сумму элементов главной диагонали. Использовать функцию для нахождения следа матрицы и функцию печати матрицы.

*Задача 9*. Составить и протестировать функцию для вычисления

𝑛

𝑓(𝑥, 𝑛, 𝑚) = ∑

𝑖=𝑚

𝑥2𝑖

(2 ∙ 𝑖)!

*Задача 10*. Написать и протестировать функцию 𝑐𝑜𝑚𝑝𝑟𝑒𝑠(),

которая «сжимает» строку, удаляя из нее все пробелы.

*Задача 11*. Написать и протестировать функцию, которая подсчитывает, сколько раз в заданной строке встретился указанный символ.

*Задача 12*. Написать и протестировать функцию, которая находит в массиве минимальный по модулю элемент и заменяет им всё элементы с нечетными номерами.

*Задача 13*. Написать и протестировать функцию, которая в прямоугольной матрице находит сумму элементов -й строки.

*Задача 14*. Написать и протестировать функцию, которая по заданному натуральному числу определяет количество цифр в нем и их сумму.

*Задача 15*. Написать и протестировать функцию, которая по заданной строке 𝑆𝑡𝑟 , содержащей буквы и цифры, формирует новую строку, состоящую только из цифр, входящих в 𝑆𝑡𝑟.

*Задача 16*. Написать и протестировать функцию, подсчитывающую количество положительных элементов в массиве.

*Задача 17*. Написать и протестировать функцию,

вычисляющую =

формулу

√𝑥 (0 < |𝑥| < 2) , используя итерационную

4

3

1

𝑦𝑖+1 = 𝑦𝑖 + 3 (𝑦𝑖 −

𝑦𝑖 )

𝑥

Начальное приближение 𝑦0 = 𝑥 . Итерации прекратить при

достижении условия

|𝑦𝑖+1 − 𝑦𝑖| < 2 ∙ 10−6

*Задача 18*. Составить и протестировать функцию для замены

cимволов «:» на «.» в заданной строке, начиная с указанной позиции.

*Задача 19*. Выяснить, сколько простых чисел находится в интервале [𝑚, 𝑛], и распечатать их. Для определения, является ли очередное число простым, составить функцию.

*Задача 20*. Написать и протестировать функцию для вычисления площади треугольника, заданного координатами вершин.

*Задача 21*. Написать и протестировать функцию для нахождения в прямоугольной матрице номера строки, имеющей максимальную сумму элементов.

*Задача 22*. Написать и протестировать функцию, которая преобразует строку двоичных цифр в эквивалентное ей целое десятичное число.

*Задача 23*. Написать и протестировать функцию, которая в строке, передаваемой ей в качестве параметра, заменяет каждый второй элемент на заданный символ.

*Задача 24*. Написать и протестировать функцию для сложения и вычитания вещественных матриц. Одним из формальных параметров должен быть признак вида операции.

*Задача 25*. Написать и протестировать функцию, которая преобразует строку шестнадцатеричных элементов числа в эквивалентное ей целое десятичное число.