Язык программирования Тривиль

Алексей Недоря 19.02.2023

Содержание

1	Назначение		
2	Обзо	р языка	5
3	Лекс	сика	6
	3.1	Комментарии	6
	3.2	Разделители синтаксических конструкций	6
	3.3	Идентификаторы	7
	3.4	Ключевые слова	7
	3.5	Знаки операций и знаки препинания	7
	3.6	Целочисленные литералы	8
	3.7	Вещественные литералы	8
	3.8	Строковые литералы	8
	3.9	Символьные литералы	8
		Модификаторы	9
		• •	
4		сания и области действия	10
	4.1	Области действия	10
	4.2	Предопределенные идентификаторы	10
	4.3	Предопределенные типы	11
	4.4	Указание типа	11
	4.5	Описание констант	12
	4.6	Описание переменных	13
	4.7	Описание типов	13
	4.8	Тип вектора	14
	4.9	Тип класса	14
		4.9.1 Инициализация полей	15
		4.9.2 Экспорт полей	15
		4.9.3 Наследование	15
	4.10	Может быть тип	15
		Описание функций	16
		4.11.1 Параметры	16
		4.11.2 Полиморфные параметры	17
		4.11.3 Внешние функции	17
	4.12	Описание методов	17
_			40
5	-	ажения	19
	5.1	Первичные выражения	19
	5.2	Операнды	19
	5.3	Доступ к полям и методам	20
	5.4	Индексация	20
	5.5	Преобразование типа	21
	5.6	Вызов функции или метода	22
		5.6.1 Развернутое выражение	22
	5.7	Конструктор вектора	23
	5.8	Конструктор экземпляра класса	23
	5.9	Подтверждение типа	24
	5.10	Унарные операции	24

	11 Бинарные операции 25 5.11.1 Приоритеты операции 25 5.11.2 Арифметические операции 25 5.11.3 Операции сравнения 26 5.11.4 Логические операции 26 5.11.5 Битовые операции 26 12 Константные выражения 26
6. 6. 6. 6. 6. 6. 6. 6.	2 Локальные описания 27 3 Выражение, как оператор 28 4 Оператор присваивания 28 5 Инкремент и декремент 29 6 Оператор 'если' 29 7 Оператор 'надо' 29 8 Оператор 'когда' или оператор выбора 30
7 C	7.1.1 Длина
8. 8. 8. 9. П 9.	2 Импорт 35 3 Вход или инициализация модуля 36 4 Инициализация и исполнения программы 36 равила совместимости 37 1 Эквивалентность типов 37
	2 Совместимость по присваиванию 37 рименять с осторожностью 38 0.1 Осторожные преобразования 38

1. Назначение

Язык программирования Тривиль разработан в рамках работы на семейством языков программирования Языки выходного дня (ЯВД) проекта Интенсивное программирования.

Тривиль является нулевым языком семейства ЯВД, предназначенным для реализации компиляторов и экосистемы других языков семейства. В рамках классификации языков, принятом в проекте Интенсивное программирования, это язык L2.

Основными требованиями к языку при разработке были поставлены

- Язык должен быть минимально достаточным для удобной разработки компиляторов. Требование это во многом субъективно, так как компиляторы можно писать существенно по разному.
- Язык должен быть русскоязычным и с синтаксисом минимизирующим переключение на латиницу в процессе разработки программ.

Название языка происходит от слова "тривиальный что означает, что при разработке языка практически везде использовались решения, проверенные в других современных языках программирования, в первую очередь "донорами"являются Go, Swift, Kotlin и Oberon.

Несмотря на узкую направленность на разработку компиляторов, Тривиль является языком программирования общего назначения, пригодным для решения широкого круга задач.

Язык (и экосистема) обладает существенными предпосылками для использование его в качестве учебного языка для обучения студентов разработке компиляторов, библиотек, средств разработки, алгоритмов оптимизации и так далее, в первую очередь это:

- Простота языка
- Современный вид и набор конструкций языка
- Простота компилятора
- Открытая лицензия.

2. Обзор языка

Тривиль - это модульный язык с явным экспортом и импортом, автоматическим управлением памятью (сборка мусора), с поддержкой ООП.

Программа на языке Тривиль состоит из модулей (единиц компиляции), исходный текст каждого модуля расположен в одном или нескольких исходных файлов.

Пример программы:

```
1 модуль х

2 импорт "стд/вывод"

4 вход {

 вывод.ф("Привет!\n")

7 }
```

Для описания языка используется EBNF в формате, близком к формату ANTLR4. Операции:

Пример:

Список-операторов: Оператор (Разделитель Оператор)*

3. Лексика

Исходный текст есть последовательность лексем: идентификаторов (§3.3), ключевых слов (§3.4), знаков операций и знаков препинания (§3.5), литералов (§3.6, §3.7, §3.8, §3.9) и модификаторов (§3.10). Каждая лексема состоит из последовательности Unicode символов (unicode code point) в кодировке UTF-8.

Пробелы (U+0020), символы табуляции (U+0009) и символы завершения строки $(U+000D,\ U+000A)$ разделяют лексемы, и, игнорируются, кроме следующих случаев:

- Символы завершения строк могут использоваться как разделители синтаксических конструкций (§3.2).
- Пробелы являются значащими символами в идентификаторах, состоящих из нескольких слов (§3.3).
- Пробелы являются значащими в строковых и символьных литералах (§3.3).

Несколько таких разделителей трактуются, как один.

Исходный текст может содержать комментарии (§3.1).

3.1. Комментарии

Есть две формы комментариев:

- Строчный комментарий начинается с последовательности символов '//' и заканчивается в конце строки.
- Блочный комментарий начинается с последовательности символов '/*' и заканчивается последовательностью символов '*/'. Блочные комментарии могут быть вложенные.

```
Комментарий
: '//' (любой символ, кроме завершения строки)*
| '/*' ( любой символ) * '*/
```

3.2. Разделители синтаксических конструкций

Некоторые синтаксические правила используют нетерминал *Разделитель* для разделения двух подряд идущих синтаксических конструкций, например:

```
Список-операторов: Оператор (Разделитель Оператор)*
```

В качестве разделителя может использоваться символ ';' или символ завершения строки.

```
Разделитель: ';' | символ-завершения-строки
```

Пример:

```
1 a := 1; 6 := 2
2 B := 1
```

В строке 1 операторы разделены символом ';', а оператор в строке 2 отделен от операторов строки 1 символом завершения строки.

Ошибка компиляции - нет разделителя:

```
1 a := 1 f := 2
```

3.3. Идентификаторы

Идентификатор - это последовательность *Слов*, разделенных пробелами или символами дефис '-' с опционально завершающим знаком препинания:

Каждое слово состоит из *Букв* и $\mu \mu p$, и начинается с Буквы. Буквой считается любой Unicode символ, имеющий признак *Letter*, и, дополнительно, символы ' ν ' и ' '.

```
Идентификатор: Слово ((' ' | '-') Слово)* Знак-препинания? Слово: Буква (Буква | Цифра)* Буква: Unicode-letter | '_' | 'N' ' Цифра: '0' .. '9' Знак-препинания: '?' | '!'
```

Примеры идентификаторов:

```
1 буква
2 буква-или-цифра
3 №-символа
4 Цифра?
5 Пора паниковать!
```

3.4. Ключевые слова

Следующие ключевые слова зарезервированы и не могут быть использованы, как идентификаторы:

авария	есть	когда	надо	прервать
вернуть	иначе	конст	осторожно	пусть
вход	импорт	мб	пока	ТИП
если	класс	модуль	позже	фн

3.5. Знаки операций и знаки препинания

Следующие последовательности символов обозначают знаки операций и знаки препинания:

```
+ - * / %
& | ~
= # < <= > >=
```

```
( ) [ ] { }
(: . ^ , : ;
```

3.6. Целочисленные литералы

```
Целочисленный-литерал: Цифра+ | '0x' Цифра16+
Цифра16: '0'..'9' | 'a'..'f' | 'A'..'F'
```

3.7. Вещественные литералы

В текущей реализации есть только одна форма записи вещественных литералов, без экспоненты.

```
Вещественный-литерал: Цифра+ '.' Цифра*
```

3.8. Строковые литералы

Строковый литерал - это последовательность символов, заключенные в двойные кавычки. Строковый литерал может содержать символы, закодированные с помощью escape-последовательности, которая начинается с символа '\'.

```
Строковый-литерал

: '"'
    (~('"' | '\\' | '\n' | '\r' | '\t') | Escape)*

""'

Escape

: '\\'
    ('u' Цифра16 Цифра16 Цифра16 Цифра16 | 'n' | 'r' | 't' |
    | """
    | """
    )
```

3.9. Символьные литералы

Символьный литерал задает значение для Unicode code point, - это последовательность символов, заключенные в двойные кавычки. Он записывается как один или несколько символов, заключенных в одинарные кавычки. Символьный литерал может быть закодирован с помощью escape-последовательности, которая начинается с символа '\'.

3.10. Модификаторы

Модификаторы использутся в исходном тексте, чтобы внести изменение в в семантику и/или синтаксис конструкции языка, см., например, §4.11.3.

```
Модификатор: '@' Буква+ Список-атрибутов?
Список-атрибутов: '(' (Атрибут (',' Атрибут)*)? ')'
Атрибут: Строковый-литерал ': 'Строковый-литерал
```

```
@внеш("имя":"print_string")
```

4. Описания и области действия

Каждый идентификатор, встречающийся в программе, должен быть описан, если только это не предопределенный идентификатор (§4.2). Идентификатор может быть описан как тип, константа, переменная, функция или метод, а так же как параметр в сигнатуре функции или метода, и как поле в типе класса.

Описание

: Описание-типов

| Описание-констант

| Описание-переменных

| Описание-функций

| Описание-метода

4.1. Области действия

Описанный идентификатор используется для ссылки на связанный объект, в тех частях программы, которые попадают в *область действия* описания. Идентификатор не может обозначать более одного объекта в пределах заданной области действия. Область действия может содержать внутри себя другие области действия, в которых идентификатор может быть переопределен.

Область действия, которая содержит в себе все исходные тексты на языке Тривиль называется *Универсум*.

Области действия:

- Областью действия предопределенного идентификатора является Универсум
- Областью действия идентификатора, описанного на верхнем уровне (вне какой-либо функции), является весь модуль (§8).
- Областью действия идентификатора импортируемого модуля является файл (часть модуля), содержащего импорт (§8.2).
- Областью действия идентификатора, обозначающего параметр функции, является тело функции (§4.11).
- Областью действия идентификатора, описанного в теле функции (§4.11) или теле входа (§8.4), является часть блока (§6.1), в котором описан идентификатор, от точки завершения описания и до завершения этого блока.

В описании верхнего уровня за идентификатором может следовать признак экспорта '*', указывающий, что идентификатор экспортирован и может использоваться в другом модуле, *импортирующем* данный (§8.2).

Правило для опционально экспортированного идентификатора:

Идент-оп: Идентификатор '*'?

4.2. Предопределенные идентификаторы

Следующие идентификаторы неявно описаны в области действия Универсум.

```
Типы (§4.3):

Байт Цел64 Слово64 Вещ64 Лог Символ Строка

Константы типа Лог (§4.3):

ложь истина

Литерал nullable (§4.10):

пусто

Стандартные функции (§7.1):

длина тег нечто

Кроме того, для векторных типов определен набор встроенных методов (§7.2).
```

4.3. Предопределенные типы

Следующие типы обозначаются предопределенными идентификаторами, значениями данных типов являются:

Тип	Множество значений	
Байт	множество целых числа от 0 до 255	
Цел64	Цел64 множество всех 64-битных знаковых целых чисел	
Слово64 множество всех 64-битных беззнаковых целых чисел		
Вещ64 множество всех 64-разрядных чисел с плавающей запятой стандарта IEEE-7		
Лог	константы ложь и истина	
Символ	множество всех Unicode символов	
Строка	множество всех строковых литералов	

Операции над значениями этих типов определены в (§5.10, §5.11).

4.4. Указание типа

Тривиль является языком со статической типизацией, что означает, что тип любого объекта языка явно или неявно указывается во время описания объекта. Неявное указание типа может быть использовано в описании констант ($\S4.5$), переменных ($\S4.5$) и полей класса ($\S4.9$).

Для явного указании типа используется имя типа, перед которым может стоять ключевое слово мб (*может быть*).

```
Указ-типа: 'мб'? Имя-типа
Имя-типа: Идентификатор | Импорт-идент
Импорт-идент: Идентификатор '.' Идентификатор
```

Если в качестве имени типа используется *Импорти-идент*, то первый идентификатор в нем должен обозначать имя импортированного модуля (§8.2), а второй идентификатор должен быть идентификатором экспортированного из этого модуля объекта, в данном контексте - идентификатор типа.

Множество значений объекта с типом мб T состоит из значения, обозначенного предопределенным идентификатором пусто и значений типа T. Тип такого объекта называется может быть T (§4.10).

4.5. Описание констант

Описание константы связывает идентификатор с постоянным значением. Значение константы может быть задано явно или неявно, в случае группового описания констант.

```
Описание-констант: 'конст' (Константа | Группа-констант)
Константа: Идент-оп (':' Указ-типа)? '=' Выражение
```

Если тип константы не указан, то он устанавливается равным типу выражения (§5). Выражение для константы должно быть константным выражением (§5.12).

```
1 конст к1: Цел64 = 1 // тип Цел64
2 конст к2: Байт = 2 // тип Байт
3 конст к3 = 3 // тип Цел64
4 конст к4 = "Привет" // тип Строка
```

Групповое описание констант позволяет опускать тип и выражение для всех констант, кроме первой константы в группе и указать признак экспорт для всех констант группы.

```
Группа-констант:
    '*'? '('
    Kонстанта (Разделитель След-константа)*
    ')'
След-константа: Идент-оп ((':' Указ-типа)? '=' Выражение)?
```

Пример группового экспорта:

```
1 конст *(
2 Счетчик = 1
3 Имя = "Вася"
4 )
```

Пример неявного задания значения для констант:

```
конст (
1
2
        / / операции
                  // тип Цел64, значение = 1
3
        ПЛЮС = 1
        МИНУС
                     // тип Цел64, значение = 2
4
5
        OCTATOK
                  // тип Цел64, значение = 3
6
        // ключевые слова
7
        ЕСЛИ = 21
                    // тип Цел64, значение = 21
                     // тип Цел64, значение = 22
        ИНАЧЕ
8
                       // тип Цел64, значение = 23
9
        ПОКА
10
```

4.6. Описание переменных

Описание переменной создает переменную, привязывает к ней идентификатор и указывает её тип и начальное значение через *Инициализацию*.

```
Описание-переменной:
    'пусть' Идент-оп (':' Указ-типа)? Инициализация
Инициализация: (':=' | '=') ('позже' | Выражение)
```

Если тип переменной не указан, то он устанавливается равным типу выражения (§5).

Каждая переменная должна быть явно проинициализирована. Если инициализация задана через лексему '=', то значение переменной не может быть изменено (переменная с единственным присваиванием), если же в инициализации используется лексема ':=', то значение может быть изменено (§6.4, §6.5), а переменная называется изменяемой.

Явное указание типа:

```
1 пусть ц1: Цел64 = 1
2 пусть ц2: Цел64 := 2
```

Неявное указание типа:

```
1 пусть ц3 = 3 // неявное указание типа
2 пусть ц4 := 4
3 пусть имя = Имя языка()
```

Компилятор выдает ошибку, при попытке изменить значение переменной с единственным присваиванием:

```
1 ц1 := 2
2 ц3++
3 имя := "C++"
```

Если в инициализации задано *Выражение*, то начальным значение переменной является значение выражения. Если при этом тип переменной явно задан, то выражение должно быть *совместимо по присваиванию* с типом переменной (§9.2).

Вместо выражения может быть указано ключевое слово позже, это *поздняя инициализация*. Такая форма инициализации разрешена только для переменных уровня модуля, при этом тип таких переменных, должен быть явно указан.

Значение переменной с поздней инициализацией должно быть задано в инициализации модуля (§8.4). Текущая реализация компилятора запрещает позднюю инициализацию для переменных с единственным присваиванием.

4.7. Описание типов

Описание типа связывает идентификатор с новым типом, который определяет структуру данных этого типа, и, как следствие, набор операций над данными этого типа.

```
Описание-типа: 'тип' Идент '=' (Тип-вектора | Тип-класса)
```

В текущей версии языка поддерживают два структурных типа: тип вектора ($\S4.8$) и тип класса ($\S4.9$).

4.8. Тип вектора

Вектор - это пронумерованная последовательность элементов одного типа, называемая типом элемента. Количество элементов называется длиной вектора, длина не может быть отрицательной. Элементы вектора доступны через операцию индексации (§5.4). Для индексации вектора используются целочисленные индексы от 0 до длина-1. Векторы всегда одномерны, но типом элемента может быть вектор, там самым формируя многомерную структуру.

```
Тип-вектора: '[' ']' Указ-типа

1 тип Байты = []Байт

тип Матрица = [][]Вещ64
```

Длина вектора может изменятся во время выполнения, другими словами, вектор - это *динамический массив*. Для получения текущей длины вектора используется стандартная функция длина (§7.1.1). Начальное значение для объекта типа вектор задается с помощью конструктора вектора (§5.7). Далее, длину вектора можно поменять с помощью стандартных методов (§7.2).

Пример инициализации вектора из трех элементов:

```
1 пусть байты = Байты[1, 2, 3]
```

4.9. Тип класса

Тип класса - это структура, состоящая из полей. С типом класса могут быть связаны функции, называемые методами (§4.12).

Описание класса состоит из опционального указания базового класса и списка полей. Если базовый класс указан, то класс *наследует* поля и методы базового класса (§4.9.3).

Для каждого поля в списке задается идентификатор, тип (явно или неявно) и начальное значение. Областью действия идентификаторов полей является само описание класса, но они могут быть доступны в операции доступа к полям и методам (§5.3).

```
Тип-класса: 'класс' Базовый-класс? '{' Список-полей? '}'
Базовый-класс: '(' Указ-типа ')'
Список-полей: Поле (Разделитель Поле)*
Поле: Идент-оп (':' Указ-типа)? Инициализация

тип Человек = класс {
имя: Строка := ""
возраст: Цел64 := 0
}
```

Если тип поля не указан, то он устанавливается равным типу выражения (§5).

4.9.1. Инициализация полей

Каждое поле класса должно быть явно проинициализировано. Если инициализация задана через лексему '=', то значение поля не может быть изменено (*none с единственным присваиванием*), если же в инициализации используется лексема ':=', то значение поля может быть изменено (§6.4, §6.5), а поле называется изменяемым.

Если в инициализации задано *Выражение*, то начальным значение поля является значение выражения. Если при этом тип поля явно задан, то выражение должно быть *совместимо по присваиванию* с типом поля (§9.2).

Вместо выражения может быть указано ключевое слово позже, это *поздняя инициализация*. Такая форма инициализации разрешена только для полей, тип которых явно указан.

Значение поля с поздней инициализацией должно быть задано при создании экземпляра класса в конструкторе экземпляра класса (§5.8).

4.9.2. Экспорт полей

Если тип класса экспортируется, его поля могут быть помечены признаком экспорта, такие поля называется экспортированными полями. Поля, которые не экспортированы, доступны только в том модуле, в котором описан тип класса.

4.9.3. Наследование

Наследование позволяет определить новый (*расширенный*) класс на основе существующего (*базового*) класса.

Базовый класс E, указанный в описании класса E, называется E прямым базовым классом. Так как E может быть, в свою очередь, расширением другого базового класса, для каждого класса определен список базовых классов, возможно, пустой. Термин базовый класс будет использовать для обозначения любого класса из этого списка. Циклы в списке базовых классов запрещены.

Расширенный класс наследует поля и методы базового класса и может добавлять новые поля и методы. Обращение к полям и методам базового класса ничем не отличается от обращения к полям и методам самого класса. Методы базового класса можно переопределять, сохраняя те же самые типы параметра и тип результата.

4.10. Может быть тип

Как правило, современные языки программирования ограничивают работу со объектами ссылочных (reference) типов для того, чтобы сделать явными все места в программе, в которых может возникнуть ошибка использования нулевой ссылки (null pointer exception). Так как русская терминология в этой области не устоялась, приходится обращаться к англоязычным терминам.

Язык Тривиль следует уже выработанному в современных языках программированию подходу (но не синтаксису):

• Если в указании типа объекта использована нотация $m \delta T$, где T - это некоторый ссылочный тип, то значением этого объекта, кроме значений типа T, может быть специальное значение 'пусто'. Тип такого объекта называется может быть T.

• Иначе, если ключевое слово мб отсутствует в указании типа объекта, значением этого объекта может быть только значений типа T.

Ссылочными типами являются Строка, типы вектора и типы класса, к остальным типам мб не может применятся.

Для объекта типа мб Т определены следующие действия:

- Присваивание объекту значения 'пусто' или значения типа Т (§9.2)
- Сравнение на равно/не равно с 'пусто' или с другим объектом типа $m \delta T$ (§5.11.3)
- Операция подтверждения типа ' ^ ' (§5.9), позволяющая перейти от значения типа $m\delta$ T к значению типа T

Пример указания типа и использования объекта:

```
1 пусть кличка: мб Строка := пусто
2 ...
3 если кличка # пусто { вывод.ф("%v\n", кличка^}
```

4.11. Описание функций

Описание функции состоит из идентификатора, сигнатуры и тела функции. Сигнатура определяет формальные параметры и тип результата (если таковой имеется). Особым видом функции является метод, см. §4.12.

```
Описание-функции: 'фн' Идент-оп Сигнатура Тело
Сигнатура: '(' Список-параметров? ')' Тип-результата?
Список-параметров: Параметр (',' Параметр)* ','?
Параметр: Идентификатор ':' '...'? ('*' | Указ-типа)
Тело: (Блок | Модификатор)?
```

Вместо тела функции может стоять модификтор (§3.10) @внеш, см. §4.11.3. Если у функции задан тип результата, тело функции должно завершаться операторами вернуть (§6.11) или авария (§6.12).

```
фн Факториал(ц: Цел64): Цел64 {
    если ц <= 1 { вернуть 1 }
    вернуть ц * Факториал(ц - 1)
}
```

4.11.1. Параметры

Формальные параметры - это идентификаторы, которые обозначают *аргументы* (фактические параметры), указанные при вызове функции (§5.6).

```
Параметр: Идентификатор ':' '...'? ('*' | Указ-типа)
```

Последний параметр функции может иметь тип с префиксом . . . - это вариативный параметр. Функция с таким параметром может быть вызвана с нулем или более аргументов для этого параметра. Если тип вариативного парметра указан как T, то в теле функции типом параметра является [] T.

4.11.2. Полиморфные параметры

Если вместо типа параметра указан символ '*', то параметр называется полиморфным. Аргумент, соотвествующий этому параметру, может быть выражением любого типа. Для вариативного полиморфного параметра каждый аргумент может быть выражением любого типа. Для работы с полиморфным параметром определены специальные стандартные функции (§7.1.2).

```
фн печать по формату(формат: Строка, аргументы: ...*) {
    ...
}
```

4.11.3. Внешние функции

Функция, в которой вместо тела стоит модификатор @внеш, является внешней функцией, то есть реализованной каким-то внешним способом. Атрибут "имя"модификатора задает внешнее имя функции.

```
фн строка(с: Строка) @внеш("имя":"print_string")
```

Если атрибут "имя" не задан, внешнее имя совпадает с идентификатором функции.

4.12. Описание методов

Метод - это функция, связанная с типом классом. Для вызова метода ($\S 5.6$) должен быть указан экземпляр этого класса или расширенного класса. В описание метода класс, с которым связан метод указывается с помощью *Привязки*.

```
Описание-метода: 'фн' Привязка Идент-оп Сигнатура Блок
Привязка: '(' Идентификатор ':' Указ-типа ')'
```

Привязка определяет идентификатор и тип, который должен быть типом класса. В теле функции идентификатор привязки является параметром указанного типа.

```
1 тип K = класс {}
2 фн (к: K) метод() {}
```

Идентификатор метода должен быть уникальным в классе среди идентификатором полей и методов.

В расширенном классе может быть определен метод с таким же идентификатором как в одном из базовых классов.

переопределен, то есть опеределен другой метод . При этом сигнатура переопределеного метода должна совпадать с сигнатуром переопределяемого метода, а именно:

• Число параметров должно совпадать

- Типы параметров должны быть эквивалентны §9.1
- Признаки вариативности и полиморфности должны совпадать

Пример ошибки при переопределении метода (разное число параметров):

```
1 тип K1 = класс (K) {}
2
3 фн (к: K1) метод(с: Строка) {}
```

5. Выражения

Выражение определяет вычисление значения путем применения операций к операндам.

```
Выражение
: Унарное-выражение
| Выражение Бинарная-операция Выражение
Унарное-выражение
: Первичное-выражение
| Унарная-операция Унарное-выражение
```

5.1. Первичные выражения

Первичные выражения - это выражения, которые являются операндами унарных ($\S 5.10$) и бинарных ($\S 5.11$) операций .

```
Первичное-выражение
: Операнд
(Доступ
| Индексация
| Преобразование
| Вызов
| Конструктор-вектора
| Конструктор-класса
| Подтверждение-типа
) *
```

5.2. Операнды

Операнды - это элементарные значения в выражении. Операнд может быть литералом, именем объекта, обозначающим константу, переменную или функцию, или выражением в скобках.

```
Операнд: Литерал | Имя-объекта | '(' Выражение ')'
Литерал
```

```
: Целочисленный-литерал
| Вещественный-литерал
| Строковый-литерал
| Символьный-литерал

Имя-объекта: Идентификатор | Импорт-идент
Импорт-идент: Идентификатор '.' Идентификатор
```

Если в качестве имени объекта используется *Импортичдент*, то первый идентификатор в нем должен обозначать имя импортированного модуля (§8.2), а второй идентификатор должен быть идентификатором экспортированного из этого модуля объекта.

Операнд	Тип выражения
Имя объекта	тип объекта
Целочисленный-литерал	Цел64
Вещественный-литерал	Вещ64
Строковый-литерал	Строка
Символьный-литерал	Символ
Выражение в скобках	тип выражения

5.3. Доступ к полям и методам

Доступ применяется к первичному выражению.

```
Доступ: '.' Идентификатор
```

В выражении что-то.имя:

- что-то должно быть выражением типа T, где T это тип класса.
- а имя должно быть идентификатором поля или метода типа T или базового класса T

Типом выражения доступа является тип поля или метода.

```
1 чел.возраст := 5
2 если чел.возраст < 18 { ... }
```

5.4. Индексация

Индексация применяется к первичному выражению.

```
Индексация: '[' Выражение '['
```

Выражении что-то [индекс] обозначает элемент вектора или вариативного параметра, индекс которого определяется выражением индекс.

• что-то должно быть идентификатором, обозначающим вариативный параметр

- или что-то должно быть выражением типа T, где T это тип вектора
- индекс должен быть выражением типа Цел 64 или Байт

Если во время исполнения индекс выходит за границы вектора или параметра (индекс $< 0 \mid$ индекс >= длина), происходит *авария* (§6.12).

Типом выражения индексации является тип элемента вектора, если индексируется вектор или тип вариативного параметра, если индексируется вариативный параметр.

Так как индексация и конструктор вектора ($\S 5.7$) не всегда различимы синтаксически, выбор между ними происходит на уровне семантики, если выражение $\protect\operatorname{uto-to}$ обозначает тип, то это конструктор вектора, иначе индексация.

5.5. Преобразование типа

Преобразование типа применяется к первичному выражению, и преобразует, если это возможно, значение выражения к *целевому типу*, указанному в операции преобразования.

```
Преобразование: '(:' 'осторожно'? Указ-типа ')'
```

Если в преобразование есть ключевое слово осторожно, то это небезопасная операция, которая рассматривается отдельно (§10.1).

Часть преобразований, например, преобразования между числовыми типами и символами, а так же строковые преобразования могут изменить представление значения первичного выражения и повлечь за собой затраты на выполнение, а так же *аварийное завершение*, если выполнить преобразование невозможно.

Преобразование классов изменяют только тип, но не представление значения.

Разрешенные преобразования:

Целевой тип	Тип выражения	Условие выполнения
Байт	Цел64, Слово64, Символ,	значение в диапазоне
	Строковый литерал длины 1	0255
	Байт, Символ, Строковый ли-	всегда
Цел64	терал длины 1	
	Слово64	значение в диапазоне
		0Мах(Цел64)
	Вещ64	всегда
Слово64	Байт, Символ, Строковый ли-	всегда
Словооч	терал длины 1	
	Цел64	не отрицательное значение
Вещ64	Цел64	всегда
Символ	Байт, Строковый литерал дли-	всегда
Символ	ны 1	
	Цел64, Слово64	значение в диапазоне
		Unicode символа
Строка	Символ, [] Символ, [] Байт	всегда
[] Байт	Строка, Символ	всегда
[] Символ	Строка	всегда
Класс	Класс	От базового класса к рас-
		ширенному

```
тип Байты = []Байт
пусть байты = "Привет"(:Байты)
пусть три = pi(:Цел64)
```

5.6. Вызов функции или метода

Вызов применяется к первичному выражению. Это выражение

- является идентификатором стандартной функции (§7.1)
- или обозначает функцию, тогда это вызов функции
- или явлется выражением типа класс, тогда это вызов метода

```
Вызов: '(' Список-аргументов? ')'
Список-аргументов: Аргумент (',' Аргумент)* ','?
Аргумент: Выражение '...'?
```

В вызове указывается список аргументов. Число аргументов в списке должно быть равно числу параметров в сигнатуре функции или метода, за исключение случая, когда в сигнатуре есть вариативный параметр. Для вариативного параметра число аргументов может быть от нуля и более.

Каждый аргумент в вызове является *выражением*. Если для вариативного параметра задан единственный аргумент, то после выражения может стоять знак операции . . . , делающий *развернутое выражение* (§5.6.1).

Для всех других аргументов, выражение должно быть *совместимо по присва- иванию* с типом параметра (§9.2).

Типом выражения вызова является тип результата. Если в сигнатуре функции или метода тип результата отсутствует, то вызов не может быть операндом для любой операции.

```
1 фн Факториал(ц: Цел64): Цел64 { /*тело*/ }
2 пусть рез = Факториал(5)
```

5.6.1. Развернутое выражение

Развернутое выражение может быть использовано только как аргумент для вариативного параметра и только, если это единственный аргумент для этого параметра. Если развернутое выражение $x\dots$ используется как аргумент для параметра $\pi:\dots$ т, то должно выполняется одно из условий:

- x является вариативным параметром того же типа ${\mathbb T}$
- x это выражение типа вектора, причем тип вектора определен, как [] ${\mathbb T}$

Пример использования развернутого вектора в качестве аргумента вариативного параметра:

```
фн Соединить (список: ...Строка) {/*тело*/}

тип Строки = []Строка

фн Сохранить (строки: Строки) {
 пусть текст = Соединить (строки...)
 /* сохранение текста */

}
```

5.7. Конструктор вектора

Конструктор вектора применяется к первичному выражению, которое должно обозначать тип вектора. Результатом выполнения конструктора является вектор.

```
Конструктор-вектора: '[' Список-элементов? ']'
Список-элементов: Элемент (',' Элемент)* ','?
Элемент: (Индекс-элемента ':')? Значение
Индекс-элемента: Выражение
Значение: Выражение
```

Пример использования конструктора вектора в вызове:

```
тип Строки = []Строка
фн Сохранить(строки: Строки) { /*тело*/ }

Сохранить(Строки["привет", "мир"])
```

Для всех элементов должно выполняться:

- Индекс-элемента должен быть выражением типа Цел 64 или Байт
- Значение должно быть выражением, совместимым по присваиванию с типом элемента вектора (§9.2).

Если хотя бы для одного элемента конструктора задан *Индекс-элемента*, то он должен быть задан для всех элементов конструктора.

Если *Индекс-элемента* задан, то в результирующем векторе в позиции *Индекс-элемента* будет стоять значение *Значение*. Если *Индекс-элемента* не задан, то элементы будет размещены в результирующем векторе последовательно, начиная с индекса 0.

Длина результирующего вектора равна максимальному индексу + 1.

TBD: Если конструктор вектора задает значения не для всех индексов, то тип элемента вектора должен быть типом, для которое есть значение по умолчанию. Или должно быть явное значение по умолчанию *: 1

5.8. Конструктор экземпляра класса

Конструктор вектора применяется к первичному выражению, которое должно обозначать тип класса. Результатом выполнения конструктора является экземпляр класса.

```
Конструктор-класса: '{' Список-пар? '}'
Список-пар: Пара (',' Пара)* ','?
Пара: (Идентификатор ':')? Значение
```

Для каждой *Пары* идентификатор должен обозначать поле конструируемого класса. В результирующем экземпляре значением этого поля будет значение выражения *Значение*. *Значение* должно быть выражением, *совместимым по присва-иванию* с типом поля (§9.2).

Указание нескольких значений для одного идентификатора поля является ошибкой.

Конструктор должен задавать значения для каждого поля класса с *поздней инициализацией* (§4.9.1).

Если класс импортирован из другого модуля, в конструкторе не могут использоваться идентификаторы не экспортированных полей. Если при этом у не экспортированного поля задана *поздняя инициализация*, конструктор такого класса может быть использован только в модуле, содержащем описание класса.

Пример ошибки - значение поля "имя"с поздней инициализацией не задано:

```
1 пусть Некто := Человек{возраст: 25}
```

5.9. Подтверждение типа

Подтверждение типа применяется к первичному выражению, которое должно быть может быть типа ($\S4.10$).

```
Подтверждение-типа: '^'
```

Если значение первичного выражение, тип которого есть мб T равно 'пусто', то происходит *авария* (§6.12), иначе тип выражения становится равным T.

```
1 пусть кличка: мб Строка = пусто
2 ...
3 кличка = "Мурка"
4 ...
5 напечатать (кличка^)
```

5.10. Унарные операции

```
Унарная-операция: '-' | '~'
```

Операция	Тип Операнда	Тип результата
	Цел64	Цел64
-	Цел64 Слово64	Цел64 Слово64
	Вещ64	Вещ64
~ (not)	Лог	Лог

Унарные операторы имеют приоритет выше, чем бинарные, то есть выражение -x + 1 выполняется как (-x) + 1.

5.11. Бинарные операции

```
Бинарная-операция

: Арифметические
| Сравнения
| Логические
| Битовые
Арифметические: '+' | '-' | '*' | '/' | '%'
Сравнения: '=' | '#' | '<' | '<=' | '>' | '>='
Логические: '|' | '&'
Битовые: ТВD
```

ТВО: Определить множество и знаки битовых операций.

5.11.1. Приоритеты операции

Язык определяется пять уровней приоритета для бинарных операций. Операции умножения имеют самый высокий приоритет, а операция логическое ИЛИ (\mid) имеет самый низкий приоритет.

Приоритет	Операция
5	* / %
4	+ -
3	= # < <= > >=
2	&
1	

Бинарные операции с одинаковым приоритетом ассоциируются слева направо. Например, х / у * z эквивалентно (х / у) * z.

5.11.2. Арифметические операции

Арифметические операторы применяются к числовым операндам, выдают результат того же типа, что и тип операндов. Типы левого и правого операнда должны совпадать.

Операция	Действие	Разрешенные типы
+	сумма	Байт, Цел64, Слово64, Вещ64
_	разница	Байт, Цел64, Слово64, Вещ64
*	произведение	Байт, Цел64, Слово64, Вещ64
/	деление	Байт, Цел64, Слово64, Вещ64
%	остаток от деления	Байт, Цел64, Слово64

5.11.3. Операции сравнения

Операции сравнения выдают результат типа Лог.

Все операции сравнения '=', '#', '<', '<=', '>', '>=' применимы к операндам типов Байт, Цел64, Слово64, Цел64 и Символ. Типы левого и правого операнда, при этом, должны совпадать.

Операции '=', '#' также применимы в следующих случаях:

Тип левого операнда	Правый операнд
Строка	операнд типа Строка
класс T	операнд типа T
мб T	'пусто' или операнд типа мб T

5.11.4. Логические операции

Операции логическое ИЛИ (|) и логическое U (&) применимы к операндам типа Лог и выдают результат типа Лог.

```
а \mid б означает если а, то истина, иначе б а \& б означает если а, то б, иначе ложь
```

5.11.5. Битовые операции

ТВО: будут определены позже.

5.12. Константные выражения

Константные выражения - это выражения, которые могут быть вычислены во время компиляции.

Минимальные требования к компилятору языка Тривиль: компилятор должен вычислять выражение, состоящее из одного операнда, который является литералом или идентификатором, обозначающим константу.

6. Операторы

Операторы задают действия.

Оператор

- : Локальное-описание
- | Простой-оператор
- | Оператор-если
- | Оператор-надо
- | Оператор-когда
- | Оператор-пока
- | Оператор-вернуть
- | Оператор-прервать
- | Оператор-авария

Простой-оператор

- : Оператор-выражение
- | Оператор-присваивания
- | Инкремент
- | Декремент

6.1. Блоки

Операторы сгрупированы в *Блоки*, которые задают последовательность действий для функций и методов ($\S4.11$, $\S4.12$) и входа в модуль ($\S8.4$). Блок - это, возможно, пустая последовательность операторов, которые могут включать локальные описания.

```
Блок: '{' Список-операторов? '}'
Список-операторов: Оператор (Разделитель Оператор)*
```

6.2. Локальные описания

Локальное описание определяет идентификатор, для которого областью действия (§4.1) является часть блока, от точки завершения описания до завершающей скобки блока, исключая вложенные блоки, в которых описан такой же идентификатор.

```
Локальное-описание: Описание-переменной
```

Для локальных переменных не может быть задана *поздняя инициализация* ($\S4.6$). ТВD: Нужны ли локальные константы? Или это задача оптимизации?

6.3. Выражение, как оператор

Выражение, являющееся вызовом функции или метода, кроме вызова стандартных функций ($\S7.1$), может быть использовано как оператор. Результат вызова при этом игнорируется.

Оператор-выражение: Выражение

```
1 напечатать ("Привет")
2 Факториал (5)
```

Не могут быть использованы в качестве оператора:

6.4. Оператор присваивания

Присваивание заменяет текущее значение, хранящееся в переменной, новым значением, заданным выражением.

```
Оператор-присваивания: Выражение ':=' Выражение
```

Выражение в левой части должно быть изменяемым:

Левое выражени	e	Условие
Имя объекта	пер или мод.пер	пер - это изменяемая перемен-
		ная (§4.6)
Доступ к полю	что-то.поле	что-то - это изменяемое выра-
		жение, поле - это изменяемое
		поле (§4.9.1)
Индексация	что-то[индекс]	что-то - это изменяемое выра-
		жение
Преобразование	что-то(:Тип)	что-то - это изменяемое выра-
		жение

Выражение в правой части должно быть coвместимо по присваиванию с типом левого выражение (§9.2).

```
1 пусть ц := 0
2 пусть байты := Байты[1, 2, 3]
3 ц := 2
4 байты[ц] = 5
```

Ошибка присваивания в неизменяемое выражение:

```
1 пусть ц = 0
2 ц := 2
```

6.5. Инкремент и декремент

Операторы *Инкремент* ('++') и *Декремент* ('--') увеличивают или уменьшают свои операнды на 1.

```
Инкремент: Выражение '++'
Декремент: Выражение '--'
```

Как и в случае присваивания (§6.4), выражение должно быть *изменяемым*. Тип выражения должен быть одного из типов: Байт, Цел64, Слово64.

```
1 пусть ц := 1
2 пусть байты := Байты[1, 2, 3]
3 байты[ц]++
```

6.6. Оператор 'если'

Оператор если определяет условное выполнение двух ветвей в соответствии со значением логического выражения. Если выражение принимает значение истина, выполняется Блок первой веткви, в противном случае, выполняется ветвь иначе, если она задана.

```
Оператор-если:
'если' Выражение Блок ('иначе' (Оператор-если | Блок))?
```

```
1 если ц > макс { ц := макс }
2 если Буква?(сим) { Имя() }
4 иначе если Цифра?(сим) { Число() }
5 иначе { Ошибка!() }
```

6.7. Оператор 'надо'

Оператор надо используется для завершения исполнения операторов некоторого контекста, если условие, заданое логическим выражением, не выполнено.

```
Оператор-надо:
    'надо' Выражение 'иначе' (Завершающий-оператор | Блок)
Завершающий-оператор
    : Оператор-вернуть
    | Оператор-прервать
    | Оператор-авария
```

Если в ветви иначе стоит $\mathit{Eлок}$, то последним оператором блока должен быть завершающий оператор.

Завершающий оператор	Действие
вернуть (§6.11)	выход из тела функции, метода или входа
прервать (§6.10)	выход из ближайшего объемлющего цикла
авария (§6.12)	аварийное завершение программы

```
1 надо делитель # 0 иначе авария("деление на ноль")
2 надо число > 1 иначе вернуть 1
```

6.8. Оператор 'когда' или оператор выбора

Оператор когда выбирает одну ветвь выполнения из нескольких вариантов, в соответствии со значением выражения.

ТВД: Добавить предикатный и оператор выбора по типу?

При выполнении оператора когда вычисляется *Выражение*, затем выражения в *Вариантах* вычисляются слева направо и сверху вниз. Как только значение выражения *Варианта* стало равным значению первого выражения, выполняется список операторов этого *Варианта*. Остальные варинты игнорируются. Если ни один из вариантов не выполнен, и есть ветка иначе, выполняется список операторов этой ветки.

Тип каждого выражения в вариантах должен быть равен типу первого выражения (§9.1).

```
когда x {
2 есть 0: вернуть "ничего"
3 есть 1: вернуть "один"
4 есть 2: вернуть "два"
5 иначе вернуть "много"
6 }
```

6.9. Оператор цикла 'пока'

Оператор пока определяет повторное выполнение блока до тех пор, пока значение логического выражение равно истина. Выражение вычисляется перед каждой итерацией.

```
Оператор-пока: 'пока' Выражение Блок
```

```
1 пока a < б {
2 a := a * 2
3 }
```

6.10. Оператор 'прервать'

Оператор прервать завершает выполнение самого внутреннего оператора цикла в рамках одной и той же функции.

```
Оператор-прервать: 'прервать'
```

```
1 пока истина {
2 a := a * 2
3 если a > б { прервать }
4 }
```

6.11. Оператор 'вернуть'

Оператор вернуть завершает выполнение тела функции, метода или входа, и, возможно, возращает значение.

```
Оператор-вернуть: 'вернуть' (Выражение | Разделитель)
```

Если оператор вернуть используется в теле входа или в теле функции или метода, в сигнатуре которых не указан тип результата, то выражение в операторе должно отсутствовать.

И наоборот, если в оператор используется в теле функции или метода, в сигнатуре которых указан тип результата, выражение должно быть и оно должно быть совместимо по присваиванию с типом результата (§9.2).

```
1 фн Факториал(ц: Цел64): Цел64 {
2 если ц <= 1 { вернуть 1 }
3 вернуть ц * Факториал(ц - 1)
4 }
```

6.12. Оператор 'авария' и аварийное завершение

Оператор авария запускает аварийную ситуация, которая, обычно, приводит к аварийному завершению программы.

```
Оператор-вернуть: 'авария' '(' Выражение ')'
```

Тип выражения должен быть Строка. Значение этого выражения используется в сообщении об аварийной ситуации.

Кроме явного запуска аварийной ситуации, аварийная ситуация может быть запущена неявно, в следующих случаях:

- Выход индекса за пределы вектора при индексации
- Операция подтверждения типа выполнена над объектом со значением 'пусто'
- Недопустимое преобразования типа
- Недопустимая последовательность байт в кодировке UTF-8
- Невозможность выделения памяти для динамического объекта
- Нереализованная возможность в системе поддержки выполнения

Язык Тривиль не содержит средств перехвата аварийной ситуации или восстановления после него. Такие средства могут быть добавлены на уровне библиотек.

7. Стандартные функции и методы

7.1. Стандартные функции

7.1.1. Длина

Стандартная функция длина возвращает длину строк, векторов и вариативных параметров. Тип результата: Цел64.

Длина строки считается в символах. Для остальных типов выдается число элементов.

```
1 пусть д1 = длина("Привет") // равно 6
2
3 пусть д2 = длина("Привет"(:Байты)) // равно 12
```

7.1.2. Функции для полиморфных параметров

Стандартные функции тег и нечто предназначены для работы с полиморфными параметрами (§4.11.2)

Аргумент функция тег может быть типом или полиморфным параметром или элементом вариативного полиморфного параметра. Для каждого типа определен его *Тег*, который является постоянным объектом времени исполнения. Постоянный, в данном контексте, контексте, означает, что Тег типа не меняется во время исполнения программы. Тег представлен значением типа Слово64.

Вид аргумента		Результат
предопределенный тип	тег(Цел64)	Тег указанного типа
описанный тип	тег(Байты)	Тег указанного типа
полиморфный параметр	тег(п)	Тег параметра
элемент вариативного по-	тег(п[номер])	Тег элемента
лиморфного параметра		

Аргумент функция нечто должен быть полиморфным параметром или элементом вариативного полиморфного параметра. Функция выдает результат типа Слово64, которые является скрытым представлением значения переданного в полиморфный параметр.

К результату функции нечто может применяется преобразование типа (§5.5) или *осторожное* преобразование типа (§10.1).

```
1 фн Что это?(п: *) {
2 если тег(п) = тег(Цел64) {
3 вывод.ф("это число: %v", нечто(п)(:Цел64))
4 }
5 иначе {
6 вывод.ф("это не число")
7 }
8 }
```

7.2. Встроенные методы для векторов

Для всех типов векторов определены стандартные методы:

- добавить
- заполнить

ТВО: Уточнить семантику и набор методов.

7.2.1. Метод добавить

Метод добавить можно применить к любому вектору. Если вектор типа $[\]$ T, то метод можно рассматривать, как описанный с заголовком:

```
фн (в: []Т) добавить(аргументы: ...Т)
```

Это означает, что в вызове метода можно указать любое количество аргументов типа Т или одно *развернутое выражение* (§5.6).

Вызов метода добавляет к вектору указанные аргументы.

```
тип Числа = []Цел64

пусть фиб = Числа[1, 1]
фиб.добавить(2, 3, 5)
```

7.2.2. Метод заполнить

Метод заполнить можно применить к любому вектору. Если вектор типа $[\]$ T, то метод можно рассматривать, как описанный с заголовком:

```
фн (в: []Т) заполнить(кол-во: Цел64, значение: Т)
```

Вызов метода добавляет к вектору кол-во элементов с указанным значением. Пример создания вектора из 10 пустых строк:

```
тип Имена = []Строка

пусть имена = Имена[]
имена.заполнить(10, "")
```

8. Модули

Программа на языке Тривиль состоит из модулей (единиц компиляции), исходный текст каждого модуля расположен в одном или нескольких исходных файлов.

Каждый исходный файл состоит из заголовка модуля, за которым следует, возможно, пустой список импорта, за которым следует, возможно, пустой набор описаний типов, констант, переменных, функций и методов и, возможно, вход в модуль.

```
Модуль: Исходный-файл+
Исходный-файл:
Заголовок-модуля
Список-импорта
Описание-или-вход*

Заголовок-модуля:
'модуль' Идентификатор Разделитель
('осторожно' Разделитель)?

Описание-или-вход: Описание | Вход
```

8.1. Заголовок модуля

Каждый исходный файл начинается с заголовка модуля. Заголовок содержит идентификтор, который определяет модуль, к которому принадлежит файл. Идентификатор не принадлежит никакой области действия и не связан ни с каким объектом.

Заголовок модуля может содержать признак осторожно, который означает, что в исходном файле можно использовать *небезопасные операции* (§10).

8.2. Импорт

Наличие импорта в исходном файле указывает, что этот исходный файл, зависит от функциональности импортируемого модуля, и обеспечивает доступ к экспортированным идентификаторам этого модуля.

Каждый импорт содержит путь импорта, указывающий на импортируемый модуль.

```
Список-импорта: Импорт*
Импорт: 'импорт' Путь-импорта Разделитель
Путь-импорта: Строковый-литерал
```

TBD: структура пути.

TBD: явное задание имени импорта.

Импорт добавляет описание идентификатора, текстуально равного последнему имени в пути импорта, который будет использоваться для доступа к экспортированным объектам импортиромого модуля.

```
1 модуль х

2 импорт "стд/вывод"

4 вход {

вывод.ф("Привет!")

7 }
```

8.3. Вход или инициализация модуля

Модуль может содержать действия, которые выполняются при инициализации модуля - *вход в модуль*.

```
1 Вход: 'вход' Блок
```

ТВО: требования к инициализации переменных с поздней инициализацией

Система выполняет должна обеспечивать следующие условия инициализации модуля M:

- инициализация модуля выполняется один раз
- инициализации M выполняется **после** инициализации всех модулей, которые импортирует M
- инициализации M выполняется до инициализации тех модулей, которые импортируют M

8.4. Инициализация и исполнения программы

Программа состоит из головного модуля и всех прямо или косвенно импортированных модулей.

Исполнение программы состоит из:

- Инициализации всех модулей, импортированных из головного модуля. Это приводит к рекурсивной инициализации всех используемых модулей программы. Для корректной инициализации граф импорта должен быть ациклическим.
- И, затем, выполнение входа головного модуля.

- 9. Правила совместимости
- 9.1. Эквивалентность типов
- 9.2. Совместимость по присваиванию

10. Применять с осторожностью

10.1. Осторожные преобразования