

Московский государственный университет им. М.В. Ломоносова
Факультет вычислительной математики и кибернетики



А.А. Вылиток

**Металингвистические формулы
и синтаксические диаграммы**

(Учебно-методическое пособие для студентов 1 курса)

Москва

2012

УДК 519.7 (075.8)

ББК 22.18я73

В92

*Печатается по решению
Редакционно-издательского совета факультета
вычислительной математики и кибернетики МГУ им. М.В. Ломоносова*

Рецензенты:

доцент кафедры системного
программирования ВМК МГУ
профессор кафедры алгоритмических
языков ВМК МГУ

Кузьменкова Е.А.

Соловьев С.Ю.

Вылиток А.А.

В92 Металингвистические формулы и синтаксические диаграммы:
Учебно-методическое пособие. - М.: Издательский отдел Факультета
ВМК МГУ им. М.В. Ломоносова (лицензия ИД № 05899 от 24.09.2001 г.);
МАКС Пресс, 2012. – 24с.

ISBN 978-5-89407-484-9

ISBN 978-5-317-04154-0

Учебно-методическое пособие «Металингвистические формулы и синтаксические диаграммы» предназначено для студентов 1 курса факультета ВМК МГУ им. М.В. Ломоносова. В пособии рассматриваются два способа описания синтаксиса языков: металингвистические формулы и синтаксические диаграммы. Предложен набор задач и примеры их решения.

УДК 519.7(075.8)

ББК 22.18я73

Учебно-методическое пособие

ВЫЛИТОК Алексей Александрович

МЕТАЛИНГВИСТИЧЕСКИЕ ФОРМУЛЫ И СИНТАКСИЧЕСКИЕ ДИАГРАММЫ

19992, ГСП-2, Москва, Ленинские Горы, МГУ им. М.В. Ломоносова, 2-й учебный корпус

ISBN 978-5-89407-484-9

ISBN 978-5-317-04154-0

© Вылиток А.А.

© Факультет вычислительной математики и
кибернетики МГУ им. М.В. Ломоносова, 2012

1. Языки и метаязыки

Обработка данных с помощью компьютеров основана на использовании искусственных языков, таких как языки программирования, управления заданиями, форматирования документов, протоколы связи и другие специальные языки.

Каждый язык имеет три составляющие: алфавит, синтаксис и семантику. *Алфавит* — это фиксированный набор символов. *Синтаксис* определяет, по каким правилам из символов образуются тексты. *Семантика* — набор правил истолкования, позволяющих по тексту понять его смысл.

Рассмотрим язык арифметических выражений. В его алфавит входят цифры, скобки, знаки арифметических операций $+$, $-$, $*$, $/$, означающие соответственно сложение, вычитание, умножение и деление. Опишем способ образования текстов на этом языке, вводя вспомогательные понятия для обозначения частей выражения. Простым выражением, или термом, будем называть последовательность цифр, т.е. число, а также любое выражение, заключенное в скобки. Слагаемое определим как последовательность термов, соединенных операциями умножения и деления. Наконец, выражение — это последовательность слагаемых, соединенных операциями сложения и вычитания. Перед первым слагаемым может стоять знак минус. Смысл выражения — это его значение, которое вычисляется по математическим правилам с учетом старшинства операций и скобок.

Язык арифметических выражений мы описали с помощью естественного языка. Язык, используемый для описания другого языка, называется *метаязыком*. В качестве метаязыка можно использовать естественный язык, однако он не вполне удобен для точного описания синтаксических конструкций искусственных языков, так как содержит двусмысленности, метафоры и т.п. Для описания синтаксиса был создан специальный метаязык — металингвистические формулы Бэкуса-Наура¹, или, сокращенно, язык БНФ.

¹ Джон Бэкус, Петер Наур – учёные в области информатики, авторы данной нотации.

2. Металингвистические формулы

Будем называть *цепочкой* последовательность из символов алфавита. *Сцеплением* двух цепочек будем называть результат приписывания второй цепочки в конец первой.

Описание на языке БНФ представляет собой набор формул. Формулы описывают способ формирования цепочек, являющихся синтаксически правильными текстами. В формулы входят *метапеременные* и символы алфавита. Символы алфавита соответствуют элементарным понятиям языка. Например, символ + в языке арифметических выражений означает операцию сложения. Метапеременные соответствуют более сложным понятиям языка. К примеру, понятие «выражение» языка арифметических выражений можно обозначить с помощью метапеременной *⟨выражение⟩*. Для записи метапеременных используются курсивный шрифт и угловые скобки.

Формула БНФ состоит из левой и правой частей, разделенных символом ::=, который обычно читается как «есть по определению». В левой части находится метапеременная, означающая определяемое понятие, в правой части – комбинации метапеременных, символов алфавита и некоторых специальных символов. Приведем формулу для понятия «число 5».

⟨число 5⟩ ::= 5

Теперь, используя сцепление символов 1, 2 и 5 в цепочку, определим более сложное понятие – «число 125».

⟨число 125⟩ ::= 125

Мы можем использовать в правой части формулы сцепление символов алфавита и метапеременных. Например, определить число 9125125 можно как сцепление символа 9 и двух чисел 125.

⟨число 9125125⟩ ::= 9 ⟨число 125⟩⟨число 125⟩

Сложное понятие может состоять из нескольких более простых понятий – альтернатив. В БНФ альтернативы разделяются вертикальной чертой. Опишем понятие «цифра».

$$\langle \text{цифра} \rangle ::= 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9$$

Теперь, с помощью сцепления цифр, можно определить число из двух знаков (цифр).

$$\langle \text{число из двух знаков} \rangle ::= \langle \text{цифра} \rangle \langle \text{цифра} \rangle$$

Данная формула выражает множество цепочек, состоящих из двух цифр: 00, 01, 02, ..., 99.

Если к числу из двух знаков приписать еще одну цифру, получим число из трех знаков.

$$\langle \text{число из трех знаков} \rangle ::= \langle \text{число из двух знаков} \rangle \langle \text{цифра} \rangle$$

Понятно, что число из n знаков можно описать через число из $n-1$ знаков, приписывая к нему еще одну цифру.

Зададимся теперь вопросом, как описать с помощью БНФ все n -знаковые числа для всех натуральных n . Мы не сможем выписать отдельную формулу для каждого n , так как множество натуральных чисел бесконечно, а число формул БНФ должно быть конечным.

Заметим, что если к некоторому числу приписать цифру, то получится новое число. Начать процесс приписывания можно с элементарных чисел, являющихся цифрами. Таким образом, получаем следующую БНФ для множества всех чисел.

$$\langle \text{число} \rangle ::= \langle \text{цифра} \rangle | \langle \text{число} \rangle \langle \text{цифра} \rangle$$

Здесь понятие числа определяется через само себя – такое определение называется *рекурсивным*. В рекурсивном определении должна быть хотя бы одна альтернатива, не содержащая определяемое понятие, иначе получится замкнутый круг. В БНФ для понятия *число* такой альтернативой является первая альтернатива – $\langle \text{цифра} \rangle$.

Убедимся, что все числа описываются данной БНФ. Доказательство проведем индукцией по количеству знаков в записи числа. Очевидно, что число из одного знака входит в множество цепочек, определяемых метапеременной $\langle \text{число} \rangle$. Пусть числа из n знаков ($n > 0$) входят в множество цепочек для понятия $\langle \text{число} \rangle$. Тогда сцепление метапеременных $\langle \text{число} \rangle \langle \text{цифра} \rangle$ определяет множество цепочек, являющихся числами из $n+1$ знака. Следовательно, все числа определяются метапеременной $\langle \text{число} \rangle$.

Существует способ определить понятие число без использования рекурсии. Для этого в БНФ используется *итерация*. Итерацией называется последовательность метапеременных и символов алфавита, заключенная в фигурные скобки. Фигурные скобки играют служебную роль и означают, что последовательность, заключенная в эти скобки, может повторяться несколько раз или отсутствовать. С использованием итерации БНФ для числа выглядит так:

$$\langle \text{число} \rangle ::= \langle \text{цифра} \rangle \{ \langle \text{цифра} \rangle \}$$

Дадим определение с помощью БНФ понятию числа в «единичной» системе счисления. В этой системе числа представляются в виде последовательности палочек (или единиц).

$$\langle \text{единичное число} \rangle ::= \{ 1 \}$$

Ноль представляется пустой последовательностью. Для обозначения пустой цепочки в БНФ будем использовать $\langle \text{пусто} \rangle$. Приписывание пустой цепочки слева или справа к произвольной цепочке x дает снова цепочку x . Другое определение единичного числа, с помощью рекурсии, выглядит так:

$$\langle \text{единичное число} \rangle ::= \langle \text{единичное число} \rangle 1 \mid \langle \text{пусто} \rangle$$

Кроме фигурных скобок особую роль в БНФ играют прямоугольные скобки: последовательность, заключенная в них может либо отсутствовать, либо присутствовать ровно один раз. Опишем понятие целое число. Перед целым числом может стоять знак минус.

$\langle \text{целое число} \rangle ::= [-] \langle \text{число} \rangle \{ \langle \text{число} \rangle \}$

Опишем с помощью БНФ более сложное понятие – выражение.

$\langle \text{выражение} \rangle ::= [-] \langle \text{слагаемое} \rangle \{ \langle \text{знак} + \text{ или } - \rangle \langle \text{слагаемое} \rangle \}$

$\langle \text{знак} + \text{ или } - \rangle ::= + \mid -$

$\langle \text{слагаемое} \rangle ::= \langle \text{терм} \rangle \{ \langle \text{знак} * \text{ или } / \rangle \langle \text{терм} \rangle \}$

$\langle \text{знак} * \text{ или } / \rangle ::= * \mid /$

$\langle \text{терм} \rangle ::= \langle \text{число} \rangle \mid (\langle \text{выражение} \rangle)$

В данном определении есть косвенная рекурсия: выражение определяется через слагаемое, слагаемое – через терм, терм – через выражение. Полностью избавиться от рекурсии при описании выражения нельзя. Доказательство существования языков, в описании которых нельзя обойтись без рекурсии, можно найти в [1].

3. Синтаксические диаграммы

Синтаксическая диаграмма (СД) представляет собой графический способ описания языка. Элементами диаграммы являются прямоугольники и овалы с надписями, соединенные стрелками. В прямоугольниках записываются названия метaperеменных, в овалах – основные символы. В диаграмме есть одна входная стрелка (не исходящая из какого-либо элемента) и одна выходная (не входящая в какой-либо элемент). Входная и выходная стрелки означают соответственно начало и конец синтаксического определения.

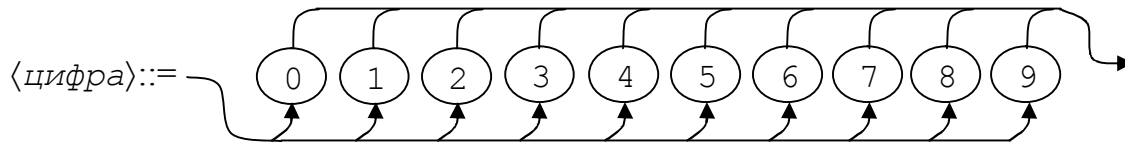
Приведем диаграмму для понятия *число 5*. Само определяемое понятие будем, как и в случае БНФ, заключать в угловые скобки и отделять от синтаксического определения знаком ::= .

$\langle \text{число } 5 \rangle ::= \rightarrow \textcircled{5} \rightarrow$

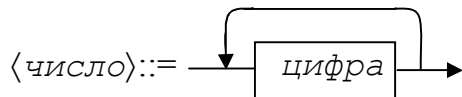
Диаграмма для числа 125 будет содержать три овала, по одному на каждую цифру.

$\langle \text{число } 125 \rangle ::= \rightarrow \textcircled{1} \rightarrow \textcircled{2} \rightarrow \textcircled{5} \rightarrow$

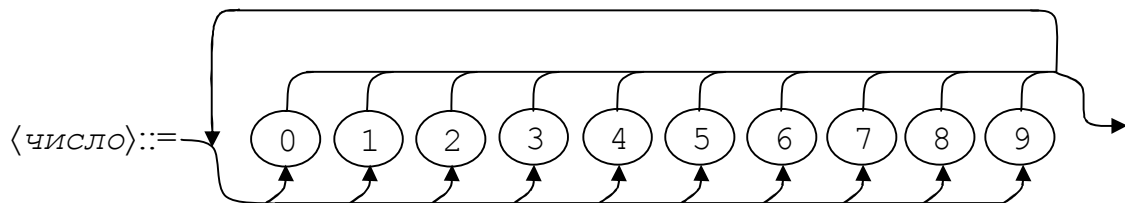
Покажем на примере понятия цифра, как с помощью диаграммы изображаются альтернативы.



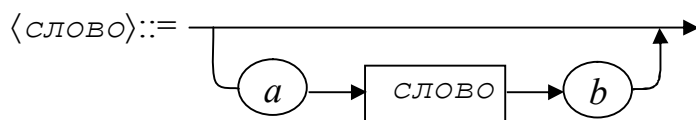
Опишем теперь понятие «число» с помощью СД. В ней будет использовано вспомогательное понятие «цифра», изображенное в виде прямоугольника.



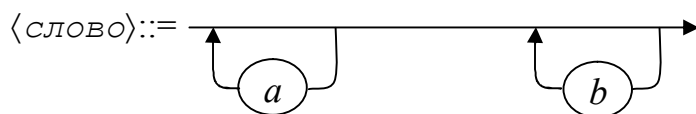
Подставим вместо прямоугольника соответствующую диаграмму и получим СД для понятия «число», не содержащую вспомогательных понятий (метапеременных).



Существуют языки, которые невозможно описать, не используя в диаграмме вспомогательных понятий. Рассмотрим, к примеру, язык, цепочки которого состоят из букв a и b , причем сначала в цепочке следуют n букв a ($n \geq 0$), а за ними столько же букв b : ab , $aabb$, $aaabbb$ и т.д. Пустая цепочка тоже принадлежит данному языку (случай $n=0$). Определим понятие «слово» для данного языка с помощью СД.

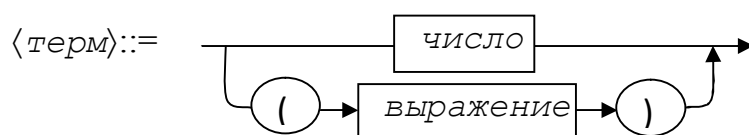
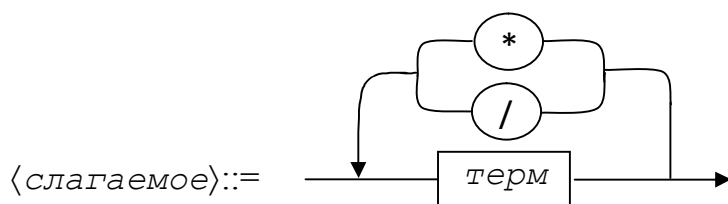
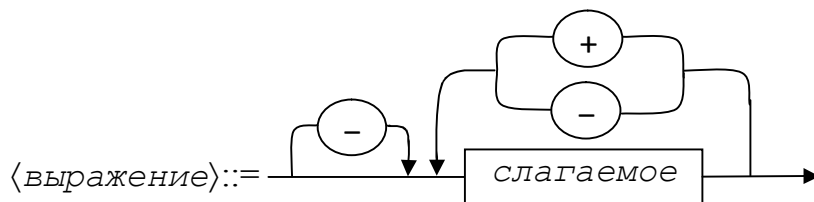


Попытаемся устранить рекурсию – построим СД, в которой нет вспомогательного понятия (метапеременной) «слово»:



Однако полученная диаграмма описывает другой язык – в его словах количество букв a может отличаться от количества букв b .

Опишем с помощью СД понятие «выражение».



Как металингвистические формулы, так и синтаксические диаграммы используются для описания синтаксиса языков программирования (см., например, [2]). Синтаксические диаграммы обеспечивают большую наглядность. Преимущество БНФ в том, что они могут использоваться в системах автоматизированной обработки языков.

4. Упражнения

В данном разделе приводятся упражнения для развития навыков работы с БНФ и СД, а также задачи, связывающие описание языков с теорией алгоритмов [3]. К некоторым задачам даются образцы решений. В условиях отдельных задач определяются вспомогательные понятия, которые могут использоваться и в последующих задачах.

1. Перечислить все цепочки, которые удовлетворяют понятию «слово», описанному с помощью БНФ:

$\langle \text{слово} \rangle ::= \langle \text{корень} \rangle \langle \text{суффикс} \rangle$

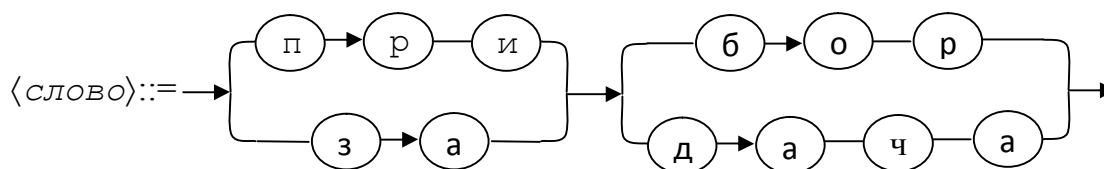
$\langle \text{корень} \rangle ::= \text{сад} \mid \text{сур}$

$\langle \text{суффикс} \rangle ::= \text{ик} \mid \text{ок}$

Решение

Искомые цепочки получаются сцеплением двух цепочек: первая должна удовлетворять понятию *корень*, вторая – понятию *суффикс*. Понятию *корень* удовлетворяет множество цепочек {сад, сур}, а понятию *суффикс* удовлетворяет множество {ик, ок}. Всевозможные сцепления цепочек из первого множества с цепочками из второго множества дают такие слова: садик, садок, сурик, сурок.

2. Перечислить все цепочки, которые удовлетворяют понятию «слово», заданному СД:



3. С помощью БНФ описано понятие «поезд». Буква П означает паровоз, буква В – вагон.

$\langle \text{поезд} \rangle ::= \langle \text{тяга} \rangle \langle \text{состав} \rangle$

$\langle \text{тяга} \rangle ::= П$

$\langle \text{состав} \rangle ::= В \langle \text{состав} \rangle \mid \langle \text{пусто} \rangle$

Удовлетворяет ли цепочка ПВВ понятию поезд ?

Решение

Данная цепочка является сцеплением двух цепочек П и ВВ. Первая непосредственно удовлетворяет понятию *тяга*. Покажем, что ВВ удовлетворяет понятию *состав*, тогда ПВВ удовлетворяет

понятию *поезд*. Сначала покажем, что цепочка В удовлетворяет понятию *состав*. Действительно, так как пустая цепочка непосредственно удовлетворяет понятию *состав* (согласно второй альтернативе в формуле), то сцепление В с пустой цепочкой, т.е. с *составом*, также удовлетворяет понятию *состав* (согласно первой альтернативе в формуле). Далее, в цепочке ВВ первая буква В соответствует вхождению в первую альтернативу, а вторая буква В, по доказанному, удовлетворяет понятию *состав*. Следовательно, сцепление В с *составом* (т.е. с еще одной В) дает *состав*. Итак, П – это *тяга*, ВВ – *состав*, а ПВВ – *поезд*.

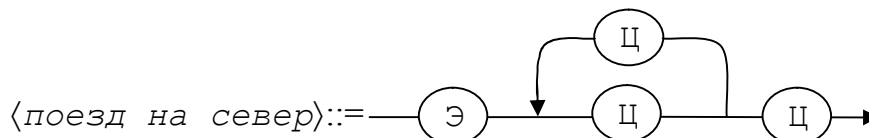
4. Дана БНФ для понятия «поезд, следующий на юг»:

$\langle \text{поезд на юг} \rangle ::= \text{ПВ}\{ВВ\}$

Какие из перечисленных ниже поездов следуют на юг?

(а) ПВ ; (б) ПВВ ; (в) ПВВВВВ ; (г) ППВВВ.

5. Дана СД для понятия «поезд, следующий на север». Буква Ц означает цистерну, Э означает электровоз.



Какие из перечисленных ниже поездов следуют на север?

(а) ЭЦ ; (б) ЭЦЦ ; (в) ЭЦЦЦЦ ; (г) ЭЦЦЦ.

6. Обезьяний язык с помощью БНФ описывается так:

$\langle \text{фраза} \rangle ::= \langle \text{слово} \rangle @ \langle \text{фраза} \rangle | \langle \text{слово} \rangle$

$\langle \text{слово} \rangle ::= \langle \text{слог} \rangle \langle \text{слог} \rangle | \langle \text{слово} \rangle \langle \text{слог} \rangle$

$\langle \text{слог} \rangle ::= \langle \text{слог} \rangle ba | \langle \text{слог} \rangle bb \langle \text{слог} \rangle | a$

Кто из перечисленных ниже ораторов является шпионом, замаскированным под обезьяну?

Бабуин: *ababbaa@abaabba@aa*

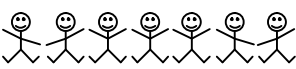
Шимпанзе: *ababa@abba@abbaaaa*

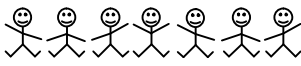
Горилла: *abaa@abbaaa@aabbaa*

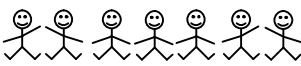
7. Язык «пляшущих человечков» задается БНФ:

$$\langle \text{предложение} \rangle ::= \text{чел} \langle \text{фраза} \rangle \text{чел} \mid \text{чел}$$
$$\langle \text{фраза} \rangle ::= \text{чел} \langle \text{предложение} \rangle \text{чел} \mid \text{чел} \langle \text{фраза} \rangle \text{чел}$$

Какие из приведенных ниже писем на языке «пляшущих человечков» фальшивые?

а) 

б) 

в) 

8. Описать с помощью БНФ и СД понятие «электричка». Электричка состоит из простых (В) и тяговых (Т) вагонов. Простые и тяговые вагоны чередуются. Первый и последний вагоны – тяговые.

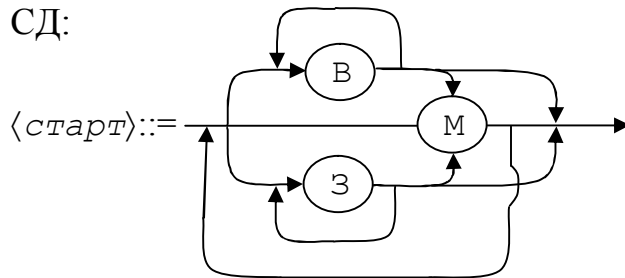
9. В лесу планируется забег зверей на длинную дистанцию. В забеге участвуют зайцы (З), волки (В) и медведи (М). По объективным причинам заяц не может бежать рядом с волком. Помогите организаторам соревнований расположить зверей на стартовой линии. Построить БНФ и СД допустимых расположений.

Решение

БНФ:

$$\langle \text{старт} \rangle ::= \text{М} \langle \text{старт} \rangle \mid \text{З} \langle \text{не волк} \rangle \mid \text{В} \langle \text{не заяц} \rangle \mid \text{М} \mid \text{З} \mid \text{В}$$
$$\langle \text{не волк} \rangle ::= \text{М} \langle \text{старт} \rangle \mid \text{З} \langle \text{не волк} \rangle \mid \text{М} \mid \text{З}$$
$$\langle \text{не заяц} \rangle ::= \text{М} \langle \text{старт} \rangle \mid \text{В} \langle \text{не заяц} \rangle \mid \text{М} \mid \text{В}$$

СД:



10. Описать с помощью БНФ с минимальным числом альтернатив язык, состоящий из цепочек длины 5 в алфавите $\{a, b\}$.

11. Описать с помощью СД с минимальным числом дуг язык, состоящий из цепочек длины 5 в алфавите $\{a, b\}$.

12. Описать с помощью БНФ и СД понятие «товарный поезд». Поезд состоит из тяги и состава. Тяга содержит от одного до трех электровозов (Э). В состав входят крытые вагоны (К) и платформы (П). Между двумя соседними крытыми вагонами находятся как минимум две платформы.

13. Описать с помощью БНФ и СД понятие «грузовой поезд». Поезд состоит из тяги и состава. Тяга содержит ненулевое четное число тепловозов (Т). В состав входят думпкары (Д) и хопперы (Х). В поезде может встречаться не более трех думпкаров подряд.

14. Описать с помощью БНФ и СД понятие «пассажирский поезд». Поезд состоит из тяги и состава. Тяга содержит нечетное число электровозов (Э). В состав входят вагоны (В) и рестораны (Р). В поезде может быть не более трех ресторанов.

15. Описать с помощью БНФ и СД понятие «скорый поезд». Поезд состоит из тяги и состава. Тяга состоит из двух электровозов (Э). В состав входят обычные вагоны (В) и рестораны (Р). Два последних вагона поезда одновременно не могут быть ресторанами.

16. Описать с помощью БНФ и СД понятие «пассажирский поезд». Поезд состоит из тяги и состава. Тяга состоит из одного

электровоза (Э). В состав входят купейные (К) и плацкартные (П) вагоны. В поезде должно быть четное число плацкартных вагонов.

17. Построить БНФ и СД для понятия «периодическая дробь». Примеры дробей: 0.5, 1.(3), 5. 34656(45665).

18. Университет состоит из факультетов. Факультеты состоят из кафедр и лабораторий. Лаборатории состоят из сотрудников (С). Кафедры состоят из преподавателей (П). Например, университет из двух факультетов, в первом из которых одна кафедра с тремя преподавателями, а во втором одна лаборатория с одним сотрудником и одна кафедра с двумя преподавателями, можно описать с помощью такой структуры: (((ППП))((С)(ПП))). Внешние скобки соответствуют университету, наиболее вложенные – кафедрам и лаборатории, скобки второго уровня вложенности соответствуют факультетам. Построить БНФ и СД для описания всех возможных структур университета.

19. Определим понятие *функциональная запись* выражения. Функциональной записью выражения, состоящего из одной цифры, является эта цифра. Функциональной записью выражения $A \text{ op } B$, где op – операция, является запись $\text{op}(A', B')$, где A' и B' – функциональные записи выражений A и B соответственно. Функциональной записью выражения в скобках является функциональная запись этого выражения без скобок. Например, для выражения $2 * (5 + 6)$ функциональная запись выглядит так: $*(2, +(5, 6))$. Построить БНФ и СД для функциональных записей выражений, содержащих цифры, скобки и знаки операций $+$, $-$, $*$, $/$.

20. Определим понятие *префиксная запись* выражения. Префиксной записью выражения, состоящего из одной цифры, является эта цифра. Префиксной записью выражения $A \text{ op } B$, где op – операция, является запись $\text{op } A' B'$, где A' и B' – префиксные записи выражений A и B соответственно. Префиксной записью выражения в скобках является префиксная запись этого выражения

без скобок. Построить БНФ и СД для префиксных записей выражений, содержащих цифры, скобки и знаки операций $+$, $-$, $*$, $/$.

21. Определим понятие *постфиксная запись* выражения. Постфиксной записью выражения, состоящего из одной цифры, является эта цифра. Постфиксной записью выражения $A \text{ op } B$, где op – операция, является запись $A' B' \text{ op}$, где A' и B' – постфиксные записи выражений A и B соответственно. Постфиксной записью выражения в скобках является постфиксная запись этого выражения без скобок. Построить БНФ и СД для постфиксных записей выражений, содержащих цифры, скобки и знаки операций $+$, $-$, $*$, $/$.

22. Для правильного понимания смысла выражения, когда скобки в нем явно не расставлены, как в случаях $a + b * c$ и $a - b + c$, важно учитывать *приоритет* операций, а также *ассоциативность* операций одинакового приоритета. Умножение имеет больший приоритет, чем сложение, поэтому в выражении $a + b * c$ операнд b относится к операции умножения и эквивалентное выражение со скобками будет таким: $a + (b * c)$. В выражении $a - b + c$ операнд b относится к левой операции, т. е. к «минусу», а не к «плюсу» (в силу левой ассоциативности операций $+$ и $-$, имеющих одинаковый приоритет), и это выражение эквивалентно выражению $(a - b) + c$. Лево-ассоциативные операции группируются с помощью скобок слева направо: $a - b + c - d$ эквивалентно $((a - b) + c) - d$. Построить БНФ для арифметических выражений в алфавите $\{+, -, *, /, ^, a, (,)\}$, где a символизирует операнд, $^$ — операцию возведения в степень. В БНФ должен отражаться естественный приоритет операций, правая ассоциативность операции возведения в степень и левая ассоциативность остальных операций.

23. *Палиндромом* называется цепочка, которая читается одинаково слева направо и справа налево. Например, цепочки a , aba , $abba$ являются палиндромами. Пустая цепочка также является палиндромом. Построить БНФ и СД для понятия «палиндром» в алфавите $\{a, b\}$.

24. Если из арифметического выражения удалить все символы, кроме скобок, получим *скобочную систему*. Пустая цепочка также является скобочной системой. Построить БНФ и СД для скобочных систем.

25. Назовем *неделимой* скобочную систему, которую нельзя представить в виде сцепления двух других скобочных систем. Например, скобочная система $((()())())$ является неделимой, а $((())())$ не является. Построить БНФ для неделимой скобочной системы.

Решение

Заметим, что если произвольную скобочную систему заключить в объемлющие скобки, то получим неделимую скобочную систему. Поэтому БНФ для неделимой скобочной системы определяется так:

$$\langle \text{неделимая скобочная система} \rangle ::= (\langle \text{скобочная система} \rangle)$$

Теперь опишем понятие *скобочная система*. Простейшей скобочной системой является пустая цепочка. Непустая скобочная система может быть представлена как сцепление неделимой скобочной системы и произвольной скобочной системы. Следовательно, БНФ для произвольной скобочной системы можно определить так:

$$\langle \text{скобочная система} \rangle ::=$$

$$(\langle \text{скобочная система} \rangle) \langle \text{скобочная система} \rangle | \langle \text{пусто} \rangle$$

26. Построить СД, описывающую неделимые скобочные системы.

27. *Глубиной* скобочной системы называется максимальный уровень вложенности скобок. Например, система $((()())())$ имеет глубину 3, а система $(())$ глубину 1. Глубина пустой системы равна нулю. Описать с помощью БНФ и СД скобочные системы с глубиной не более трех.

28. Построить БНФ и СД для скобочных систем с глубиной не менее двух.

29. *Протяжением* скобочной системы называется количество неделимых скобочных систем, сцепление которых дает данную систему. Например, система $() (() ()) ()$ имеет протяжение 3. Протяжение пустой системы равно нулю. Построить БНФ и СД для скобочных систем с протяжением от трех до пяти.

30. Построить БНФ и СД для скобочных систем с протяжением не менее четырех.

31. *Шириной* скобочной системы называется максимальное протяжение среди всех подсистем, входящих в данную систему. Например, ширина системы $() (() () ())$ равна трем, а ширина $() (())$ равна двум. Ширина пустой системы равна нулю. Построить БНФ и СД для скобочных систем с шириной не более трех.

32. Построить БНФ и СД для скобочных систем с шириной не менее двух.

33. Построить БНФ и СД для скобочных систем, удовлетворяющих ограничениям:

- (а) глубина и ширина не превосходят двух;
- (б) глубина не более двух, а ширина не менее двух;
- (в) глубина не менее двух, а ширина не более двух;
- (г) глубина и ширина не менее двух;
- (д) протяжение не более двух, ширина не менее трех;
- (е) протяжение, глубина и ширина не менее двух.

34. Описать с помощью БНФ понятие «пассажирский поезд». В начале поезда находится электровоз (Э). Далее следуют вагоны.

Вагоны бывают плацкартные (П) и купейные (К). В поезде должно быть одинаковое количество купейных и плацкартных вагонов.

Решение

Поезд представляет собой сцепление электровоза (символа Э) и последовательности вагонов. Последовательность вагонов обозначим с помощью метапеременной $\langle \text{состав} \rangle$.

$\langle \text{пассажирский поезд} \rangle ::= \text{Э} \langle \text{состав} \rangle$

Разобьем состав (т.е. последовательность вагонов) на непустые участки так, что в каждом участке купейных и плацкартных вагонов поровну и никакой участок нельзя представить в виде сцепления двух меньших участков, в которых вагонов обоих типов поровну. Например, состав КПППККККПП разбивается на три участка: КП, ППКК, ККППП. Каждый такой участок можно рассматривать как неделимую скобочную систему со скобками К и П. Если участок начинается символом К и заканчивается символом П, то он представляет собой неделимую скобочную систему, в которой символ К – открывающая скобка, а П – закрывающая. Обозначим такой участок метапеременной $\langle \text{участок КП} \rangle$. Если участок начинается символом П и заканчивается символом К, то он представляет собой неделимую скобочную систему, в которой символ П – открывающая скобка, а К – закрывающая. Обозначим такой участок метапеременной $\langle \text{участок ПК} \rangle$. Состав можно рассматривать как последовательность участков двух видов:

$\langle \text{состав} \rangle ::= \langle \text{участок КП} \rangle \langle \text{состав} \rangle \mid$

$\langle \text{участок ПК} \rangle \langle \text{состав} \rangle \mid$

$\langle \text{пусто} \rangle$

Осталось определить участки двух видов аналогично тому, как это делалось в упражнении 25 для неделимых скобочных систем.

$\langle \text{участок КП} \rangle ::= \text{К} \langle \text{система КП} \rangle \text{П}$

$\langle \text{система КП} \rangle ::= \text{К} \langle \text{система КП} \rangle \text{П} \langle \text{система КП} \rangle \mid \langle \text{пусто} \rangle$

$\langle \text{участок ПК} \rangle ::= \text{П} \langle \text{система ПК} \rangle \text{К}$

$\langle \text{система ПК} \rangle ::= \text{П} \langle \text{система ПК} \rangle \text{К} \langle \text{система ПК} \rangle \mid \langle \text{пусто} \rangle$

35. Описать понятие «пассажирский поезд» из упражнения 34 с помощью БНФ, содержащей всего три альтернативы.

36. Построить СД для понятия «пассажирский поезд» из упражнения 34.

37. Описать с помощью БНФ и СД понятие «пассажирский поезд». Поезд состоит из тепловоза (Т) и вагонов. Вагоны бывают плацкартные (П) и купейные (К). В поезде должно быть неравное количество купейных и плацкартных вагонов.

38. Описать с помощью БНФ и СД понятие «пассажирский поезд». Поезд состоит из электровоза (Э) и вагонов. Вагоны бывают плацкартные (П) и купейные (К). В поезде купейных вагонов должно быть больше, чем плацкартных.

39. Построить БНФ и СД для «товарных поездов». Поезд состоит из тяги, за которой следуют вагоны. Тяга — это один или несколько тепловозов (Т). Вагоны бывают двух типов: цистерны (Ц) и рефрижераторы (Р). На каждый рефрижератор в поезде приходится по меньшей мере две цистерны.

40. Построить нормальный алгоритм Маркова (НАМ), область применимости которого относительно алфавита $\{a, b\}$ совпадает с множеством слов, описываемых с помощью БНФ:

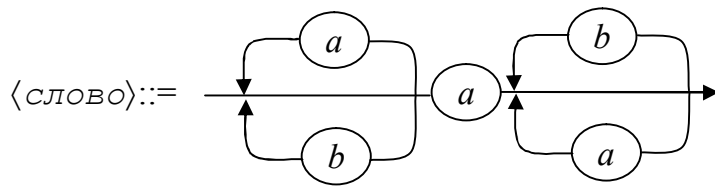
$\langle \text{слово} \rangle ::= a \langle \text{окончание} \rangle$

$\langle \text{окончание} \rangle ::= a \langle \text{окончание} \rangle \mid b \langle \text{окончание} \rangle \mid \langle \text{пусто} \rangle$

41. По заданному НАМ построить БНФ, описывающую область применимости данного НАМ относительно алфавита $\{a, b\}$.

$$\begin{cases} ab \rightarrow \\ b \rightarrow a \\ a \rightarrow b \end{cases}$$

42. Построить несамоприменимый НАМ, область применимости которого относительно алфавита $\{a, b\}$ задается СД:



43. Построить самоприменимый НАМ, область применимости которого относительно алфавита $\{a, b\}$ задается БНФ:

$$\langle \text{СЛОВО} \rangle ::= a \langle \text{СЛОВО} \rangle b \mid ab$$

44. Построить БНФ, описывающую множество слов, в которые заданный НАМ переводит слова вида $||| \dots || \# || \dots |||$.

$$\begin{cases} || \# | \rightarrow \\ \# \rightarrow \end{cases}$$

45. Построить самоприменимый НАМ, область применимости которого состоит из слов четной длины в алфавите $\{a, b\}$. Построить БНФ и СД для области применимости.

Литература

1. И.А. Волкова, А.А. Вылиток, Т.В. Руденко. Формальные грамматики и языки. Элементы теории трансляции. — М.: МГУ, 2009.
2. В.Г. Абрамов, Н.П. Трифонов, Г.Н. Трифонова. Введение в язык Паскаль. — М.: Наука, 1988.
3. В. Н. Пильщиков, В. Г. Абрамов, А. А. Вылиток, И. В. Горячая. Машины Тьюринга и алгоритмы Маркова. Решение задач. — М.: МГУ, 2006.
4. В.Г. Абрамов, Н.П. Трифонов, Г.Н. Трифонова. Введение в язык Паскаль. — М.: КНОРУС, 2011.

Содержание

1. Языки и метаязыки	3
2. Металингвистические формулы	4
3. Синтаксические диаграммы	7
4. Упражнения.....	9
Литература	21