**Контекстно-свободные грамматики и языки**

К классу контекстно-свободных (КС) относятся грамматики, у которых не накладывается никаких ограничений на вид правых частей их правил, а левая часть каждого правила — единственный нетерминал.

С помощью КС-грамматик задают синтаксис языков программирования.

**Левосторонние и правосторонние выводы в КС-грамматике**

Рассмотрим (однозначную) грамматику G8, задающую синтаксис арифметических выражений.

G8: E→T| E+T | E-T

T→ M | T\*M | T/M

M→a | b | c | (Е)

Построим два различных вывода цепочки **а + b\*с** в этой грамматике.

В первом случае, если сентенциальная форма содержит более одного нетерминала, будем выполнять подстановку (замену нетерминала правой частью одного из правил) для самого левого нетерминала этой сентенциальной формы:

Е⇒Е+Т⇒Т+Т⇒М+Т⇒а+Т⇒а+T\*М⇒а+М\*М⇒

а + b\*М⇒а + b\*с.

Такой вывод называется левосторонним.

Аналогично, вывод, в ходе которого замене всегда подвергается самый правый нетерминал сентенциальной формы, называется правосторонним. Для цепочки **а+b\*с** в грамматике G8 он будет таким:

Е⇒Е+ Т⇒ Е + T\*М ⇒Е + Т\*с ⇒ Е + М\*с ⇒

Е + b\*с ⇒ T+ b\*с ⇒ М + b\*с ⇒ а + b\*с.

Нетрудно убедиться, что обе эти последовательности подстановок соответствуют одному и тому же дереву вывода, хотя и разному порядку его построения.

Для цепочки а + b\*с в грамматике G8 и левосторонний, и правосторонний выводы строятся однозначно. Это же справедливо и для любой другой цепочки, порождаемой G8.

*КС-грамматика называется однозначной, если для каждого предложения языка, порождаемого этой грамматикой, существует единственный левосторонний вывод.*

**Алгоритмы распознавания КС-языков**

Существует алгоритм, позволяющий для любой КС-грамматики проверить принадлежность произвольной цепочки терминальных символов языку, порождаемому этой грамматикой, и получить вывод этой цепочки.

Наличие такого алгоритма и принцип его устройства следуют из того простого соображения, что если имеется конечная цепочка терминалов, то для проверки ее принадлежности языку достаточно построить все возможные сентенциальные формы грамматики, имеющие длину, совпадающую с длиной этой цепочки.

Количество этих сентенциальных форм конечно.

Такого рода алгоритм действует по принципу полного перебора. Объем перебора может быть сокращен, если процесс порождения сентенциальных форм будет организован так, что станет возможным определить тупики в ходе перебора еще до получения сентенциальной формы нужной длины.

Переборные алгоритмы работают с возвратами и вследствие своей неэффективности мало пригодны для практического использования. Для организации перебора с возвратами используют стек — структуру данных, подчиняющуюся дисциплине «последним пришел — первым ушел» (LIFO — Last In First Out).

В то же время для достаточно обширных подклассов КС-грамматик, подчиняющихся некоторым дополнительным ограничениям, существуют эффективные алгоритмы распознавания, которые мы будем в дальнейшем рассматривать и применять.

Для синтаксического анализа КС-языков используются как нисходящие (строящие дерево разбора от корня к листьям), так и восходящие алгоритмы.

**Распознающий автомат для КС-языков**

Для автоматных языков роль распознавателя может выполнять детерминированный конечный автомат.

Существует ли универсальный автоматный распознаватель КС-языков? Да, существует.

*Для произвольной КС-грамматики может быть построен недетерминированный автомат с магазинной памятью, принимающий язык, порождаемый этой грамматикой.*

Автомат с магазинной памятью (магазинный автомат, МП-автомат) подобен конечному автомату, оснащенному дополнительным запоминающим устройством со стековой дисциплиной «последним пришел — первым ушел».

Переходы МП- автомата определяются не только входным символом и текущим состоянием, но и значением вершины стека — элемента, поступившего в магазин последним.

В литературе на русском языке стек часто называют «магазином» по аналогии с магазином автоматического оружия: патрон, заряженный в магазин последним, выстреливается первым.

Недетерминированный МП-автомат — это не что иное, как устройство, реализующее перебор с возвратами. В этом смысле он эквивалентен обсуждавшемуся выше общему алгоритму распознавания КС-языков и так же неэффективен.

Для КС-грамматик, подчиняющихся определенным ограничениям, могут быть построены эффективные детерминированные магазинные автоматы, которые могут использоваться на практике для трансляции языков программирования.

**Самовложение в КС-грамматиках**

Если в грамматике G есть нетерминал А, для которого

А **α1** А **α2**

**G**

то есть из А нетривиально выводится цепочка **α1**А **α2**, где **α1**, **α2** — непустые цепочки терминалов и нетерминалов, то говорят, что такая грамматика содержит самовложение.

Например, грамматика арифметических выражений G8 содержит самовложение, поскольку из ее начального нетерминала E выводится цепочка (E).

Е⇒Т⇒М⇒(Е).

Содержит самовложение и грамматика регулярных выражений, поскольку

R (R).

Самовложение — характерный признак КС-грамматик.

КС-грамматика, не содержащая самовложения, эквивалентна автоматной грамматике.

Языки арифметических и регулярных выражений являются контекстно-свободными и не могут быть заданы автоматными грамматиками.