Лингвистические основы информатики Лекция 3 Однозначные грамматики

И. А. Михайлова, Ю. В. Нагребецкая

Уральский федеральный университет
Институт естественных наук и математики
Департамент математики, механики и компьютерных наук
Направления: Математика и компьютерные науки
Компьютеная безопасность (6 семестр)

Деревья вывода

Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, $S \Rightarrow \alpha_1 \Rightarrow \ldots \Rightarrow \alpha_n = w \in \Sigma^*$ — вывод, тогда деревом вывода цепочки w называется корневое дерево, построенное по следующим правилам:

Деревья вывода

Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, $S \Rightarrow \alpha_1 \Rightarrow \ldots \Rightarrow \alpha_n = w \in \Sigma^*$ — вывод, тогда деревом вывода цепочки w называется корневое дерево, построенное по следующим правилам:

- Корень дерева помечен аксиомой,

Деревья вывода

Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, $S \Rightarrow \alpha_1 \Rightarrow \ldots \Rightarrow \alpha_n = w \in \Sigma^*$ — вывод, тогда деревом вывода цепочки w называется корневое дерево, построенное по следующим правилам:

- Корень дерева помечен аксиомой,
- Все внутренние узлы дерева помечены нетерминалами, при этом если внутренний узел помечен нетерминалом A и $A \to X_1 X_2 \ldots X_s$ — правило вывода, которое было применно в выводе, тогда у этого узла сыновьями будут узлы, помеченные $X_1, X_2, \ldots X_s$ в указанном порядке слева направо,

Деревья вывода

Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, $S \Rightarrow \alpha_1 \Rightarrow \ldots \Rightarrow \alpha_n = w \in \Sigma^*$ — вывод, тогда деревом вывода цепочки w называется корневое дерево, построенное по следующим правилам:

- Корень дерева помечен аксиомой,
- Все внутренние узлы дерева помечены нетерминалами, при этом если внутренний узел помечен нетерминалом A и $A \rightarrow X_1 X_2 \ldots X_s$ — правило вывода, которое было применно в выводе, тогда у этого узла сыновьями будут узлы, помеченные $X_1, X_2, \ldots X_s$ в указанном порядке слева направо,
- Листья дерева помечены терминалами или ε, при этом если узел помечен ε, то у него нет братьев.

Деревья вывода

Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, $S \Rightarrow \alpha_1 \Rightarrow \ldots \Rightarrow \alpha_n = w \in \Sigma^*$ — вывод, тогда деревом вывода цепочки w называется корневое дерево, построенное по следующим правилам:

- Корень дерева помечен аксиомой,
- Все внутренние узлы дерева помечены нетерминалами, при этом если внутренний узел помечен нетерминалом A и $A \to X_1 X_2 \ldots X_s$ — правило вывода, которое было применно в выводе, тогда у этого узла сыновьями будут узлы, помеченные $X_1, X_2, \ldots X_s$ в указанном порядке слева направо,
- Листья дерева помечены терминалами или ε, при этом если узел помечен ε, то у него нет братьев.

Пример 1

- Рассмотрим грамматику
  $S \to AB, A \to aA \mid \varepsilon, B \to Bb \mid \varepsilon.$
- Построим вывод цепочки abb:

$S \Rightarrow AB$

Деревья вывода

Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, $S \Rightarrow \alpha_1 \Rightarrow \ldots \Rightarrow \alpha_n = w \in \Sigma^*$ — вывод, тогда деревом вывода цепочки w называется корневое дерево, построенное по следующим правилам:

- Корень дерева помечен аксиомой,
- Все внутренние узлы дерева помечены нетерминалами, при этом если внутренний узел помечен нетерминалом A и $A \to X_1 X_2 \ldots X_s$ — правило вывода, которое было применно в выводе, тогда у этого узла сыновьями будут узлы, помеченные $X_1, X_2, \ldots X_s$ в указанном порядке слева направо,
- Листья дерева помечены терминалами или ε, при этом если узел помечен ε, то у него нет братьев.

Пример 1

- Рассмотрим грамматику
  $S \to AB, A \to aA \mid \varepsilon, B \to Bb \mid \varepsilon.$
- Построим вывод цепочки abb:

$$S \Rightarrow AB \Rightarrow aAB$$

Деревья вывода

Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, $S \Rightarrow \alpha_1 \Rightarrow \ldots \Rightarrow \alpha_n = w \in \Sigma^*$ — вывод, тогда деревом вывода цепочки w называется корневое дерево, построенное по следующим правилам:

- Корень дерева помечен аксиомой,
- Все внутренние узлы дерева помечены нетерминалами, при этом если внутренний узел помечен нетерминалом A и $A \to X_1 X_2 \ldots X_s$ — правило вывода, которое было применно в выводе, тогда у этого узла сыновьями будут узлы, помеченные $X_1, X_2, \ldots X_s$ в указанном порядке слева направо,
- Листья дерева помечены терминалами или ε, при этом если узел помечен ε, то у него нет братьев.

Пример 1

- Рассмотрим грамматику
  $S \to AB, A \to aA \mid \varepsilon, B \to Bb \mid \varepsilon.$
- Построим вывод цепочки abb:

$$S \Rightarrow AB \Rightarrow aAB \Rightarrow aB$$

Деревья вывода

Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, $S \Rightarrow \alpha_1 \Rightarrow \ldots \Rightarrow \alpha_n = w \in \Sigma^*$ — вывод, тогда деревом вывода цепочки w называется корневое дерево, построенное по следующим правилам:

- Корень дерева помечен аксиомой,
- Все внутренние узлы дерева помечены нетерминалами, при этом если внутренний узел помечен нетерминалом A и $A \to X_1 X_2 \ldots X_s$ — правило вывода, которое было применно в выводе, тогда у этого узла сыновьями будут узлы, помеченные $X_1, X_2, \ldots X_s$ в указанном порядке слева направо,
- Листья дерева помечены терминалами или ε, при этом если узел помечен ε, то у него нет братьев.

Пример 1

- Рассмотрим грамматику
  $S \to AB, A \to aA \mid \varepsilon, B \to Bb \mid \varepsilon.$
- Построим вывод цепочки abb:

$$S \Rightarrow AB \Rightarrow aAB \Rightarrow aB \Rightarrow aBb$$

Деревья вывода

Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, $S \Rightarrow \alpha_1 \Rightarrow \ldots \Rightarrow \alpha_n = w \in \Sigma^*$ — вывод, тогда деревом вывода цепочки w называется корневое дерево, построенное по следующим правилам:

- Корень дерева помечен аксиомой,
- Все внутренние узлы дерева помечены нетерминалами, при этом если внутренний узел помечен нетерминалом A и $A \to X_1 X_2 \ldots X_s$ — правило вывода, которое было применно в выводе, тогда у этого узла сыновьями будут узлы, помеченные $X_1, X_2, \ldots X_s$ в указанном порядке слева направо,
- Листья дерева помечены терминалами или ε, при этом если узел помечен ε, то у него нет братьев.

Пример 1

- Рассмотрим грамматику
  $S \to AB, A \to aA \mid \varepsilon, B \to Bb \mid \varepsilon$.
- Построим вывод цепочки abb:

$$S \Rightarrow AB \Rightarrow aAB \Rightarrow aB \Rightarrow aBb \Rightarrow aBbb$$

Деревья вывода

Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, $S \Rightarrow \alpha_1 \Rightarrow \ldots \Rightarrow \alpha_n = w \in \Sigma^*$ — вывод, тогда деревом вывода цепочки w называется корневое дерево, построенное по следующим правилам:

- Корень дерева помечен аксиомой,
- Все внутренние узлы дерева помечены нетерминалами, при этом если внутренний узел помечен нетерминалом A и $A \to X_1 X_2 \ldots X_s$ — правило вывода, которое было применно в выводе, тогда у этого узла сыновьями будут узлы, помеченные $X_1, X_2, \ldots X_s$ в указанном порядке слева направо,
- Листья дерева помечены терминалами или ε, при этом если узел помечен ε, то у него нет братьев.

Пример 1

- Рассмотрим грамматику $S \to AB, A \to aA \mid \varepsilon, B \to Bb \mid \varepsilon$.
- Построим вывод цепочки abb:

$$S \Rightarrow AB \Rightarrow aAB \Rightarrow aB \Rightarrow aBb \Rightarrow aBbb \Rightarrow abb$$

Деревья вывода

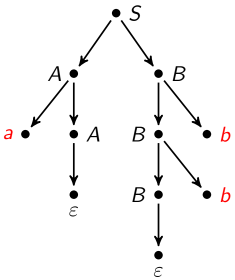
Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, $S \Rightarrow \alpha_1 \Rightarrow \ldots \Rightarrow \alpha_n = w \in \Sigma^*$ — вывод, тогда деревом вывода цепочки w называется корневое дерево, построенное по следующим правилам:

- Корень дерева помечен аксиомой,
- Все внутренние узлы дерева помечены нетерминалами, при этом если внутренний узел помечен нетерминалом A и $A \to X_1 X_2 \ldots X_s$ — правило вывода, которое было применно в выводе, тогда у этого узла сыновьями будут узлы, помеченные $X_1, X_2, \ldots X_s$ в указанном порядке слева направо,
- Листья дерева помечены терминалами или ε, при этом если узел помечен ε, то у него нет братьев.

Пример 1

- Рассмотрим грамматику
  $S \to AB, A \to aA \mid \varepsilon, B \to Bb \mid \varepsilon$.
- Построим вывод цепочки abb:

$$S \Rightarrow AB \Rightarrow aAB \Rightarrow aB \Rightarrow aBb \Rightarrow aBbb \Rightarrow abb$$



Деревья вывода

Пусть $G = (\Gamma, \Sigma, P, S)$ — КС грамматика, $S \Rightarrow \alpha_1 \Rightarrow \ldots \Rightarrow \alpha_n = w \in \Sigma^*$ — вывод, тогда деревом вывода цепочки w называется корневое дерево, построенное по следующим правилам:

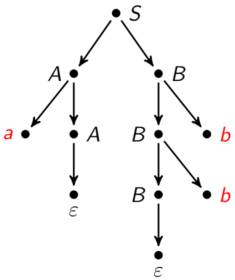
- Корень дерева помечен аксиомой,
- Все внутренние узлы дерева помечены нетерминалами, при этом если внутренний узел помечен нетерминалом A и $A \to X_1 X_2 \ldots X_s$ — правило вывода, которое было применно в выводе, тогда у этого узла сыновьями будут узлы, помеченные $X_1, X_2, \ldots X_s$ в указанном порядке слева направо,
- Листья дерева помечены терминалами или ε, при этом если узел помечен ε, то у него нет братьев.

Пример 1

- Рассмотрим грамматику $S \to AB, A \to aA \mid \varepsilon, B \to Bb \mid \varepsilon.$
- Построим вывод цепочки abb:

$$S \Rightarrow AB \Rightarrow aAB \Rightarrow aB \Rightarrow aBb \Rightarrow aBbb \Rightarrow abb$$

- Теперь цепочку abb можно прочитать на листьях дерева при их обходе слева направо (точнее, в глубину от корня S, записывая только листья, не являющиеся ε).



Пример 1 неоднозначной грамматики

(1) Однозначно ли можно восстановить вывод по дереву вывода? Почему?

(2) Однозначно ли можно восстановить дерево вывода по цепочке?

Ответим на вопрос (1).
Для одного дерева вывода для данной цепочки из $(\Gamma \cup \Sigma)^*$ может быть нескольких выводов. Все зависит от последовательности применения правил вывода. Далее мы будем рассматривать правосторонние (левосторонний) выводы и др.

Ответим на вопрос (2).
Рассмотрим грамматику

$$S \to S + S \mid S * S \mid (S) \mid x$$

Пример 1 неоднозначной грамматики

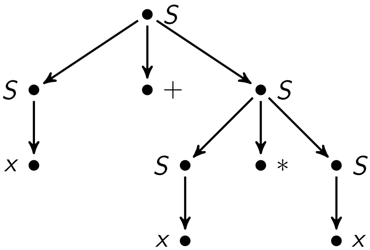
- Постройте дерево вывода цепочки $x + x * x$.

Пример 1 неоднозначной грамматики

- Постройте дерево вывода цепочки $x + x * x$.
- Какое дерево получилось?

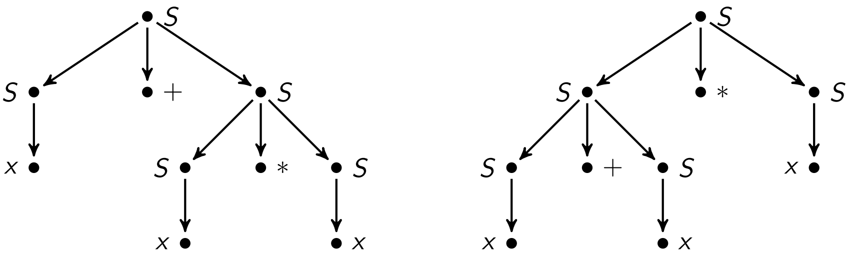
Пример 1 неоднозначной грамматики

- Постройте дерево вывода цепочки $x + x * x$.
- Какое дерево получилось?



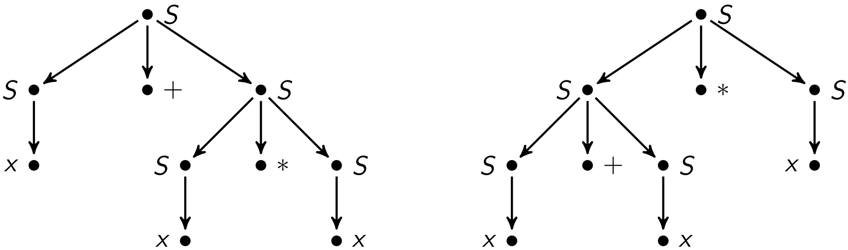
Пример 1 неоднозначной грамматики

- Постройте дерево вывода цепочки $x + x * x$.
- Какое дерево получилось?



Пример 1 неоднозначной грамматики

- Постройте дерево вывода цепочки $x + x * x$.
- Какое дерево получилось?



- Почему левое дерево “правильное”?

Пример 1 неоднозначной грамматики (продолжение)

Объяснение.

Грамматика не учитывает привычный нам приоритет операций, не является однозначной.

Левое дерево можно построить по выводу

$$S \Rightarrow$$

Пример 1 неоднозначной грамматики (продолжение)

Объяснение.

Грамматика не учитывает привычный нам приоритет операций, не является однозначной.

Левое дерево можно построить по выводу

$$S \Rightarrow S + S$$

Пример 1 неоднозначной грамматики (продолжение)

Объяснение.

Грамматика не учитывает привычный нам приоритет операций, не является однозначной.

Левое дерево можно построить по выводу

$$S \Rightarrow S + S \Rightarrow x + S$$

Пример 1 неоднозначной грамматики (продолжение)

Объяснение.

Грамматика не учитывает привычный нам приоритет операций, не является однозначной.

Левое дерево можно построить по выводу

$$S \Rightarrow S + S \Rightarrow x + S \Rightarrow x + S * S$$

Пример 1 неоднозначной грамматики (продолжение)

Объяснение.

Грамматика не учитывает привычный нам приоритет операций, не является однозначной.

Левое дерево можно построить по выводу

$$S \Rightarrow S + S \Rightarrow x + S \Rightarrow x + S * S \Rightarrow x + x * S$$

Объяснение.

Грамматика не учитывает привычный нам приоритет операций, не является однозначной.

Левое дерево можно построить по выводу

$$S \Rightarrow S + S \Rightarrow x + S \Rightarrow x + S * S \Rightarrow x + x * S \Rightarrow x + x * x$$

Пример 1 неоднозначной грамматики (продолжение)

Объяснение.

Грамматика не учитывает привычный нам приоритет операций, не является однозначной.

Левое дерево можно построить по выводу

$$S \Rightarrow S + S \Rightarrow x + S \Rightarrow x + S * S \Rightarrow x + x * S \Rightarrow x + x * x$$

(Можно ли построить это дерево еще по какому нибудь выводу?)

Пример 1 неоднозначной грамматики (продолжение)

Объяснение.

Грамматика не учитывает привычный нам приоритет операций, не является однозначной.

Левое дерево можно построить по выводу

$$S \Rightarrow S + S \Rightarrow x + S \Rightarrow x + S * S \Rightarrow x + x * S \Rightarrow x + x * x$$

(Можно ли построить это дерево еще по какому нибудь выводу?)

Этот вывод (и дерево вывода) соответствуют привычному нам приоритету операций в выражении $w = x + x * x$.

Пример 1 неоднозначной грамматики (окончание)

Правое дерево можно построить по выводу

$$S \Rightarrow$$

Пример 1 неоднозначной грамматики (окончание)

Правое дерево можно построить по выводу

$$S \Rightarrow S * S$$

Пример 1 неоднозначной грамматики (окончание)

Правое дерево можно построить по выводу

$$S \Rightarrow S * S \Rightarrow S + S * S$$

Пример 1 неоднозначной грамматики (окончание)

Правое дерево можно построить по выводу

$$S \Rightarrow S * S \Rightarrow S + S * S \Rightarrow x + S * S$$

Пример 1 неоднозначной грамматики (окончание)

Правое дерево можно построить по выводу

$$S \Rightarrow S * S \Rightarrow S + S * S \Rightarrow x + S * S \Rightarrow x + x * S$$

Пример 1 неоднозначной грамматики (окончание)

Правое дерево можно построить по выводу

$$S \Rightarrow S * S \Rightarrow S + S * S \Rightarrow x + S * S \Rightarrow x + x * S \Rightarrow x + x * x$$

Пример 1 неоднозначной грамматики (окончание)

Правое дерево можно построить по выводу

$$S \Rightarrow S * S \Rightarrow S + S * S \Rightarrow x + S * S \Rightarrow x + x * S \Rightarrow x + x * x$$

Этот вывод (и дерево вывода) не соответствуют привычному нам приоритету операций и на самом деле соответствуют выражению $w = (x + x) * x$.

Однозначные грамматики

Таким образом возникает естественный вопрос об однозначности произвольной грамматики.

- КС грамматика $G = (\Gamma, \Sigma, P, S)$ называется однозначной, если для любого слова $w \in L(G)$ существует ровно одно дерево вывода для w.
- **Плохие новости :-(** нет алгоритма, который строил бы однозначную грамматику, распознающую тот же язык.

Получается, для однозначной грамматики для каждой цепочки из $(\Gamma \cup \Sigma)^*$ выводов из может быть и несколько, но дерево вывода всегда одно.

Однозначные грамматики. Пример 2

Однозначные грамматики. Пример 2

- Пример 2 Построим однозначную грамматику, распознающую тот же язык, что и грамматика из примера 1:

$$S \to S + S \mid S * S \mid (S) \mid x$$

Однозначные грамматики. Пример 2

- Пример 2 Построим однозначную грамматику, распознающую тот же язык, что и грамматика из примера 1:

$$S \to S + S \mid S * S \mid (S) \mid x$$

- Эта грамматика распознает язык всех арифметических выражений с операндом x и операторами $+, *$, где скобки нужны для изменения порядка вычисления выражения. Будем записывать выражение с учетом приоритета операций.

Однозначные грамматики. Пример 2

- Пример 2 Построим однозначную грамматику, распознающую тот же язык, что и грамматика из примера 1:

$$S \to S + S \mid S * S \mid (S) \mid x$$

- Эта грамматика распознает язык всех арифметических выражений с операндом x и операторами $+, *$, где скобки нужны для изменения порядка вычисления выражения. Будем записывать выражение с учетом приоритета операций.
- Каждое такое выражение имеет вид $u_1 + u_2 + \ldots + u_k$, либо вид u_1, где каждая строка u_i если и содержит знак $+$, то он находится внутри скобок. Тогда $S \to S + A \mid A$ выводит сумму, а каждое из u_i выведем из A.

Однозначные грамматики. Пример 2 (продолжение)

- Каждое из выражений, выводимых из A имеет вид $v_1 * v_2 \ldots * v_s$, либо вид v_1, где каждое v_i — это либо x, либо выражение в скобках, тогда $A \to A * B \mid B$

Однозначные грамматики. Пример 2 (продолжение)

- Каждое из выражений, выводимых из A имеет вид $v_1 * v_2 \dots * v_s$, либо вид v_1, где каждое v_i — это либо x, либо выражение в скобках, тогда $A \to A * B \mid B$
- Выражение, стоящее в скобках, может быть любым, поэтому $B \to x \mid (S)$

Однозначные грамматики. Пример 2 (продолжение)

- Каждое из выражений, выводимых из A имеет вид $v_1 * v_2 \ldots * v_s$, либо вид v_1, где каждое v_i — это либо x, либо выражение в скобках, тогда $A \to A * B \mid B$
- Выражение, стоящее в скобках, может быть любым, поэтому $B \to x \mid (S)$
- В итоге получим грамматику

$$
\begin{aligned}
S &\to S + A \mid A \\
A &\to A * B \mid B \\
B &\to x \mid (S)
\end{aligned}
$$

Однозначные грамматики. Пример 2 (окончание)

Напишем, например, вывод выражения $x * (x + x * x)$:

Однозначные грамматики. Пример 2 (окончание)

Напишем, например, вывод выражения $x * (x + x * x)$:

$S \Rightarrow$

Однозначные грамматики. Пример 2 (окончание)

Напишем, например, вывод выражения $x * (x + x * x)$:

$S \Rightarrow A \Rightarrow$

Однозначные грамматики. Пример 2 (окончание)

Напишем, например, вывод выражения $x * (x + x * x)$:

$S \Rightarrow A \Rightarrow A * B \Rightarrow$

Однозначные грамматики. Пример 2 (окончание)

Напишем, например, вывод выражения $x * (x + x * x)$:

$S \Rightarrow A \Rightarrow A * B \Rightarrow x * B \Rightarrow$

Однозначные грамматики. Пример 2 (окончание)

Напишем, например, вывод выражения $x * (x + x * x)$:

$S \Rightarrow A \Rightarrow A * B \Rightarrow x * B \Rightarrow x * (S) \Rightarrow$

Однозначные грамматики. Пример 2 (окончание)

Напишем, например, вывод выражения $x * (x + x * x)$:

$S \Rightarrow A \Rightarrow A * B \Rightarrow x * B \Rightarrow x * (S) \Rightarrow x * (S + A) \Rightarrow$

Однозначные грамматики. Пример 2 (окончание)

Напишем, например, вывод выражения $x * (x + x * x)$:

$S \Rightarrow A \Rightarrow A * B \Rightarrow x * B \Rightarrow x * (S) \Rightarrow x * (S + A) \Rightarrow x * (x + A) \Rightarrow$

Напишем, например, вывод выражения $x * (x + x * x)$:

$S \Rightarrow A \Rightarrow A * B \Rightarrow x * B \Rightarrow x * (S) \Rightarrow x * (S + A) \Rightarrow x * (x + A) \Rightarrow$
$x * (x + A * B) \Rightarrow$

Однозначные грамматики. Пример 2 (окончание)

Напишем, например, вывод выражения $x * (x + x * x)$:

$S \Rightarrow A \Rightarrow A * B \Rightarrow x * B \Rightarrow x * (S) \Rightarrow x * (S + A) \Rightarrow x * (x + A) \Rightarrow$
$x * (x + A * B) \Rightarrow x * (x + B * B) \Rightarrow$

Однозначные грамматики. Пример 2 (окончание)

Напишем, например, вывод выражения $x * (x + x * x)$:

$S \Rightarrow A \Rightarrow A * B \Rightarrow x * B \Rightarrow x * (S) \Rightarrow x * (S + A) \Rightarrow x * (x + A) \Rightarrow$
$x * (x + A * B) \Rightarrow x * (x + B * B) \Rightarrow x * (x + x * B) \Rightarrow$

Однозначные грамматики. Пример 2 (окончание)

Напишем, например, вывод выражения $x * (x + x * x)$:

$S \Rightarrow A \Rightarrow A * B \Rightarrow x * B \Rightarrow x * (S) \Rightarrow x * (S + A) \Rightarrow x * (x + A) \Rightarrow$
$x * (x + A * B) \Rightarrow x * (x + B * B) \Rightarrow x * (x + x * B) \Rightarrow x * (x + x * x)$

Однозначные грамматики. Пример 2 (окончание)

Напишем, например, вывод выражения $x * (x + x * x)$:

$S \Rightarrow A \Rightarrow A * B \Rightarrow x * B \Rightarrow x * (S) \Rightarrow x * (S + A) \Rightarrow x * (x + A) \Rightarrow$
$x * (x + A * B) \Rightarrow x * (x + B * B) \Rightarrow x * (x + x * B) \Rightarrow x * (x + x * x)$

Заметим, что доказательство того, что рассматриваемая грамматика порождает нужный нам язык, легко доказать индукцией по длине вывода.