Категориальные грамматики

Теория формальных языков.

Степан Кузнецов, Мати Пентус, Алексей Сорокин

МГУ им. М. В. Ломоносова, межфакультетский курс,
весенний семестр 2015–2016 учебного года

Регулярные выражения

Пусть зафиксирован конечный алфавит Σ.

Регулярные выражения

Пусть зафиксирован конечный алфавит Σ.

- Базовые регулярные выражения: элементы алфавита,

Регулярные выражения

Пусть зафиксирован конечный алфавит Σ.

- Базовые регулярные выражения: элементы алфавита,
- Также есть константы 0 (пустой язык) и 1 (язык, содержащий только пустое слово ε).

Регулярные выражения

Пусть зафиксирован конечный алфавит Σ.

- Базовые регулярные выражения: элементы алфавита,
- Также есть константы 0 (пустой язык) и 1 (язык, содержащий только пустое слово ε).
- Бинарные операции: $|$ (объединение) и $\cdot$ (конкатенация или приписывание): $u \cdot v = uv$,

Регулярные выражения

Пусть зафиксирован конечный алфавит Σ.

- Базовые регулярные выражения: элементы алфавита,
- Также есть константы 0 (пустой язык) и 1 (язык, содержащий только пустое слово ε).
- Бинарные операции: $|$ (объединение) и $\cdot$ (конкатенация или приписывание): $u \cdot v = uv$,
- Унарная операция $*$ (итерация, взять любое количество раз): L^* состоит из слов вида $u_1 \ldots u_r$, где $r \in \mathbb{N}$, $u_1, \ldots, u_r \in L$.

Регулярные выражения

Пусть зафиксирован конечный алфавит Σ.

- Базовые регулярные выражения: элементы алфавита,
- Также есть константы 0 (пустой язык) и 1 (язык, содержащий только пустое слово ε).
- Бинарные операции: $|$ (объединение) и $\cdot$ (конкатенация или приписывание): $u \cdot v = uv$,
- Унарная операция $*$ (итерация, взять любое количество раз): L^* состоит из слов вида $u_1 \ldots u_r$, где $r \in \mathbb{N}$, $u_1, \ldots, u_r \in L$.
- Если α — регулярное выражение, то $L(\alpha)$ — задаваемый им язык.

Регулярные выражения

Пусть зафиксирован конечный алфавит Σ.

- Базовые регулярные выражения: элементы алфавита,
- Также есть константы 0 (пустой язык) и 1 (язык, содержащий только пустое слово ε).
- Бинарные операции: $|$ (объединение) и $\cdot$ (конкатенация или приписывание): $u \cdot v = uv$,
- Унарная операция $*$ (итерация, взять любое количество раз): L^* состоит из слов вида $u_1 \ldots u_r$, где $r \in \mathbb{N}$, $u_1, \ldots, u_r \in L$.
- Если α — регулярное выражение, то $L(\alpha)$ — задаваемый им язык.
- Например, $L((a|b)^*) = \{\varepsilon, a, b, aa, ab, ba, bb, \ldots\}$.

Регулярные выражения

Пусть зафиксирован конечный алфавит Σ.

- Базовые регулярные выражения: элементы алфавита,
- Также есть константы 0 (пустой язык) и 1 (язык, содержащий только пустое слово ε).
- Бинарные операции: $|$ (объединение) и $\cdot$ (конкатенация или приписывание): $u \cdot v = uv$,
- Унарная операция $*$ (итерация, взять любое количество раз): L^* состоит из слов вида $u_1 \ldots u_r$, где $r \in \mathbb{N}$, $u_1, \ldots, u_r \in L$.
- Если α — регулярное выражение, то $L(\alpha)$ — задаваемый им язык.
- Например, $L((a|b)^*) = \{\varepsilon, a, b, aa, ab, ba, bb, \ldots\}$.
- Приоритет операций: итерация, конкатенация, объединение. При этом значок конкатенации можно опускать.

Регулярные выражения

Пусть зафиксирован конечный алфавит Σ.

- Базовые регулярные выражения: элементы алфавита,
- Также есть константы 0 (пустой язык) и 1 (язык, содержащий только пустое слово ε).
- Бинарные операции: $|$ (объединение) и $\cdot$ (конкатенация или приписывание): $u \cdot v = uv$,
- Унарная операция $*$ (итерация, взять любое количество раз): L^* состоит из слов вида $u_1 \dots u_r$, где $r \in \mathbb{N}$, $u_1, \dots, u_r \in L$.
- Если α — регулярное выражение, то $L(\alpha)$ — задаваемый им язык.
- Например, $L((a|b)^*) = \{\varepsilon, a, b, aa, ab, ba, bb, \dots\}$.
- Приоритет операций: итерация, конкатенация, объединение. При этом значок конкатенации можно опускать.

Язык регулярный, если он задаётся регулярным выражением.

Примеры регулярных языков

- Все слова в алфавите $\{a, b\}$: $(a|b)^*$,

Примеры регулярных языков

- Все слова в алфавите $\{a, b\}$: $(a|b)^*$,
- Слова в алфавите $\{a, b, c\}$, где предпоследняя буква — b: $(a|b|c)^* b(a|b|c)$.

Примеры регулярных языков

- Все слова в алфавите $\{a, b\}$: $(a|b)^*$,
- Слова в алфавите $\{a, b, c\}$, где предпоследняя буква — b: $(a|b|c)^* b(a|b|c)$.
- Слова в алфавите $\{a, b, c\}$, содержащие ровно 2 буквы a: $(b|c)^* a(b|c)^* a(b|c)^*$.

Примеры регулярных языков

- Все слова в алфавите $\{a, b\}$: $(a|b)^*$,
- Слова в алфавите $\{a, b, c\}$, где предпоследняя буква — b: $(a|b|c)^* b(a|b|c)$.
- Слова в алфавите $\{a, b, c\}$, содержащие ровно 2 буквы a: $(b|c)^* a(b|c)^* a(b|c)^*$.
- Слова нечётной длины в алфавите $\{a, b\}$: $(a|b)((a|b)(a|b))^*$.

Примеры регулярных языков

- Все слова в алфавите $\{a, b\}$: $(a|b)^*$,
- Слова в алфавите $\{a, b, c\}$, где предпоследняя буква — b: $(a|b|c)^*b(a|b|c)$.
- Слова в алфавите $\{a, b, c\}$, содержащие ровно 2 буквы a: $(b|c)^*a(b|c)^*a(b|c)^*$.
- Слова нечётной длины в алфавите $\{a, b\}$: $(a|b)((a|b)(a|b))^*$.
- Слова в алфавите $\{a, b, c\}$, содержащие чётное число букв a:

Примеры регулярных языков

- Все слова в алфавите $\{a, b\}$: $(a|b)^*$,
- Слова в алфавите $\{a, b, c\}$, где предпоследняя буква — b: $(a|b|c)^* b (a|b|c)$.
- Слова в алфавите $\{a, b, c\}$, содержащие ровно 2 буквы a: $(b|c)^* a (b|c)^* a (b|c)^*$.
- Слова нечётной длины в алфавите $\{a, b\}$: $(a|b)((a|b)(a|b))^*$.
- Слова в алфавите $\{a, b, c\}$, содержащие чётное число букв a: $((b|c)^* a (b|c)^* a)^* (b|c)^*$

Примеры регулярных языков

- Все слова в алфавите $\{a, b\}$: $(a|b)^*$,
- Слова в алфавите $\{a, b, c\}$, где предпоследняя буква — b: $(a|b|c)^* b(a|b|c)$.
- Слова в алфавите $\{a, b, c\}$, содержащие ровно 2 буквы a: $(b|c)^* a(b|c)^* a(b|c)^*$.
- Слова нечётной длины в алфавите $\{a, b\}$: $(a|b)((a|b)(a|b))^*$.
- Слова в алфавите $\{a, b, c\}$, содержащие чётное число букв a: $((b|c)^* a(b|c)^* a)^*(b|c)^*$.
- Слова в алфавите $\{a, b, c\}$, где перед a идёт только b:

Примеры регулярных языков

- Все слова в алфавите $\{a, b\}$: $(a|b)^*$,
- Слова в алфавите $\{a, b, c\}$, где предпоследняя буква — b: $(a|b|c)^*b(a|b|c)$.
- Слова в алфавите $\{a, b, c\}$, содержащие ровно 2 буквы a: $(b|c)^*a(b|c)^*a(b|c)^*$.
- Слова нечётной длины в алфавите $\{a, b\}$: $(a|b)((a|b)(a|b))^*$.
- Слова в алфавите $\{a, b, c\}$, содержащие чётное число букв a: $((b|c)^*a(b|c)^*a)^*(b|c)^*$.
- Слова в алфавите $\{a, b, c\}$, где перед a идёт только b: $((b|c)^*ba)^*(b|c)^*$

Примеры регулярных языков

- Все слова в алфавите $\{a, b\}$: $(a|b)^*$,
- Слова в алфавите $\{a, b, c\}$, где предпоследняя буква — b: $(a|b|c)^*b(a|b|c)$.
- Слова в алфавите $\{a, b, c\}$, содержащие ровно 2 буквы a: $(b|c)^*a(b|c)^*a(b|c)^*$.
- Слова нечётной длины в алфавите $\{a, b\}$: $(a|b)((a|b)(a|b))^*$.
- Слова в алфавите $\{a, b, c\}$, содержащие чётное число букв a: $((b|c)^*a(b|c)^*a)^*(b|c)^*$.
- Слова в алфавите $\{a, b, c\}$, где перед a идёт только b: $((b|c)^*ba)^*(b|c)^*$.
- Непустые слова в алфавите $\{a, b\}$, в которых одинаковые буквы не идут подряд:

Примеры регулярных языков

- Все слова в алфавите $\{a, b\}$: $(a|b)^*$,
- Слова в алфавите $\{a, b, c\}$, где предпоследняя буква — b: $(a|b|c)^* b(a|b|c)$.
- Слова в алфавите $\{a, b, c\}$, содержащие ровно 2 буквы a: $(b|c)^* a(b|c)^* a(b|c)^*$.
- Слова нечётной длины в алфавите $\{a, b\}$: $(a|b)((a|b)(a|b))^*$.
- Слова в алфавите $\{a, b, c\}$, содержащие чётное число букв a: $((b|c)^* a(b|c)^* a)^*(b|c)^*$.
- Слова в алфавите $\{a, b, c\}$, где перед a идёт только b: $((b|c)^* ba)^*(b|c)^*$.
- Непустые слова в алфавите $\{a, b\}$, в которых одинаковые буквы не идут подряд: $(a(ba)^*(b|1))|(b(ab)^*(a|1))$

Примеры регулярных языков

- Все слова в алфавите $\{a, b\}$: $(a|b)^*$,
- Слова в алфавите $\{a, b, c\}$, где предпоследняя буква — b: $(a|b|c)^*b(a|b|c)$.
- Слова в алфавите $\{a, b, c\}$, содержащие ровно 2 буквы a: $(b|c)^*a(b|c)^*a(b|c)^*$.
- Слова нечётной длины в алфавите $\{a, b\}$: $(a|b)((a|b)(a|b))^*$.
- Слова в алфавите $\{a, b, c\}$, содержащие чётное число букв a: $((b|c)^*a(b|c)^*a)^*(b|c)^*$.
- Слова в алфавите $\{a, b, c\}$, где перед a идёт только b: $((b|c)^*ba)^*(b|c)^*$.
- Непустые слова в алфавите $\{a, b\}$, в которых одинаковые буквы не идут подряд: $(a(ba)^*(b|1))|(b(ab)^*(a|1))$.
- Непустые слова в алфавите $\{a, b, c\}$, в которых одинаковые буквы не идут подряд:

Примеры регулярных языков

- Все слова в алфавите $\{a, b\}$: $(a|b)^*$,
- Слова в алфавите $\{a, b, c\}$, где предпоследняя буква — b: $(a|b|c)^*b(a|b|c)$.
- Слова в алфавите $\{a, b, c\}$, содержащие ровно 2 буквы a: $(b|c)^*a(b|c)^*a(b|c)^*$.
- Слова нечётной длины в алфавите $\{a, b\}$: $(a|b)((a|b)(a|b))^*$.
- Слова в алфавите $\{a, b, c\}$, содержащие чётное число букв a: $((b|c)^*a(b|c)^*a)^*(b|c)^*$.
- Слова в алфавите $\{a, b, c\}$, где перед a идёт только b: $((b|c)^*ba)^*(b|c)^*$.
- Непустые слова в алфавите $\{a, b\}$, в которых одинаковые буквы не идут подряд: $(a(ba)^*(b|1))|(b(ab)^*(a|1))$.
- Непустые слова в алфавите $\{a, b, c\}$, в которых одинаковые буквы не идут подряд: $(H|1)(cH)^*(1|c|cH)$, где H — ответ на предыдущий пункт.

Определение конечного автомата

Пусть Σ — конечный алфавит.

Определение конечного автомата

Конечный автомат: помеченный граф $M = \langle Q, \Sigma, \Delta, q_0, F \rangle$, где

Определение конечного автомата

Пусть Σ — конечный алфавит.

Определение конечного автомата

Конечный автомат: помеченный граф $M = \langle Q, \Sigma, \Delta, q_0, F \rangle$, где

- Q — конечное множество состояний (вершин графа),

Определение конечного автомата

Пусть Σ — конечный алфавит.

Определение конечного автомата

Конечный автомат: помеченный граф $M = \langle Q, \Sigma, \Delta, q_0, F \rangle$, где

- Q — конечное множество состояний (вершин графа),
- Δ — конечное множество переходов (рёбер с метками) вида $\langle q_1, w \rangle \to q_2$, $q_1, q_2 \in Q$, $w \in \Sigma^*$.

Определение конечного автомата

Пусть Σ — конечный алфавит.

Определение конечного автомата

Конечный автомат: помеченный граф $M = \langle Q, \Sigma, \Delta, q_0, F\rangle$, где

- Q — конечное множество состояний (вершин графа),
- Δ — конечное множество переходов (рёбер с метками) вида $\langle q_1, w\rangle \to q_2$, $q_1, q_2 \in Q$, $w \in \Sigma^*$.
- $q_0 \in Q$ — стартовое состояние
- $F \subseteq Q$ — завершающие состояния.

Определение конечного автомата

Пусть Σ — конечный алфавит.

Определение конечного автомата

Конечный автомат: помеченный граф $M = \langle Q, \Sigma, \Delta, q_0, F \rangle$, где

- Q — конечное множество состояний (вершин графа),
- Δ — конечное множество переходов (рёбер с метками) вида $\langle q_1, w \rangle \to q_2$, $q_1, q_2 \in Q$, $w \in \Sigma^*$.
- $q_0 \in Q$ — стартовое состояние
- $F \subseteq Q$ — завершающие состояния.

Метка пути — конкатенация всех меток на входящих в него рёбрах.

Определение конечного автомата

Пусть Σ — конечный алфавит.

Определение конечного автомата

Конечный автомат: помеченный граф $M = \langle Q, \Sigma, \Delta, q_0, F \rangle$, где

- Q — конечное множество состояний (вершин графа),
- Δ — конечное множество переходов (рёбер с метками) вида $\langle q_1, w \rangle \to q_2$, $q_1, q_2 \in Q$, $w \in \Sigma^*$.
- $q_0 \in Q$ — стартовое состояние
- $F \subseteq Q$ — завершающие состояния.

Метка пути — конкатенация всех меток на входящих в него рёбрах.
Язык $L(M)$ — метки путей из начального состояния в завершающие.

Определение конечного автомата

Пусть Σ — конечный алфавит.

Определение конечного автомата

Конечный автомат: помеченный граф $M = \langle Q, \Sigma, \Delta, q_0, F \rangle$, где

- Q — конечное множество состояний (вершин графа),
- Δ — конечное множество переходов (рёбер с метками) вида $\langle q_1, w \rangle \to q_2$, $q_1, q_2 \in Q$, $w \in \Sigma^*$.
- $q_0 \in Q$ — стартовое состояние
- $F \subseteq Q$ — завершающие состояния.

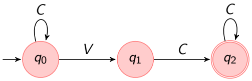
Метка пути — конкатенация всех меток на входящих в него рёбрах.
Язык $L(M)$ — метки путей из начального состояния в завершающие.
Язык — автоматный, если задаётся некоторым конечным автоматом.

Примеры конечных автоматов

- Закрытый слог

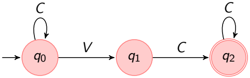
Примеры конечных автоматов

- Закрытый слог



Примеры конечных автоматов

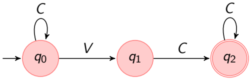
- Закрытый слог



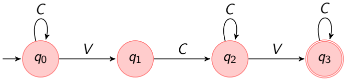
- Слово с 2 гласными, разделёнными хотя бы одним согласным:

Примеры конечных автоматов

- Закрытый слог

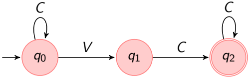


- Слово с 2 гласными, разделёнными хотя бы одним согласным:

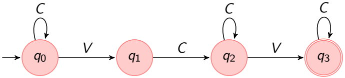


Примеры конечных автоматов

- Закрытый слог



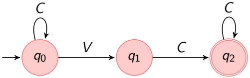
- Слово с 2 гласными, разделёнными хотя бы одним согласным:



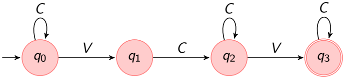
- Слова в алфавите a, b, c, где перед a обязательно идёт b:

Примеры конечных автоматов

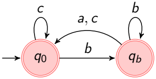
- Закрытый слог



- Слово с 2 гласными, разделёнными хотя бы одним согласным:



- Слова в алфавите a, b, c, где перед a обязательно идёт b:



Свойства автоматных языков

Теорема

Любой автоматный язык распознаётся автоматом с однобуквенными переходами.

Свойства автоматных языков

Теорема

Любой автоматный язык распознаётся автоматом с однобуквенными переходами.

Определение

Автомат с однобуквенными переходами — детерминированный, если ни из какого состояния нет двух переходов по одной и той же букве.

Теорема

Любой автоматный язык распознаётся детерминированным конечным автоматом.

Свойства автоматных языков

Теорема

Любой автоматный язык распознаётся автоматом с однобуквенными переходами.

Определение

Автомат с однобуквенными переходами — детерминированный, если ни из какого состояния нет двух переходов по одной и той же букве.

Теорема

Любой автоматный язык распознаётся детерминированным конечным автоматом.

Теорема (Клини, 1956)

Классы автоматных и регулярных языков совпадают.

Доказательства неавтоматности языков

Теорема

Пусть найдётся бесконечное множество слов $u_1, v_1, \ldots, u_n, v_n, \ldots$ *таких, что* $\forall i (u_i v_i \in L)$, *если* $i \neq j$, *то* $u_i v_j \notin L$. *Тогда язык* L *— не автоматный.*

Доказательства неавтоматности языков

Теорема

Пусть найдётся бесконечное множество слов $u_1, v_1, \ldots, u_n, v_n, \ldots$ *таких, что* $\forall i (u_i v_i \in L)$, *если* $i \neq j$, *то* $u_i v_j \notin L$. *Тогда язык* L *— не автоматный.*

Пример

Язык $L = \{a^n b^n \mid n \in \mathbb{N}\}$ *— не автоматный.*

Доказательства неавтоматности языков

Теорема

Пусть найдётся бесконечное множество слов $u_1, v_1, \ldots, u_n, v_n, \ldots$ *таких, что* $\forall i (u_i v_i \in L)$, *если* $i \neq j$, *то* $u_i v_j \notin L$. *Тогда язык* L *— не автоматный.*

Пример

Язык $L = \{a^n b^n \mid n \in \mathbb{N}\}$ *— не автоматный.*

Теорема (Лемма о разрастании)

Для всякого автоматного языка L *найдётся число* p, *такое что для всякого* $w \in L$, *такого что* $|w| \geqslant p$, *верно что* $\exists x, y, z\ (w = xyz,\ |xy| \leqslant p, |y| > 0)$ *и* $\forall k \in \mathbb{N} (xy^k z \in L)$.

Доказательства неавтоматности языков

Теорема

Пусть найдётся бесконечное множество слов $u_1, v_1, \ldots, u_n, v_n, \ldots$ таких, что $\forall i (u_i v_i \in L)$, если $i \neq j$, то $u_i v_j \notin L$. Тогда язык L — не автоматный.

Пример

Язык $L = \{a^n b^n \mid n \in \mathbb{N}\}$ — не автоматный.

Теорема (Лемма о разрастании)

Для всякого автоматного языка L найдётся число p, такое что для всякого $w \in L$, такого что $|w| \geqslant p$, верно что $\exists x, y, z \ (w = xyz, \ |xy| \leqslant p, |y| > 0)$ и $\forall k \in \mathbb{N} (xy^k z \in L)$.

Упражнение

Доказать, что $L = \{a^n b^n \mid n \in \mathbb{N}\}$ — не автоматный язык, с помощью леммы о разрастании.

Неавтоматность естественных языков

Замечание: естественные языки неавтоматны.

Неавтоматность естественных языков

Замечание: естественные языки неавтоматны.

- The dog chases the cat.

Неавтоматность естественных языков

Замечание: естественные языки неавтоматны.

- The dog chases the cat.
- The dog whom a dog chases chases the cat.

Неавтоматность естественных языков

Замечание: естественные языки неавтоматны.

- The dog chases the cat.
- The dog whom a dog chases chases the cat.
- The dog whom a dog whom a dog chases chases the cat.

Неавтоматность естественных языков

Замечание: естественные языки неавтоматны.

- The dog chases the cat.
- The dog whom a dog chases chases the cat.
- The dog whom a dog whom a dog chases chases the cat.
- ...
- The dog (whom a dog)^m chases^{m+1} the cat.

Неавтоматность естественных языков

Замечание: естественные языки неавтоматны.

- The dog chases the cat.
- The dog whom a dog chases chases the cat.
- The dog whom a dog whom a dog chases chases the cat.
- ...
- The dog (whom a dog)^m chases^{m+1} the cat.
- Тот же шаблон, что в $a^n b^n$.

Неавтоматность естественных языков

Замечание: естественные языки неавтоматны.

- The dog chases the cat.
- The dog whom a dog chases chases the cat.
- The dog whom a dog whom a dog chases chases the cat.
- ...
- The dog (whom a dog)^m chases^{m+1} the cat.

- Тот же шаблон, что в $a^n b^n$.
- На самом деле регулярных языков не хватит даже для полного описания морфологии (редупликация):

Пример (Редупликация в языке бамбара, семья манде)

wulu	*“собака”*
wulu-o-wulu	*“любая собака”*
wulunyinina	*“видящий собаку”*
wulunyinina-o-wulunyinina	*“любой, кто видит собаку”*

Контекстно-свободные грамматики

Определение

Контекстно-свободная грамматика: $G = \langle N, \Sigma, P, S \rangle$, *где*

- Σ — *алфавит,*
- N — *множество вспомогательных символов,*
- P — *множество правил* $A \to \alpha,\ A \in N, \alpha \in (\Sigma \cup N)^*$,
- $S \in N$ — *стартовый символ.*

Контекстно-свободные грамматики

Определение

Контекстно-свободная грамматика: $G = \langle N, \Sigma, P, S \rangle$, *где*

- Σ — *алфавит,*
- N — *множество вспомогательных символов,*
- P — *множество правил* $A \to \alpha$, $A \in N, \alpha \in (\Sigma \cup N)^*$,
- $S \in N$ — *стартовый символ.*

Отношение перехода $\vdash'$ на множестве $(\Sigma \cup N)^*$ — наименьшее отношение, такое что

$$(A \to \alpha \in P) \Leftrightarrow \forall \gamma, \eta \in (\Sigma \cup N)^* \; (\gamma A \eta \vdash' \gamma \alpha \eta)$$

Контекстно-свободные грамматики

Определение

Контекстно-свободная грамматика: $G = \langle N, \Sigma, P, S \rangle$, *где*

- Σ — *алфавит,*
- N — *множество вспомогательных символов,*
- P — *множество правил* $A \to \alpha$, $A \in N, \alpha \in (\Sigma \cup N)^*$,
- $S \in N$ — *стартовый символ.*

Отношение перехода $\vdash'$ на множестве $(\Sigma \cup N)^*$ — наименьшее отношение, такое что

$$(A \to \alpha \in P) \Leftrightarrow \forall \gamma, \eta \in (\Sigma \cup N)^* \; (\gamma A \eta \vdash' \gamma \alpha \eta)$$

$$(\alpha \vdash \beta) \leftrightarrow (\exists \alpha_1, \ldots, \alpha_m \; (\alpha \vdash' \alpha_1 \ldots \alpha_2 \ldots \alpha_m \vdash' \beta))$$

Контекстно-свободные грамматики

Определение

Контекстно-свободная грамматика: $G = \langle N, \Sigma, P, S \rangle$, *где*

- Σ — *алфавит,*
- N — *множество вспомогательных символов,*
- P — *множество правил* $A \to \alpha$, $A \in N, \alpha \in (\Sigma \cup N)^*$,
- $S \in N$ — *стартовый символ.*

Отношение перехода $\vdash'$ на множестве $(\Sigma \cup N)^*$ — наименьшее отношение, такое что

$$(A \to \alpha \in P) \Leftrightarrow \forall \gamma, \eta \in (\Sigma \cup N)^* \; (\gamma A \eta \vdash' \gamma \alpha \eta)$$

$$(\alpha \vdash \beta) \leftrightarrow (\exists \alpha_1, \ldots, \alpha_m \; (\alpha \vdash' \alpha_1 \ldots \alpha_2 \ldots \alpha_m \vdash' \beta))$$

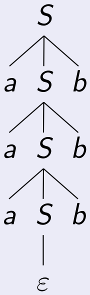
$L(G) = \{w \in \Sigma^* \mid S \vdash w\}$ — язык, порождаемый грамматикой.

Примеры контекстно-свободных грамматик

Пример

$L(G) = \{a^n b^n \mid n \in \mathbb{N}\}$: $S \to aSb$, $S \to \varepsilon$.

Пример (Дерево вывода)



Примеры контекстно-свободных грамматик

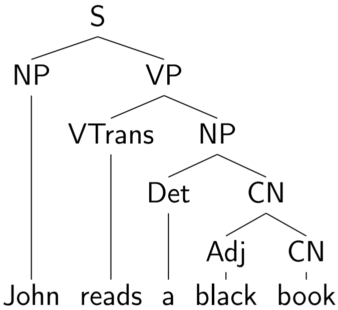
Пример

$S \to NP\ VP$	$NP \to Det\ CN$	$NP \to John$
$VP \to VTrans\ NP$	$VTrans \to reads$	$Det \to a$
$CN \to Adj\ CN$	$Adj \to black$	$CN \to book$

Примеры контекстно-свободных грамматик

Пример

$$
\begin{array}{lll}
S \to NP\ VP & NP \to Det\ CN & NP \to John \\
VP \to VTrans\ NP & VTrans \to reads & Det \to a \\
CN \to Adj\ CN & Adj \to black & CN \to book
\end{array}
$$

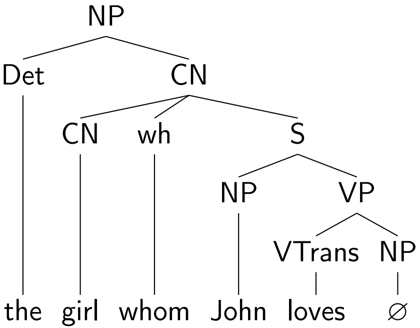


Примеры контекстно-свободных грамматик

The girl whom John loves

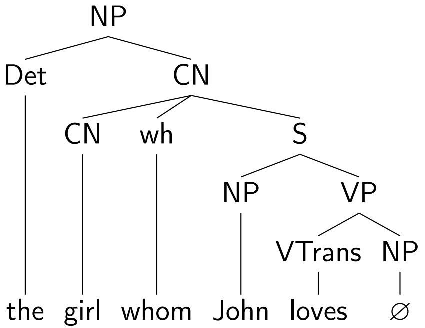
Примеры контекстно-свободных грамматик

The girl whom John loves



Примеры контекстно-свободных грамматик

The girl whom John loves



Однако мы не можем разрешить переход NP в пустое слово!

Примеры контекстно-свободных грамматик

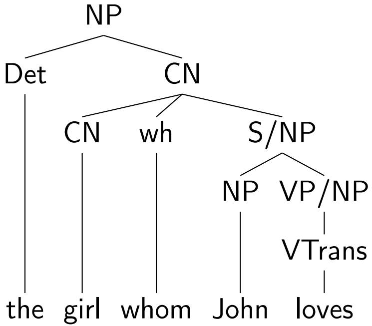
Пример

$NP \to Det\ CN$	$Det \to the$	$NP \to John$
$CN \to CN\ wh\ S/NP$	$wh \to whom$	$CN \to girl$
$S/NP \to NP\ VP/NP$	$VP/NP \to VTrans\ \cancel{NP}$	$VTrans \to loves$

Примеры контекстно-свободных грамматик

Пример

$$NP \to Det\ CN \qquad Det \to the \qquad NP \to John$$
$$CN \to CN\ wh\ S/NP \qquad wh \to whom \qquad CN \to girl$$
$$S/NP \to NP\ VP/NP \qquad VP/NP \to VTrans\ \cancel{NP} \qquad VTrans \to loves$$



Свойства контекстно-свободных языков

- Всякий автоматный язык является контекстно-свободным.

Свойства контекстно-свободных языков

- Всякий автоматный язык является контекстно-свободным.

Идея доказательства

Положить $N = Q$, $S = q_0$, заменить переходы $\langle q_1, a\rangle \to q_2$ на правила $q_1 \to aq_2$, а для всех завершающих состояний $q \in F$ добавить правила $q \to \varepsilon$.

- всякий контекстно-свободный язык порождается грамматикой только с правилами вида $A \to BC$, $A \to a$ и $S \to \varepsilon$, причём S не содержится в правой части правил.

Свойства контекстно-свободных языков

- Всякий автоматный язык является контекстно-свободным.

Идея доказательства

Положить $N = Q$, $S = q_0$, заменить переходы $\langle q_1, a \rangle \to q_2$ на правила $q_1 \to aq_2$, а для всех завершающих состояний $q \in F$ добавить правила $q \to \varepsilon$.

- всякий контекстно-свободный язык порождается грамматикой только с правилами вида $A \to BC$, $A \to a$ и $S \to \varepsilon$, причём S не содержится в правой части правил.
- Пересечение контекстно-свободного языка с автоматным является контекстно-свободным.

Свойства контекстно-свободных языков

- Всякий автоматный язык является контекстно-свободным.

Идея доказательства

Положить $N = Q$, $S = q_0$, заменить переходы $\langle q_1, a\rangle \to q_2$ на правила $q_1 \to aq_2$, а для всех завершающих состояний $q \in F$ добавить правила $q \to \varepsilon$.

- всякий контекстно-свободный язык порождается грамматикой только с правилами вида $A \to BC$, $A \to a$ и $S \to \varepsilon$, причём S не содержится в правой части правил.
- Пересечение контекстно-свободного языка с автоматным является контекстно-свободным.
- Если отображение ϕ заменяет букву a на некоторое фиксированное слово w_a, то из того, что L — контекстно-свободный язык, следует, что $\phi(L)$ – тоже контектсно-свободный язык.