Контекстно-свободные грамматики и языки, иерархия Хомского

Михайлова Инна Анатольевна

Институт математики и естественных наук.
Кафедра алгебры и фундаментальной информатики.

Контекстно-свободные грамматики и языки

- Грамматика $G = (\Gamma, \Sigma, P, S)$ называется контекстно-свободной (КС), если все правила вывода в ней имеют вид $A \to \alpha$, где $A \in \Gamma$, $\alpha \in (\Sigma \cup \Gamma)^*$.

Контекстно-свободные грамматики и языки

- Грамматика $G = (\Gamma, \Sigma, P, S)$ называется контекстно-свободной (КС), если все правила вывода в ней имеют вид $A \to \alpha$, где $A \in \Gamma$, $\alpha \in (\Sigma \cup \Gamma)^*$.
- Язык L называется контекстно-свободным, если **существует** КС грамматика, порождающая этот язык.

Контекстно-свободные грамматики и языки

- Грамматика $G = (\Gamma, \Sigma, P, S)$ называется контекстно-свободной (КС), если все правила вывода в ней имеют вид $A \to \alpha$, где $A \in \Gamma$, $\alpha \in (\Sigma \cup \Gamma)^*$.
- Язык L называется контекстно-свободным, если **существует** КС грамматика, порождающая этот язык.

Пример Язык $\{a^n b^n \mid n \geq 0\}$ является КС языком, так как порождается КС грамматикой $S \to aSb \mid \varepsilon$.

Контекстно-свободные грамматики и языки

- Грамматика $G = (\Gamma, \Sigma, P, S)$ называется контекстно-свободной (КС), если все правила вывода в ней имеют вид $A \to \alpha$, где $A \in \Gamma$, $\alpha \in (\Sigma \cup \Gamma)^*$.
- Язык L называется контекстно-свободным, если **существует** КС грамматика, порождающая этот язык.

Пример Язык $\{a^n b^n \mid n \geq 0\}$ является КС языком, так как порождается КС грамматикой $S \to aSb \mid \varepsilon$.
А какой язык порождает грамматика $S \to aaAb \mid \varepsilon, aA \to S$?

Связь с рациональными языками

КС грамматика $G = (\Gamma, \Sigma, P, S)$ называется правoлинейной, если ее правила вывода имеют вид либо $A \to aB$, где $a \in \Sigma, A, B \in \Gamma$, либо $A \to u$, где $u \in \Sigma^*$.

Связь с рациональными языками

КС грамматика $G = (\Gamma, \Sigma, P, S)$ называется праволинейной, если ее правила вывода имеют вид либо $A \to aB$, где $a \in \Sigma, A, B \in \Gamma$, либо $A \to u$, где $u \in \Sigma^*$.

Теорема

Язык рационален тогда и только тогда, когда он порождается некоторой праволинейной грамматикой

Связь с рациональными языками

КС грамматика $G = (\Gamma, \Sigma, P, S)$ называется правolинейной, если ее правила вывода имеют вид либо $A \to aB$, где $a \in \Sigma, A, B \in \Gamma$, либо $A \to u$, где $u \in \Sigma^*$.

Теорема

Язык рационален тогда и только тогда, когда он порождается некоторой праволинейной грамматикой

- **Необходимость.** Пусть язык L рационален, тогда существует ДКА $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$, который распознает этот язык.

Связь с рациональными языками

КС грамматика $G = (\Gamma, \Sigma, P, S)$ называется праволинейной, если ее правила вывода имеют вид либо $A \to aB$, где $a \in \Sigma, A, B \in \Gamma$, либо $A \to u$, где $u \in \Sigma^*$.

Теорема

Язык рационален тогда и только тогда, когда он порождается некоторой праволинейной грамматикой

- **Необходимость.** Пусть язык L рационален, тогда существует ДКА $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$, который распознает этот язык.
- Построим грамматику $G = (\Gamma, \Sigma, P, S)$, где $\Gamma = Q, S = q_0$, а правила вывода будут следующими:

$$\begin{aligned} \delta(q, a) = r &\Leftrightarrow q \to ar \\ q \in F &\Leftrightarrow q \to \varepsilon \end{aligned}$$

Связь с рациональными языками

КС грамматика $G = (\Gamma, \Sigma, P, S)$ называется праволинейной, если ее правила вывода имеют вид либо $A \to aB$, где $a \in \Sigma, A, B \in \Gamma$, либо $A \to u$, где $u \in \Sigma^*$.

Теорема

Язык рационален тогда и только тогда, когда он порождается некоторой праволинейной грамматикой

- **Необходимость.** Пусть язык L рационален, тогда существует ДКА $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$, который распознает этот язык.
- Построим грамматику $G = (\Gamma, \Sigma, P, S)$, где $\Gamma = Q, S = q_0$, а правила вывода будут следующими:

$$\delta(q, a) = r \Leftrightarrow q \to ar$$
$$q \in F \Leftrightarrow q \to \varepsilon$$

- Проверим, что $L(\mathcal{A}) = L(G)$:

$$\begin{aligned} w = a_1 \dots a_n \in L(\mathcal{A}) &\Leftrightarrow \delta(q_0, a_1) = q_1, \delta(q_1, a_2) = q_2, \dots, \delta(q_{n-1}, a_n) = q_n \in F \\ &\Leftrightarrow q_0 \to a_1 q_1, q_1 \to a_2 q_2 \dots q_{n-1} \to a_n q_n, q_n \to \varepsilon \\ &\Leftrightarrow q_0 \Rightarrow a_1 q_1 \Rightarrow a_1 a_2 q_2, \dots, \Rightarrow a_1 a_2 \dots a_n q_n \Rightarrow a_1 a_2 \dots a_n = w \\ &\Leftrightarrow w \in L(G) \end{aligned}$$

Связь с рациональными языками (окончание)

КС грамматика $G = (\Gamma, \Sigma, P, S)$ называется праволинейной, если ее правила вывода имеют вид либо $A \to aB$, где $a \in \Sigma, A, B \in \Gamma$, либо $A \to u$, где $u \in \Sigma^*$.

Теорема

Язык рационален тогда и только тогда, когда он порождается некоторой праволинейной грамматикой

- **Достаточность.** Пусть $G = (\Gamma, \Sigma, P, S)$ — праволинейная грамматика, пусть в ней есть правило вида $A \to u \in \Sigma^*$, где $u \neq \varepsilon$.

Связь с рациональными языками (окончание)

КС грамматика $G = (\Gamma, \Sigma, P, S)$ называется праволинейной, если ее правила вывода имеют вид либо $A \to aB$, где $a \in \Sigma$, $A, B \in \Gamma$, либо $A \to u$, где $u \in \Sigma^*$.

Теорема

Язык рационален тогда и только тогда, когда он порождается некоторой праволинейной грамматикой

- **Достаточность.** Пусть $G = (\Gamma, \Sigma, P, S)$ — праволинейная грамматика, пусть в ней есть правило вида $A \to u \in \Sigma^*$, где $u \neq \varepsilon$.
- Пусть $u = a_1 \dots a_m$. Рассмотрим грамматику $G' = (\Gamma', \Sigma, P', S)$ такую, что $\Gamma' = \Gamma \cup \{A_1, \dots A_m\}$ ($A_1, \dots A_m$ — новые нетерминалы) и $P' = P \setminus \{A \to u\} \cup \{A \to a_1 A_1, \dots, A_{m-1} \to a_m A_m, A_m \to \varepsilon\}$. Очевидно, что G' и G порождают один и тот же язык.

Связь с рациональными языками (окончание)

КС грамматика $G = (\Gamma, \Sigma, P, S)$ называется правволинейной, если ее правила вывода имеют вид либо $A \to aB$, где $a \in \Sigma$, $A, B \in \Gamma$, либо $A \to u$, где $u \in \Sigma^*$.

Теорема

Язык рационален тогда и только тогда, когда он порождается некоторой праволинейной грамматикой

- **Достаточность.** Пусть $G = (\Gamma, \Sigma, P, S)$ — праволинейная грамматика, пусть в ней есть правило вида $A \to u \in \Sigma^*$, где $u \neq \varepsilon$.
- Пусть $u = a_1 \dots a_m$. Рассмотрим грамматику $G' = (\Gamma', \Sigma, P', S)$ такую, что $\Gamma' = \Gamma \cup \{A_1, \dots A_m\}$ ($A_1, \dots A_m$ — новые нетерминалы) и $P' = P \setminus \{A \to u\} \cup \{A \to a_1 A_1, \dots, A_{m-1} \to a_m A_m, A_m \to \varepsilon\}$. Очевидно, что G' и G порождают один и тот же язык.
- Поэтому можно считать, что если $A \to u \in \Sigma^*$ — правило вывода, то $u = \varepsilon$. Построим НКА $\mathcal{A} = (Q, \Sigma, E, q_0, F)$, где $Q = \Gamma$, $q_0 = S$.

Связь с рациональными языками (окончание)

КС грамматика $G = (\Gamma, \Sigma, P, S)$ называется правoлинейной, если ее правила вывода имеют вид либо $A \to aB$, где $a \in \Sigma$, $A, B \in \Gamma$, либо $A \to u$, где $u \in \Sigma^*$.

Теорема

Язык рационален тогда и только тогда, когда он порождается некоторой праволинейной грамматикой

- **Достаточность.** Пусть $G = (\Gamma, \Sigma, P, S)$ — праволинейная грамматика, пусть в ней есть правило вида $A \to u \in \Sigma^*$, где $u \neq \varepsilon$.
- Пусть $u = a_1 \ldots a_m$. Рассмотрим грамматику $G' = (\Gamma', \Sigma, P', S)$ такую, что $\Gamma' = \Gamma \cup \{A_1, \ldots A_m\}$ ($A_1, \ldots A_m$ — новые нетерминалы) и $P' = P \setminus \{A \to u\} \cup \{A \to a_1 A_1, \ldots, A_{m-1} \to a_m A_m, A_m \to \varepsilon\}$. Очевидно, что G' и G порождают один и тот же язык.
- Поэтому можно считать, что если $A \to u \in \Sigma^*$ — правило вывода, то $u = \varepsilon$. Построим НКА $\mathcal{A} = (Q, \Sigma, E, q_0, F)$, где $Q = \Gamma$, $q_0 = S$.
- Определим множество переходов и множество конечных состояний:

$$A \to aB \Leftrightarrow (A, a, B) \in E$$
$$A \to \varepsilon \Leftrightarrow A \in F$$

Связь с рациональными языками (окончание)

КС грамматика $G = (\Gamma, \Sigma, P, S)$ называется праволинейной, если ее правила вывода имеют вид либо $A \to aB$, где $a \in \Sigma, A, B \in \Gamma$, либо $A \to u$, где $u \in \Sigma^*$.

Теорема

Язык рационален тогда и только тогда, когда он порождается некоторой праволинейной грамматикой

- **Достаточность.** Пусть $G = (\Gamma, \Sigma, P, S)$ — праволинейная грамматика, пусть в ней есть правило вида $A \to u \in \Sigma^*$, где $u \neq \varepsilon$.
- Пусть $u = a_1 \dots a_m$. Рассмотрим грамматику $G' = (\Gamma', \Sigma, P', S)$ такую, что $\Gamma' = \Gamma \cup \{A_1, \dots A_m\}$ ($A_1, \dots A_m$ — новые нетерминалы) и $P' = P \setminus \{A \to u\} \cup \{A \to a_1 A_1, \dots, A_{m-1} \to a_m A_m, A_m \to \varepsilon\}$. Очевидно, что G' и G порождают один и тот же язык.
- Поэтому можно считать, что если $A \to u \in \Sigma^*$ — правило вывода, то $u = \varepsilon$. Построим НКА $\mathcal{A} = (Q, \Sigma, E, q_0, F)$, где $Q = \Gamma$, $q_0 = S$.
- Определим множество переходов и множество конечных состояний:

$$A \to aB \Leftrightarrow (A, a, B) \in E$$
$$A \to \varepsilon \Leftrightarrow A \in F$$

- Равенство $L(\mathcal{A}) = L(G)$ доказывается аналогично предыдущему.

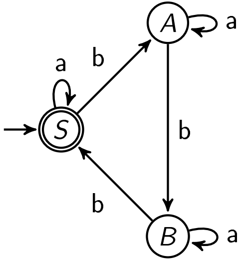
Пример

Постройте грамматику, которая распознает язык всех слов над алфавитом $\{a, b\}$, в которых число букв b делится на 3.

Пример

Постройте грамматику, которая распознает язык всех слов над алфавитом $\{a, b\}$, в которых число букв b делится на 3.

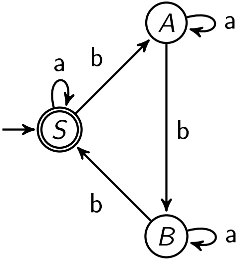
1. Нарисуем автомат, распознающий язык (отметим, что для решения задачи он необязательно должен быть детерминированным)



Пример

Постройте грамматику, которая распознает язык всех слов над алфавитом $\{a, b\}$, в которых число букв b делится на 3.

1. Нарисуем автомат, распознающий язык (отметим, что для решения задачи он необязательно должен быть детерминированным)
2. Напишем грамматику, используя теорему.



$$S \to aS \mid bB \mid \varepsilon$$
$$A \to aA \mid bS$$
$$B \to aB \mid bA$$

Иерархия Хомского