Нека $[d] = \{0, 1, \ldots, d\}$. За думи $\alpha, \beta \in [d]^*$, използваме $\alpha \preceq_{lex} \beta$, съответно $\alpha \prec_{lex} \beta$, за да отбележим, че α, тоест:

$$\alpha \prec_{lex} \beta \iff \beta \in \{\alpha\}[d]^+ \text{ или } \exists i(\alpha[0..i-1] = \beta[0..i-1] \& \alpha(i) < \beta(i)).$$

Задача 0.1. *1. Нека $\tau \subseteq [d]^*$ е наредено кореново дърво, а $\alpha \in \tau$ и $\beta = \alpha[0..i-1]$ за някое $i < |\alpha|$. Да се докаже, че β е връх в τ, който не е листо.*

2. Нека $T : \tau \to \Sigma \cup \mathcal{N} \cup \{e\}$ е дърво на извод в контекстносвободна граматика $G = \langle \Sigma, \mathcal{N}, P, S \rangle$. Да означим с $\overline{T}\tau \to \Sigma \cup \mathcal{N} \cup \{\varepsilon\}$ функцията, за която:

$$\overline{T}(\alpha) = \begin{cases} T(\alpha) \text{ ако } T(\alpha) \in \Sigma \cup \mathcal{N} \\ \varepsilon \text{ ако } T(\alpha) = e. \end{cases}$$

Да се докаже, че ако $\alpha_1 \prec_{lex} \alpha_2 \cdots \prec_{lex} \alpha_N$ са всички листа на τ в нарастващ лексикографски ред на τ, то:

$$w(T) = \overline{T}(\alpha_1)\overline{T}(\alpha_2)\ldots\overline{T}(\alpha_N).$$

3. Нека $T' : \tau' \to \Sigma \cup \mathcal{N} \cup \{e\}$ и $T'' : \tau'' \to \Sigma \cup \mathcal{N} \cup \{e\}$ са дървета на извод в в контекстносвободна граматика $G = \langle \Sigma, \mathcal{N}, P, S \rangle$. Ако $\alpha_1 \prec_{lex} \alpha_2 \cdots \prec_{lex} \alpha_N$ са всички листа на τ в нарастващ лексикографски ред на τ и $T'(\alpha_i) = T''(\varepsilon)$ за някое $1 \leq i \leq N$, да се докаже, че:

$$w(T' +_{\alpha_i} T'') = \overline{T}(\alpha_1)\overline{T}(\alpha_2)\ldots\overline{T}(\alpha_{i-1})w(T'')\overline{T}(\alpha_{i+1})\ldots\overline{T}(\alpha_N)$$

Задача 0.2. *Да се докаже, че ако $G = \langle \Sigma, \mathcal{N}, P, S \rangle$ е контекстносвободна граматика, и P' са правилата:*

$$P' = \{A \to \alpha^{rev} \mid A \to \alpha \in P\},$$

то $G' = \langle \Sigma, \mathcal{N}, P', S \rangle$ е контекстносвободна граматика с език $\mathcal{L}(G') = \mathcal{L}(G)^{rev}$.

Задача 0.3. *Казваме, че контекстносвободна граматика $G = \langle \Sigma, \mathcal{N}, P, S \rangle$ е* еднозначна, *ако за всяка дума $w \in \mathcal{L}(G)$ има единствено дърво на извод T в G, за което едновременно: (i) $T(\varepsilon) = S$ и (ii) $w(T) = w$.*

1. Да се докаже, че ако $\mathcal{A} = \langle \Sigma, Q, s, \delta, F \rangle$ е краен детерминиран автомат, за който $Q \cap \Sigma = \emptyset$, то граматиката:

$$G_{\mathcal{A}} = \langle \Sigma, Q, P, s \rangle, \text{ където } P = \{p \to aq \mid \delta(p, a) = q\} \cup \{f \to \varepsilon \mid f \in F\}$$

е еднозначна и има език $\mathcal{L}(G_{\mathcal{A}}) = \mathcal{L}(G)$.

Да се даде пример за недетерминиран краен автомат, за който съответната му автоматна граматика не е еднозначна.

Да се докаже, че граматиката:

$$G = \langle \{a, b\}, \{S\}, \{S \to aSbS | \varepsilon\}, S \rangle$$

не е еднозначна.

Задача 0.4. *Нека $k \in \mathbb{N}$ е естествено числа, а $x_1, \ldots, x_k$ и $y_1, \ldots, y_k$ са думи над Σ, които удовлетворяват следните две условия:*

1. $x_1 \cdot x_2 \cdot \dots \cdot x_k \preceq_{pref} y_1 \cdot y_2 \cdot \dots \cdot y_k$,

2. *за всяко* $i \leq k$, *ако* $x_i \preceq_{pref} y_i$ *или* $y_i \preceq_{pref} x_i$, *то* $x_i = y_i$.

Да се докаже, че $x_i = y_i$ *за всяко* $i \leq k$.

Задача 0.5. *Нека* $G = \langle \Sigma, \mathcal{N}, P, S \rangle$ *е контекстносвободна граматика, която има следните две свойства:*

1. *ако* $A \to \alpha \in P$, *то* $\alpha \in \Sigma(\Sigma \cup \mathcal{N})^*$,

2. *ако* $A \to a\alpha' \in P$ *и* $A \to a\alpha'' \in P$, *то* $\alpha' = \alpha''$.

Да се докаже, че G *е еднозначна.*

Задача 0.6. *Нека* $G = \langle \Sigma, \mathcal{N}, P, S \rangle$ *е контекстносвободна граматика, която има следните две свойства:*

1. *ако* $A \to \alpha \in P$, *то* $\alpha \in \Sigma(\Sigma \cup \mathcal{N})^*$,

2. *ако* $A \to a\alpha' \in P$ *и* $A \to a\alpha'' \in P$, *за които* $\alpha'(0) = \alpha''(0)$, *то*
   - $|\alpha'| = |\alpha''|$ *и*
   - $\alpha'(j) = \alpha''(j)$ *за всяко* j, *за което* $\alpha'(j) \in \mathcal{N}$.

Да се докаже, че G *е еднозначна.*

Задача 0.7. *Да се докаже, че следните граматики са еднозначни:*

1. $G = \langle \{x, y, z, +, *\}, \{S\}, \{S \to +SS| * SS|x|y|z\}, S \rangle$.

2. $G = \langle \{x, y, z, +, *, (,)\}, \{S\}, \{S \to (S + S)|(S * S)|x|y|z\}, S \rangle$.

3. $G = \langle \{x, y, z, \exists, \vee, \&\}, \{S\}, \{S \to \exists S| \vee SS|\&SS|x|y|z\}, S \rangle$.

4. $G = \langle \{x, y, z, \exists, \vee, \&(,)\}, \{S\}, \{S \to \exists S|(S \vee S)|(S\&S)|x|y|z\}, S \rangle$.

Забележка: Обикновено езици като горните биха били описани във формата на Backus-Naur:

$$E := +EE| * EE|x|y|z \text{ или } E := (E + E)|(E * E)|x|y|z.$$

Упътване 0.1. 1. Използвайте, че τ е затворено относно префикси, съответно β е префикс на α и $i < |\alpha$, съответно $\alpha[0..i]$ също е префикс на α.

2. Използвайте индукция по големината на $|\tau|$. Първо разгледайте случаите $\tau = \{\varepsilon\}$ и $\tau = \{\varepsilon, 0\}$, като $T(0) = e$. В общия случай, $T(\varepsilon) \to T(0) \dots T(i)$ е правило, като $T(0), \dots, T(i) \in \Sigma \cup \mathcal{N}$. Означете с $Lvs(\tau)$ листата на дърво τ и съобразете, че:

$$Lvs(\tau) = \{0\}Lvs(\tau_0) \cup \{1\}Lvs(\tau_1) \dots \{i\}Lvs(\tau_i).$$

Заключете, че в лексикографската наредба на $Lvs(\tau)$, листата са наредени в $i + 1$ блока, $\{j\}Lvs(\tau_j)$, за $0 \leq j \leq i$ като във всеки блок наредбата е същата като в $Lvs(\tau_j)$. Използвайте индуктивната хипотеза за T_j и довършете.

3. Използвайте 2 и това, че $Lvs(T' +_{\alpha_i} T'') = Lvs(T') \setminus \{\alpha_i\} \cup \{\alpha_i\}Lvs(T'')$. Обосновете, че $\{\alpha_i\}Lvs(T'')$ са лексикографски по-големи от α_{i-1} и лексикографски по-малки от α_{i+1}. Довършете като заместите в 2.

Упътване 0.2. Използвайте, че $(w_1 \dots w_k)^{rev} = w_k^{rev} \dots w_1^{rev}$ за всеки k думи $w_1, \dots, w_k$.

Упътване 0.3. 1. Може да използвате индукция по дължината на път $p \overset{w}{\to}{}^{*}_{\mathcal{A}} q$, за да установите, че има единствено дърво на извод T в G с $T(\varepsilon) = p$, $w(T) = wq$.

2. Разгледайте недетерминиран автомат, при който някоя дума от езика му допуска поне две успешни изпълнения. Разгледайте съотевтните изводи и дървета на извод в граматиката.

3. Забележете, че всяка дума, която се извежда от тази граматика има равен брой a и b. Експериментирайте с къси думи с това двойство, за да разберете как изглеждат дърветата на извод в тази граматика и съответно да получите пример за нееднозначност.

Упътване 0.4. Използвайте индукция по k. При индуктивния преход покажете, че $|x_1| \leq |y_1|$ влече, че $x_1 \preceq_{pref} y_1$ и съответно $|y_1| \leq |x_1|$ влече $y_1 \preceq_{pref} x_1$. Използвайте второто условие, за да заключите, че $x_1 = y_1$ и съответно да сведете към индуктивната хипотеза.

Упътване 0.5. Разгледайте свойството за две дървета на извод T', T'', $\mathfrak{p}(T', T'')$, което казва следното:

$$T'(\varepsilon) = T''(\varepsilon) \& w(T') \preceq_{pref} w(T'') \in \Sigma^* \Rightarrow T' = T''.$$

Използвайте пълна математическа индукция по n, за да покажете, че $\phi(n)$:

$$\phi(n) \overset{def}{\iff} \forall T', T''(|dom(T')| \leq n \& |dom(T'')| \leq n \Rightarrow \mathfrak{p}(T', T'').$$

При индуктивната стъпка, от това, че $T'(\varepsilon) = T''(\varepsilon)$ и всяко правило започва с терминал, заключете, че $T'(0) = T''(0)$ е първата буква на $w(T') = w(T'')$. Използвайте условието, за да заключите, че ε има едни и същи синове в $\tau' = dom(T')$ и $\tau'' = dom(T'')$ и съответно, ако те са $0, 1, \dots, i$, то $T'(j) = T''(j)$ за $j \leq i$.

Използвайте индуктивната хипотеза за T'_j и T''_j и предишната задача, за да заключите, че $T'_j = T''_j$ за всяко $j = 0, 1 \dots, i$. Довършете.

Упътване 0.6. Модифицирайте решението на предишната задача.

Упътване 0.7. Приложете горните две задачи.