Jméno: Noblesní Plšice **UČO:** 1234567

2. [3 body] Uvažme problém pro danou gramatiku $\mathcal{G}$ typu 0 (tedy bez omezení) a dané $k \in \mathbb{N}$ rozhodnout, zda lze v této gramatice odvodit prázdné slovo ε v nejvýše k krocích. Problém nazveme *bounded epsilon derivation* a ztotožníme s množinou

$$BED = \{\langle \mathcal{G}, k\rangle \mid \mathcal{G} \text{ je gramatika splňující } S \overset{\leq k}{\Longrightarrow}_{\mathcal{G}} \varepsilon, \text{ kde } S \text{ je kořenem } \mathcal{G}\}.$$

Dokažte, že problém *BED* je NP-těžký.

NP-těžkost problému *BED* dokážeme redukcí z problému *3SAT*, tedy redukcí $\mathit{3SAT} \leq_p \mathit{BED}$.

Redukční funkce $f(w)$ je definovaná následovně. Není-li w kódem výrokové formule v 3CNF, pak klademe $f(w) = \langle \mathcal{G}_0, 1\rangle$, kde $\mathcal{G}_0 = (\{S\}, \{a\}, \{S \to a\}, S)$ je gramatika, ve které nelze odvodit ε. Ekvivalence $w \in \mathit{3SAT} \iff f(w) \in \mathit{BED}$ pro tato w jistě platí, neboť $w \notin \mathit{3SAT}$ a $f(w) \notin \mathit{BED}$.

Nyní se zaměříme na případ, kdy $w = \langle \varphi \rangle$ pro nějakou formuli φ v 3CNF. Nechť n je počet klauzulí ve φ a $X_1, \ldots, X_p$ jsou výrokové proměnné, které se vyskytují ve φ. Jistě platí $p \leq |\varphi|$. Nyní uvažme následující gramatiku $\mathcal{G} = (N, \Sigma, P, S)$:

- $N = \{S, T, F\} \cup \{X_i, X_i^?, X_i^T, X_i^F \mid 1 \leq i \leq p\}$, kde T, F reprezentují pravdivostní hodnoty *true*, *false* a pro každou výrokovou proměnnou máme neterminál X_i pro zápis této proměnné ve formuli, neterminál $X_i^?$ sloužící k nedeterministickému výběru valuace této proměnné a neterminály X_i^T, X_i^F značící, že proměnné X_i byla přiřazena hodnota *true*, resp. *false*
- $\Sigma = \{\wedge, \vee, \neg, (,)\}$
- $P = P_1 \cup P_2 \cup P_3 \cup P_4 \cup P_5 \cup P_6$, kde
  - Pravidlo přepisující S na φ (zapsanou pomocí znaků z $N \cup \Sigma$) a řetězec neterminálů sloužících k nedeterministickému výběru valuace:
  $$P_1 = \{S \to \varphi X_1^? X_2^? \ldots X_p^?\}$$
  - Pravidla pro nedeterministický výběr valuace:
  $$P_2 = \{X_i^? \to X_i^T,\ X_i^? \to X_i^F \mid 1 \leq i \leq p\}$$
  - Pravidla umožňující neterminálům X_i^T, X_i^F "projet" doleva přes celou větnou formu, každý výskyt proměnné X_i ve formuli přepsat na T nebo F dle zvolené valuace a následně se přepsat na ε:
  $$\begin{aligned} P_3 = \{&AX_i^T \to X_i^T A,\ AX_i^F \to X_i^F A \mid 1 \leq i \leq p,\ A \in (N \cup \Sigma) \setminus \{X_i\}\} \cup \\ \{&X_i X_i^T \to X_i^T T,\ X_i X_i^F \to X_i^F F \mid 1 \leq i \leq p\} \cup \\ \{&X_i^T \to \varepsilon,\ X_i^F \to \varepsilon \mid 1 \leq i \leq p\} \end{aligned}$$
  - Pravidla umožnující vyhodnotit negované literály:
  $$P_4 = \{\neg T \to F,\ \neg F \to T\}$$
  - Pravidla umožnující přepsat na T každou klauzuli s alespoň jedním literálem vyhodnoceným na T:
  $$\begin{aligned} P_5 = \{&(T \vee T \vee T) \to T,\ (T \vee T \vee F) \to T,\ (T \vee F \vee T) \to T, \\ &(T \vee F \vee F) \to T,\ (F \vee T \vee T) \to T,\ (F \vee T \vee F) \to T,\ (F \vee F \vee T) \to T\} \end{aligned}$$

– Pravidla umožnující přepsat libovolnou sekvenci tvaru $T \wedge T \wedge \ldots \wedge T$ na ε:

$$P_6 = \{T \wedge T \to T,\ T \to \varepsilon\}$$

Pravidla z $P_1 \cup P_2$ umožnují v $p+1$ krocích přepsat počáteční neterminál S na řetězec $\varphi\alpha$, kde $\alpha \in \{X_1^T, X_1^F\}.\{X_2^T, X_2^F\}\ldots\{X_p^T, X_p^F\}$ jednoznačně určuje valuaci proměnných. Každý neterminál X_i^T, X_i^F pak může pomocí pravidel z P_3 v nejvýše $|\varphi|+1$ krocích "projet" doleva přes celou formuli, přepsat výskyty X_i na odpovídající hodnotu T nebo F a zmizet. Tato fáze zabere nejvýše $p\cdot(|\varphi|+1)$ kroků. Pravidla v P_4, P_5, P_6 následně umožňují vyhodnotit negace, splněné klauzule přepsat na T a konjunkci libovolně mnoha T přepsat na ε, to vše v nejvýše $|\varphi|$ krocích. Je tedy zřejmé, že pokud existuje valuace splňující φ, pak lze v gramatice $\mathcal{G}$ odvodit ε v nejvýše $(|\varphi|+1)+p\cdot(|\varphi|+1)+|\varphi| \leq |\varphi|^2+3|\varphi|+1$ krocích.

Nyní předpokládejme, že v $\mathcal{G}$ lze odvodit ε. Jediným pravidlem pro S je pravidlo z P_1, po jehož aplikaci se do větné formy dostane n řetězců tvaru $\{(\}.(N \cup \{\vee, \neg\})^+.\{)\}$ (jeden pro každou klauzuli formule φ). Jelikož závorky lze odstranit jen pomocí pravidel z P_5, musel se každý řetězec výše uvedeného tvaru přepsat na řetězec tvaru $(V_1 \vee V_2 \vee V_3)$, kde $V_1, V_2, V_3 \in \{T, F\}$ a aspoň jedno z V_1, V_2, V_3 je T. To však mohlo nastat jen po aplikaci p pravidel z P_2 a dalších pravidel z P_3 a P_4. Aplikovaná pravidla z P_2 pak kódují valuaci, která je splňující pro φ.

Tím jsme dokázali, že platí

$$\langle\varphi\rangle \in \mathit{3SAT} \iff \langle\mathcal{G}, |\varphi|^2+3|\varphi|+1\rangle \in \mathit{BED}.$$

Pokud tedy definujeme $f(\langle\varphi\rangle) = \langle\mathcal{G}, |\varphi|^2+3|\varphi|+1\rangle$, vidíme, že ekvivalence $w \in \mathit{3SAT} \iff f(w) \in \mathit{BED}$ je splněna pro všechna slova w, která jsou kódem nějaké formule v 3CNF, a dohromady s úvodním odstavcem tak dostáváme, že je tato ekvivalence splněna pro všechna slova w.

Zbývá dokázat, že funkce f je vyčíslitelná deterministickým strojem v polynomiálním čase. Rozhodnout, zda vstupní slovo w je kódem nějaké formule φ v 3CNF, a zkonstruovat gramatiku $\mathcal{G}_0$ jistě lze deterministicky v polynomiálním čase. Nyní zanalyzujeme gramatiku $\mathcal{G}$ konstruovanou v případě, že $w = \langle\varphi\rangle$: Množina N je velikosti $\mathcal{O}(p)$, abeceda Σ je konstantní. Délka pravidla v P_1 je $\mathcal{O}(|\varphi|+p)$ a množiny $P_2, \ldots, P_6$ obsahují pouze pravidla konstantních délek, kterých je $\mathcal{O}(p)$ v případě P_2, $\mathcal{O}(p^2)$ v případě P_3 a konstantní počet v případě P_4, P_5, P_6. Jelikož $p \leq |\varphi|$, je velikost konstruované gramatiky v $\mathcal{O}(|\varphi|^2)$. Její konstrukce je přitom přímočará.