Ilustrační materiál k závěrečnému testu z předmětu "Teorie jazyků a automatů"

Při zahájení zkoušky obdržíte zadání s následující hlavičkou:

============

PŘÍJMENÍ a JMÉNO:
Login studenta:
DATUM:

Závěrečný test z předmětu "Teorie jazyků a automatů"

Doba trvání: **100 minut** Max. zisk: **90 bodů**

Obecné pokyny:
Po obdržení testu ihned do pravého horního rohu napište čitelně své příjmení a jméno, svůj login a datum.
K řešení úloh používejte pomocné papíry. Požadovaný *výsledek* zapište odpovídajícím způsobem do odpovídajícího místa přímo *na zadání testu*. V případě, že se vám (mimořádně) stane, že se řešení nevejde do vyhrazeného místa, pište jej na *zvláštní list* – u příslušné úlohy (např. s číslem [8]) pak napište výrazné ZP (tj. zvláštní papír) a výsledek na zvláštním papíru označte zřetelně číslem úlohy (v našem příkladu tedy číslem 8). (Kromě doplněného zadání a případných označených zvláštních listů nakonec odevzdáte i pomocné papíry, aby se odstranily případné pochybnosti o tom, jak jste k výsledkům dospěli; k obsahu pomocných papírů může být při bodovém hodnocení přihlédnuto.)
Součet maxim. bodových hodnocení příkladů 1 – 11 je 90. Příklad 12 je doplňkový, jeho úspěšným vyřešením můžete bodové hodnocení příkladů 1 – 11 zvednout max. o 15 *procent* (celkově však nelze překročit 90).
Úlohy řešte v libovolném pořadí, rozvrhněte si dobře čas.
Poznámka. Řazení slov v dané abecedě. Pro libovolnou abecedu $\Sigma = \{a_1, a_2, \ldots, a_n\}$, jejíž prvky chápeme jako uspořádané, $a_1 < a_2 < \ldots < a_n$, seřazujeme všechna slova ze Σ^* podle délky a v rámci stejné délky "abecedně" (tj. v souladu s uspořádáním prvků abecedy Σ). Např. pro $\Sigma = \{a, b\}$ vypadá posloupnost slov sestavená podle vzestupného pořadí takto: $\varepsilon, a, b, aa, ab, ba, bb, aaa, aab, \ldots$. Každému slovu je přiřazeno pořadové číslo; ε má číslo 1, a má číslo 2 atd. Uspořádání slov v Σ^* určuje uspořádání slov libovolného jazyka $L \subseteq \Sigma^*$; je takto jednoznačně určeno 1. slovo jazyka L (tj. slovo z L s nejmenším pořadovým číslem v rámci seřazení slov ze Σ^*), 2. slovo jazyka L, 3. slovo jazyka L, atd.

==========================

Pak bude následovat zadání 12 příkladů, z nichž poslední je doplňkový. Následující text vám má dát **přibližnou** představu, co se asi dá očekávat – **rozhodně ale není v žádném smyslu vyčerpávající**.

První 4 příklady (celk. max. 16 bodů) patří do pomyslného bloku prověřujícího

- Porozumění pojmu formální jazyk a především operacím s jazyky.
- Porozumění pojmům souvisejícím s konečnými automaty (determ. i nedeterm. verze), regulárními výrazy a regulárními gramatikami.
- Porozumění pojmům souvisejícím s bezkontextovými gramatikami, zásobníkovými automaty, Turingovými stroji.

Do tohoto bloku patří např. příklady jako

=====================

Doplňte níže uvedené definice (L^R označuje zrcadlový obraz jazyka L)

- $L_1 \cdot L_2 = \{w \mid \exists$
- $L^0 =$
- $L^* =$
- $L^R = \{w \mid \exists$

————————-

Jazyky L_1, L_2 jsou definovány takto ($|w|_a$ značí počet výskytů symbolu a ve slově w; připomeňme také, že 0 je sudé číslo):

$$L_1 = \{w \in \{a,b\}^* \mid |w|_a \text{ je sudé }\},$$
$$L_2 = \{w \in \{a,b\}^* \mid |w|_a = |w|_b\}.$$

Zjistěte: (zde $L_2 \backslash L_1$ označuje levý kvocient jazyka L_1 podle L_2)

- 5. slovo v $L_1 \cup L_2$
- 4. slovo v $L_1 \cap L_2$
- Kolik slov délky nejvýš 4 je v $\overline{L_1}$
- 5. slovo v $L_2 \backslash L_1$

————————————————-

Mějme ZNKA $A = (Q, \Sigma, \delta, I, F)$. Definujte zobrazení $E : \mathcal{P}(Q) \to \mathcal{P}(Q)$, které přiřazuje každé $K \subseteq Q$ množinu $E(K) = \{q' \mid$ z něj. $q \in K$ vede v grafu automatu A cesta do q' obsahující pouze hrany označené $\varepsilon\}$; konkrétně definujte $E(K)$ jako nejmenší množinu splňující tyto dvě podmínky (ty doplňte):

1.
2.

————————

K NKA zadanému následující tabulkou

	a	b
1	$\{5,7\}$	$\{2,3\}$
2	$\{3,6\}$	$\{1,5,7\}$
3	$\emptyset$	$\{6\}$
4	$\{1\}$	$\{3,7\}$
5	$\{2,4\}$	$\emptyset\}$
6	$\{1,2\}$	$\{6\}$
7	$\{6,7\}$	$\emptyset$

- zjistěte $\delta(1, baab) = \{$
- zjistěte jedno ze slov w délky 7 s max. možnou mohutností množiny $\delta(7, w)$
- Když $\{6\}$ je množina počátečních stavů a $\{4,5\}$ množina přijímajících stavů, a vzniklý NKA označíme A, jaké bude 3. slovo v $L(A)$?

Mějme bezkontextovou gramatiku $G = (\Pi, \Sigma, S, P)$. Definujte na množině $(\Pi \cup \Sigma)^*$ relaci $\Rightarrow$ (odvození v jednom kroku):

$\alpha \Rightarrow \beta$ právě když

Mějme gramatiku (s počátečním neterminálem E)

1: $E \to T + E$
2: $E \to T$
3: $T \to T * F$
4: $T \to F$
5: $F \to (E)$
6: $F \to a$

V derivačním stromě pro $(a + a) * a + a$ vyjmenujte ohodnocení všech vrcholů se vzdáleností 3 od kořene zleva doprava (délka cesty v grafu je počet hran, kterými prochází; vzdálenost dvou uzlů v grafu je délka nejkratší cesty vedoucí z jednoho do druhého):

Jaké číslo má pravidlo, které se použije v 5. kroku levé derivace slova $(a + a) * a + a$?

==============================

Další pomyslný blok sestávající ze 2 příkladů (max. 26 bodů) má prověřit

- Schopnost návrhu konečných automatů, regulárních výrazů, regulárních gramatik.
- Schopnost návrhu bezkontextových gramatik, zásobníkových automatů, Turingových strojů.

Do tohoto bloku patří např. příklady jako

================================

Sestrojte redukovaný KA rozpoznávající jazyk

$$L = \{w \in \{a, b\}^* \mid w \text{ začíná a končí stejným symbolem } \}$$

Zadejte jej tabulkou v normovaném tvaru.

Napište (co nejjednodušší) regulární výraz tak, aby reprezentoval jazyk

$L = \{w \in \{a, b\}^* \mid |w|_a$ je sudé a w obsahuje podslovo $bab\}$

Sestrojte zásobníkový automat rozpoznávající jazyk $L = \{ u \in \{a, b, c\}^* \mid$ po vynechání všech výskytů symbolu c z u dostaneme slovo ve tvaru $w(w)^R \}$.

==========================

Další pomyslný blok sestávající ze 2 příkladů (max. 24 bodů) má prověřit zejména

znalost algoritmů používaných v (konstruktivních) důkazech tvrzení, speciálně např. algoritmů

- prokazujících uzávěrové vlastnosti třídy regulárních či bezkontextových jazyků,
- používaných při převodech do normálních (redukovaných) tvarů, nedeterministické verze na deterministickou, či obecně při převodu jednoho prostředku na jiný (ekvivalentní)

Do tohoto bloku patří např. příklady jako

=======================================

K ZNKA zadanému následující tabulkou, kde $\{1,5\}$ je množina počátečních stavů a $\{4,6\}$ je množina přijímajících stavů,

	a	b	ε
1	$\{1,2\}$	$\{1\}$	$\emptyset$
2	$\emptyset$	$\{3\}$	$\emptyset$
3	$\emptyset$	$\{4\}$	$\emptyset$
4	$\{4\}$	$\{4\}$	$\emptyset$
5	$\{7\}$	$\emptyset$	$\{6\}$
6	$\emptyset$	$\{8\}$	$\emptyset$
7	$\{5\}$	$\emptyset$	$\emptyset$
8	$\emptyset$	$\{6\}$	$\emptyset$

sestrojte ekvivalentní deterministický redukovaný KA v normovaném tvaru.

Mějme jazyky

$L_1 = \{w \in \{a,b\}^* \mid w \text{ obsahuje podslovo } bab \text{ nebo } abba\}$

$L_2 = \{w \in \{a,b\}^* \mid |w|_a \text{ je sudé }\}$.

Níže uvedený ZNKA, kde $\{1\}$ je množina počátečních stavů a $\{12\}$ je (jednoprvková) množina přijímajících stavů, nerozpoznává jazyk $L_1 \cdot L_2$. Uveďte slovo s nejmenším pořadovým číslem, které to dokazuje. Dále proveďte minimální změny v tabulce, které zajistí, že upravený automat rozpoznává $L_1 \cdot L_2$; množinu přijímajících stavů přitom můžete změnit, množina počátečních stavů musí zůstat stejná.

	a	b	ε
1	$\emptyset$	$\emptyset$	$\{2,3\}$
2	$\{2\}$	$\{2,4\}$	$\emptyset$
3	$\{3,7\}$	$\{3\}$	$\emptyset$
4	$\{5\}$	$\emptyset$	$\emptyset$
5	$\emptyset$	$\{6\}$	$\emptyset$
6	$\{6\}$	$\{6\}$	$\emptyset$
7	$\emptyset$	$\{8\}$	$\emptyset$
8	$\emptyset$	$\{9\}$	$\emptyset$
9	$\{10\}$	$\emptyset$	$\emptyset$
10	$\{10\}$	$\{10\}$	$\emptyset$
11	$\{12\}$	$\{11\}$	$\emptyset$
12	$\{11\}$	$\{12\}$	$\emptyset$

Níže jsou uvedená pravidla bezkontextové gramatiky. Neterminál X nazveme *dobrý*, jestliže gramatika generuje neprázdný jazyk v případě, že X je prohlášen počátečním neterminálem. Vyjmenujte všechny dobré neterminály:

$$
\begin{aligned}
A_1 &\to aA_1b \mid aA_2bb \mid \varepsilon \\
A_2 &\to aA_2A_3 \mid bA_3 \\
A_3 &\to aA_2b \mid A_3A_3 \\
A_4 &\to A_4A_4 \mid cA_1 \\
A_5 &\to aA_5 \mid A_6A_7 \mid A_8A_9 \\
A_6 &\to aA_9b \mid A_6A_9 \mid A_7A_8 \\
A_7 &\to bA_5b \mid A_7A_7 \\
A_8 &\to A_7A_6 \mid A_6A_5b \\
A_9 &\to A_6A_7A_8A_9 \mid \varepsilon
\end{aligned}
$$

===============================

Další blok obsahuje 3 příklady (max. 24 bodů) a má prověřit

porozumění studovaným pojmům a vztahům mezi nimi (porozumění větám, tvrzením, lemmatům, poznámkám, ...); speciálně např. schopnost prokázat nepříslušnost daného jazyka k regulárním či bezkontextovým jazykům (porozumění tvrzení, důkazu a použití pumping lemmat).

Do tohoto bloku patří např. příklady jako

==================================

Zjistěte, které z daných jazyků

jsou regulární:

jsou bezkontextové, ale ne regulární:

nejsou bezkontextové:

$L_1 = \{w \in \{a,b\}^* \mid |w|_a = |w|_b\}$
$L_2 = \{w \in \{a,b\}^* \mid |w|_a \text{ je sudé }\}$
$L_3 = \{w \in \{a,b\}^* \mid w \text{ obsahuje podslovo } abba\}$
$L_4 = \{w \in \{a,b,c\}^* \mid |w|_a = |w|_b = |w|_c\}$

$L_5 = \{w \in \{a,b\}^* \mid |w|_a \text{ je prvočíslo}\}$
$L_6 = \{0^m1^n \mid m \leq 2n\}$
$L_7 = \{0^m1^n0^m \mid m = 2n\}$

Připomeňme si hru dvou hráčů pro daný jazyk L:

1. A zvolí $n \in \mathbb{N}$
2. B zvolí slovo z tž. $z \in L$ a $|z| \geq n$
3. A zvolí u, v, w, x, y tž. $z = uvwxy$, $|vwx| \leq n$ a $vx \neq \varepsilon$
4. B zvolí $i \geq 0$
5. *Výsledek:* je-li $uv^iwx^iy \in L$, pak vyhrál A, v případě $uv^iwx^iy \notin L$ vyhrál B.

Víme, že jestliže B má vítěznou strategii, pak L není bezkontextový. Popište vítěznou strategii B pro jazyk $L = \{w \in \{a, b, c\}^* \mid w = a^n b^n c^n$ pro nějak. $n \geq 0\}$:

1. A zvolí $n \in \mathbb{N}$
2. B zvolí slovo $z =$
3. A zvolí u, v, w, x, y tž. $z = uvwxy$, $|vwx| \leq n$ a $vx \neq \varepsilon$
4. B zvolí $i =$
5. *Výsledek:* B vyhrál, protože ....

Dokažte tuto variantu pumping lemmatu pro regulární jazyky:

Nechť L je regulární jazyk. Pak nutně existuje n tž. pro každé slovo $z \in L$ lze pro každé jeho rozdělení $z = uvw$, kde v je délky alespoň n, psát $v = v_1 v_2 v_3$, kde $|v_2| \geq 1$ a pro vš. $i \geq 0$ je $uv_1 v_2^i v_3 w \in L$.

==

K příkladům 1 – 11 z předchozích bloků bude přiřazen doplňkový příklad 12; jeho úspěšným vyřešením můžete bodové hodnocení příkladů 1 – 11 zvednout max. o 15 *procent* (celkově však nelze překročit 90).

Příklad má prověřit schopnost sestavit důkaz (speciálně induktivní důkaz) a může vypadat podobně jako např. následující příklad:

==

K bezkontextové gramatice $G = (\Pi, \Sigma, S, P)$ uvažujme ZA $M = (\{q_0\}, \Sigma, \Pi \cup \Sigma, \delta, q_0, S)$, kde pro každé $X \in \Pi$ je $\delta(q_0, \varepsilon, X) = \{(q_0, \alpha) \mid (X \to \alpha) \in P\}$ a pro každé $a \in \Sigma$ je $\delta(q_0, a, a) = \{(q_0, \varepsilon)\}$; jiným argumentům přiřazuje δ prázdnou množinu.

Nechť nyní $S \Rightarrow_G^n \beta$ označuje, že řetězec β ($\in (\Pi \cup \Sigma)^*$) lze z S odvodit levou derivací délky nejvýše n. Zápis $K_1 \vdash_M^n K_2$ označuje, že z konfigurace K_1 (automatu M) lze odvodit konfiguraci K_2 výpočtem, v němž se použije nejvýše n ε-kroků.

Dokažte indukcí podle n toto

TVRZENÍ: pro slovo $u\alpha$, kde $u \in \Sigma^*$, $\alpha \in \Pi(\Pi \cup \Sigma)^*$ nebo $\alpha = \varepsilon$, platí $S \Rightarrow_G^n u\alpha$ právě tehdy, když platí $(q_0, u, S) \vdash_M^n (q_0, \varepsilon, \alpha)$.

Základ indukce, $n = 0$:

- Implikace "$\Rightarrow$":

  Jestliže $S \Rightarrow_G^0 u\alpha$, pak nutně $u = \varepsilon$, $\alpha = S$; jelikož platí $(q_0, \varepsilon, S) \vdash_M^0 (q_0, \varepsilon, S)$, máme v našem případě $(q_0, u, S) \vdash_M^0 (q_0, \varepsilon, \alpha)$.

- Implikace "$\Leftarrow$":

  Jestliže $(q_0, u, S) \vdash_M^0 (q_0, \varepsilon, \alpha)$, pak nutně

Indukční krok. Předpokládáme, že tvrzení platí pro $0, 1, 2, \ldots, n$. Ukažme, že platí pro $n + 1$:

- Implikace "$\Rightarrow$":

  Jestliže $S \Rightarrow_G^{n+1} u\alpha$, pak existují u', X, α', v, β, kde $(X \to v\beta) \in P$, $\beta \in \Pi(\Pi \cup \Sigma)^*$ nebo $\beta = \varepsilon$, takové, že $S \Rightarrow_G^n u'X\alpha' \Rightarrow_G u'v\beta\alpha'$, kde $u'v\beta\alpha' = u\alpha$ (tj. $u'v = u$, $\beta\alpha' = \alpha$). Podle indukčního předpokladu platí $(q_0, u', S) \vdash_M^n (q_0, \varepsilon, X\alpha')$, tedy $(q_0, u'v, S) \vdash_M^n (q_0, v, X\alpha')$. Jelikož dále platí dostáváme $(q_0, u, S) \vdash_M^{n+1} (q_0, \varepsilon, \alpha)$.

- Implikace "$\Leftarrow$":

Ukažte, že z uvedeného TVRZENÍ plyne $L(M) = L(G)$.

================================