Programovací techniky TS, TS s omezeními a rozšířeními, Univerzální TS

Jan Konečný

1. října 2013

Doplnění:
„TS zastaví pro w" – znamená, že TS slovo w přijme nebo zamítne; prostě necyklí.

„TS T bere na vstup slova ve tvaru x." – znamená, že stroj zamítá všechna slova, která v tomto tvaru nejsou. Vždycky půjde o jednoduchý tvar (k otestování stačí KA).

Dnes:

- programovací techniky (konečně hodnotové proměnné, modularita, více stop na jedné pásce)
- TS s omezeními a TS s rozšířeními (stroj s unikátními koncovými konfiguracemi, TS který se nepokusí přejet levý okraj pásky, TS nikdy nezapisuje ␣)
- Varianty TS (stroj s více páskami, stroj s oboustranně nekonečnou páskou, nedeterminismus)
- Univerzální TS,
- enumerator, lineárně ohraničený automat (LBA).

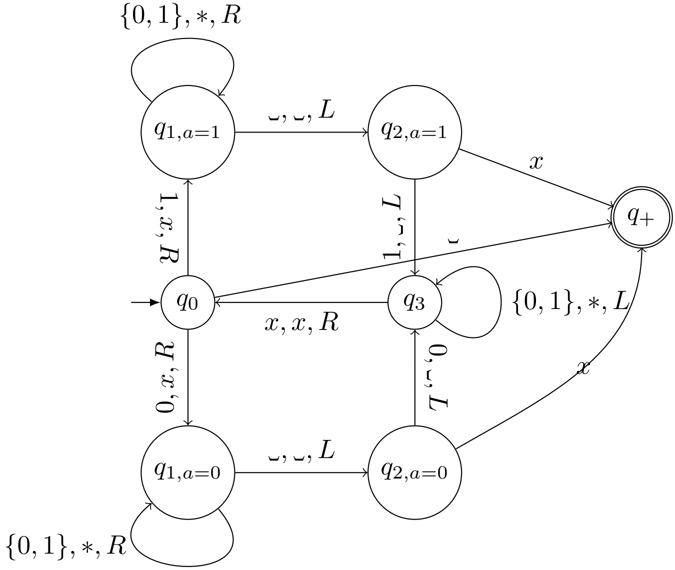
Hned po ránu Konečný cosi chce.

Začněme 'minicvičením':

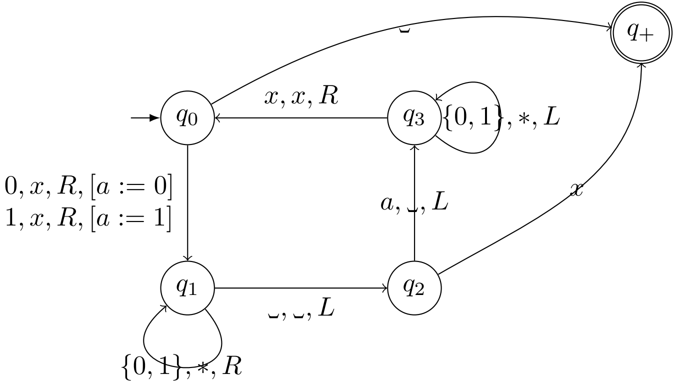
Navrhněte TS, který přijímá jazyk palindromů nad abecedou $0, 1$ (= slova, která se čtou zepředu stejně jako zezadu:
$\varepsilon, 0, 1, 00, 11, 000, 010, 101, 111, 0000, 0110, 1001, 1111, \dots)$

Programovací techniky – konečně hodnotové proměnné

Stavy můžeme používat k zapamatování si *konečně mnoha* hodnot:

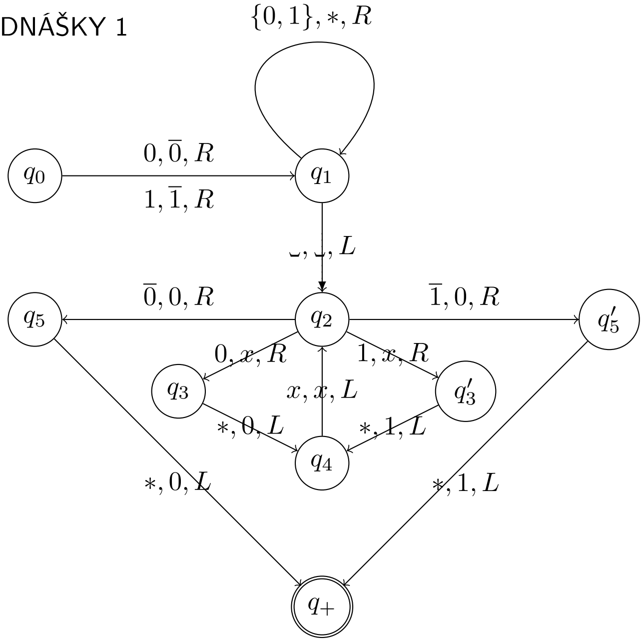


Totéž s použitím konečně hodnotové proměnné a.



Toto ukazuje základní myšlenku použití konečně hodnotových proměnných v TS.

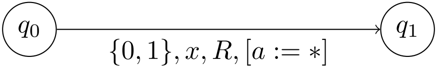
Stroj z PŘEDNÁŠKY 1



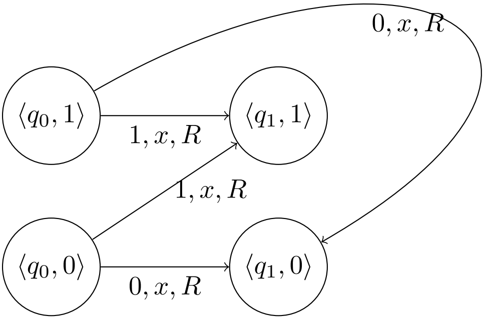
Jaktože si to můžu dovolit:
V tomto případě stačí uvažovat, že množina stavů je

$$Q := \{q_0, \ldots, q_n\} \times \{0, 1\}$$

(pořád je to konečné množství).
a následně uvažovat, že například hrana



znamená

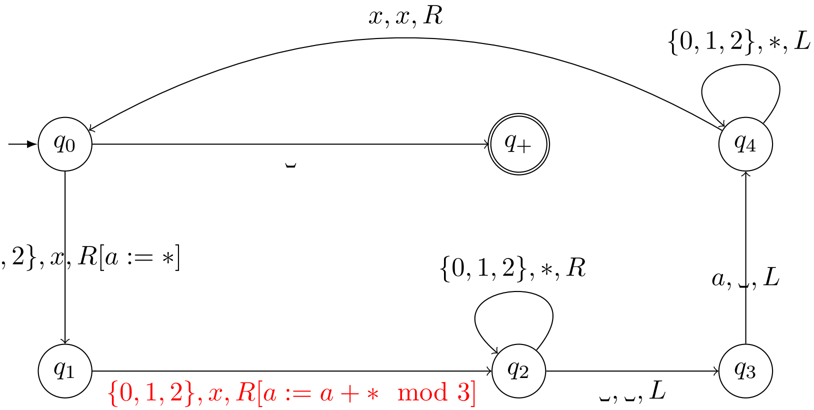


. . . z tohoto už jde snadno vytvořit obecný případ.

Zajímavější příklad:

TS rozpoznávající jazyk L nad $\{0, 1, 2\}$, t.ž.

- $\varepsilon \in L$,
- $ab\omega c \in L$, pokud $c = a + b \mod 3$ a $\omega \in \Sigma^*$.



Obecně můžeme uvažovat TS s libovolným konečným množstvím konečně hodnotových proměnných.

Můžeme tedy uvažovat, že když máme množinu Q_{c} 'kontrolních stavů', konečně mnoho proměnných $A_1, \dots, A_n$ a k nim příslušné konečné domény $D_1 \dots, D_n$, je množina stavů TS $Q_{\mathrm{c}} \times D_1 \times \cdots \times D_n$.

Programovací techniky TS: Modularita, procedury

TS mohou být využity jako moduly v dalších TS.

Programovací techniky: Více stop na jedné pásce

Uvažujme prvky množiny Γ jako prvky kartézského součinu $\Gamma_1 \times \cdots \times \Gamma_n$, kde Γ_i jsou abecedy. Přechodovou funkci δ pak můžeme uvažovat jako funkci

$$Q \times (\Gamma_1 \times \cdots \times \Gamma_n) \to Q \times (\Gamma_1 \times \cdots \times \Gamma_n) \times \{L, R\}$$

Každé políčko pásky Turingova stroje tak obsahuje n-tici prvků z $(\Gamma_1 \times \cdots \times \Gamma_n)$. Můžeme ji tedy chápat jako pásku, která se skládá z n stop; pokud je na políčku pásky zapsána n-tice $\langle x_1, \ldots, x_i, \ldots, x_n \rangle$, říkáme, že na i-té stopě je zapsán symbol x_i.
Mluvíme pak o Turingově stroji s vícestopou páskou (resp. přesněji o Turingově stroji s n-stopou páskou).

Příklad uvidíme později v této lekci.

Turingovy stroje s omezeními

Ekvivalence TS

Definice

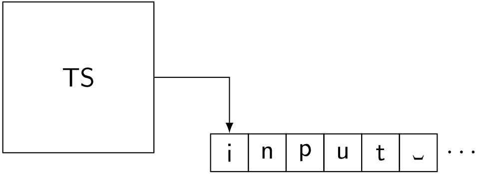
Dva TS T a S nazveme *ekvivalentní* pokud pro každé slovo w

- T přijímá w p.k. S přijímá w;
- T zamítá w p.k. S zamítá w;
- (T cyklí pro w p.k. S cyklí pro w).

Ukážeme, ke každému TS existuje ekvivalentní

- TS, který se nikdy nepokusí přejet levý okraj pásky,
- TS, který nikdy nezapíše na pásku $\llcorner\!\lrcorner$,
- TS, který po sobě uklízí.

TS, který se nikdy nepokusí přejet levý okraj pásky



Definice

Turingovým strojem, který se nikdy nepokusí přejet levý okraj pásky, rozumíme TS, u kterého při výpočtu nad jakýmkoli slovem nedojde k tomu že,

- je v konfiguraci $(q, a\omega, 0)$,
- existuje přechod $\delta(q, a) = (q', b, L)$.

Tedy nikdy nenastane situace, kdy by byla hlava nad nejlevějším políčkem pásky a přechodová fuknce by určovala pohyb vlevo.

Věta

Ke každému TS T existuje ekvivalentní TS T', který se při výpočtu nikdy nepokusí přejet levý okraj pásky.

Idea důkazu.

Nejdříve k TS $T = \langle Q, \Sigma, \Gamma, \delta, q_0, q_+, q_- \rangle$ zkonstruujeme ekvivalentní TS $X = \langle Q, \Sigma \cup \{\triangleright\}, \Gamma \cup \{\triangleright\}, \delta_X, q_0, q_+, q_- \rangle$, který se nikdy nepokusí přejet levý okraj pásky, ale na vstup bere jen slova ve speciálním tvaru.

Uvažujme, že všechna vstupní slova pro X' jsou ve tvaru $\triangleright\omega$, kde $\omega \in \Sigma^*, \triangleright \notin \Sigma$.

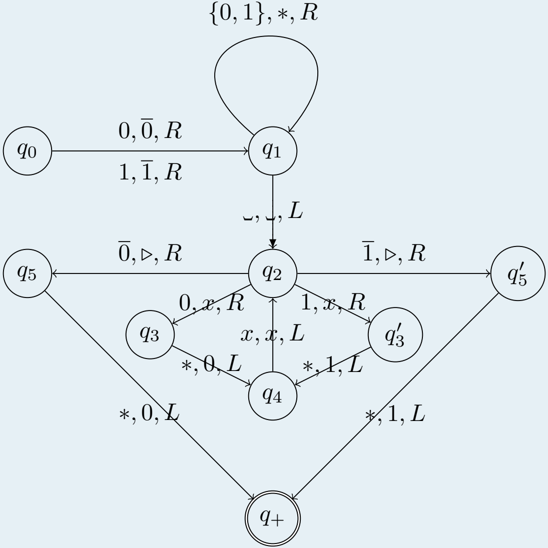
$$\delta_X(q, a) = \begin{cases} (q, \triangleleft, R) & \text{pokud } a = \triangleleft, \\ \delta(q, a) & \text{v ostatních případech,} \end{cases}$$

pro všechna $\langle q, a \rangle \in Q \times \Gamma \cup \{\triangleright\}$.

Zjevně X' se nikdy nepokusí přejet levý okraj pásky, protože na nejlevějším políčku je vždy $\triangleright$, nikdy to není přepsáno, a hlava se při čtení tohoto symbolu může pohnout jen doprava. $\square$

Idea důkazu (pokračování).

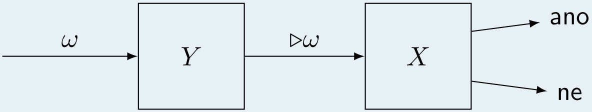
Máme tedy X a taky známe stroj vyčislující přidání počáteční nuly (z první přednášky). Označme jeho lehkou úpravu Y.



Idea důkazu (pokračování).

Všimněte si, že Y se také nikdy nepokusí přejet levý okraj pásky.

T' sestrojíme jako kombinaci X a Y:



□

Zjevně tento stroj T' bude přijímat tentýž jazyk jako T, a je jisté, že se nikdy nepokusí překročit levý okraj pásky.

Proč je to důležité: zjednodušení definice kroku výpočtu; není potřeba zvlášť řešit, co se děje, když je hlava nad nejlevějším políčkem pásky a má se pohnout doleva.
(viz následující slajd)

Definice

Krok výpočtu TS je definován jako binární relace na množině konfigurací: Nechť $(q, a_1 \dots a_n, i)$ je taková konfigurace T, kde $q \neq q_\pm, n \in \mathbb{N}_0, a_0, \dots, a_n \in \Gamma, i \leq n$.

(a) Je-li $1 \leq i \leq n$ a $\delta(q, a_i) = (q', b, \mathrm{L})$, pak

$$(q, a_0, \dots a_n, i) \vdash (q', a_0 \dots a_{i-1} b a_{i+1} \dots a_n, i-1).$$

(b) Je-li $\delta(q, a_0) = (q', b, \mathrm{L})$, pak

$$(q, a_0 \dots a_n, 0) \vdash (q', b a_1 \dots a_n, 0).$$

(c) Je-li $\delta(q, a_i) = (q', b, \mathrm{R})$, pak

$$(q, a_0 \dots a_n, i) \vdash (q', a_0 \dots a_{i-1} b a_{i+1} \dots a_n, i+1).$$

Definice (zjednodušená)

Krok výpočtu TS je definován jako binární relace na množině konfigurací: Nechť $(q, a_1 \dots a_n, i)$ je taková konfigurace T, kde $q \neq q_\pm, n \in \mathbb{N}_0, a_0, \dots, a_n \in \Gamma, i \leq n$.

$$(q, a_0, \dots a_n, i) \vdash (q', a_0 \dots a_{i-1} b a_{i+1} \dots a_n, i + d),$$

pokud $\delta(q, a_i) = (q', b, \mathrm{D})$, kde $d = -1$ pro $\mathrm{D} = L$, $d = +1$ pro $\mathrm{D} = R$.

Důležitá věc: funkce f, která provádí transformaci $[T] \mapsto [T']$ popsanou v předchozí idei důkazu je vyčíslitelná.

TS, který nikdy nezapíše na pásku ␣

Turingovým strojem, který nikdy nezapíše na pásku ␣ rozumíme TS, který nemá žádný přechod ve tvaru

$$q \xrightarrow[a,\, \textvisiblespace,\, D]{} q'$$

Věta

Ke každému TS T existuje ekvivalentní TS T', který nikdy nezapíše na pásku $\text{\textvisiblespace}$.

Idea důkazu.

K TS T sestrojíme T' tak, že do páskové abecedy Γ přidáme $\overline{\text{\textvisiblespace}}$, a upravíme přechodovou funkci tak, že každý přechod ve tvaru:

$$\delta(q, a) = (q', \text{\textvisiblespace}, D)$$

nahradíme přechodem

$$\delta(q, a) = (q', \overline{\text{\textvisiblespace}}, D).$$

Poté, ke každému přechodu ve tvaru

$$\delta(q, \text{\textvisiblespace}) = (q', b, D)$$

vytvoříme nový přechod

$$\delta(q, \overline{\text{\textvisiblespace}}) = (q', b, D).$$

Výsledný TS nikdy nebude zapisovat $\text{\textvisiblespace}$; místo něj zapisuje na pásku $\overline{\text{\textvisiblespace}}$, který má stejný význam. $\square$

Proč je to důležité? Prázdný symbol odteď bude reprezentovat příznak začátku dosud nepoužité části pásky.

Důležitá věc (opět): funkce f, která provádí transformaci $[T] \mapsto [T']$ popsanou v předchozí idei důkazu je vyčíslitelná.

Turingův stroj, „který po sobě uklízí“

Věta

Ke každému TS T existuje TS T', který zastaví pouze v konfiguraci $\langle q_+, \varepsilon, 0\rangle$ nebo $\langle q_-, \varepsilon, 0\rangle$.

Idea důkazu.

K TS T sestavíme nejdříve stroj S, který se nikdy nepokusí přejet levý okraj pásky a který nikdy nezapíše na pásku ␣. U TS S můžeme snadno detekovat začátek pásky podle zarážky $\triangleright$ a konec použitého úseku pásky ␣. TS S zkombinujeme s procedurami M_+ a M_-, které provedou mazání (tj. dojede na konec nepoužitého úseku pásky, pak jede doleva a vše přepisuje prázdnými symboly ␣, dokud nenarazí na ␣, ten smaže a zastaví v q_+ (resp. v q_-). □

Proč je to důležité? Odteď můžeme uvažovat ze existuje nejvýše jedna přijímající konfigurace a nejvýše jedna zamítající konfigurace, do které se TS může během výpočtu dostat.

TS s rozšířeními

Varianty TS: TS s instrukcí stop

Definice

Turingův stroj s instrukcí stop je struktura $\langle Q, \Sigma, \Gamma, \delta, q_0, q_+, q_- \rangle$ daná:

1. neprázdnou konečnou množinou stavů Q,
2. vstupní abecedou Σ, t.ž. $\llcorner\!\lrcorner \notin \Sigma$,
3. páskovou abecedou Γ, t.ž. $\Sigma \subset \Gamma$, $\llcorner\!\lrcorner \in \Gamma$,
4. přechodovou funkcí $\delta : Q \times \Gamma \to Q \times \Gamma \times \{L, S, R\}$,
5. počátečním stavem $q_0 \in Q$,
6. přijímacím stavem $q_+ \in Q$
7. zamítacím stavem $q_- \in Q$. $q_+ \neq q_-$.

Poznámka

Odpovídajícím způsobem se upraví definice kroku výpočtu, vlastně se přidá bod

(d) Je-li $\delta(q, a_i) = (q', b, \mathrm{S})$, pak

$$(q, a_0 \dots a_n, i) \vdash (q', a_0 \dots a_{i-1} b a_{i+1} \dots a_n, i).$$

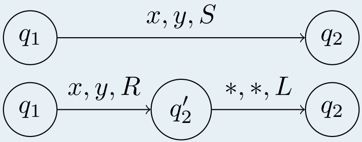
Varianty TS: TS s instrukcí stop

Věta

Ke každému TS T s instrukcí STOP existuje ekvivalentní (klasický) TS T'.

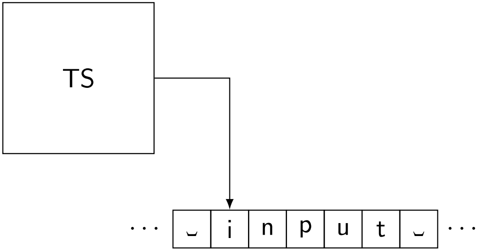
Důkaz.

Ke každému TS M s instrukcí STOP existuje ekvivalentní TS M' (bez STOP):
Pro $M = \langle Q, \Sigma, \Gamma, \delta, q_0, q_+, q_- \rangle$ sestavíme $M' = \langle Q', \Sigma, \Gamma, \delta, q_0, q_+, q_- \rangle$, t.ž.
$Q' = \{q, q' \mid q \in Q\}$



Ke každému TS M' (bez STOP) existuje ekvivalentní TS M s instrukcí STOP: *triviální*. □

Varianty TS: TS s oboustranně nekonečnou páskou



Poznámka

Definice TS s oboustranně nekonečnou páskou, jakožto struktury $\langle Q, \Sigma, \Gamma, \delta, q_0, q_+, q_- \rangle$, je totožná s definicí klasického TS; rozdíl bude v definici konfigurace (nevystačíme si s $\langle q, \omega, i \rangle \in Q \times \Gamma^* \times \mathbb{N}_0$) a výpočtu (není zarážení o levý okraj).

Konfigurace – použijeme alternativu z PŘEDNÁŠKY 1:

> *Alternativně budeme konfiguraci zapisovat jako řetězec $\alpha q \beta \in \Gamma^* Q \Gamma^*$. Konfigurace $\alpha q \beta$ přestavuje status stroje, který má na pásce zapsán řetězec $\alpha\beta$, hlava je nad prvním symbolem řetězce β, řídicí jednotka je ve stavu q.*

U TS s oboustranně nekonečnou páskou považujeme $\alpha q \beta$ za totožnou s $\alpha q \beta \llcorner\!\lrcorner$ a s $\llcorner\!\lrcorner \alpha q \beta$.

Definice

Krok výpočtu TS (s oboustranně nekonečnou páskou) je definován jako binární relace na množině konfigurací: Nechť $\alpha a q b \beta$ je taková konfigurace T, kde $q \neq q_{\pm}, \alpha, \beta \in \Gamma^*, a, b \in \Gamma$.

(a) Je-li $\delta(q, a) = (q', x, \mathrm{L})$,

$$\alpha a q b \beta \vdash \alpha q' a x \beta.$$

(b) Je-li $\delta(q, a_1) = (q', b, \mathrm{R})$, pak

$$\alpha a q b \beta \vdash \alpha a x q' \beta.$$

Věta

Ke každému TS T s oboustranně nekonečnou páskou existuje ekvivalentní TS T' (klasický, se zleva omezenou páskou).

Idea důkazu.

Sestavíme TS T' s dvoustopou páskou; z levého 'kusu' pásky uděláme horní stopu. V proměnné si budeme pamatovat, jestli je relevatní horní stopa, nebo dolní stopa (tedy jestli původní stroj byl na levém kusu pásky, nebo na pravém). $\square$

Věta

Ke každému TS T existuje ekvivalentní TS T' s oboustranně nekonečnou páskou.

Důkaz.

Dokázáno existencí ekvivalentního TS, který se nikdy nepokusí přejet levý okraj pásky. □

TS s více páskami

Definice

Turingův stroj s k páskami je struktura $\langle Q, \Sigma, \Gamma, \delta, q_0, q_+, q_- \rangle$ daná:

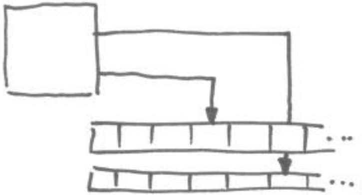
1. neprázdnou konečnou množinou stavů Q,
2. vstupní abecedou Σ, t.ž. $\llcorner\!\lrcorner \notin \Sigma$,
3. páskovou abecedou Γ, t.ž. $\Sigma \subset \Gamma$, $\llcorner\!\lrcorner \in \Gamma$,
4. přechodovou funkcí $\delta : Q \times \Gamma^k \to Q \times [\Gamma \times \{L, R\}]^k$,
5. počátečním stavem $q_0 \in Q$,
6. přijímacím stavem $q_+ \in Q$
7. zamítacím stavem $q_- \in Q$. $q_+ \neq q_-$.

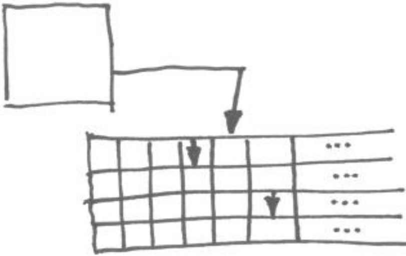
Věta

Ke každému TS s více páskami existuje ekvivalentní (klasický) TS.

Idea důkazu.

Ke TS s k páskami sestavíme klasický TS s páskou o $2k$ stopách; na lichých stopách budeme uchovávat obsah původních k pásek. Na sudých stopách budeme uchovávat pozice hlav příslušným k těm k páskám. □





Varianty TS: Nedeterministický TS

Poznámka

Obdobný rozdíl jako mezi KDA a KNA): přechodová funkce ve tvaru

$$\delta : Q \times \Gamma \to 2^{Q \times \Gamma \times \{L,R\}}$$

Definice

Turingův stroj s instrukcí stop je struktura $\langle Q, \Sigma, \Gamma, \delta, q_0, q_+, q_- \rangle$ daná:

1. neprázdnou konečnou množinou stavů Q,
2. vstupní abecedou Σ, t.ž. $\sqcup \notin \Sigma$,
3. páskovou abecedou Γ, t.ž. $\Sigma \subset \Gamma$, $\sqcup \in \Gamma$,
4. přechodovou funkcí $\delta : Q \times \Gamma \to 2^{Q \times \Gamma \times \{L,R\}}$,
5. počátečním stavem $q_0 \in Q$,
6. přijímacím stavem $q_+ \in Q$
7. zamítacím stavem $q_- \in Q$. $q_+ \neq q_-$.

Definice

Krok výpočtu TS je definován jako binární relace na množině konfigurací: Nechť $(q, a_1 \dots a_n, i)$ je taková konfigurace T, kde $q \neq q_\pm, n \in \mathbb{N}_0, a_0, \dots, a_n \in \Gamma, i \leq n$.

(a) Je-li $1 \leq i \leq n$ a $\delta(q, a_i) \ni (q', b, \mathrm{L})$, pak

$$(q, a_0, \dots a_n, i) \vdash (q', a_0 \dots a_{i-1} b a_{i+1} \dots a_n, i-1).$$

(b) Je-li $\delta(q, a_0) \ni (q', b, \mathrm{L})$, pak

$$(q, a_0 \dots a_n, 0) \vdash (q', b a_1 \dots a_n, 0).$$

(c) Je-li $\delta(q, a_i) \ni (q', b, \mathrm{R})$, pak

$$(q, a_0 \dots a_n, i) \vdash (q', a_0 \dots a_{i-1} b a_{i+1} \dots a_n, i+1).$$

Poznámka

Výpočet a přijetí slova jsou definovány stejně jako u deterministických TS. Zamítnutí (a cyklení) ale ne.

Definice

Výpočet $\vdash^*$ TS je reflexivní, tranzitivní uzávěr relace $\vdash$.

Zápis $\mathcal{C} \vdash \mathcal{C}'$ čteme takto:

konfigurace $\mathcal{C}'$ je odvoditelná *z konfigurace* $\mathcal{C}$ *jedním* krokem výpočtu.

Zápis $\mathcal{C} \vdash^n \mathcal{C}'$ čteme takto:

konfigurace $\mathcal{C}'$ je odvoditelná *z konfigurace* $\mathcal{C}$ n kroky výpočtu.

Zápis $\mathcal{C} \vdash^* \mathcal{C}'$ čteme takto:

konfigurace $\mathcal{C}'$ je odvoditelná *z konfigurace* $\mathcal{C}$.

Definice

Nechť T je TS a $w \in \Sigma^*$:

T *přijímá* w, pokud $\mathcal{C}_0^w \vdash^* \mathcal{C}_+$, kde $\mathcal{C}_0^w$ je iniciální konfigurace T pro vstup w a $\mathcal{C}_+$ je přijímající konfigurace stroje T.

T *cyklí* pro w, pokud pro libovolné $n \in \mathbb{N}_0$ existuje $\mathcal{C}$, t.ž. $\mathcal{C}_0^w \vdash^n \mathcal{C}$.
T *zamítá* w, pokud w nepřijímá ani pro w necyklí.

Ekvivalence TS a NTS

Lemma

Ke každému (deterministickému) TS existuje ekvivalentní nedeterministický TS.

Důkaz.

Triviálně. □

Lemma

Ke každému nedeterministickému TS existuje ekvivalentní (deterministický) TS.

Idea důkazu.

Budeme procházet 'strom dosažitelných konfigurací' do šířky. □

Univerzální Turingův stroj.

Univerzální TS

Univerzální TS U je TS, který bere na vstup slova ve tvaru $[T, w]$, kde T je TS, a w je slovo nad jeho vstupní abecedou stroje T, a který

- přijímá, pokud T přijímá w,
- zamítá, pokud T zamítá w,
- jinak cyklí.

Říkáme taky, že TS U *simuluje výpočet* TS T pro w.

Poznámka (Univerzální TS jsme schopni sestrojit)

(3páskový) TS U pro $[T, w]$

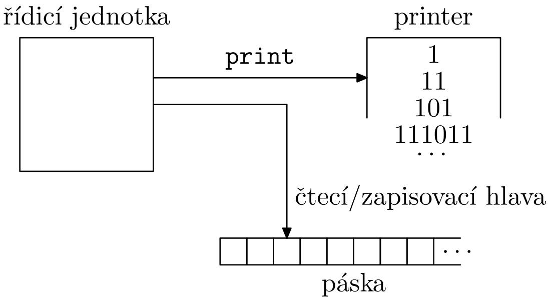
1. na druhou pásku zapíše $\mathcal{C}_0^w$ – počáteční konfiguraci TS T.
2. testne, jestli není v koncové kofiguraci, pokud
   - ano: přijme/zamítne, pokud je to $\mathcal{C}^+/\mathcal{C}^-$.
   - ne: pokračujeme krokem 3
3. podle obsahu první (kód TS T) a obsahu druhé (konfigurace C TS T) zapíše na pásku $\mathcal{C}'$, t.ž. $\mathcal{C} \vdash \mathcal{C}'$.
4. překopíruje obsah třetí pásky do druhé a pokračuje krokem 2.

Enumerator

Enumerator je alternativa TS, která nemá žádný vstup, ale má navíc instrukci `print`, která, když je vyvolána vytiskne obsah pásky (až po první prázdný symbol). Jazykem enumeratoru rozumíme množinu všech slov, které vytiskne. Můžeme jej formalizovat třeba tak, že přechodová funkce bude mít tvar

$$\delta : Q \times \Gamma \to Q \times \Gamma \times \{R, L\} \times \{\texttt{print}, \emptyset\}$$

nebo tak, že určíme množinu stavů, do kterých, když enumerator přejde, nastane `print`.



Konkrétní formalizace nás zajímat nemusí. Důležitější pro nás bude fakt, že enumerator je výpočetně stejně silný jako klasický TS:

Věta

Nechť $L \subseteq \Sigma^*$ *je jazyk.*
(a) Jazyk L *je rekurzivní, právě když existuje enumerator, který jej tiskne lexikografickém uspořádání.*
(b) Jazyk L *je částečně rekurzivní, právě když existuje enumerator, který jej tiskne.*

Důkaz.

Na cvičení □

Lineárně ohraničený automat (LBA)

LBA (linearly bounded automaton); lineárně ohraničený automat.

LBA je nedeterministický TS, který bere vstupy ve tvaru

$$\triangleright\omega\triangleleft, \quad \omega \in \Sigma^*$$

a který

- symboly $\triangleright$ a $\triangleleft$ nikdy nepřepíše jiným symbolem,
- při čtení $\triangleright$ pohne čtecí/zapisovací hlavou vždy doprava,
- při čtení $\triangleleft$ pohne čtecí/zapisovací hlavou vždy doleva.

Poznámka

Vlastně podobná funkce jako zarážka u TS, který se nikdy nepokusí přejet levý okraj pásky. Zde se nepokusí přejet ani jeden z okrajů pásky.

Jazyky Chomského hierarchie

regulární jazyky	=jazyky KA
bezkontextové jazyky	=jazyky nedeterministických ZA
kontextové jazyky	=jazyky LBA
jazyky generované gramatikami bez omezení	=jazyky přijímané TS