# 1.12 Pravděpodobnostní algoritmy

**1.12.1 Randomizovaný Turingův stroj.** RTM je, zhruba řečeno, Turingův stroj $M$ se dvěma nebo více páskami, kde první páska má stejnou roli jako u deterministického Turingova stroje, ale druhá páska obsahuje náhodnou posloupnost 0 a 1, tj. na každém políčku se 0 objeví s pravděpodobností $\frac{1}{2}$ a 1 také s pravděpodobností $\frac{1}{2}$.

Na začátku práce:

- stroj $M$ se nachází v počátečním stavu $q_0$;
- první páska obsahuje vstupní slovo $w$, zbytek pásky pak blanky $B$;
- druhá páska obsahuje náhodnou posloupnost 0 a 1;
- případné další pásky obsahují $B$;
- všechny hlavy jsou nastaveny na prvním políčku dané pásky.

Na základě stavu $q$, ve kterém se stroj $M$ nachází, a na základě obsahu políček, které jednotlivé hlavy čtou, přechodová funkce $\delta$ určuje, zda se $M$ zastaví nebo přejde do nového stavu $p$, přepíše obsah první pásky (**nikoli ale obsah druhé pásky**) a hlavy posune doprava, doleva nebo zůstanou stát (posuny hlav jsou nezávislé).

Formálně, je-li $M$ ve stavu $q$, hlava na první pásce čte symbol $X$, na druhé pásce je číslo $a$ a

$$\delta(q, X, a) = (p, Y, D_1, D_2), \quad q, p \in Q, a \in \{0, 1\}, X, Y \in \Gamma, D_1, D_2 \in \{L, R, S\},$$

pak $M$ se přesune do stavu $p$, na první pásku napíše $Y$ a $i$-tá hlava se posune doprava pro $D_i = R$, doleva pro $D_i = L$ nebo zůstane na místě pro $D_i = S$.

Jestliže $\delta(q, X, a)$ není definováno, $M$ se zastaví.

$M$ se úspěšně zastaví právě tehdy, když se přesune do koncového (přijímacího) stavu $q_f$.

**1.12.2 Poznámka.** Rozdíl mezi RTM a obyčejným TM je v roli druhé pásky. Dvoupáskový TM může přepisovat i obsah druhé pásky a to je v případě RTM zakázáno. Navíc při dvou bězích RTM může být průběh práce RTM různý (záleží na náhodně vygenerovaném obsahu druhé pásky). To se u vícepáskového deterministického TM stát nemůže.

Může se zdát, že tento model je nerealistický — nemůžeme před začátkem práce naplnit nekonečnou pásku. Toto je ale „realizováno“ tak, že v okamžiku, kdy druhá hlava čte dosud nenavštívené políčko druhé pásky, náhodně se vygeneruje 0 nebo 1 každé s pravděpodobností $\frac{1}{2}$ a tento symbol už se nikdy během jednoho průběhu práce TM nezmění.

**1.12.3 Příklad.** Je dán RTM $M$, kde $Q = \{q_0, q_1, q_2, q_3, q_f\}$, $\Gamma = \{0, 1, B\}$ a přechodová funkce $\delta$ je definována:

$$\begin{aligned}
\delta(q_0, 0, 0) &= (q_1, 0, R, S), & \delta(q_0, 1, 0) &= (q_2, 1, R, S),\\
\delta(q_1, 0, 0) &= (q_1, 0, R, S), & \delta(q_1, B, 0) &= (q_f, B, S, S),\\
\delta(q_2, 1, 0) &= (q_2, 1, R, S), & \delta(q_2, B, 0) &= (q_f, B, S, S),\\
\delta(q_0, a, 1) &= (q_3, a, S, R), & \delta(q_3, a, a) &= (q_3, a, R, R),\\
\delta(q_3, B, a) &= (q_f, B, S, S), & &\text{pro } a \in \{0, 1\}.
\end{aligned}$$

Předpokládejme, že na vstupu má RTM $M$ slovo w, pak:

- Jestliže první symbol druhé pásky je 0 (tj. náhodně jsme vygenerovali 0), $M$ zkontroluje, zda $w = 0^n$ nebo $w = 1^n$ pro nějaké $n > 0$.

- Jestliže první symbol druhé pásky je 1 (tj. náhodně jsme vygenerovali 1), hlava na druhé pásce se posune doprava a $M$ zkontroluje, zda se obsah druhé pásky od druhého políčka shoduje se vstupem $w$.

Nenastane-li ani jeden z předchozích případů, $M$ se neúspěšně zastaví.

V případě RTM je třeba spočítat pravděpodobnost s jakou se $M$ pro dané vstupní slovo $w$ úspěšně zastaví, tj. zastaví v „přijímacím“ stavu $q_f$. V našem příkladě je odpověď tato:

- Jestliže $w$ je prázdné slovo, $M$ se v $q_f$ nikdy nezastaví (tj. pro žádný náhodný obsah druhé pásky).

- Jestliže $w = 0^n$ nebo $w = 1^n$ pro $n > 0$, $M$ se zastaví v $q_f$ s pravděpodobností

$$\frac{1}{2} + \frac{1}{2}\left(\frac{1}{2}\right)^n = \frac{1}{2} + 2^{-(n+1)}.$$

- Jestliže $w$ je jiného tvaru, tj. obsahuje jak 0, tak 1, pak pravděpodbnost, že se $M$ zastaví v $q_f$ je

$$\frac{1}{2}\left(\frac{1}{2}\right)^{|w|} = 2^{-(|w|+1)}.$$

**1.12.4 Třída $\mathcal{RP}$.** Jazyk $L$ patří do třídy $\mathcal{RP}$ právě tehdy, když existuje RTM $M$ takový, že:

1. Jestliže $w \notin L$, stroj $M$ se ve stavu $q_f$ zastaví s pravděpodobností 0.

2. Jestliže $w \in L$, stroj $M$ se ve stavu $q_f$ zastaví s pravděpodobností, která je alespoň rovna $\frac{1}{2}$.

3. Existuje polynom $p(n)$ takový, že každý běh $M$ (tj. pro jakýkoli obsah druhé pásky) trvá maximálně $p(n)$ kroků, kde $n$ je délka vstupního slova.

Millerův test prvočíselnosti je příklad algoritmu, který splňuje všechny tři podmínky (utvoříme-li k němu odpovídající RTM) a proto jazyk $L$, který se skládá ze všech složených čísel, patří do třídy $\mathcal{RP}$.

**1.12.5 Turingův stroj typu Monte-Carlo.** RTM splňující podmínky 1 a 2 z předchozí definice 1.12.4, se nazývá TM typu *Monte-Carlo*.

Uvědomte si, že RTM typu Monte-Carlo obecně nemusí pracovat v polynomiálním čase.

**1.12.6 Příklad.** Jazyk $L$ se skládá ze všech slov odpovídajících těm neorientovaným grafům, které obsahují trojúhleník (tj. úplný graf na třech vrcholech) jako podgraf.

Náš algoritmus v jednom kroku náhodně vybere jednu hranu, označme ji $\{x, y\}$, náhodně vybere jiný vrchol $z$ a zkontroluje, zda podgraf indukovaný $\{x, y, z\}$ je úplný. Jestliže ano, úspěšně skončí, jestliže ne, opakuje postup znovu (tj. opět vybere hranu a k ní jeden vrchol a provede kontrolu). Výběr hrany, vrcholu a kontrolu provádí $k$-krát. Jestliže ani po $k$-tém opakování nenajde trojúhelník, neúspěšně skončí. Jedná se o algoritmus typu Monte-Carlo? A pro jaké $k$?

**1.12.7** Po $k$ výběrech a testech platí:

- Jestliže graf neobsahuje trojúhelnik, algoritmus se úspěšně zastaví s pravděpodobností 0.
- Jestliže graf obsahuje trojúhelnik, algoritmus se úspěšně zastaví s pravděpodobností
$$1 - \left(1 - \frac{3}{m(n-2)}\right)^k.$$
kde $n$ je počet vrcholů grafu a $m$ je počet hran.
- Jeden běh algoritmu trvá $\mathcal{O}(k\,n\,m)$.

Abychom dostali algoritmus typu Monte-Carlo, musí být $k \geq \frac{m(n-2)}{3}$.

**1.12.8 Tvrzení.** Je dán jazyk $L \in \mathcal{RP}$, pak pro každou kladnou konstantu $0 < c < \frac{1}{2}$ je možné sestrojit RTM $M$ (algoritmus) s polynomiální složitostí a takový, že:

1. Jestliže $w \notin L$, stroj $M$ se úspěšně zastaví (tj. zastavi se ve stavu $q_f$) pravděpodobností 0.
2. Jestliže $w \in L$, stroj $M$ se úspěšně zastaví (tj. zastaví se ve stavu $q_f$) s pravděpodobností aspoň $1 - c$.

**1.12.9 Třída $\mathcal{ZPP}$.** Jazyk $L$ patří do třídy $\mathcal{ZPP}$ právě tehdy, když existuje RTM $M$ takový, že:

1. Jestliže $w \notin L$, stroj $M$ se úspěšně zastaví (tj. zastaví se ve stavu $q_f$) s pravděpodobností 0.
2. Jestliže $w \in L$, stroj $M$ se úspěšně zastaví (tj. zastaví se ve stavu $q_f$) s pravděpodobností 1.
3. Střední hodnota počtu kroků $M$ v jednom běhu je $p(n)$, kde $p(n)$ je polynom a $n$ je délka vstupního slova.

To znamená: $M$ neudělá chybu, ale nezaručujeme vždy polynomiální počet kroků při jednom běhu, pouze střední hodnota počtu kroků je polynomiální.

**1.12.10 Turingův stroj typu Las-Vegas.** RTM splňující podmínky z předchozí definice 1.12.9, se nazývá typu *Las-Vegas*.

**1.12.11 Tvrzení.** Jestliže jazyk $L$ patří do třídy $\mathcal{ZPP}$, pak i jeho doplněk $\overline{L}$ patří do třídy $\mathcal{ZPP}$.

Stejný RTM $M$ typu Las-Vegas slouží k přijetí jak jazyka $L$, tak i jeho doplňku $\overline{L}$; stačí koncové (přijímající) stavy RTM $M$ prohlásit za nekoncové a z všech nekoncových stavů $M$ udělat koncové.

**1.12.12 Poznámka.** Pro jazyky ze třidy $\mathcal{RP}$ se tvrzení obdobné 1.12.11 neumí dokázat. To motivuje následující třídu jazyků.

**1.12.13 Třída co-$\mathcal{RP}$.** Jayzk $L$ patří do třídy co-$\mathcal{RP}$ právě tehdy, když jeho doplněk $\overline{L}$ patří do třídy $\mathcal{RP}$.

**1.12.14 Věta.**

$$\mathcal{ZPP} = \mathcal{RP} \cap \text{ co-}\mathcal{RP}.$$

**Nástin důkazu.** Ukážeme nejprve $\mathcal{RP} \cap \text{ co-}\mathcal{RP} \subseteq \mathcal{ZPP}$.

Předpokládejme, že jazyk $L$ leží v obou třídách $\mathcal{RP}$ i co-$\mathcal{RP}$. Existují proto dva RTM $M_1$ a $M_2$ typu Monte Carlo pracující v polynomiálním čase a takové, že

$M_1$ — přijímá jazyk $L$;

$M_2$ — přijímá jazyk $\overline{L}$.

Označme $p(n)$ ten větší z polynomů, které určují počet kroků $M_1$ a $M_2$. Setrojíme RTM $M$ typu Las-Vegas, který přijímá $L$ takto: Pro dané vstupní slovo $w$

1. $M$ nechá pracovat $M_1$ po dobu $p(n)$ kroků. Jestliže $M_1$ přijme, $M$ skončí a také přijme.
2. $M$ nechá pracovat $M_2$ po dobu $p(n)$ kroků. Jestliže $M_2$ přijme, $M$ skončí a nepřijme.
3. Jestliže $M$ neskončí ani v kroku 1 ani v kroku 2, $M$ pokračuje krokem 1.

Dá se dokázat, že RTM $M$ je typu Las-Vegas.

Nyní ukážeme, že $\mathcal{ZPP} \subseteq \mathcal{RP} \cap \text{ co-}\mathcal{RP}$.

Předpokládejme, že jazyk $L$ leží ve třídě $\mathcal{ZPP}$, existuje tedy RTM $M_1$ typu Las-Vegas, který přijímá jazyk $L$. Označme $p(n)$ polynom, který udává střední hodnotu počtu kroků RTM $M_1$ pro vstupní slovo délky $n$. Vytvoříme RTM $M$ typu Monte Carlo pracující polynomiálně dlouho a přijímající jazyk $L$.

$M$ nechá na vstupu $w$ pracovat RTM $M_1$ po dobu $2p(n)$. Jestliže $M_1$ úspěšně skončí, $M$ úspěšně skončí; ve všech ostatních případech RTM $M$ skončí neúspěšně.

Dá se dokázat, že $M$ splňuje všechny podmínky pro RTM typu Monte Carlo. Protože pracuje v čase $2p(n)$, jedná se o polynomiální RTM typu Monte Carlo. Proto je jazyk $L$ ve třídě $\mathcal{RP}$.

Protože třída $\mathcal{ZPP}$ je uzavřena na doplňky, je každý jazyk ze třídy $\mathcal{ZPP}$ také ve třídě co-$\mathcal{RP}$.

**1.12.15 Věta.** Platí

$$\mathcal{P} \subseteq \mathcal{ZPP}, \quad \mathcal{RP} \subseteq \mathcal{NP}, \quad \text{co-}\mathcal{RP} \subseteq \text{co-}\mathcal{NP}.$$

První inkluze je zřejmá, každý polynomiální Turingův stroj můžeme považovat za randomizovaný Turingův stroj typu Las-Vegas.

Druhá inkluze je složitější. Její důkaz spočívá v tom, že pro daný polynomiální RTM $M$ typu Monte Carlo pracující v polynomiálním čase zkonstruujeme nedeterministický Turingův stroj, který přijímá stejný jazyk jako $M$.

Třetí inkluze jednoduše vyplývá z definic tříd co-$\mathcal{RP}$, co-$\mathcal{NP}$ a z druhé inkluze.