Kapitola 3

Třídy složitosti algoritmů

Co si zde procvičíme?

- Zopakujeme si dva základní grafové algoritmy - algoritmus prohledávání do hloubky a algoritmus prohledávání do šířky
- Budeme se zabývat odhadem růstu funkcí. K tomu nám budou složit asymptotické notace, které nám umožní zjednodušovat výrazy odhadující složitosti algoritmů
- Seznámíme se se základními asymptotickými notacemi $\mathcal{O}, o, \Omega, \omega, \Theta$ a jejich definicemi
- Vysvětlíme si význam jednotlivých notací a ukážeme si možnosti jejich využití
- Ukážeme si dva způsoby, jak pro jednotlivé funkce dokázat, že patří do dané asymptotické třídy
  - z definice asymptotické meze
  - s využitím znalostí limit

3.1 Prohledávání do hloubky a do šířky

Algoritmy prohledávání do hloubky a do šířky jsou grafové algoritmy. Slouží k procházení grafů.

Prohledávání do hloubky (**DFS - depth-first search**) pracuje tak, že vždy navštíví dalšího možného následníka každého vrcholu. Pokud narazí na vrchol, z nějž už nelze dále pokračovat, vrací se metodou backtrackingu zpět.

Prohledávání do šířky (**BFS - breadth-first search**) prochází všechny nenavštívené sousedy daného vrcholu, které jsou vloženy do fronty a postupně stejným způsobem zpracovány. Takto se postupuje, dokud fronta není prázdná.

Úloha 3.1.1. Prohledávání do hloubky a do šířky
Je dán graf na obrázku (viz další strana).
Graf prohledáváme z vrcholu 1. Číslo vrcholu určuje prioritu jeho zpracování. Řešte následující úlohy:

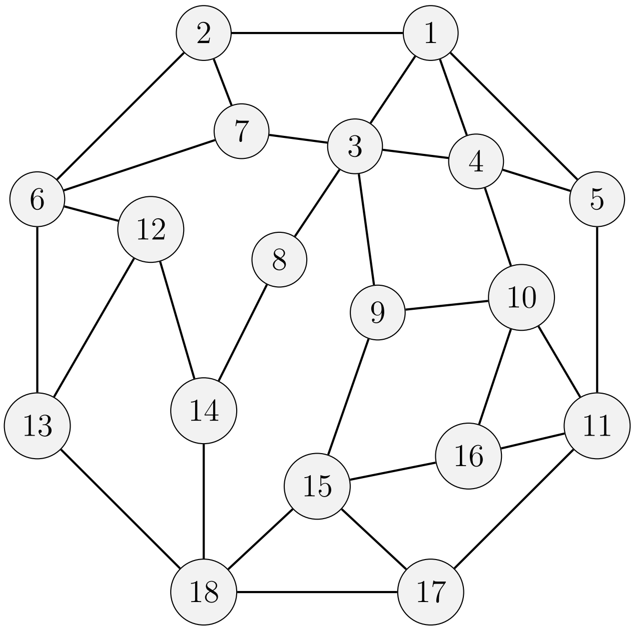
a) Sepište posloupnost, v níž budou vrcholy zpracovány při využití prohledávání do hloubky.

b) Sepište seznam navštívených vrcholů při prohledávání grafu do šířky.

c) Jsou dány následující posloupnosti vrcholů, jak byly navštíveny při prohledávání grafu. Zjistěte, které z posloupností odpovídají prohledávání do hloubky a které prohledávání do šířky.

- $1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18$
- $1, 2, 6, 13, 18, 17, 11, 5, 4, 3, 7, 8, 14, 12, 9, 10, 16, 15$
- $1, 2, 6, 13, 18, 17, 11, 5, 4, 3, 7, 8, 14, 12, 9, 15, 16, 10$
- $1, 2, 6, 13, 18, 17, 11, 5, 4, 3, 7, 8, 14, 12, 10, 9, 15, 16$

Řešení: Odpověď po řadě BFS, DFS, DFS, ani jedno.



3.2 Odhad růstu funkcí

Dále se budeme zabývat funkcemi, které vyjadřují složitost algoritmu, které vyčíslují počet vykonaných operací nebo čas potřebný pro výpočet. Takové funkce jsou z definice **nezáporné**. Jsou definovány s ohledem na velikost vstupu $n, n \geq 0$. Zajímat nás bude průběh funkce v prvním kvadrantu.

3.2.1 Asymptotická horní mez

Říkáme, že funkce $g(n)$ **asymptoticky shora ohraničuje** funkci $f(n)$, když

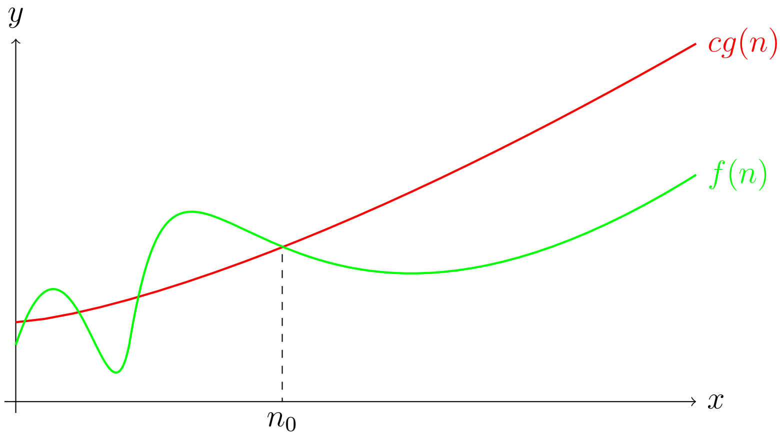
$$\exists c > 0, \exists n_0 \geq 0, \forall n \geq n_0 : 0 \leq f(n) \leq cg(n).$$

Množinu všech funkcí $f(n)$, které jsou asymptoticky shora ohraničeny funkcí $g(n)$ označujeme $\mathcal{O}(g(n))$. Tj.

$$\mathcal{O}(g(n)) = \{f(n) : \exists c > 0, \exists n_0 \geq 0, \forall n \geq n_0 : 0 \leq f(n) \leq cg(n)\}$$

Skutečnost, že funkce $f(n)$ patří do třídy funkcí $\mathcal{O}(g(n))$ zapisujeme zkráceně $f(n) = \mathcal{O}(g(n))$.
$\mathcal{O}$-výrazy používáme pro odhady složitostí algoritmů **v nejhorších případech**.

Situaci znázorňuje obrázek, na kterém je vidět, že funkce $cg(n)$ pro všechna $n > n_0$ ohraničuje shora funkci $f(n)$. Funkce $f(n)$ neroste rychleji než funkce $g(n)$, roste naopak stejně rychle nebo pomaleji.



Úloha 3.2.1. Asymptotická horní mez

a) Podle definice ověřte, že $40n + 100 = \mathcal{O}(n^2 + 10n + 300)$.

Řešení: Podle definice potřebujeme najít hodnotu c tak, aby platilo:

$$0 \le 40n + 100 \le c(n^2 + 10n + 300) \tag{1}$$

Zvolíme $c = 1$.

$$40n + 100 = n^2 + 10n + 300$$

$$n^2 - 30n + 200 = 0, n_1 = 20, n_2 = 10$$

Tedy například pro $c = 1$ a $n_0 \ge 20$ bude vztah (1) platit.

b) Zjistěte, zda platí $\sin(n) = \mathcal{O}(n)$.

Řešení: $|\sin(n)| \le n, \forall n \ge n_0 = 1$, proto $\sin n = \mathcal{O}(n)$. Navíc $|\sin(n)| \le c \cdot 1$, například pro $c = 1$, tedy $\sin(n) = \mathcal{O}(1)$.

c) Ověřte, že platí:

- $5n + 4 = \mathcal{O}(n)$
- $\frac{3}{2}n^2 + \frac{7}{2}n - 4 = \mathcal{O}(n^2)$

Určete hodnoty c a n_0 v definici asymptotické horní meze.

Asymptotickou horní mez funkce lze obvykle vypočíst **s využitím limity**:

Pokud existuje $\lim_{n\to\infty} \frac{f(n)}{g(n)} = L, 0 \le L < \infty$, pak $f(n) = \mathcal{O}(g(n))$.

Úloha 3.2.2. Sestavte a vypočtěte potřebné limity a ověřte, že platí:

- $5n + 4 = \mathcal{O}(n)$
- $\frac{3}{2}n^2 + \frac{7}{2}n - 4 = \mathcal{O}(n^2)$

Asymptotická striktně větší horní mez - $f(n) = o(g(n))$, kde

$$o(g(n)) = \{f(n) : \forall c > 0, \exists n_0 \geq 0, \forall n \geq n_0 : 0 \leq f(n) < cg(n)\}$$

Funkce $f(n)$ roste pomaleji než funkce $g(n)$. Platí, že $f(n) = o(g(n))$ právě tehdy, když $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$.Současně platí $f(n) = o(g(n)) \Rightarrow f(n) = \mathcal{O}(g(n))$.

Úloha 3.2.3. Rozhodněte o platnosti následujících vztahů a případně určete hodnoty c a n_0 podle definice.

- $n^2 = o(n^5)$
- $23n = o(n^{0,02})$
- $14,706\sqrt{n} = o(\frac{n}{2})$
- $\sin(n) = o(n)$
- $\sin(n) = o(1)$
- $\frac{1}{n} = o(1)$
- $\ln(n) = \mathcal{O}(n)$
- $n \cdot \ln(n)^2 = \mathcal{O}(n^2)$
- $\ln(n) = o(n^{0,02})$

Řešení: Zastavíme se u úlohy $\sin(n) = o(n)$. Pokud by tvrzení mělo platit, pak podle definice

$$\forall c > 0, \exists n_0 \geq 0, \forall n \geq n_0 : 0 \leq \sin(n) < cn$$

$\sin(n) < cn, \sin(n) \leq 1 < cn$, tedy $\frac{1}{c} < n$.
Nechť $c > 0$ libovolné, vyberme n_0 tak, aby $\frac{1}{c} < n_0$. Tj. platí, že $\sin(n) = o(n)$.

3.2.2 Asymptotická dolní mez

Asymptotickou dolní mez označujeme $\Omega(g(n))$ a používáme pro odhady složitostí algoritmů **v nejlepších případech**.

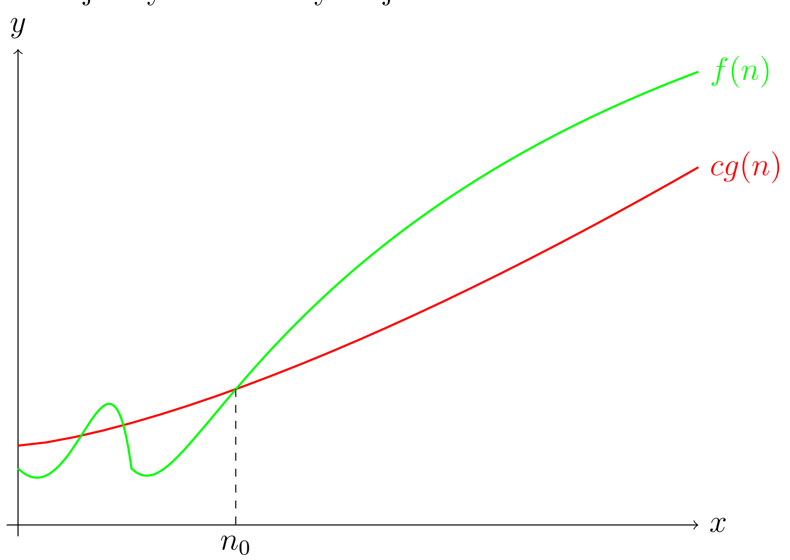
$$\Omega(g(n)) = \{f(n) : \exists c > 0, \exists n_0 \geq 0, \forall n \geq n_0 : 0 \leq cg(n) \leq f(n)\}$$

Pokud existuje limita $\lim_{n\to\infty} \frac{f(n)}{g(n)} = L, 0 < L \leq \infty$, pak $f(n) = \Omega(g(n))$.
Asymptotická striktně menší dolní mez - $f(n) = \omega(g(n))$

$$\omega(g(n)) = \{f(n) : \forall c > 0, \exists n_0 \geq 0, \forall n \geq n_0 : 0 \leq cg(n) < f(n)\}$$

Pokud limita $\lim_{n\to\infty} \frac{f(n)}{g(n)} = \infty$, pak $f(n) = \omega(g(n))$.

Situaci znázorňuje obrázek, na kterém je vidět, že funkce $cg(n)$ pro všechna $n > n_0$ ohraničuje zdola funkci $f(n)$. Funkce $f(n)$ neroste pomaleji než funkce $g(n)$, roste naopak stejně rychle nebo rychleji.



Úloha 3.2.4. Rozhodněte o platnosti následujících vztahů. Svoje tvrzení zdůvodněte.

- $n^5 2^n = \omega(n^{10})$
- $2^n = \Omega(n^2)$
- $n^{0{,}05} = \omega(\ln(n))$
- $3^n = \Omega(2^n)$
- $3^n = \omega(2^n)$
- $4^{n-1} = \Omega(3^n)$
- $(2\log_2 n)^{10} + 2n^{1{,}1} = \Omega(n \log_2 n)$
- $\left(\sqrt[3]{4n^2+3} + n\sqrt{2}\right)^6 = \omega(n^4)$

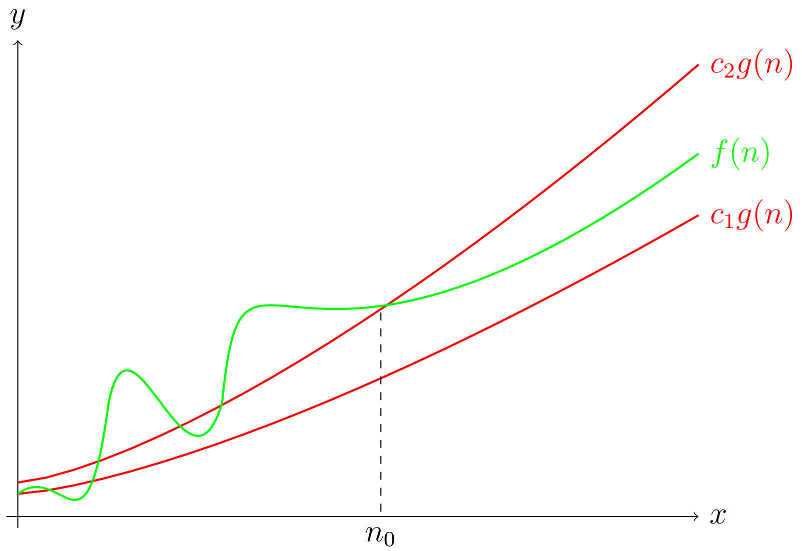
3.2.3 Asymptotická těsná mez

Asymptotickou těsnou mez označujeme $\Theta(g(n))$. Θ-notace vyjadřuje fakt, že dvě funkce jsou asymptoticky stejně až na multiplikativní konstantu.

$$\Theta(g(n)) = \{f(n) : \exists c_1 > 0, \exists c_2 > 0, \exists n_0 \geq 0, \forall n \geq n_0 : 0 \leq c_1 g(n) \leq f(n) \leq c_2 g(n)\}$$

Pokud existuje limita $\lim_{n\to\infty} \frac{f(n)}{g(n)} = L, 0 < L < \infty$, pak $f(n) = \Theta(g(n))$. L je multiplikativní konstanta zmiňovaná v definici asymptotické těsné meze.

Situaci znázorňuje obrázek, na kterém je vidět, že od jistého n_0 se funkce $f(n)$ dostane mezi $c_1 g(n)$ a $c_2 g(n)$ a v těchto mezích zůstane.

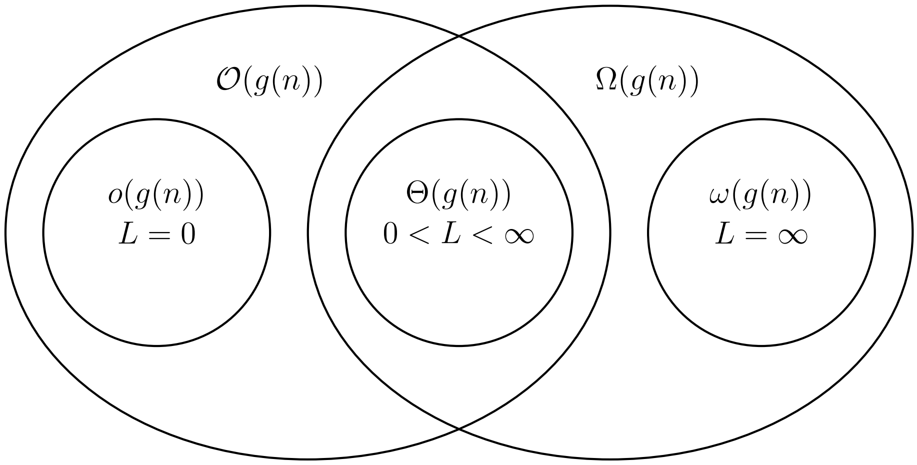


Úloha 3.2.5. Dokažte následující tvrzení.

- $(n+1)^2 = \Theta(3n^2)$
- $\frac{n^2+5n+7}{5n^3+7n+2} = \Theta(\frac{1}{n})$
- $\sqrt{3+\sqrt{2n}} = \Theta(n^{\frac{1}{4}})$
- $\left(1+\frac{3}{n}\right)^n = \Theta)(1)$
- $\sqrt{n+10} = \Theta(\sqrt{n})$
- $(n+a)^b = \Theta(n^b)$

Návod: $\lim_{n\to\infty}\left(1+\frac{3}{n}\right)^n = \lim_{n\to\infty}\left(\left(1+\frac{1}{\frac{n}{3}}\right)^{\frac{n}{3}}\right)^3 = e^3$

Na následujícím obrázku jsou znázorněny vztahy mezi jednotlivými asymptotickými notacemi a jejich limitami.



Úloha 3.2.6. Dokažte, že

- $(n^2 + 3n + 1)^3 = \Theta(n^6)$
- $e^{\frac{1}{n}} = \Theta(1)$
- $\sqrt{\log(n) + 1} = \Omega(\log(\log(n)))$
- $\frac{(\sqrt{n}+1)^3}{n^2+1} = o(1)$
- $\frac{x^2+5n+7}{5n^3+7n+2} = \Theta(\frac{1}{n})$
- $2^{n-1} = \Theta(2^n)$
- $3^{n+1} = \omega(2^n)$
- $2^n = o(3^n)$
- $3\log_2 n + 2\log_3 n = \mathcal{O}(\log_3 n)$
- $(10n)^6 + 0,99^n = \Omega(0,99^n)$
- $\frac{\cos(n)}{n} = \mathcal{O}(1)$
- $3 \cdot n^4 + 2 \cdot 1,1^n = \omega(n^4)$
- $2n = \omega(\ln(n))$

Úloha 3.2.7. Uspořádejte funkce podle jejich asymptotického růstu. Svoje tvrzení zdůvodněte.

a) $n^{0,01}, \ln(n)$

b) $\ln(\ln(n)), n, \ln(n), 2^n, n!$

c) $\ln(\ln(n)), n, 2n^{10}, n!, n^{0,01n}, \ln(n)$

d) $n^3\log(n), (\log(\log(n)))^3, n^5 2^n, (n+4)^{12}$

Úloha 3.2.8. $f(n)$ a $g(n)$ jsou asymptoticky nezávislé funkce. S využitím definice Θ-notace dokažte

$$\max(f(n), g(n)) = \Theta(f(n) + g(n)).$$

Úloha 3.2.9. Dokažte, že platí $\log_a(n) = \mathcal{O}(\log_b(n))$.

Návod: $\log_b(n) = \frac{\log_a(n)}{\log_a(b)}$