7 Barevnost a další těžké problémy

Pro motivaci této lekce se podíváme hlouběji do historie počátků grafů v matematice. Kromě slavného problému *sedmi mostů v Královci* (dnešním Kaliningradě) je za další historický milník vývoje teorie grafů považován *problém čtyř barev* pocházející z poloviny 19. století: Kolik nejméně barev je třeba použít na obarvení politické mapy pro rozlišení sousedních států?

Na rozdíl od sedmi mostů, problém čtyř barev zůstal nevyřešený po více než 100 let a stimuloval rozvoj skoro všech moderních oblastí teorie grafů.

Někteří však kladou prapočátky grafů v matematice daleko před Eulerovými sedmi mosty či problémem čtyř barev, až ke středověké otázce, zda lze šachovým *koněm obejít celou šachovnici* bez opakování. Tato otázka vede v moderním pojetí k dalšímu zajímavému a obtížnému problému tzv. *Hamiltonovské kružnice* v grafu. . . □

Stručný přehled lekce

- Definice barevnosti grafu, základní vlastnosti.
- Varinaty problému barvení.
- Další „obtížné“ problémy jako Hamiltonovská kružnice.
- Algoritmická složitost (NP-úplnost) základních grafových problémů.

7.1 Barevnost grafu

Nejprve si uveďme pojem barevnosti – představme si, že hrany grafu nám říkají, že jejich koncové vrcholy musí být barevně odlišené (třeba proto, že reprezentují sousední státy, nebo proto, že jinak jsou si příliš podobné a je třeba je jinak rozlíšit, atd).□ Samozřejmě bychom mohli každému vrcholu grafu dát jinou barvu, ale k čemu by pak takový problém byl? My bychom chtěli použít barev celkem co nejméně. □

Definice: *Obarvením grafu* G pomocí k barev myslíme libovolné zobrazení

$$c : V(G) \to \{1, 2, \ldots, k\}$$

takové, že každé dva vrcholy spojené hranou dostanou různé barvy, tj. $c(u) \neq c(v)$ pro všechny $\{u, v\} \in E(G)$.

Definice 7.1. Barevnost grafu G
je nejmenší číslo $\chi(G)$ pro které existuje obarvení grafu G pomocí $\chi(G)$ barev. □

Čísla $1, 2, \ldots, k$ z předchozí definice tak nazýváme barvami vrcholů (je to pohodlnější, než popisovat barvy běžnými jmény jako bílá, červená, atd).

Poznámka: Uvědomme si, že barevnost lze definovat pouze pro graf bez smyček, protože oba konce smyčky mají vždy stejnou barvu a nic víc s tím „nenaděláme“.

Lema 7.3. *Mějme jednoduchý graf (bez smyček) G a jeho libovolný podgraf $H \subseteq G$. Pak $\chi(G) \leq \chi(H)$.* □

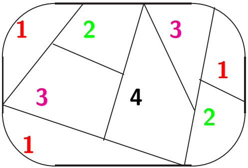
Důkaz: Zřejmě postačí použít restrikci obarvení grafu G na vrcholy podgrafu H. □

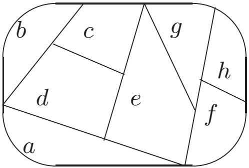
Říkáme, že barevnost grafu je *monotónní na podgrafy*. □

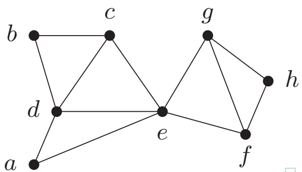
Lema 7.4. *Nechť G je jednoduchý graf (bez smyček) na n vrcholech. Pak $\chi(G) \leq n$ a rovnost nastává právě když $G \simeq K_n$ je úplný graf.* □

Důkaz: Stačí každý vrchol obarvit jinou barvou a máme skutečné obarvení n barvami dle definice. Navíc pokud některá dvojice u, v vrcholů není spojená hranou, můžeme volit lepší obarvení $c(u) = c(v) = 1$ a zbylé vrcholy různými barvami $2, 3, \ldots, n-1$, tj. pak $\chi(G) < n$. □

Příklad 7.5. *Vraťme se k příkladu „barvení“ mapy z úvodu lekce a ukažme si, jak mapy souvisejí s grafy a jejich barevností.*



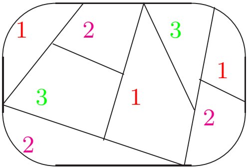


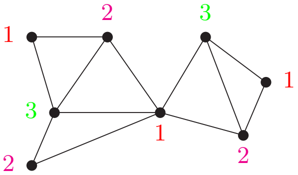


□

Jednotlivé oblasti na mapě (předpokládáme, že každý stát má souvislé území, tj. státy = oblasti) prohlásíme za vrcholy našeho grafu a sousední dvojice států spojíme hranami. Nezapomeňme, že „sousední“ znamená sdílení celého úseku hranice, ne jen jednoho rohu. □

Při troše snahy také najdeme lepší obarvení uvedené mapy využívající pouhých tří barev:



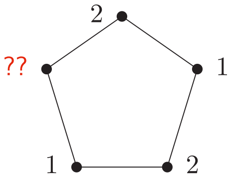


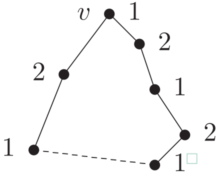
□

Věta 7.6. *Neprázdný graf G má barevnost 1 právě když nemá žádné hrany. G má barevnost ≤ 2 právě když nemá žádnou kružnici liché délky jako podgraf.* $\square$

Důkaz: Pokud graf nemá hrany, můžeme všechny vrcholy obarvit stejnou barvou 1. Naopak pokud mají všechny vrcholy stejnou barvu, nemůže graf mít žádnou hranu. $\square$

Druhá část: Na jednu stranu, lichou kružnici nelze obarvit dvěma barvami, viz obrázek. Na druhou stranu si představme, že zvolíme libovolný vrchol v grafu G s barvou 1 a ostatní vrcholy obarvíme takto: Vrcholy, jejichž vzdálenost od v je lichá, obarvíme 2. Vrcholy, jejichž vzdálenost od v je sudá, obarvíme 1.





$\square$

Pokud bychom tak získali třeba dva vrcholy spojené hranou f v sudé vzdálenosti od v, získáme uzavřený sled S liché délky přes f a v. Stejně tak pro dva vrcholy v liché vzdálenosti. Ponecháme-li ze sledu S ty hrany, které se opakují lichý počet krát, dostaneme Eulerovský podgraf T lichého počtu hran. Jak již víme (Oddíl 5.1), T pak obsahuje kružnici a tudíž jej lze induktivně sestrojit jako hranově-disjunktní sjednocení kružnic. Avšak sjednocení kružnic sudé délky nevytvoří T liché velikosti, spor. Proto naše obarvení za daných předpokladů nemůže dát stejnou barvu sousedním vrcholům, a tudíž dvě barvy stačí. $\square$

7.2 Jak obarvit graf

Hladové obarvování

Definice: Graf G je *k-degenerovaný*, pokud každý podgraf G obsahuje vrchol stupně nejvýše k.

Příkladem k-degenerovaného grafu je každý graf stupně nejvýše k, ale na druhou stranu k-degenerované grafy mohou mít vysoké stupně. (Nestačí však mít jen nízký nejmenší stupeň!) □

Věta 7.7. *Každý k-degenerovaný graf lze správně hladově obarvit $k+1$ barvami.* □

Důkaz: Jelikož graf G je k-degenerovaný, vybereme libovolný jeho vrchol v_1 stupně nejvýše k a rekurzivní aplikací tohoto postupu obarvíme podgraf $G - v_1$, který je podle definice také k-degenerovaný. Nakonec si všimneme, že $\leq k$ sousedé vrcholu v_1 dostanou nejvýše k různých barev, takže v_1 dobarvíme zbylou barvou. □

Důležité aplikace této věty uvidíme v příští lekci, avšak jedno zajímavé zesílení (*Brooksovu větu*) si uvedeme nyní:

Věta 7.8. *Nechť G je souvislý jednoduchý graf maximálního stupně $k \geq 2$. Pak $\chi(G) \leq k$ až na případy, kdy G je úplný graf nebo lichá kružnice.*

Důkaz (náznak): Pro $k = 2$ plyne tvrzení z Věty 7.6. Nechť tedy $k \geq 3$. V jednom směru je jasné, že $\chi(K_{k+1}) = k + 1$. Naopak tedy předpokládejme, že G není úplný. Zároveň se omezme jen na případ, že G má všechny stupně rovné k, neboť jinak lze aplikovat postup z Věty 7.7. □

- Prvním krokem nahlédneme, že pak G obsahuje dva nespojené vrcholy u, v se společným sousedem w. Pokud ale je graf $G - \{u, v\}$ nesouvislý, pak graf příslušně rozdělíme a indukcí po částech obarvíme. □

- Přidejme tedy předpoklad, že $G - \{u, v\}$ je souvislý. Druhým krokem nahlédneme, že graf H vzniklý z $G - w$ ztotožněním u s v do jednoho vrcholu je $(k - 1)$-degenerovaný. □

- Tudíž graf H hladově obarvíme k barvami podle Věty 7.7. Po opětovném „rozpojení" vrcholů u, v získáme obarvení $G - w$ k barvami takové, že u, v mají stejnou barvu. Nyní w má v sousedství nejvýše $k - 1$ barev a G celý obarvíme.

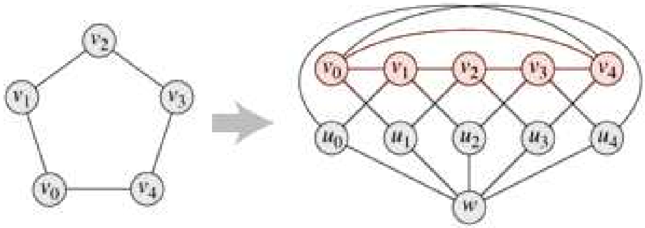
□

Grafy vysoké barevnosti

Ke správnému pochopení barevnosti grafu je nezbytné se zamyslet, které grafy mají vysokou barevnost. Jedná se například o grafy obsahující velké kliky (úplné podgrafy). Je to však vše? □

Není! Lze nalézt grafy s libovolně vysokou barevností neobsahující ani trojúhelníky. Třeba známá Mycielského konstrukce nám dává toto:

Tvrzení 7.9. *Graf získaný z grafu G následující konstrukcí (viz obrázek) má barevnost $\chi(G)+1$ a neobsahuje trojúhelníky, pokud je neobsahuje ani G.*



□

Nejobecněji lze říci následující překvapivě silné tvrzení nalezené Erdősem:

Věta 7.10. *Pro každá $c, r > 0$ existuje graf s barevností alespoň c a neobsahující kružnice kratší než r.*

7.3 Variace na barevnost a jiné

Definice 7.11. **Hranová barevnost** grafu G.
Hledáme obarvení $c_e(E(G)) \to \{1, 2, \ldots, k\}$ takové, že žádné dvě hrany se společným vrcholem nedostanou stejnou barvu.

Nejmenší možný počet barev k, pro které hranové obarvení existuje, se nazývá *hranová barevnost* $\chi_e(G)$ grafu. □

Na rozdíl od běžné barevnosti umíme hranovou barevnost velmi dobře aproximovat za použití této Vizingovy věty.

Věta 7.12. *Pro každý jednoduchý graf platí* $\Delta(G) \le \chi_e(G) \le \Delta(G) + 1$. □

Platí, že většina grafů splňuje $\Delta(G) = \chi_e(G)$. Umíte jednoduše sestrojit (a dokázat) příklady pro druhý případ?

Problém přesného určení hranové barevnosti grafu však stále zůstává algoritmicky velmi obtížný a také úzce souvisí s problémem čtyř barev.

Definice 7.13. **Výběrová barevnost** grafu G.
Je dán graf G spolu s přiřazenými „seznamy barev" $L : V(G) \rightarrow \binom{\mathbb{N}}{k}$ (k-prvkové podmnožiny). Nyní hledáme obarvení $c_{ch}(V(G)) \rightarrow \mathbb{N}$ takové, že žádné dva sousední vrcholy nedostanou stejnou barvu a navíc $c_{ch}(v) \in L(v)$ pro každý vrchol v.

Nejmenší možná délka k seznamů barev, pro kterou výběrové obarvení vždy existuje (tj. pro každou možnou takovou volbu seznamů), se nazývá *výběrová barevnost* $ch(G)$ grafu. □

Výběrová barevnost může (kupodivu!) být libovolně „vzdálena" běžné barevnosti.

Tvrzení 7.14. *Pro každé k nalezneme bipartitní graf s výběrovou barevností větší než k.* □

Fakt: Hranová výběrová barevnost úplných bipartitních grafů úzce souvisí se známým problémem tzv. "latinských obdélníků".

Hamiltonovské grafy

Definice: Kružnice C obsažená v grafu G se nazývá *Hamiltonovská*, pokud C prochází všemi vrcholy G. Obdobně mluvíme o *Hamiltonovské cestě* P v G, pokud cesta $P \subset G$ prochází všemi vrcholy G.

Graf G je *Hamiltonovský*, pokud obsahuje Hamiltonovskou kružnici. □

Možná to zní překvapivě, ale i problém Hamiltonovské kružnice úzce souvisel s řešením problému čtyř barev. To je však mimo rámec našeho textu.

Místo toho si ukážeme následující krásný výsledek Diraca:

Věta 7.15. *Každý graf na $n \geq 3$ vrcholech s minimálním stupněm $\geq n/2$ je Hamiltonovský.* □

Důkaz (náznak): Nechť P je nejdelší cesta v grafu G s vrcholy po řadě $u_0, u_1, \ldots, u_k$. Podle její maximality leží každý soused u_0 i u_k na P. Pak existuje $0 < i < k$ takové, že $u_0u_{i+1} \in E(G)$ a zároveň $u_ku_i \in E(G)$. Pak $u_0u_{i+1}\, P\, u_ku_i\, P$ tvoří kružnici v G a snadno plyne, že se jedná o Hamiltonovskou kružnici. □

7.4 $\mathcal{NP}$-úplnost grafových problémů

Definice složitostní třídy $\mathcal{NP}$ se týká výhradně rozhodovacích problémů (s odpovědí „ANO/NE“). Dá se neformálně říci, že problém patří do třídy $\mathcal{NP}$, pokud jeho odpověď ANO lze prokázat (ve smyslu „uhodnout a ověřit“) výpočtem, který běží v polynom. čase.

$\mathcal{NP}$-úplné problémy jsou zhruba řečeno ty, které ve třídě $\mathcal{NP}$ mají nejvyšší obtížnost řešení.

Od jednoho $\mathcal{NP}$-úplného problému A se dostaneme k jinému B tzv. *polynomiálním převodem*: Ukážeme, jak bychom ze známého postupu řešení B efektivně nalezli řešení lib. instance A. □

Nyní si ukážeme vhodnými převody, že oněch „nejobtížnějších“ ($\mathcal{NP}$-úplných) problémů je v teorii grafů mnoho, bohužel by se dalo říci většina. To ostatně ukazuje, proč jsme zatím v praxi tak málo úspěšní při počítačovém řešení mnohých praktických problémů – přesné a efektivní řešení $\mathcal{NP}$-úplných úloh se totiž považuje za nemožné.

Problém 7.16. 3-SAT (splnitelnost logických formulí ve spec. verzi)
Následující problém je $\mathcal{NP}$-úplný:

Vstup: Logická formule Φ v konjunktivním normálním tvaru taková, že každá klauzule obsahuje nejvýše 3 literály.

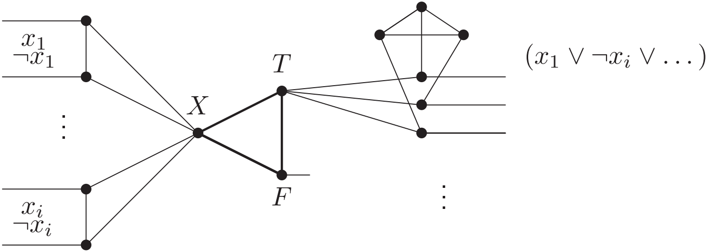
Výstup: Existuje logické ohodnocení proměnných tak, aby výsledná hodnota Φ byla 1 (pravda)?

Příkladem formule problému 3-SAT je třeba $\Phi \equiv (x_1 \vee \neg x_3 \vee x_4) \wedge (x_2 \vee \neg x_4) \wedge (\neg x_1 \vee x_2 \vee x_3)$.

Problém 7.17. 3-COL (3-obarvení grafu)
Následující problém je $\mathcal{NP}$-úplný:

Vstup: Graf G.
Výstup: Lze vrcholy G korektně obarvit 3 barvami?

Důkaz (náznak): Ukážeme si polynomiální převod z problému 3-SAT.
Sestrojíme graf G pro danou formuli Φ. Základem grafu je trojúhelník, jehož vrcholy označíme X, T, F. Každé proměnné x_i ve Φ přiřadíme dvojici vrcholů spojených s X. Každé klauzuli ve Φ přiřadíme podgraf na 6 vrcholech (z nichž tři jsou spojené s T), jako na obrázku. Nakonec volné „půlhrany" z obrázku pospojujeme dle toho, jaké literály vystupují v klauzulích.

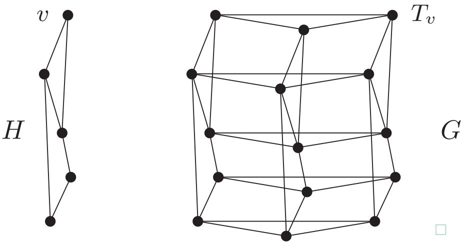


Pak G má 3-obarvení právě když je Φ splnitelná, jak si lze ověřit na obrázku. $\square$

Problém 7.18. IS (nezávislá množina)
Následující problém je $\mathcal{NP}$-úplný:

Vstup: Graf G a přirozené číslo k.
Výstup: Lze v G najít nezávislou podmnožinu velikosti (aspoň) k? □

Důkaz: Ukážeme polynomiální převod z problému 3-COL.
Nechť H je graf na n vrcholech, který je za úkol obarvit třemi barvami. Položíme $k = n$ a graf G sestrojíme ze tří disjunktních kopií grafu H takto:



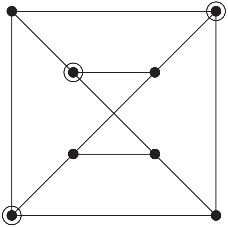
Pokud $c : V(H) \to \{1, 2, 3\}$ je obarvení H třemi barvami, v grafu G lze vybrat $k = n$ nezávislých vrcholů tak, že pro každý $v \in V(H)$ vezmeme $c(v)$-tou kopii vrcholu v v grafu G. Naopak pokud I je nezávislá množina v grafu G o velikosti $k = n$, pak z každého trojúhelníku T_v, $v \in V(H)$ náleží do I právě jeden vrchol. Podle toho již určíme jednu ze tří barev pro vrchol v v H. □

Problém 7.19. VC (vrcholové pokrytí)
Následující problém je $\mathcal{NP}$-úplný:

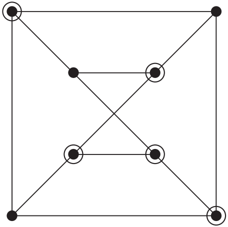
Vstup: Graf G a přirozené číslo k.

Výstup: Lze v G najít *vrcholové pokrytí*, tj. množinu $C \subseteq V(G)$ takovou, že každá hrana G má alespoň jeden konec v C, o velikosti nejvýše k? □

Důkaz: Ukážeme polynomiální převod z problému IS. Nechť G je graf na n vrcholech, v němž máme najít nezávislou množinu I velikosti ℓ. Všimněme si, že doplněk $C = V(G) \setminus I$ nezávislé množiny I je vlastně vrcholovým pokrytím. Takže v našem převodu stačí použít stejný graf G a $k = n - \ell$. □



nezávislá množina

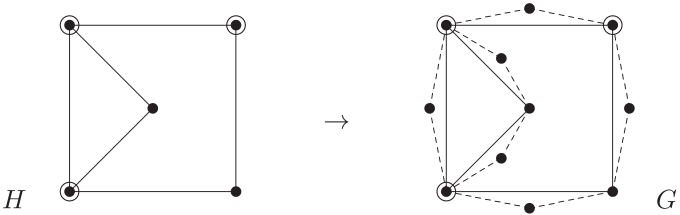


vrcholové pokrytí

Problém 7.20. DOM (dominující množina)
Následující problém je $\mathcal{NP}$-úplný:

Vstup: Graf G a přirozené číslo k.
Výstup: Lze v G najít dominující množinu, tj. množinu $D \subseteq V(G)$ takovou, že každá vrchol G má některého souseda v D, o velikosti nejvýše k? □

Důkaz (náznak): Problém dominující množiny jasně patří do $\mathcal{NP}$ a jeho úplnost je dokázána následujícím schematickým polynomiálním převodem.



Pro daný graf H vytvoříme graf G přidáním, pro každou hranu $e \in E(H)$, nového vrcholu v_e spojeného hranami do obou koncových vrcholů hrany e. (Tak se vlastně z každé hrany stane trojúhelník s třetím novým vrcholem, viz naznačený obrázek.) Číselný parametr k zůstane tentokrát nezměněn. Nyní zbývá dokázat, že G má vrcholové pokrytí velikosti k, právě když H má dominující množinu velikosti také k, což není obtížné. □

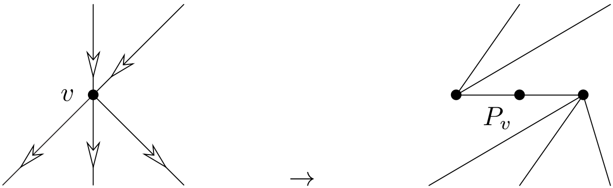
Problém 7.21. HC (Hamiltonovský cyklus)
Následující problém je $\mathcal{NP}$-úplný:

Vstup: Orientovaný graf G.
Výstup: Lze v G najít orientovanou kružnici (cyklus) procházející všemi vrcholy? □

Problém 7.22. HK (Hamiltonovská kružnice)
Následující problém je $\mathcal{NP}$-úplný:

Vstup: Graf G.
Výstup: Lze v G najít kružnici procházející všemi vrcholy? □

Důkaz:



Použijeme snadný převod z předchozího problému HC. Každý vrchol v orientovaného grafu H nahradíme třemi vrcholy tvořícími cestu P_v délky 2 v grafu G. Orientované hrany grafu H přicházející do v pak přivedeme do prvního vrcholu cesty P_v, hrany odcházející z v naopak vedeme z posledního vrcholu cesty P_v. □

7.5 Příběh problému vrcholového pokrytí

Ač se to nezdá, někdy i zcela okrajová a poněkud banální otázka může nakonec vést k dalekosáhlým závěrům a novým teoriím. . . Co třeba proč zdánlivě velmi „podobné“ problémy jako vrcholové pokrytí VC a dominující množina DOM mají (přestože oba $\mathcal{NP}$-úplné) tak rozdílné algoritmické chování? Vysvětlení [R. Downey and M. Fellows, Parameterized complexity, Springer 1999] by nás dovedlo až ke zcela novému pohledu na výpočetní složitost problémů, který jde „jaksi mimo“ klasickou polynomiální hierarchii a umožňuje v jistých situacích docela rozumně řešit i některé problémy, které jsou jinak $\mathcal{NP}$-těžké. □

Takže v čem spočívá zásadní rozdíl v našich znalostech o řešení problémů *dominující množiny* a *vrcholového pokrytí*?

- Pokud se v analýze zaměříme na hodnotu parametru k vstupu, tak dominující množinu velikosti k stejně nedokážeme nalézt rychleji než probráním prakticky všech k-tic vrcholů grafu G.

  To je i pro malé fixní hodnoty k, třeba $k = 10, 20$, v praxi neproveditelné. □

- Avšak vrcholové pokrytí velikosti k dokážeme nalézt jednoduchým algoritmem v čase $O(2^k \cdot n)$, což pro malé fixní hodnoty k, třeba opět $k = 10, 20$, dává skvěle použitelný lineární algoritmus!

Algoritmus 7.24. k-**VC (vrcholové pokrytí)**
Pro fixní k „poměrně rychle“ vyřešíme následující problém.

Vstup: Graf G.
Výstup: Lze v G najít vrcholové pokrytí o velikosti nejvýše k? □

Pro inicializaci položíme $C = \emptyset$ a $F = E(G)$.

- Pokud $F = \emptyset$, vracíme C jako vrcholové pokrytí.
  Jinak pokud $|C| \geq k$, vracíme odpověď NE.
- Vybereme libovolnou hranu $f = uv \in F$ a pro oba její konce $x = u, v$ uděláme:
  - $C' = C \cup \{x\}$ a nová množina hran F' vzniká z F odebráním všech hran vycházejících z vrcholu x v G.
  - Rekurzivně zavoláme tento algoritmus pro G, C' a F'. □

Kolik tento algoritmus provede rekurzivních volání celkem? Každý průchod generuje dvě další volání, ale jen do fixní hloubky k rekurze, takže ve výsledku bude čas výpočtu asymptoticky jen $O(2^k \cdot n)$.

Poznámka: Dnes je již známo, že faktor 2^k lze promyšlenějším přístupem „vylepšit“ na mnohem menší základ mocniny. (2006: 1.2738^k)