Barvení grafu (Wilf 1984)

Nechť $G \in \mathcal{G}_n$ je neorientovaný graf s n vrcholy, $k \le n$ je počet barev.

Obarvit graf znamená přiřadit jeho vrcholům barvy (čísla $1, 2, \dots, k$) tak, aby koncové vrcholy libovolné hrany byly obarveny různě.

Chromatické číslo grafu $\chi(G)$ je minimální počet barev, kterými lze obarvit graf G.

Existují jednoduchá řešení pro $k = 1$ a $k = 2$.

Nechť $k = 1$. Obarvení jednou barvou existuje, pokud graf nemá žádné hrany, jinak ne.

Nechť $k = 2$ a graf je souvislý (tj. mezi každými dvěma vrcholy grafu vede cesta). Zvolíme libovolný bod, obarvíme ho barvou 1, jeho sousedy barvou 2, jejich dosud neobarvené sousedy barvou 1 atd. Toto obarvení je jediné možné a pokud neexistuje, nedá se graf dvěma barvami obarvit.

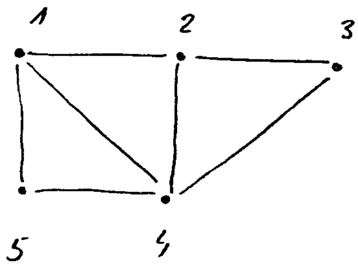
Je-li graf nesouvislý, dělá se to pro každou komponentu zvlášť (komponenta = největší souvislý úplný podgraf; úplný podgraf ... obsahuje všechny hrany grafu G, které mají oba koncové vrcholy v podgrafu).

Časová složitost: O(počet vrcholů + počet hran)

Pro $k = 3$ je barvení grafu NP-úplný problém.

Použijeme tzv. metodu backtracking (metoda větví a mezí).

Příklad. Graf G:



$k = 3$

Budu vytvářet posloupnost částečných řešení postupným přidáváním dat a uspořádám je do stromu.

0) Začnu nejmenším možným podgrafem grafu G, tj. grafem, který neobsahuje žádný vrchol. Tomu odpovídá jedno obarvení (prázdné) a bude tvořit kořen stromu.

1) Vezmu podgraf G tvořený vrcholem 1. Dá se obarvit celkem třemi způsoby a tato obarvení budou tvořit 1. hladinu stromu.

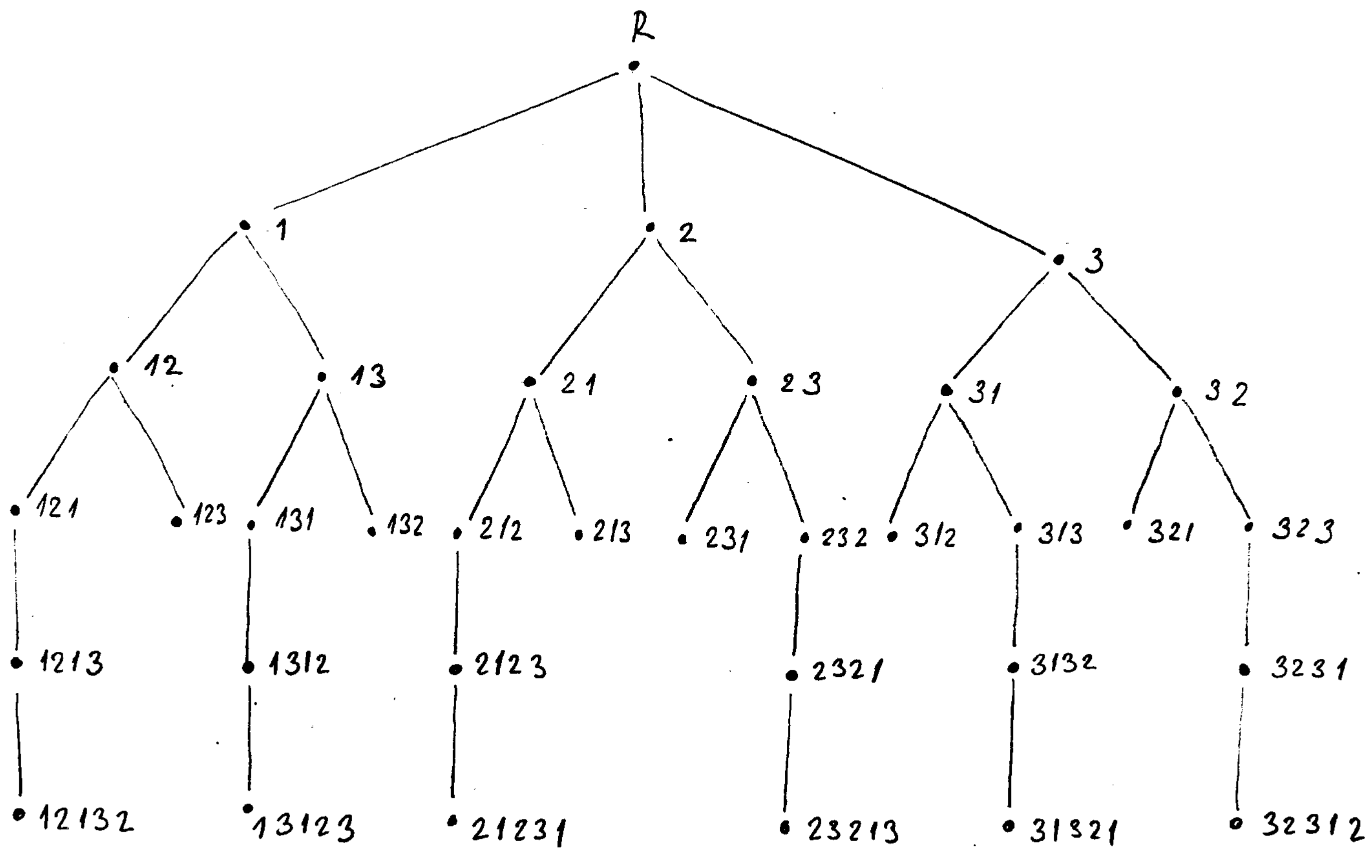
2) Vezmu podgraf G tvořený vrcholy 1 a 2. Pro každé obarvení vrcholu 1 existují dvě obarvení vrcholu 2, celkem tedy šest obarvení, která budou tvořit 2. hladinu

⋮

l) Vezmu podgraf $H_l(G)$ grafu G tvořený vrcholy $1, 2, \ldots, l$. Možná obarvení budou tvořit l-tou hladinu.

⋮

n) V poslední hladině budou všechna přípustná obarvení grafu G, pokud existují.



Částečná řešení, která se dají prodloužit na <u>přípustná řešení</u>, se nazývají <u>přípustná částečná řešení</u>.

Vznikl <u>strom problému</u> – kořenový strom, jehož kořenem je prázdné řešení, vnitřními vrcholy přípustná částečná řešení a koncovými vrcholy (listy) přípustná řešení a částečná řešení, která již nelze prodlužovat.

<u>Backtracking</u> je prohledávání stromu problému do hloubky.

Doba výpočtu algoritmu závisí na počtu vrcholů v prohledávacím stromu. Je nejmenší pro úplný graf, a to má hodnotu

$$f_n(k) = 1 + k + k(k-1) + k(k-1)(k-2) + \dots + k!$$

a největší pro kompletně nesouvislý graf (tj. graf, kde nejsou žádné hrany), kdy

$$F_n(k) = 1 + k + k^2 + \dots + k^n.$$

Označme $\beta(G)$ počet vrcholů v prohledávacím stromu. Zřejmě

$$f_n(k) \leq \beta(G) \leq F_n(k), \qquad f_n(k) = O(1)$$

$F_n(k)$ roste exponenciálně s n.

Bude nás zajímat střední hodnota $\beta(G)$, pro $G \in \mathcal{G}_n$, tj:

$$\beta_n = \sum_{G \in \mathcal{G}_n} \beta(G)\, 2^{-\binom{n}{2}}.$$

(předpokládáme, že všechny grafy jsou stejně pravděpodobné; $2^{\binom{n}{2}}$ je celkový počet grafů s n vrcholy)

Nejprve pomocná tvrzení.

<u>Lemma 1</u>: Nechť C je jedno ze k^l možných obarvení abstraktních bodů $1, 2, \dots, l$ k barvami (nepožadujeme, aby sousední body byly obarveny různě). Označme $Q(l)$ počet grafů G s l vrcholy takových, že C je vlastní(*) obarvení vrcholů G. Pak

$$Q(l) \leq 2^{l^2(1-\frac{1}{k})/2}$$

(*) vlastní = grafařsky správné, tj. sousední vrcholy mají různé barvy

Dk: obarvení C rozdělyje body grafu G do k tříd, označme jejich velikosti $s_1, \dots, s_k$. C je obarvením grafu G právě tehdy, když hrany vedou pouze mezi navzájem různými třídami. Maximální počet takových hran je

$$\sum_{1 \le i < j \le k} s_i s_j = \frac{1}{2}\left(L^2 - \sum_{i=1}^{k} s_i^2\right) \qquad = \frac{1}{2}\left(\sum_{i,j}\sum - \sum_{i=j}\right)$$

a počet takových grafů je

$$2^{\frac{1}{2}\left(L^2 - \sum_{i=1}^{k} s_i^2\right)} \le 2^{\frac{1}{2}L^2\left(1-\frac{1}{k}\right)},$$

přičemž maximum nastává, je-li $s_i = \frac{L}{k}$ pro vš. $i = 1, \dots, k$.

Lemma 2: Nechť $P(k, G)$ je počet vlastních obarvení vrcholů G k barvami, nechť $M_L(k) = \sum_{G \in \mathcal{G}_L} P(k, G)$. Pak

$$M_L(k) \le k^L 2^{\frac{1}{2}L^2\left(1-\frac{1}{k}\right)}.$$

Dk: Nechť C má význam jako v předch. lemmatu.

Definujme fci $f(C, G) = \begin{cases} 1 & \text{když } C \text{ je vlastním obarvením } G \\ 0 & \text{jinak.} \end{cases}$

Pak

$$M_L(k) = \sum_{G \in \mathcal{G}_L} P(k, G) = \sum_{G \in \mathcal{G}_L} \sum_{C} f(C, G) = \sum_{C} \sum_{G \in \mathcal{G}_L} f(C, G) \le$$

$$\le k^L \max_{C} \{ \underbrace{\sum_{G \in \mathcal{G}_L} f(C, G)}_{Q(L)} \} \le k^L 2^{\frac{1}{2}L^2\left(1-\frac{1}{k}\right)} \quad \text{podle předch. lemmatu}$$

Věta: Jestliže $\beta_n = \beta_n(k)$ je střední hodnota veličiny $\beta(G)$ přes všechny grafy $G \in \mathcal{G}_n$ pro pevné k, pak

$$\lim_{n\to\infty} \beta_n = h(k) < \infty \quad \text{pro každé } k = 1, 2, \dots$$

Dk:

$$\beta_n = 2^{-\binom{n}{2}} \sum_{G\in\mathcal{G}_n} \beta(G) = 2^{-\binom{n}{2}} \sum_{G\in\mathcal{G}_n} \sum_{l=0}^{n} P(k, H_l(G)) =$$

$$= 2^{-\binom{n}{2}} \sum_{l=0}^{n} \sum_{H\in\mathcal{G}_l} \sum_{\{G : H_l(G) = H\}} P(k; H) \qquad (*)$$

$\sum$ = počet grafů $G \in \mathcal{G}_n$, které mají graf $H \in \mathcal{G}_l$ jako svůj $H_l(G)$

G má $\binom{n}{2}$ možných hran, H má $\binom{l}{2}$ možných hran,

tedy možností je $2^{\binom{n}{2} - \binom{l}{2}}$

$$\beta_n = 2^{-\binom{n}{2}} \sum_{l=0}^{n} \sum_{H\in\mathcal{G}_l} P(k, H)\, 2^{\binom{n}{2} - \binom{l}{2}} = \sum_{l=0}^{n} 2^{-\binom{l}{2}} \sum_{H\in\mathcal{G}_l} P(k, H) =$$

$$= \sum_{l=0}^{n} 2^{-\binom{l}{2}} M_l(k) \le \sum_{l=0}^{\infty} k^l\, 2^{\frac{l}{2}}\, 2^{-\frac{l^2}{2k}} = h(k) < \infty$$

konvergentní řada

a β_n je monotonní

(konvergence se ověří např. podílovým kritériem: $\frac{a_{n+1}}{a_n} < 1$)

Pozn. pro $k = 3$ $h(k) = 197$

$k = 5$ $h(k) = 755\,000$

(*) sčítání přes všechny n-vrcholové grafy provedu tak, že budu sčítat přes l-vrcholové grafy, každý l-vrcholový graf vezmu tolikrát, kolik je grafů G, které ho mají jako svůj podgraf na hladině l

Závěr: Algoritmus založený na prohledávacím stromu tedy rozhodne v průměrném čase $O(1)$, zda existuje obarvení grafu k barvami. Tento výsledek je však způsoben tím, že ve většině případů dostáváme velmi brzy odpověď "ne", a těch málo případů, kdy výpočet trvá exponenciálně dlouho, celkový průměr neovlivní.

Platí dokonce:

Věta. Nechť $\varepsilon > 0$ je reálné číslo. Pak pro skoro všechny grafy platí: má-li graf G n vrcholů, pak

$$\chi(G) > \frac{n}{2(1+\varepsilon)\log n}.$$

dk: Kučera - Kombinatorické algoritmy

Poznámka: Při výpočtu jsme předpokládali, že každý z $2^{\binom{n}{2}}$ možných grafů o n vrcholech je stejně pravděpodobný, tj. vyskytuje se s pravděpodobností rovnou $2^{-\binom{n}{2}}$. Místo toho jsme mohli předpokládat, že G je náhodný graf s pravděpodobností hrany $\frac{1}{2}$.

Náhodný graf: pro každou dvojici vrcholů $i \neq j$ je

$P(\text{mezi } i, j \text{ vede hrana}) = p$

$P(\text{mezi } i, j \text{ nevede hrana}) = 1-p$.

a to nezávisle pro jednotlivé ~~dvojice~~ dvojice vrcholů

Označme $\mathcal{G}_{n,r}$ množinu grafů s n vrcholy a r hranami.

Nechť $G \in \mathcal{G}_{n,r}$ je náhodný graf. Pak

$$P(G) = p^r (1-p)^{\binom{n}{2} - r}.$$

Pro $p = \frac{1}{2}$ je

$$P(G) = \left(\frac{1}{2}\right)^r \left(\frac{1}{2}\right)^{\binom{n}{2} - r} = 2^{-\binom{n}{2}}$$ stejná pro všechna ;

$= P(\text{náhodně vybraného grafu z } \mathcal{G}_n)$.

(tedy všechny grafy jsou stejně pravděpodobné).

Počítat s náhodnými grafy má své výhody – dají se tím postihnout situace, kdy všechny grafy nejsou stejně pravděpodobné.

Navíc se tak dá udělat analýza nejen vzhledem k počtu vrcholů, ale i k počtu hran. Graf, který má n vrcholů a r hran lze považovat za náhodný graf s pravděpodobností hrany $p = \frac{r}{\binom{n}{2}}$.

Zajímá-li nás jenom počet hran v grafu:

$$P(\text{náhodně vybraný } G \in \mathcal{G}_n \text{ má } r \text{ hran}) = \frac{\text{počet grafů } G_{n,r}}{\text{počet grafů } G_n} = \frac{\binom{\binom{n}{2}}{r}}{2^{\binom{n}{2}}} =$$

$= P(\text{náhodný graf s } p = \frac{1}{2} \text{ má } r \text{ hran})$

$$P(\text{náhodný graf má } r \text{ hran}) = \binom{\binom{n}{2}}{r} p^r (1-p)^{\binom{n}{2} - r}$$

binomické rozdělení

./.

pro $p = \frac{r}{\binom{n}{2}}$:

$$P(G \text{ má } r \text{ hran}) = \binom{\binom{n}{2}}{r} \left(\frac{r}{\binom{n}{2}}\right)^{r} \left(1 - \frac{r}{\binom{n}{2}}\right)^{\binom{n}{2}-r}$$

očekávaný počet hran $= \binom{n}{2} \cdot \frac{r}{\binom{n}{2}} = r$