KOMBINATORIKA A GRAFY I.

Tomáš Valla

Katedra aplikované matematiky,
Matematicko-fyzikální fakulta Univerzity Karlovy

13. května 2004

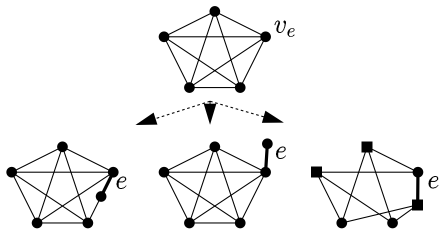
Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

K přednášce Kombinatorika a grafy I. dlouhá léta neexistoval žádný učební text, studenti se museli spoléhat na své poznámky z přednášek. Ujal jsem se proto na popud prof. Jiřího Matouška sepsání toho, co se do Kapitol z diskrétní matematiky nevešlo. Mnohokrát děkuji Michalu Zerolovi (`zerola@matfyz.cz`), který napsal kapitolu o Kuratowského větě. Prof. Matouškovi vděčím za nesčetné rady, opravy a úpravy. Informace jsem čerpal především z přednášky prof. Jana Kratochvíla a prof. Matouška, v menší míře pak z rozličných dalších přednášek, knih, článků či vlastní hlavy.

Jakékoli zaslání chyb (zejména faktických, pravopisných a typografických), zlepšovacích návrhů či názorů na mou adresu je vřele vítáno.

Tomáš Valla
`tom@ucw.cz`

(3) Vrchol v_e je krajním bodem dělení K_5. První dvě možnosti ukazující situaci před kontrakcí by vedly k dělení K_5 v grafu G, zbývající třetí možnost k dělení $K_{3,3}$. Jelikož G neobsahuje ani jeden ze zakázaných grafů, opět jsme dospěli ke sporu.



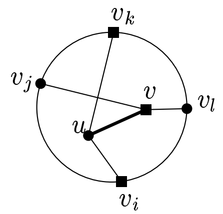
$\square$

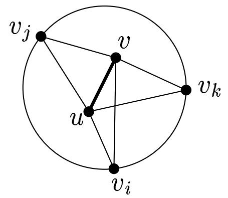
Teď už můžeme dokončit celý důkaz.

Protože graf G' má méně vrcholů než G a neobsahuje dělení zakázaných grafů (využíváme předchozí fakt), je podle indukčního předpokladu rovinný. Vezměme si jeho rovinné nakreslení a vynechme vrchol v_e (vznikl kontrakcí hrany e). Graf $G' \setminus \{v_e\}$ je vrcholově 2-souvislý, a tedy podle lemmatu 3 je hranicí každé stěny kružnice. Vrátíme vrchol v_e do stěny s, ve které ležel a poslední věc, kterou musíme ukázat je, že v_e lze roztáhnout zpět na původní hranu e a zachovat přitom rovinnost.

Označme si sousedy vrcholu v_e na hranici stěny s jako $v_1, \dots, v_m$, uvažujeme pořadí ve směru otáčení hodinových ručiček. Některé z těchto vrcholů jsou v G sousedé vrcholu u, jiní sousedé vrcholu v. Vrchol v_e nelze roztáhnout na hranu e a zachovat přitom rovinnost pouze pokud nastane jeden z následujících případů:

(a) Existují indexy $i < j < k < l$ takové, že v_i, v_k jsou sousedé vrcholu u a v_j, v_l jsou sousedé vrcholu v. Situace je schématicky znázorněna na obrázku vlevo.
(b) Existují vrcholy v_i, v_j, v_k, které jsou sousedé jak u, tak i v (obrázek vpravo).

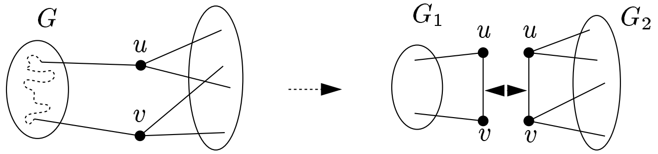




V obou případech bychom v grafu G našli dělení jednoho ze zakázaných grafů (v prvním případě $K_{3,3}$ a v druhém K_5). $\square$

uvážíme grafy $G_1 = G[C_1 \cup \{u,v\}] + \{u,v\}$ a $G_2 = G[C_2 \cup \ldots \cup C_k \cup \{u,v\}] + \{u,v\}$ (přidáme hranu $\{u,v\}$, která v původním grafu vůbec nemusela být).

Mohlo v některém G_i vzniknout dělení zakázaného grafu? Jestliže ano, muselo toto dělení obsahovat hranu $\{u,v\}$. Ale protože v původním G existovala cesta mezi u a v přes $G_j, j \neq i$, dostali bychom i v G dělení zakázaného grafu. Oba grafy tedy splňují podmínky indukčního předpokladu a můžeme je nakreslit bez křížení. Opět podle lemmatu 1 grafy G_1 i G_2 překreslíme tak, aby hrana $\{u,v\}$ byla nakreslená na hranici vnější stěny. Jednotlivá nakreslení ztotožníme v u a v, čímž dostaneme rovinné nakreslení G.



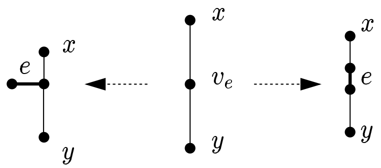
(4) $k_v(G) \geq 3$. Musíme ještě rozebrat případ, kdy je graf G vrcholově alespoň 3-souvislý. Tím se budeme zabývat po zbytek důkazu.

Podle lemmatu 2 v grafu G existuje hrana $e = \{u,v\}$ taková, že $G' = G\%e$ je také vrcholově 3-souvislý.

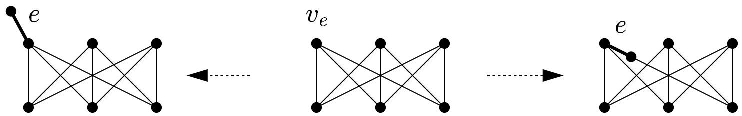
Fakt. *G' rovněž neobsahuje dělení K_5 ani $K_{3,3}$*

Důkaz. Pokud by graf G' nějaké takové dělení obsahoval, určitě by v něm ležel vrchol v_e, jenž vznikl kontrakcí hrany e. Rozebereme si případy, jak mohla vypadat situace v G před kontrakcí hrany e, v závislosti na umístění vrcholu v_e.

(1) Vrchol v_e leží v některé hraně dělení K_5 nebo $K_{3,3}$ (je jejím vnitřním vrcholem). Je zřejmé, že v takovém případě by i v G muselo existovat dělení některého ze zakázaných grafů, což by byl spor s předpokladem věty.



(2) Vrchol v_e je krajním bodem dělení $K_{3,3}$. Jak je vidět i z obrázků, opět by v G bylo dělení $K_{3,3}$ – spor.



Obsah:

1. Úvod do Ramseyovy teorie . . . 5
2. Toky v sítích . . . 9
 2.1 Definice a formulace hlavní věty . . . 9
 2.2 Důkaz hlavní věty . . . 10
 2.3 Fordův–Fulkersonův algoritmus . . . 14
 2.4 Toky v sítích a lineární programování . . . 16
 2.5 Existence maximálního toku . . . 16
3. Míra souvislosti grafů . . . 18
4. Systémy různých reprezentantů . . . 24
 4.1 Důsledky Hallovy věty . . . 26
5. Párování v obecných grafech . . . 29
 5.1 Perfektní párování a Tutteova věta . . . 29
 5.2 Maximální párování a Edmondsův algoritmus . . . 32
6. Hamiltonovské kružnice . . . 38
 6.1 Definice hamiltonovských grafů a Chvátalova věta . . . 38
 6.2 Problém obchodního cestujícího . . . 39
7. Rovinné grafy a Kuratowského věta . . . 42

Důkaz Kuratowského věty. Začneme tou lehčí implikací, zleva doprava. To, že grafy K_5 a $K_{3,3}$ nejsou rovinné (mají příliš mnoho hran), plyne například z důsledků Eulerova vzorce. Konkrétně, K_5 by musel mít maximálně $3v - 6$ hran a $K_{3,3}$ jen $2v - 4$, protože neobsahuje trojúhelníky. Jelikož ani jeden z nich příslušný vztah nesplňuje, K_5 a $K_{3,3}$ nejsou rovinné, a tedy ani jejich dělení nemůže být rovinné.

Tu opačnou, podstatnou implikaci si ukážeme nyní. Budeme postupovat indukcí podle počtu vrcholů (dále značíme n). V indukčním kroku rozlišíme několik případů podle vrcholové souvislosti grafu.

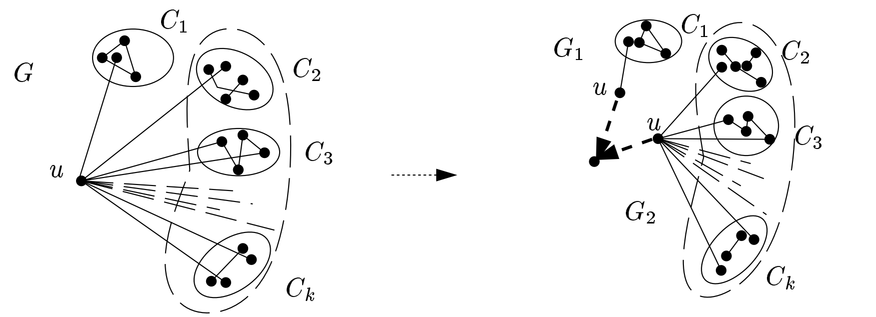
Každý graf na maximálně čtyřech vrcholech je rovinný. Jinými slovy, K_4 je možné nakreslit bez křížení, a tedy i grafy, které z něj vzniknou odebráním hran, zůstanou rovinnými. První krok indukce (pro $n \leq 4$) máme tedy hotový.

Od nepřítele jsme dostali graf G s $n \geq 5$ vrcholy, který neobsahuje dělení K_5 ani $K_{3,3}$. Předpokládejme, že pro menší grafy věta platí, tj. víme, že graf na méně než n vrcholech bez dělení K_5 a $K_{3,3}$ je rovinný. Podíváme se, jaký má G stupeň vrcholové souvislosti a jednotlivé případy dokážeme zvlášť.

(1) $k_v(G) = 0$ aneb graf G je nesouvislý. Na každou komponentu souvislosti použijeme zvlášť indukční předpoklad a výsledky nakreslíme dostatečně daleko od sebe.

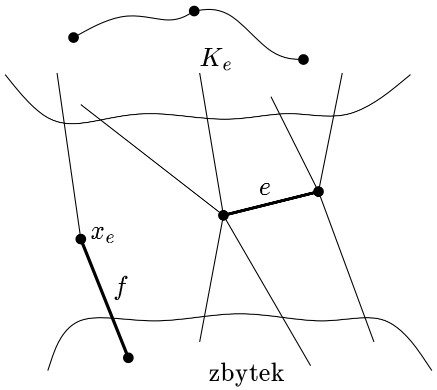
(2) $k_v(G) = 1$. Vrcholová souvislost je rovná 1, takže G obsahuje artikulaci u. Nechť se po odebrání artikulace u graf G rozpadne na komponenty souvislosti $C_1, \ldots, C_k$.

Uvážíme grafy $G_1 = G[C_1 \cup \{u\}]$ a $G_2 = G[C_2 \cup \ldots \cup C_k \cup \{u\}]$. Podle předpokladu žádný z nich neobsahuje dělení K_5 ani $K_{3,3}$ a každý má méně než n vrcholů. Použitím indukčního předpokladu je můžeme nakreslit bez křížení hran. Nyní užijeme lemma 1 a graf G_1 i G_2 překreslíme tak, aby vrchol u ležel na hranici vnější stěny. Jednotlivá nakreslení ztotožníme ve vrcholu u a dostáváme rovinné nakreslení původního grafu G.



(3) $k_v(G) = 2$. Nechť vrcholový řez je množina $\{u, v\}$ a nechť komponenty souvislosti, které vzniknou po jejím odebrání, jsou $C_1, \ldots, C_k$. Nyní

Shrnutím předešlých úvah dostáváme, že pro každou hranu e existuje vrchol x_e, tak že $e \cup \{x_e\}$ je řez v G. Zvolme si nyní hranu e a k ní vrchol x_e tak, abychom po jejich odebrání dostali největší možnou komponentu souvislosti, označíme si ji K_e. Naším cílem je najít nějakou jinou hranu (s příslušným vrcholem), jejímž odebráním vznikne větší komponenta než K_e.



Pokud budeme vrcholy neležící v komponentě K_e nazývat zbytek, z předešlého pozorování nám plyne, že x_e je spojený hranou (označme si ji f) s některým vrcholem ze zbytku. I k této hraně f existuje vrchol x_f, který s ní tvoří vrcholový řez. Rozebereme možnosti, kde všude může ležet vrchol x_f. V každém případě dostaneme větší komponentu souvislosti.

(1) Vrchol x_f leží ve zbytku. Po jeho odebrání komponenta K_e zůstane neporušená a navíc k ní přibude hrana e.
(2) Vrchol x_f je jedním z krajních vrcholů hrany e. Komponenta K_e opět zůstane neporušená a navíc je s ní spojen ještě druhý vrchol hrany e.
(3) Vrchol $x_f \in V(K_e)$. Není úplně jasné, zda-li se komponenta K_e nemůže odebráním x_f rozpadnout. V takovém případě bychom však mohli použít vrcholy x_f a x_e jako vrcholový řez v původním grafu, což předpoklad zakazuje. Odebráním x_f se sice komponenta K_e zmenší o jeden vrchol, nicméně hrana e přispěje dvěma novými.

Každá možnost tedy vede ke sporu. □

Lemma 3. *V libovolném rovinném nakreslení rovinného vrcholově 2-souvislého grafu je hranicí každé stěny (grafová) kružnice.*

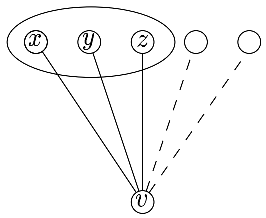
Důkaz. Nepřítele necháme, aby nám zvolil rovinné nakreslení. Pomocí ušatého lemmatu umíme toto nakreslení vytvořit z kružnice přidáváním uší. Každým přilepeným uchem se nějaká stěna rozpadne na dvě nové. □

Nyní nás už od důkazu hlavní věty nic nedělí a můžeme se do něj směle pustit.

1. Úvod do Ramseyovy teorie

Hříčka. *Ve společnosti šesti lidí se vždy vyskytují 3 lidé, kteří se navzájem znají, nebo 3 lidé, kteří se navzájem neznají.*

Důkaz. Společnost si namodelujeme jako graf G, jehož vrcholy budou lidé a hrana mezi nimi povede, pokud se navzájem znají. Zvolme libovolný vrchol v grafu G, Z vrcholu v vedou alespoň 3 hrany nebo 3 nehrany. Ať jsou to nejprve hrany a ať vedou do vrcholů x, y, z.



Pokud mezi těmito třemi vrcholy existuje byť jen jediná hrana, nalezli jsme trojúhelník. A pokud mezi nimi žádná hrana není, máme z nich nezávislou množinu. Když uvažujeme nehrany, je postup shodný. □

Od hříčky přejdeme k plnohodnotnému tvrzení.

Věta (Ramseyova pro grafy). *Pro každé přirozené číslo n existuje přirozené číslo N tak, že libovolný graf na N vrcholech obsahuje úplný podgraf na n vrcholech nebo nezávislou množinu n vrcholů.*

Trochu obecněji: Pro každá přirozená čísla n a r existuje přirozené číslo N takové, že je-li každá hrana grafu K_N obarvena některou z r barev, potom existuje jednobarevný podgraf K_n, tedy podgraf s hranami stejné barvy.

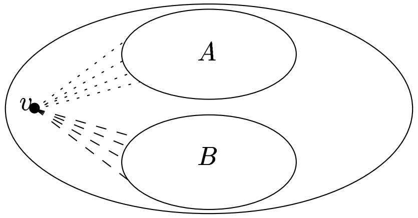
Důkaz. Nejprve dokážeme větu pro dvě barvy ($r = 2$) a až z ní odvodíme větu pro libovolný počet barev.

Definujme číslo $R(k, \ell)$ takto:

$$R(k,\ell) := \min \left\{ N;\ \begin{array}{l}\text{každý } K_N \text{ s hranami obarvenými červeně a modře ob-}\\ \text{sahuje červený } K_k \text{ nebo modrý } K_\ell\end{array} \right\}$$

Potřebujeme dokázat, že $R(n, n) < \infty$, ale ve skutečnosti ukážeme, že dokonce $R(k, \ell)$ je konečné pro každé k, ℓ. Půjdeme na to indukcí podle $k + \ell$.

V případě $k = 1$ nebo $\ell = 1$ stačí vybrat jeden libovolný vrchol. Máme tedy $R(1, \ell) = 1$, $R(k, 1) = 1$. V druhém kroku indukce pro $R(k, \ell)$ určíme, že graf bude mít $R(k-1, \ell) + R(k, \ell-1) + 1$ vrcholů (tedy dvě z indukce hotové velikosti a jeden vrchol navíc, říkejme mu v) a ukážeme, že v takto velkém grafu nutně nalezneme jako podgraf červený K_k nebo modrý K_ℓ.



Z vrcholu v vedou červené hrany do množiny A o alespoň $R(k-1,\ell)$ vrcholech *nebo* modré hrany do množiny B o alespoň $R(k,\ell-1)$ vrcholech. Připomeňme si, že $R(k,\ell)=R(\ell,k)$, definice je symetrická.

(a) Uvažujme množinu A (připojenou červeně). Když A obsahuje červený K_{k-1}, připojíme k němu vrchol v a máme červený K_k, v opačném případě A obsahuje celý modrý K_ℓ.
(b) Uvažujme množinu B (připojenou modře). Když B obsahuje modrý $K_{\ell-1}$, připojíme k němu vrchol v a dostaneme modrý K_ℓ, v opačném případě B obsahuje celý červený K_k.

Všechny možnosti tedy vedou k existenci K_k nebo K_ℓ, $R(k,\ell)\le R(k-1,\ell)+R(k,\ell-1)+1$ (jinými slovy, pro $R(k,\ell)$ mi stačí tolik vrcholů) a z indukce víme, že je tedy $R(k,\ell)$ konečné pro každé k,ℓ. Číslu $R(k,\ell)$ se říká *Ramseyovo číslo* pro grafy o dvou barvách.

Mají-li hrany grafu G tři barvy, červenou, modrou a žlutou, spojíme nejdříve barvy modrou a žlutou do jediné, zelené, a na výsledný graf G' aplikujeme větu pro dvě barvy. Věta nám vrátí potřebnou velikost $N=R(n,n)$ grafu G'. Číslo N pro modro-žlutý graf teď použijeme jako číslo n pro graf červeno-zelený. Jako finální velikost grafu G tedy vezmeme $R(N,N)$. To postačí i pro existenci červeného nebo zeleného podgrafu.

Pro čtyři barvy převedeme stejným postupem problém na větu pro tři barvy, a tak dále, pro libovolný konečný počet barev tedy věta platí. □

Poznámka. Lze také nahlédnout, že pro libovolný konečný počet barev uspěje stejná konstrukce, jakou jsme použili pro dvě barvy. Zadefinujeme si číslo R pro více barev a v indukci stavíme graf z více částí.

Větu si ještě dále zobecníme, namísto hran grafu (tedy dvojic vrcholů) budeme mít obarvené p-tice prvků.

Ramseyova věta pro systémy p-tic. *Pro všechna přirozená čísla n,r,p existuje přirozené číslo N tak, že kdykoli X je N-prvková množina a každá p-prvková podmnožina X je obarvena některou z barev $1,2,\ldots,r$, pak existuje n-prvková množina $Y\subset X$ taková, že všechny p-prvkové podmnožiny Y mají stejnou barvu.*

Důkaz. Budeme postupovat indukcí podle p, pro $p=1$ jde o Dirichletův princip (prvků lze vždy vzít dost na to, aby zaplnily barevné přihrádky), pro $p=2$ je to Ramseyova věta pro grafy (dvojice prvků jsou hrany grafu).

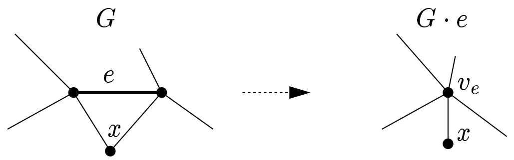
□

Pozorování. *Nechť A je minimální vrcholový řez v grafu G. Potom z každého vrcholu z A vede alespoň jedna hrana do každé komponenty grafu $G\setminus A := G[V(G)\setminus A]$.*

Důkaz. Pro spor předpokládejme, že v A existuje vrchol x, který není spojený s některou komponentou souvislosti grafu $G\setminus A$. Potom ale množina $A\setminus\{x\}$ bude tvořit též vrcholový řez, který má menší počet vrcholů než A, což je spor s její minimalitou. □

Definice. Nechť G je neorientovaný graf a $e=\{u,v\}$ jeho hrana. Operací kontrakce hrany $G\cdot e$ rozumíme „smrštění“ hrany e do jednoho vrcholu, neboli formálně

$$V(G\cdot e)=(V(G)\setminus e)\cup\{v_e\}$$
$$E(G\cdot e)=\Big(E(G)\cap\binom{V(G)\setminus e}{2}\Big)\cup\big\{\{v_e,x\};\ \{x,u\}\in E(G)\vee\{x,v\}\in E(G)\big\}.$$



Lemma 2. *V každém vrcholově 3-souvislém grafu G o alespoň 5 vrcholech existuje hrana e, kontrakcí které se 3-souvislost neporuší. Důsledkem čehož je možné vytvořit každý vrcholově 3-souvislý graf z K_4 několikanásobným opakováním zpětné operace.*

Důkaz. Postupujeme sporem: nechť existuje vrcholově 3-souvislý graf G o alespoň 5 vrcholech a nechť pro každou jeho hranu e platí, že $G\cdot e$ už není vrcholově 3-souvislý (vrchol, který jsme obdrželi kontrakcí hrany e, budeme nazývat v_e).

Z definice souvislosti potom pro každou takovou hranu existuje množina vrcholů $A\subseteq V(G\cdot e)$, s mohutností $|A|<3$, po jejímž odebrání se $G\cdot e$ stane nesouvislým. Pokud by A obsahovala jen jeden vrchol (buď v_e anebo nějaký jiný), nebyl by původní graf G vrcholově 3-souvislý, protože

- když $A=\{v_e\}$, vrcholy hrany e by tvořily řez v G.
- když $A=\{x\}$, $x\neq v_e$, množina A je řez v grafu G.

Nutně tedy $|A|=2$. Mohl by vrchol v_e chybět v množině A? V takovém případě by ji opět bylo možné použít jako řez v G, takže A obsahuje vrchol v_e a nějaký další.

7. Rovinné grafy a Kuratowského věta

Definice. Graf je *rovinný*, pokud existuje jeho nakreslení v Euklidovské rovině, v kterém se nakreslení žádných dvou hran nekříží. *Stěny* nakreslení jsou souvislé oblasti, které vzniknou z roviny po odebrání hran.[6]

V následujícím textu si ukážeme důkaz jedné z nejslavnějších vět z oblasti rovinných grafů. Pochází od polského matematika Kazimierza Kuratowského (1896–1980), po kterém je i pojmenována. Jak uvidíme, z této věty plyne možná až překvapivě jednoduchá kombinatorická charakteristika rovinných grafů.

Definice. Nechť $G = (V, E)$ je graf a $e = \{x, y\} \in E$ jeho hrana. Grafem $G\%e$ nazveme graf

$$G\%e = \left(V \cup \{z\},\ \left(E \setminus \{\{x,y\}\}\right) \cup \{\{x,z\},\{z,y\}\}\right),$$

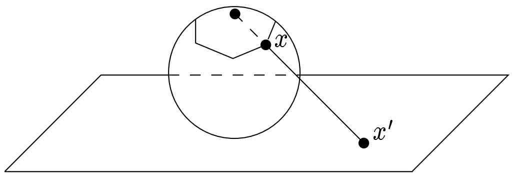
kde $z \notin V$ je nový vrchol. (Na hranu $\{x, y\}$ „přikreslíme“ nový vrchol z.) Operaci % budeme nazývat *dělení hrany* a jejím postupným opakováním dostaneme *dělení grafu*.

Věta (Kuratowského). *Graf je rovinný právě tehdy, když neobsahuje dělení grafu K_5 ani $K_{3,3}$ jako podgraf.*

Původní Kuratowského důkaz byl ale značně rozsáhlý, rozebíral množství možných případů. V nedávné době přišel s elegantnějším důkazem dánský matematik Carsten Thomassen (* 1948), který si předvedeme i my. Než ale přejdeme k samotnému důkazu, ukážeme si základní definice a věty, které v něm budeme potřebovat.

Lemma 1. *Každý rovinný graf je možné nakreslit tak, že předem zvolený vrchol v (resp. předem zvolená hrana e) leží na hranici vnější stěny.*

Důkaz. Využijeme stereografickou projekci, zobrazení, pomocí kterého umíme rovinný graf překreslit na sféru tak, že ani na ní se žádné dvě hrany nekříží. Jedná se o bijekci, takže nakreslení je možné převést i zpětně do roviny. Aby nám vrchol v (případně hrana e) ležel na hranici vnější stěny, stačí nám vhodně pootočit sférické nakreslení před projekcí do roviny. Konkrétně tak, aby „severní pól“ ležel ve stěně incidentní s vrcholem v (případně hranou e).



[6]Tato definice je spíše intuitivní, nám ale bude postačovat. Pokud by čtenáře zajímalo její formální znění, může ho najít například v knize *Kapitoly z diskrétní matematiky* od *J. Matouška* a *J. Nešetřila*.

Na chvíli si mysleme, že potřebnou velikost N množiny X už známe a p-tice prvků X jsou obarveny $r = 2$ barvami. Položíme $X_0 := X$ a uděláme tento krok pro $i = 1, \ldots, 2n - 1$:

- Zvolíme libovolný prvek $x_i \in X_{i-1}$ a obarvíme každou $(p-1)$-prvkovou podmnožinu S množiny $X_{i-1} \setminus \{x_i\}$ takovou barvou, jakou měla p-tice $S \cup \{x_i\}$. Podle indukčního předpokladu existuje dostatečně velká podmnožina X_i množiny $X_{i-1} \setminus \{x_i\}$, v níž všechny $(p-1)$-tice mají stejnou barvu b_i.

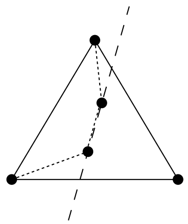
Některá z barev se objeví (z přihrádkového principu) mezi barvami b_i alespoň n-krát, dejme tomu červená. Potom $\{x_i;\ b_i = \text{červená}\}$ je hledanou n-prvkovou množinou Y.

Pro $r > 2$ barev se důkaz téměř nezmění, budeme pouze provádět více kroků, pro $i = 1, 2, \ldots, rn - 1$. Velikost N množiny X tedy lze zvolit dostatečně obrovskou (ale stále ještě konečnou) tak, aby množina X postačila na všechny kroky. □

Použití Ramseyovy věty pro p-tice si ukážeme v následujícím tvrzení.

Věta (Erdős–Szekeres). *Pro každé k existuje N tak, že libovolná N-prvková množina bodů v rovině v obecné poloze (žádné 3 body neleží na přímce) obsahuje množinu vrcholů konvexního k-úhelníku.*

Důkaz. Větu dokážeme indukcí podle k. Pro $k = 4$ je $N = 5$. Sestrojíme si konvexní obal oněch pěti bodů, což může být buď pětiúhelník nebo čtyřúhelník (v tom případě je to v pořádku) anebo trojúhelník. Tehdy však vždy můžeme sestrojit čtyřprvkovou konvexní množinu vynecháním jednoho vrcholu trojúhelníku: Přímka spojující dva vnitřní body vždy odřízne jeden vrchol trojúhelníku.



V libovolné $k > 4$ prvkové množině bodů si obarvíme každou čtveřici červeně, je-li její konvexní obal čtyřprvkový, nebo modře, je-li její konvexní obal tříprvkový (čili když čtveřice netvoří konvexní množinu).

červeně modře

Ramseyova věta s $p = 4$, $r = 2$ a $n = k$ říká, že existuje N-bodová množina X a její k-bodová podmnožina Y taková, že všechny čtveřice na ní mají tutéž barvu. Protože už na pěti bodech existuje nějaká červená čtveřice (případ $k = 4$), všechny čtveřice na Y budou červené, nebude s v ní vyskytovat ani jedna nekonvexní čtveřice, takže celá Y je konvexní. □

Poznámka. Větu je také možné dokázat metodami kombinatorické geometrie, odhad nutné velikosti množiny je jejich použitím daleko nižší, nicméně důkaz pomocí Ramseyovy věty je podstatně jednodušší.

(5) Vrátíme S.

Tvrzení. *Algoritmus je polynomiální a vrátí trasu obchodního cestujícího S, jejíž délka je*

$$w(S) \leq 2 \cdot (\textit{optimální řešení TSP}).$$

Důkaz. Kroky 1 až 3 jistě umíme provést v polynomiálním čase. V každé iteraci kroku 4 snížíme $\deg^+(u)$ a $\deg^-(u)$ o jednu, krok 4 se tedy provede nejvýše tolikrát, kolik je počet hran grafu G. Algoritmus je tak polynomiální.

Při běhu algoritmu platí invariant, že sled S prochází všemi vrcholy grafu G. Na konci neexistuje vrchol, ze kterého by vycházely dvě hrany, a sled S tedy prochází každým vrcholem právě jednou.

V kroku 1 je jistě $w(T) \leq \mathrm{OPT}$. V kroku 3 se díky orientaci $\vec{T}$ započítá každá hrana dvakrát, tedy

$$w(S) = w(\vec{T}) = 2w(T) \leq 2 \cdot \mathrm{OPT}.$$

Protože ohodnocení hran splňuje trojúhelníkovou nerovnost, hodnota $w(S)$ již opakováním kroku 4 nevzroste. □

Poznámka. Problém obchodního cestujícího je vzhledem ke své praktické využitelnosti intenzivně studován, sluší se proto říci i něco o současných známých výsledcích.

Obecný TSP je NP-těžké c-aproximovat s konstantním faktorem c. U symetrického TSP (tj. $w(x, y) = w(y, x)$) není situace lepší, c-aproximace je stále NP-těžká.

Splňuje-li vzdálenostní funkce trojúhelníkovou nerovnost a je-li symetrická, problém sice zůstává NP-úplný, nicméně je znám algoritmus aproximující s faktorem $3/2$. Tento algoritmus byl zformulován v sedmdesátých letech a od té doby se nikomu nepodařilo faktor $3/2$ vylepšit. Lze dokázat, že existuje konstanta c, $1 < c < 1.1$ taková, že je již NP-těžké symetrický trojúhelníkový TSP c-aproximovat. Pro asymetrický TSP s trojúhelníkovou nerovností je znám algoritmus aproximující s faktorem $O(\log n)$.

Další odrůdou je euklidovský TSP, tedy body rozmístěné do $\mathbb{R}^d$ a s euklidovskou metrikou. Problém lze aproximovat s faktorem $1 + \varepsilon$ pro každé pevné $\varepsilon > 0$. (Říkáme, že existuje *polynomiální aproximační schéma.*)

6.2 Problém obchodního cestujícího.

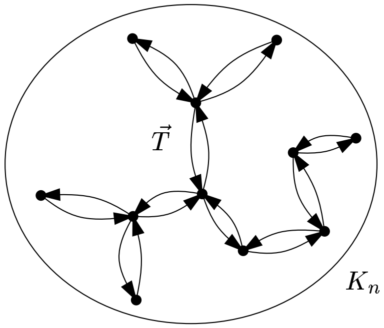
Problém obchodního cestujícího (Traveling Salesman Problem – TSP)[5] je další slavnou NP-úplnou úlohou. Máme daný graf G s ohodnocenými hranami, a úkolem je najít hamiltonovskou kružnici s minimálním součtem vah. Motivací pro TSP je, jak název napovídá, rozvrhnout obchodnímu cestujícímu program návštěv jednotlivých měst, aby v každém byl právě jednou, skončil opět doma, a co nejvíce ušetřil na benzínu.

Jelikož je TSP algoritmicky obtížný, ukážeme si alespoň algoritmus, který dává přibližný výsledek. Takovým algoritmům se říká *aproximační*. Náš algoritmus však funguje *pouze* na úplných grafech splňujících trojúhelníkovou nerovnost.

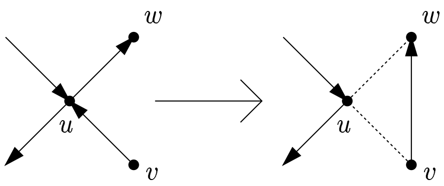
Aproximační algoritmus. Vstupem je úplný graf K_n s nezáporným ohodnocením hran w. Pro w platí trojúhelníková nerovnost, tj. pro každé tři hrany tvaru $\{x,y\}$, $\{y,z\}$, $\{x,z\}$ je

$$w(\{x,y\}) + w(\{y,z\}) \geq w(\{x,z\}).$$

(1) Najdeme v K_n minimální kostru T.
(2) Graf G (včetně kostry T) symetricky zorientujeme, zachováme při tom váhy hran.



(3) Z kostry T vznikne symetrickou orientací podgraf $\vec{T}$, který lze nakreslit jedním tahem. Sledu hran, který tím dostaneme, říkejme S.
(4) Dokud existuje vrchol u s alespoň dvěma výstupními hranami, nahrazujeme v sledu S hrany (u,w) a (v,u) hranou (v,w).



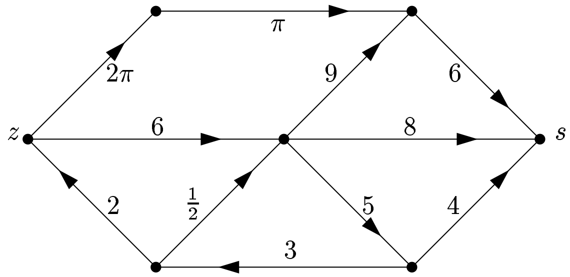
[5]Politicky korektní ovšem je „Traveling Salesperson Problem“.

2. Toky v sítích

2.1 Definice a formulace hlavní věty.

Definice. *Síť* nazveme čtveřici (G,z,s,c), kde $G=(V,E)$ je orientovaný graf, z a s dva různé vrcholy grafu G (říkáme jim *zdroj* a *stok*) a *kapacita* $c\colon E \to \mathbb{R}_0^+$ je funkce ohodnocující hrany nezápornými reálnými čísly.

Příklad sítě:



Síť si tedy lze představovat jako soustavu jednosměrných vodovodních trubek, kde každá má předepsanou maximální propustnost, a dvě význačná místa – zdroj a stok. Ve zdroji pouštíme vodu dovnitř a ve stoku nám voda vytéká ven. Tok je potom rozvržení, jak skutečně v našich trubkách voda proudí. Trubkou nesmí téci více, než je její kapacita, a vodovod „není děravý“ – to, co vteče do určitého vrcholu sítě (s výjimkou zdroje a stoku), zase vyteče ven. (Druhá podmínka je také známá pod názvem *Kirchhoffovy zákony*.) Velikostí toku se rozumí množství vody, které přebývá ve stoku, nebo ekvivalentně kolik vody je třeba posílat do zdroje. Ačkoli se to z náhledu jeví zřejmé, v definici si velikost toku zavedeme jako množství vody posílané do zdroje a rovnost obou veličin ukážeme formálně později.

Definice. *Tok v síti* je každá funkce $f\colon E \to \mathbb{R}_0^+$ splňující

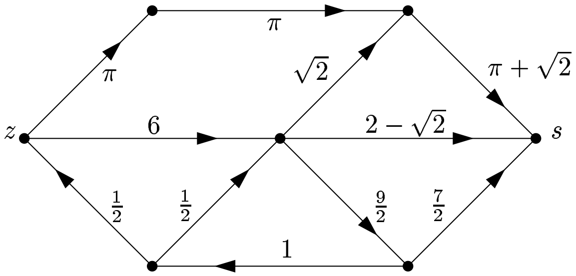
(1) Pro každou hranu $e \in E$ je $0 \leq f(e) \leq c(e)$.
(2) Pro každý vrchol $u \in V$ mimo zdroj a stok platí

$$\sum_{(x,u)\in E} f(x,u) = \sum_{(u,y)\in E} f(u,y).$$

Velikost toku je

$$w(f) = \sum_{(z,x)\in E} f(z,x) - \sum_{(x,z)\in E} f(x,z).$$

V našem příkladu tedy může existovat následující tok:



Budeme zkoumat, jaký maximální tok může danou sítí procházet. Naším předchozím příkladem prochází tok velikosti $\pi + \frac{11}{2}$, který však zřejmě maximální není. Začneme následujícím:

Tvrzení. *Pro každou síť existuje maximální tok.*

Toto tvrzení není vůbec samozřejmé. Toků je nekonečně mnoho a jejich velikosti jsou obecně reálná čísla, a tedy je třeba opatrnost (například interval $I = (0, 1)$ na reálné ose nemá maximum). Důkaz tvrzení používá metody matematické analýzy a pro nás je spíše postranním tématem, takže ho odložíme až na konec kapitoly. Pro celočíselné a racionální kapacity navíc dostaneme existenci maximálního toku jednodušším způsobem, taktéž uvedeným později.

Teď zavedeme další důležitý pojem: řez.

Definice. *Řezem mezi zdrojem z a stokem s* v síti $G = (V, E, z, s, c)$ nazveme množinu hran $R \subseteq E$ takovou, že v síti $G' = (V, E \setminus R, z, s, c)$ neexistuje žádná orientovaná cesta ze zdroje do stoku.

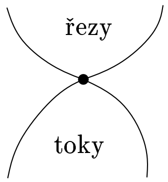
Kapacita řezu je $c(R) = \sum_{e \in R} c(e)$.

V této sekci budeme pro krátkost říkat řezu mezi z a s prostě řez.[1]

Narozdíl od toků je řezů konečně mnoho (nejvýš tolik jako všech podmnožin množiny E), a tudíž zjevně existuje řez minimální kapacity. Nyní vyslovíme nejdůležitější výsledek o tocích.

Hlavní věta o tocích. *Pro každou síť se velikost maximálního toku rovná kapacitě minimálního řezu:*

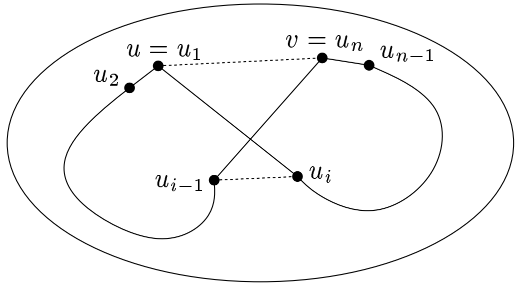
$$\max_{f \text{ tok}} w(f) = \min_{R \text{ řez}} c(R)$$



[1]Pojem *řez* má ve skutečnosti mnohem širší význam, zde však nebude hrozit nebezpečí zmatení.

Zpětným obíráním Chvátalova uzávěru $[G]$ o hrany se nakonec vrátíme zpět ke grafu G, který bude mít hamiltonovskou kružnici.

Hamiltonovskou kružnici v $G + \{u, v\}$ si označíme K. Pokud $\{u, v\} \notin K$, je vše v pořádku, K je hamiltonovskou kružnicí i v G. Nechť nadále $\{u, v\} \in K$. Předvedeme si, že hranu $\{u, v\}$ lze „přemostit“ jako zde:



Bez újmy na obecnosti $K = (u = u_1, u_2, \ldots, u_n = v)$. Množina A bude obsahovat čísla i kandidátů na počáteční vrchol u_i, množina B čísla i kandidátů na předcházející vrchol u_{i-1}:

$$A = \{i;\ \{u, u_i\} \in E(G)\} \qquad B = \{i;\ \{v, u_{i-1}\} \in E(G)\}.$$

Neprázdný průnik A a B bude znamenat existenci hrany $\{u_i, u_{i-1}\}$ jako na obrázku.

Zřejmě $|A| = \deg u$ a $|B| = \deg v$, a také určitě $A \subseteq \{2, 3, \ldots, n-1\}$ a $B \subseteq \{3, 4, \ldots, n\}$. Proto $|A \cup B| \leq n - 1$. Nyní ověříme, že $A \cap B \neq \emptyset$:

$$|A \cap B| = |A| + |B| - |A \cup B| = \deg u + \deg v - |A \cup B| \geq n - |A \cup B| \geq n - (n-1) \geq 1.$$

Tedy pro nějaké $i \in A \cap B$ se v grafu G vyskytují hrany $\{u, u_i\}$ a $\{v, u_{i-1}\}$. Hamiltonovská kružnice K' grafu G tudíž bude tvaru

$$K' = (u_1, u_2, \ldots, u_{i-1}, v = u_n, u_{n-1}, \ldots, u_i, u).$$

□

Důsledek. *Nechť v grafu G pro každé dva vrcholy $u, v \in V(G)$, $\{u, v\} \notin E(G)$ platí $\deg u + \deg v \geq n$. Pak G je hamiltonovský.*

Důkaz. Graf $[G]$ je úplný graf a ten jistě má hamiltonovskou kružnici. Tedy je hamiltonovský i G. □

Důsledek (Diracova věta). *Nechť v grafu G pro každý vrchol $v \in V(G)$ platí $\deg v = \frac{n}{2}$. Potom G je hamiltonovský.*

Důkaz. Zřejmé z předchozího důsledku. □

6. Hamiltonovské kružnice

6.1 Definice hamiltonovských grafů a Chvátalova věta.

Definice. *Hamiltonovská kružnice* v grafu G je kružnice obsahující všechny vrcholy.

Otázka, zda v grafu existuje hamiltonovská kružnice, je slavný NP-úplný problém. Proto se hledají nejrůznější postačující podmínky pro její existenci. Ukážeme si zajímavou vlastnost hamiltonovských grafů, kterou objevil český matematik Václav Chvátal. Jako důsledek z ní vyplyne například i Diracova věta. Chvátalova věta sice nedává algoritmicky použitelnou metodu, jak hamiltonovské grafy rozpoznávat, nicméně je teoreticky hezká a dá se využít v nejrůznějších heuristických vylepšeních jednoduchého backtrackového algoritmu, která běží rychleji na hustších grafech.

Definice. Mějme graf $G = (V, E)$. Na počátku položíme $E_0 := E$ a zavedeme graf $[G]^0 := (V, E_0)$. Budeme opakovat následující krok pro $i = 0, 1, 2, \ldots$ tak dlouho, dokud to půjde:

- Když existují dva vrcholy $u, v \in V$, $\{u, v\} \notin E_i$, splňující
$$\deg_{[G]^i} u + \deg_{[G]^i} v \geq |V|,$$
pak položíme $E_{i+1} := E_i \cup \{u, v\}$ a $[G]^{i+1} := (V, E_{i+1})$.

Graf $[G] := [G]^k$ pro nejvyšší dosažené k nazveme *Chvátalovým uzávěrem* grafu G.

Pozorování. *Chvátalův uzávěr je určen jednoznačně.*

Důkaz. Mějme graf G a pro spor vezměme dva různé Chvátalovy uzávěry $[G]_1$ a $[G]_2$. Do $[G]_1$ jsme postupně přidávali hrany $e_1, \ldots, e_k$, do $[G]_2$ hrany $f_1, \ldots, f_r$. Zvolíme hranu $e_i = \{u, v\}$ první takovou, že $e_i \notin E([G]_2)$.

V čase $i - 1$ byly stupně vrcholů u a v o jednu menší než v čase i a jejich součet byl alespoň $|V|$. Proto jistě
$$\deg_{[G]_2} u + \deg_{[G]_2} v \geq \deg_{[G]_1^{i-1}} u + \deg_{[G]_1^{i-1}} v \geq |V|.$$
Hrana e_i tak nutně musela být zařazena i do $[G]_2$. □

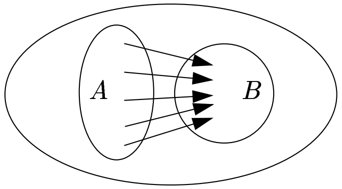
Věta (Chvátal). *Graf G má hamiltonovskou kružnici, právě když hamiltonovskou kružnici má graf $[G]$.*

Důkaz. Implikace zleva doprava je zřejmá. Pro opačný směr nám stačí ukázat následující tvrzení:

- Nechť pro dva vrcholy $u, v \in V(G)$, $\{u, v\} \notin E(G)$ takové, že $\deg_G u + \deg_G v \geq n$, je graf $G + \{u, v\}$ hamiltonovský. Potom je hamiltonovský i G.

2.2 Důkaz hlavní věty.

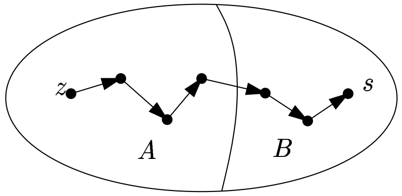
Začneme podrobnějším studiem řezů. Zavedeme si konvenci pro značení množiny hran, které vedou z části sítě A do části B:



$$S(A, B) = \{(x, y);\ x \in A,\ y \in B,\ (x, y) \in E\}$$
Pro kapacitu a velikost toku analogicky
$$c(A, B) = c(S(A, B))$$
$$f(A, B) = f(S(A, B)) = \sum_{e \in S(A,B)} f(e).$$

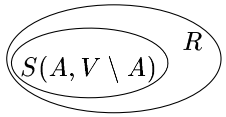
Je dobré si uvědomit následující více či méně viditelné vlastnosti řezů.

Vlastnost 1. *Rozdělíme-li síť na dvě disjunktní podmnožiny vrcholů A a B tak, že $z \in A$ a $s \in B$, potom množina $S(A, B)$ je řez (říkáme mu elementární řez).*



Důkaz. Kdyby existovala orientovaná cesta ze zdroje do stoku, museli bychom na ní někde narazit na rozhraní množin A a B. Taková hrana však podle definice patří do $S(A, B)$, množina $S(A, B)$ tedy tvoří řez. □

Vlastnost 2. *Každý řez obsahuje elementární řez.*



Důkaz. Vezmeme množinu A vrcholů, do kterých vede v rozříznuté síti orientovaná cesta ze zdroje. Řez R potom obsahuje elementární řez $S(A, V \setminus A)$, protože kdyby hrana $e = (u, v)$ z $S(A, V \setminus A)$ nepatřila do řezu R (tedy $u \in A$, ale $v \notin A$), existovala by orientovaná cesta $z, \ldots, u, v$ v rozříznuté síti, takže by byl $v \in A$. □

Vlastnost 3. *Každý v inkluzi minimální řez R je elementární; minimální v inkluzi znamená, že $R \setminus \{e\}$ není řez pro žádnou hranu $e \in R$.*

Důkaz. Vezmeme množinu A vrcholů, do kterých vede v říznuté síti orientovaná cesta ze zdroje. Z předchozího víme, že elementární řez $S(A, V \setminus A)$ je obsažen v R. Jenže z R nelze odebrat žádnou hranu, tedy $R = S(A, V \setminus A)$. □

Lemma. *Pro každou $A \subseteq V$ obsahující zdroj a ne stok a pro libovolný tok f platí $w(f) = f(A, V \setminus A) - f(V \setminus A, A)$, speciálně tedy $w(f) = \sum_{(x,s)\in E} f(x,s) - \sum_{(s,x)\in E} f(s,x)$.*

Důkaz. Pro každý vrchol u z A mimo zdroj platí Kirchhoffův zákon

$$\sum_{(u,x)\in E} f(u,x) - \sum_{(x,u)\in E} f(x,u) = 0,$$

pro zdroj pak

$$\sum_{(z,x)\in E} f(z,x) - \sum_{(x,z)\in E} f(x,z) = w(f).$$

Sečteme dohromady do jediné rovnice levé strany a pravé strany předchozích rovnic, a po úpravě dostaneme

$$\sum_{\substack{u\in A\\ v\notin A}} f(u,v) - \sum_{\substack{u\notin A\\ v\in A}} f(u,v) = w(f),$$

což je přesně $f(A, V \setminus A) - f(V \setminus A, A) = w(f)$. □

Když si dáme dohromady zjištěná fakta, dostaneme následující nerovnost mezi maximálním tokem a minimálním řezem, jež je snadnější částí hlavní věty. Z trubkové analogie se sice zdá být zřejmou, nicméně chceme-li ji dokázat poctivě, dá to nějakou práci.

Důsledek. *Pro každý tok f a každý řez R je $w(f) \le c(R)$. Tudíž maximální tok je nejvýš roven minimálnímu řezu.*

Důkaz. V daném řezu R je obsažen nějaký elementární řez $S(A, V \setminus A)$ takový, že $z \in A$. Pro velikost toku f potom platí

$$w(f) = f(A, V \setminus A) - f(V \setminus A, A) \le f(A, V \setminus A) \le c(A, V \setminus A).$$

V řezu R jsou oproti elementárnímu řezu $S(A, V \setminus A)$ nějaké hrany dokonce navíc, takže máme $c(A, V \setminus A) \le c(R)$. □

Teď zavedeme klíčový pojem nasycené cesty. Dosud jsme mluvili o *orientovaných* cestách (a také jsme to pečlivě zdůrazňovali). Po zbytek této kapitoly budeme *cestou* rozumět posloupnost $(v_0, e_1, v_1, e_2, \ldots, v_{m-1}, e_m, v_m)$, kde

Poznámka. Jak vidno, nejvíce času jsme trávili při operaci $G^{(i)} \to G^{(i+1)}$. Klíčem k lepší implementaci algoritmu je naprogramování šikovné datové struktury, která bude zaznamenávat pouze rozdíly mezi G a G'. To jde v čase (velikost květu) $\cdot\, O(n)$, což dává výslednou časovou složitost $O(n^3)$.

V současnosti nejrychlejší známý algoritmus na hledání maximálního párování pochází od Micali Vaziraniho a pracuje v čase $O(\sqrt{n} \cdot m)$, kde m je počet hran grafu G.

Některé vrcholy grafu G se v Edmondsově lese vůbec neobjeví, říkejme jim třeba *kompost*. Uvnitř kompostu jsou ale všechny vrcholy nějak spárovány, neboť všechny volné vrcholy jsou kořeny Edmondsova lesa. Takže nás kompost nadále nijak nezajímá.

(2) Pokud existuje hrana e grafu G mezi sudými hladinami různých stromů, nalezli jsme volnou střídavou cestu. Přidáme hranu e do párování a jdeme na 1. krok.

(3) Pokud existuje hrana mezi sudými hladinami jednoho stromu, nalezli jsme květ C. Tedy, ještě vlastně nenalezli, ale když se od obou konců hrany vydáme vzhůru, ve vrcholu v (nejpozději v kořeni) se obě cesty setkají. Cesta od v nahoru do kořene je stonkem květu C. Zkontrahujeme květ C do jednoho vrcholu (květ „ustřihneme") a rekurzivně zavoláme algoritmus na graf $G \cdot C$ s párováním $M \setminus E(C)$. Co mohl algoritmus vrátit:

   a) Staré párování $M \setminus E(C)$ beze změny. V tom případě vrátíme M (kteréžto párování je maximální v G) a skončíme.

   b) Nějaké větší párování M'. Položíme tedy $M := M' \cup E_M(C)$ a jdeme na 1. krok.

(4) Když neexistuje hrana mezi sudými hladinami (a tedy ani volná střídavá cesta), je $M' := M$ maximální párování.

Tvrzení. *Edmondsův algoritmus najde v polynomiálním čase maximální párování v grafu G.*

Důkaz. Při důkazu správnosti si uvědomíme, kdy může algoritmus skončit:

(1) V kroku 4, to když už v grafu neexistuje žádná volná střídavá cesta. Podle lemmatu 1 je takové párování maximální.

(2) V kroku 3a, to když algoritmus nalezl maximální párování v kontrahovaném grafu. Potom podle lemmatu 2 je nalezené párování maximální i v grafu původním.

Dosud jsme nijak blíže nepopsali důležitou operaci – kontrakci květu C a vytvoření menšího grafu G'. Vyzkoušíme nejjednodušší metodu:

(1) Nechť k je vrchol nahrazující C.

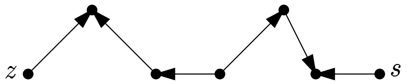
(2) Zapíšeme do G' každou hranu nemající ani jeden vrchol společný s C.

(3) Zapíšeme do G' každou hranu, která je napojena do C jedním vrcholem.

Časová náročnost takového postupu je $O(n^2)$, kde n je počet vrcholů. Transformace $G \to G'$ se v rekurzivním volání algoritmu promítne jako posloupnost tvorby grafů $G, G', G'', \dots, G^{(l)}$, nebude jich však více než $n/2$, neboť při každé kontrakci ubudou alespoň dva vrcholy.

Návrat z rekurze od $G^{(i)}$ zpět k $G^{(i-1)}$ a rekonstrukci střídavé cesty zvládneme v čase $O(n)$. Nalezení volné střídavé cesty tedy zabere $O(n^3)$, celkový čas tak bude $O(n^4)$, protože volných střídavých cest je v grafu nejvýše n. $\square$

$v_0, \dots, v_m$ jsou navzájem různé vrcholy uvažované sítě, a pro každé i je $e_i = (v_{i-1}, v_i)$ nebo $e_i = (v_i, v_{i-1})$. Ignorujeme tedy orientaci hran, a například cesta ze zdroje do stoku (o takových budeme nyní hlavně hovořit) může vypadat třeba takto:



Definice. Cesta v právě uvedeném smyslu se nazývá *nasycená* (vzhledem k danému toku f), pokud pro nějakou hranu $e_i = (v_{i-1}, v_i)$ orientovanou po směru je $f(e_i) = c(e_i)$ nebo pro nějakou hranu $e_i = (v_i, v_{i-1})$ orientovanou proti je $f(e) = 0$. Cestě, která není nasycená, budeme říkat *nenasycená*.

Tok nazveme nasycený, když každá cesta ze zdroje do stoku je nasycená.

Nasycená cesta je tedy taková, po níž teče největší možný tok.

Nyní si ukážeme důležité vlastnosti maximálního toku.

Tvrzení. *Tok f je maximální, právě když je nasycený. Pro každý maximální tok f existuje řez R takový, že $w(f) = c(R)$.*

Důkaz. Nejdřív sporem dokážeme, že maximální tok je nasycený. Nechť je f maximální a přesto nenasycený. Potom existuje nenasycená cesta P ze zdroje do stoku. Nyní najdeme možné vylepšení toku ve směru ze zdroje do stoku

$$\varepsilon_1 = \min\{c(e) - f(e);\ e \in P \text{ orientovaná po směru}\}$$

a vylepšení proti směru

$$\varepsilon_2 = \min\{f(e);\ e \in P \text{ orientovaná proti směru}\}$$

a vezmeme menší z nich: $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$. Cesta P je nenasycená, proto $\varepsilon_{1,2} > 0$, a tím spíše $\varepsilon > 0$. Stávající tok změníme na

$$f'(e) = \begin{cases} f(e) + \varepsilon & e \in P \text{ orientovaná ve směru ze zdroje do stoku} \\ f(e) - \varepsilon & e \in P \text{ orientovaná proti směru} \\ f(e) & e \notin P. \end{cases}.$$

Díky volbě ε se nepřekročí kapacity hran, a protože v rámci jednoho vrcholu z cesty P se číslo ε přičte i odečte, zůstanou v platnosti i Kirchhoffovy zákony. Máme teď nový tok a zároveň $w(f') = \sum f'(z, x) - \sum f'(x, z) = w(f) + \varepsilon$, tedy dokonce lepší tok než f. Tok f nebyl maximální, což je spor.

Nyní dokážeme, že je-li tok f nasycený, pak je maximální. Zvolíme množinu A takových vrcholů v, že existuje nenasycená cesta ze zdroje do v. Tok je nasycený, bude tedy $z \in A$, ale $s \notin A$ (tok je nasycený), pro každou hranu e z $S(A, V \setminus A)$ je $f(e) = c(e)$ a druhým směrem pro hranu e z $S(V \setminus A, A)$ je $f(e) = 0$.

Velikost toku je

$$w(f) = f(A, V \setminus A) - f(V \setminus A, A) = c(A, V \setminus A) - 0 = c(A, V \setminus A).$$

Víme, že vždy $w(f) \le c(R)$ pro každý tok f i řez R, a my jsme zde dokonce nyní našli takový řez R, že $w(f) = c(R)$. Tok f tedy určitě bude největší možný. Zároveň jsme zkonstruovali řez R, pro který platí rovnost $w(f) = c(R)$. □

Pozorný čtenář si možná všiml, že už jsme dokázali celou hlavní větu o tocích: Víme, že pro danou síť existuje maximální tok f, a podle druhé části právě dokázaného tvrzení k němu existuje řez o kapacitě $w(f)$, a ten je nutně minimální.

2.3 Fordův–Fulkersonův algoritmus.

Důkaz hlavní věty o tocích dává jednoduchý algoritmus na hledání maximálního toku.

Algoritmus (Ford, Fulkerson).

(1) Polož $f(e) = 0$ pro všechny hrany e.
(2) Pokud existuje nenasycená cesta P ze zdroje do stoku, najdi ε, vylepši cestu P a opakuj tento bod tak dlouho, dokud nějaká nenasycená cesta existuje.
(3) Stávající f je maximálním tokem.

Pozorování. *Jsou-li všechny kapacity v síti racionální čísla, je Fordův–Fulkersonův algoritmus konečný. Pro síť s racionálními kapacitami tedy dokazuje existenci maximálního toku.*

Důkaz. Po případném vynásobení všech kapacit jejich společným jmenovatelem můžeme předpokládat, že kapacity jsou dokonce celočíselné. Potom ε v algoritmu bude vždy alespoň 1, a velikost toku se tudíž zvýší v jednom kroku alespoň o 1. Algoritmus proto nutně skončí, a tedy najde nasycený tok, jenž je i maximální. □

Bohužel existuje iracionálně ohodnocená síť, na které Fordův–Fulkersonův algoritmus nedoběhne, a dokonce ani nebude konvergovat ke správnému výsledku.

Protipříklad. Příklad sítě, kde se Fordův–Fulkersonův algoritmus zacyklí, jako první podali sami autoři algoritmu. My si ukážeme podstatně jednodušší síť navrženou Uri Zwickem, na které lze simulovat výpočet členů posloupnosti $\{a_n\}$ definované $a_0 = 1$, $a_1 = r$ a $a_{n+2} = a_n - a_{n+1}$, kde $r = \frac{\sqrt{5}-1}{2}$. Potom ovšem $r^2 = 1 - r$, přenásobením obou stran rovnice číslem r dostaneme $r^{n+2} = r^n - r^{n+1}$, platí tedy $a_n = r^n$. Vezmeme následující síť a cesty p_1, p_2 a p_3:

(2) Když P' končí v k, u výsledné cesty v G potřebujeme zajistit, aby byl koncový vrchol volný. Vrchol k není v G' spárovaný a tudíž má květ C stonek nulové délky. Z cesty P tedy můžeme udělat volnou střídavou cestu obejitím květu po správné straně a zakončením ve vrcholu, kde sousedí dvě nepárovací hrany.
(3) Když k je vnitřní vrchol P', potom P' vrchol k právě z jedné strany opouští párovací hranou. V grafu G tedy nutně musí příslušná cesta P vstupovat do květu C nepárovací hranou a opouštět květ stonkem. Stačí jen správně zvolit směr, kterým obejít květ C, a propojit oba úseky cesty.

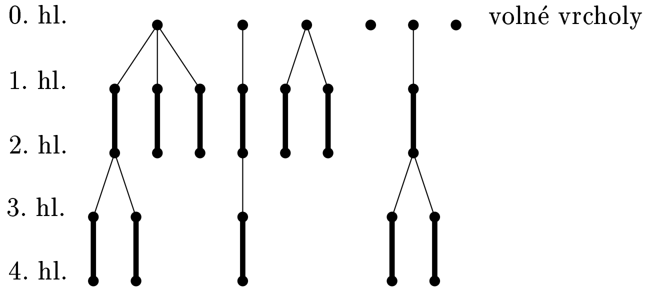
□

Idea algoritmu je poměrně jednoduchá. Vyzbrojeni lemmatem 1, chodíme po grafu a vyhledáváme volné střídavé cesty, které vylepšujeme. Komplikace nastane, když se v grafu vyskytne květ. Tehdy není jasné, jestli se po něm máme vydat zleva nebo zprava. Situaci vyřešíme zkontrahováním květu do jediného vrcholu, rekurzivním voláním algoritmu na zmenšený graf a následnou rekonstrukcí maximálního párování v původním grafu.

Nyní podrobněji:

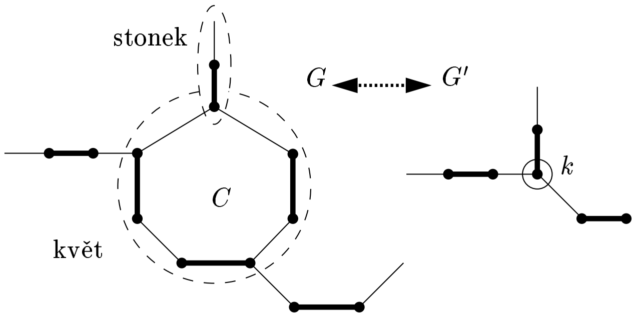
Edmondsův „zahradní“ algoritmus. Vstupem algoritmu je graf G a jeho libovolné párování M, třeba prázdné. Výstupem je maximální párování M'.

(1) Zkonstruujeme maximální možný *Edmondsův les* vzhledem k aktuálnímu M:



Co to znamená? Do nulté hladiny umístíme všechny volné vrcholy. Z každého z nich pustíme prohledání grafu do hloubky a sestrojíme maximální strom takový, že mezi hladinami se pravidelně střídají párovací a nepárovací hrany.

Edmondsův les je vhodný pro znázornění, při skutečném programování bychom z volného vrcholu spustili prohledání do hloubky a ke každému vrcholu si pamatovali jeho hladinu a z jakého volného vrcholu byl navštívený. Existence hrany uvnitř jednoho stromu Edmondsova lesa je totéž, jako když při prohledávání narazíme na již navštívený vrchol. Podobné je to s hranou mezi různými stromy.



Důkaz. Podle lemmatu 1 stačí ukázat, že v G existuje volná střídavá cesta, právě když volná střídavá cesta existuje v $G \cdot C$. Označíme si $G' = G \cdot C$ a vrchol, který vznikne kontrakcí C, jako k.

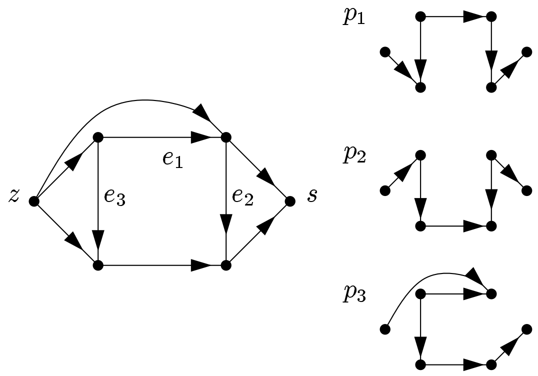
Mějme volnou střídavou cestu P v G, ukážeme, že v G' existuje volná střídavá cesta P'. Jeden konec P jistě není koncovým bodem stonku. Od něj tedy začneme zkoumat cestu P směrem k druhému konci a díváme se, jestli se při tom nestane něco zajímavého:

(1) Nepotkáme-li po cestě květ C ani jeho stonek, kontrakce se cesty P zjevně nijak nedotkne.
(2) Narazíme na květ C. Změníme tedy cestu P tak, že květ C po správné straně obejdeme (tj. podle toho, jestli následující hrana má být párovací či nepárovací), opustíme ho stonkem a cestu zakončíme v posledním vrcholu stonku („kořeni“). Toto je jiná volná střídavá cesta, která po kontrakci květu C bude volnou střídavou cestou i v G'.
(3) Narazíme na stonek v nějakém vrcholu, který si označíme v. Hodilo by se nám, kdybychom se mohli květu C úplně vyhnout. Ve vrcholu v vede stonek na jednu stranu směrem ke květu a na druhou stranu směrem od květu. Do stonku jsme přišli po nepárovací hraně, je-li tedy ve směru od květu párovací hrana, vydáme se tímto směrem až na konec stonku a dostaneme tak cestu P, které se kontrakce nedotkne.
(4) Narazíme na stonek, ale pokračovat se dá pouze směrem ke květu. Tehdy květu C přisoudíme jiný stonek (květ „přesadíme“). Novým stonkem bude celý úsek cesty P, který jsme doposud stihli prozkoumat, a dále část původního stonku směrem ke květu.

Nyní celý argument zopakujeme, avšak po cestě P se tentokrát vydáme z druhého konce. Pokud nastane znovu stejný případ, opět vyměníme stonek. Tento proces však určitě někdy skončí, neboť při každé výměně stonku se vždy alespoň o krok přiblížíme ke květu. Nakonec nastane jeden z předchozích případů.

Nyní mějme volnou střídavou cestu P' v G', ukážeme existenci cesty P v G. Možnosti:

(1) Když cesta P' neobsahuje k, hledaná $P = P'$.



Kapacity nastavíme takto: $c(e_1) = a_0 = 1$, $c(e_2) = a_1 = r$, $c(e_3) = 1$ a ostatním hranám přiřadíme nějakou dostatečně velkou kapacitu M. Maximální tok v naší síti je zřejmě $2M$.

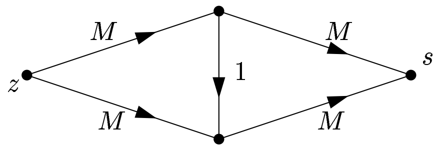
Začneme s nulovým tokem a budeme opakovaně používat posloupnost zlepšujících cest p_1, p_2, p_1, p_3. Jak se v průběhu každé iterace mění reziduální kapacity (tj. to, co chybí k plnému nasycení hrany) hran e_1, e_2, e_3:

$$(a_n, a_{n+1}, 0) \to (a_{n+2}, 0, a_{n+1}) \to (a_{n+2}, a_{n+1}, 0) \to$$

$$\to (0, a_{n+3}, a_{n+2}) \to (a_{n+2}, a_{n+3}, 0)$$

V každé iteraci se tak tok zvětší o $2a_n + 2a_{n+1}$, vylepšování toku se nezastaví a jeho velikost navíc bude konvergovat k hodnotě $1 + 2\sum_{n=2}^{\infty} a_n = 3$.

Protipříklad. Ani v celočíselných sítích Fordův–Fulkersonův algoritmus nepoběží příliš rychle, pokud v následující síti (kde za M zvolíme nějaké hodně velké číslo) bude střídavě vylepšovat obě cesty délky tři, tak potom poběží v čase úměrném M.



Existují samozřejmě mnohem rafinovanější algoritmy, konečné a rychlejší, ty však nepatří do tohoto textu. Poznamenejme jen, že pokud v každém kroku Fordova–Fulkersonova algoritmu použijeme *nejkratší* možnou nenasycenou cestu ze zdroje do stoku, skončí algoritmus pro libovolné reálné kapacity, a dokonce v počtu kroků, omezeném polynomiální funkcí počtu hran sítě.

Pozorování. *Když jsou kapacity celočíselné, resp. racionální, je i maximální tok vypočtený Fordovým–Fulkersonovým algoritmem celočíselný, resp. racionální.*

Důkaz. Algoritmus v průběhu své činnosti čísla pouze sčítá, takže ve výsledném toku nemohou vzniknout jiná čísla než celá. Pro racionální kapacity lze udělat podobnou úvahu, nebo převést na předchozí případ vynásobením vhodným jmenovatelem. □

Tato vlastnost úlohy hledání maximálního toku se výborně hodí pro nejrůznější aplikace existenčního charakteru, jako například hledání párování v bipartitních grafech, a podobně.

2.4 Toky v sítích a lineární programování.

S toky v sítích se dá zacházet také z pohledu lineární algebry. *Cirkulace* v orientovaném grafu $G = (V, E)$ je libovolná funkce $f \colon E \to \mathbb{R}$ splňující v každém vrcholu Kirchhoffův zákon:

$$\sum_{(x,v)\in E} f(x,v) - \sum_{(v,x)\in E} f(v,x).$$

Tuto podmínku můžeme stručně zapsat jako $D_G f = 0$. Zde D_G je *matice incidence* orientovaného grafu G: řádky jsou indexovány vrcholy, sloupce hranami, a prvek matice na pozici (v, e) je $+1$ pokud v je počáteční vrchol e, -1 pokud v je koncový vrchol e, a 0 jinak. Přitom f zde chápeme jako sloupcový vektor indexovaný hranami.

Tok v síti obecně není cirkulace, protože Kirchhoffův zákon se nepožaduje pro zdroj ani pro stok. Jednoduchým trikem ale můžeme z každého toku udělat cirkulaci v upraveném grafu. Mějme síť (G, z, s, c), kde $G = (V, E)$. Pro jednoduchost předpokládejme, že $(s, z) \notin E$. Definujeme nový orientovaný graf $G' = (V, E')$, kde $E' = E \cup \{(s, z)\}$ vznikne z E přidáním hrany (s, z). Položíme také kapacitu $c(s, z)$ rovnu „dostatečně velkému číslu“, čili například součtu kapacit všech ostatních hran. Je-li dán libovolný tok f v původní síti, definujeme cirkulaci f' na G': $f'(e) = f(e)$ pro $e \in E$, a $f'(s, z) = w(f)$. Je snadné si rozmyslet, že hledání maximálního toku v síti je ekvivalentní hledání takové cirkulace f' v grafu G', jejíž hodnota na každé hraně je mezi 0 a $c(e)$, a jež maximalizuje $f'(s, z)$.

Tato nová formulace nám umožňuje zapsat úlohu maximálního toku jako úlohu *lineárního programování*, tj. úlohu nalézt řešení dané soustavy lineárních rovnic a nerovností, které maximalizuje danou lineární funkci. Konkrétně, úloha maximálního toku je ekvivalentní úloze

$$\max\{f'(s, z);\ f' \in \mathbb{R}^{|E'|},\ D_{G'} f' = 0,\ f' \le c,\ f' \ge 0\}.$$

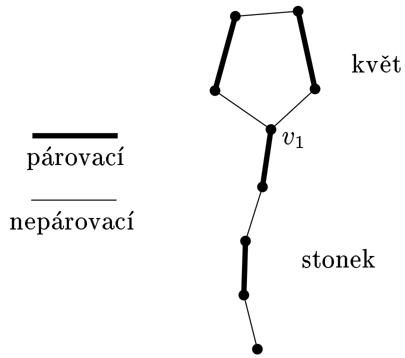
Zde opět interpretujeme f' a c jako $|E'|$-složkové reálné vektory, a rovnost či nerovnost mezi vektory znamená rovnost či nerovnost po složkách.

Řešením obecných úloh lineárního programování se zabývá dobře vyvinutá teorie. Z vět této teorie se dá odvodit i hlavní věta o tocích jako důsledek. I spousta dalších výsledků v kombinatorice, zejména v kombinatorické optimalizaci, souvisí s lineárním programováním.

(2) hrany $\{v_{2i+1}, v_{2i+2}\} \notin M$,
(3) hrany $\{v_{2i}, v_{2i+1}\} \in M$,
(4) má alespoň jeden *stonek*.

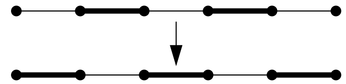
Stonkem nazveme cestu, která začíná ve vrcholu v_1, první hrana je párovací, pravidelně se na ní (jako ve volné střídavé cestě) střídají párovací a nepárovací hrany a poslední vrchol je volný. Stonek může mít i nulovou délku, to když je vrchol v_1 volný v celém grafu G. Stonek také nemusí vůbec existovat, například když je v_1 obsažen v párovací hraně, která není součástí květu C a z níž dále nepokračuje žádná další nepárovací hrana.

Ilustrace:



Lemma 1. *Párování M je maximální, právě když v grafu neexistuje volná střídavá cesta.*

Důkaz. Když v grafu volná střídavá cesta existuje, párování M jistě není největší:

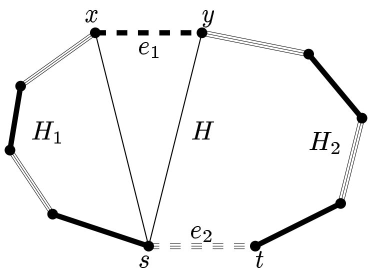


Nechť naopak M není největší, čili existuje párování M' takové, že $|M'| > |M|$. Zadefinujeme si množinu hran $\widetilde{M}$ jako

$$\widetilde{M} = M \div M'.$$

V $\widetilde{M}$ má každý vrchol stupeň 0, 1 nebo 2, komponenty souvislosti této konfigurace jsou kružnice sudé délky a cesty, kde se pravidelně střídají hrany M a M'. Jelikož $|M'| > |M|$, musí existovat komponenta tvaru liché cesty, v níž je hran M' o jednu více, než hran M. Ta tvoří volnou střídavou cestu. □

Lemma 2. *Nechť C je květ v grafu G. Potom párování M v G je maximální, právě když $M \setminus E(C)$ je maximální párování v grafu $G \cdot C$, tj. s květem C zkontrahovaným do jediného vrcholu.*



Úsek na kružnici H mezi x a s neobsahující y označíme H_1, úsek mezi y a t neobsahující x označíme H_2. Mimo kružnici H lze použít obě párování, užijeme třeba znovu M_1. V úseku H_1 zanecháme pouze hrany M_2, v úseku H_2 pouze hrany M_1. Zbývají spárovat už jen vrcholy y a s, mezi nimi ovšem vede hrana, můžeme ji proto zahrnout do párování. Perfektní párování $\widetilde{M}$ tedy je

$$\widetilde{M} := (M_1 \setminus H) \cup (H_1 \cap M_2) \cup (H_2 \cap M_1) \cup \{y, s\}.$$

$\square$

5.2 Maximální párování a Edmondsův algoritmus.

V minulém oddílu jsme si charakterizovali grafy mající perfektní párování. Protože ne v každém grafu perfektní párování existuje, bude nás zajímat alespoň párování maximální, tedy pokrývající největší možný počet vrcholů.[4] Ukážeme si algoritmus, který pro libovolný graf maximální párování nalezne v polynomiálním čase.

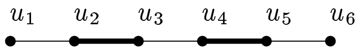
Připravíme si nejprve několik technických prostředků.

Definice. Nechť M je párování grafu G. Vrchol $u \in V(G)$ nazveme *volný*, pokud $\deg_M u = 0$. Vrchol, který není volný, nazveme *spárovaný*.

Cestu $u_1, u_2, \ldots, u_{2k}$ v grafu G nazveme *volnou střídavou cestou*, pokud

(1) vrcholy u_1 a u_{2k} jsou volné,
(2) hrany $\{u_{2i+1}, u_{2i+2}\} \notin M$,
(3) hrany $\{u_{2i}, u_{2i+1}\} \in M$.

Ilustrace:



Lichou kružnici $C_{2k+1} = \{v_1, v_2, \ldots, v_{2k+1}\}$ grafu G nazveme *květem*, jestliže

(1) vrchol v_1 je na množině hran květu C volný,

[4] V anglické terminologii se důsledně rozlišuje (a to nejen u párování) „maximum matching“, tedy párování pokrývající největší možný počet vrcholů, a „maximal matching“ pro největší párování ve smyslu inkluze, tedy párování, ke kterému již nelze přidat žádnou další hranu.

2.5 Existence maximálního toku.

Na závěr ještě slíbený důkaz existence maximálního toku i v iracionálně ohodnocených sítích.

Důkaz existence maximálního toku. Většinu tíhy důkazu (který bude spíše technického rázu) převedeme na větu ze základního kurzu matematické analýzy, která říká, že spojitá funkce na kompaktní množině nabývá svého maxima. Stačí tedy ověřit, že množina všech toků

$$\mathcal{F} = \{f;\ f \text{ je tok}\} \subseteq \mathbb{R}^{|E|}$$

je kompaktní a že funkce $w\colon \mathcal{F} \to \mathbb{R}$ je spojitá. Množina $\mathcal{F}$ je podmnožinou $|E|$-rozměrného euklidovského prostoru, protože každý tok se dá zapsat jako $|E|$-složkový vektor.

Spojitost w. Zadefinujeme si *projekce* $\pi_e\colon \mathcal{F} \to \mathbb{R}$ jako $\pi_e(f) = f(e)$, čili sadu funkcí pro každou hranu vracejících hodnotu $f(e)$. Proč je projekce spojitá? Vyjdeme z definice spojitosti na obecném metrickém prostoru. Když zvolíme nějakou velikost toku $w(f)$ a (malou) odchylku od něj, musí okolo příslušného toku f (jakožto bodu euklidovského prostoru) existovat nějaké okolí tak, že po změnách toku provedených v rámci něj se vejdeme do povolené odchylky velikosti toku. To skutečně lze udělat, viz vylepšování toků o ε v předchozích důkazech. Na funkci

$$w(f) = \sum_{(z,x)\in E} f(z,x) - \sum_{(x,z)\in E} f(x,z) = \sum_{(z,x)\in E} \pi_{(z,x)}(f) - \sum_{(x,z)\in E} \pi_{(x,z)}(f)$$

se lze pro danou síť dívat jako na lineární kombinaci konečného počtu projekcí, a je tedy též spojitá.

Kompaktnost $\mathcal{F}$. Připomeňme si, že v euklidovském prostoru je množina kompaktní, právě když je uzavřená a omezená. Množina $\mathcal{F}$ se dá uzavřít do krabičky $\prod_{e\in E}[0, c(e)]$, je tedy omezená. Pro vrcholy u si zapíšeme množiny $\mathcal{F}_u$ toků $f \in \mathbb{R}^{|E|}$ a analogicky vrcholové projekce π_u jako

$$\mathcal{F}_u = \left\{ f;\ \overbrace{\sum_{(u,x)\in E} f(u,x) - \sum_{(x,u)\in E} f(x,u)}^{\pi_u(f)} = 0 \right\} = \{f;\ \pi_u(f) = 0\} = \pi_u^{-1}(\{0\}).$$

Projekce π_u je opět spojitá, protože je lineární kombinací konečného počtu spojitých hranových projekcí. Nyní použijeme trik z matematické analýzy. Platí věta, že vzor uzavřené množiny při spojitém zobrazení je opět uzavřená množina. Projekce π_u spojitá je a jednoprvková množina $\{0\}$ skutečně je uzavřená. A množina $\mathcal{F}$ je uzavřená, neboť

$$\mathcal{F} = \bigcap_{u\in V\setminus\{z,s\}} \mathcal{F}_u \cap \prod_{e\in E} [0, c(e)]$$

a průniky i konečná sjednocení uzavřených množin jsou opět uzavřené množiny. $\square$

3. Míra souvislosti grafů

Graf můžeme rozdělit (znesouvislit) jak odebráním hran, tak také vrcholů.

Definice. *Hranový řez* v grafu $G = (V, E)$ je množina hran $F \subseteq E$ taková, že graf $G' = (V, E \setminus F)$ je nesouvislý.

Vrcholový řez v grafu $G = (V, E)$ je množina vrcholů $A \subseteq V$ taková, že graf $G'' = (V \setminus A, E \cap \binom{V \setminus A}{2})$ (čili *indukovaný podgraf* $G[V \setminus A]$) je nesouvislý.

Pomocí řezů si zavedeme dva stupně souvislosti, hranovou a vrcholovou.

Definice. *Hranová souvislost* je

$$k_e(G) = \min\{|F|;\ F \subseteq E \text{ je hranový řez}\}.$$

Vrcholová souvislost je

$$k_v(G) = \begin{cases} \min\{|A|;\ A \subseteq V \text{ je vrcholový řez}\} & \text{pro } G \neq K_n \\ n-1 & \text{pro } G \simeq K_n. \end{cases}$$

U vrcholové souvislosti je třeba zvlášť ošetřit úplné grafy, protože v nich neexistuje vrcholový řez (odebráním libovolné množiny vrcholů se graf nikdy nerozpadne). V souladu s dřívější definicí graf tedy bude hranově (resp. vrcholově) k-souvislý, když je $k_e(G) \geq k$ (resp. $k_v(G) \geq k$). Pokud se odebráním *libovolného* vrcholu (resp. hrany) sníží stupeň souvislosti, říká se také, že graf je *kriticky* vrcholově (resp. hranově) k-souvislý.

Nabízí se přirozená otázka, jak spolu $k_v(G)$ a $k_e(G)$ souvisí. Ukážeme si, že pro libovolný graf G je $k_v(G) \leq k_e(G)$. První nápad, vezmi za každou hranu z hranového řezu nějaký její vrchol, bohužel nefunguje, není totiž úplně jasné, který konec vybrat, aby vznikl korektní řez. (Viz třeba graf C_4, kde bychom zvolili dva sousední vrcholy.)

K důkazu si nejprve připravíme dvě technická tvrzení říkající v podstatě to, že odebráním hrany se hranová i vrcholová souvislost sníží nejvýše o jednu.

Lemma. *Pro každý graf G a libovolnou jeho hranu e platí*

$$k_e(G) - 1 \leq k_e(G - e) \leq k_e(G).$$

Důkaz. Vezmeme minimální řez F. Množina hran $F' = F \setminus \{e\}$ je potom řezem v grafu $G - e$, ne nutně minimálním, ale každopádně platí

$$k_e(G - e) \leq |F'| \leq |F| = k_e(G).$$

Dále vezmeme minimální hranový řez B v grafu $G - e$. Potom $B' = B \cup \{e\}$ bude řezem v grafu G, čili

$$k_e(G) \leq |B'| = |B| + 1 = k_e(G - e) + 1.$$

□

perfektní párování M_1 a v $G + e_2$ existuje perfektní párování M_2. Na obrázcích si hrany pocházející z M_1 označíme barvou hrany e_1, hrany pocházející z M_2 barvou hrany e_2. Zadefinujeme množinu hran M jako

$$M := M_1 \div M_2,$$

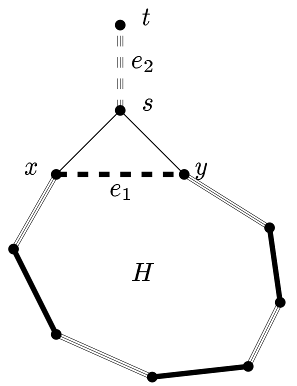
kde $\div$ značí symetrickou diferenci množin.[3]

S takovou množinou hran platí ovšem pro každý vrchol u

$$\deg_M u = \begin{cases} 0 & \text{když z vrcholu } u \text{ vedly stejné párovací hrany} \\ 2 & \text{když z vrcholu } u \text{ vedly různé párovací hrany.} \end{cases}$$

Každá komponenta souvislosti v takovém grafu je potom buďto izolovaný vrchol, nebo kružnice sudé délky, kde se pravidelně střídají hrany obou barev. Označíme H hrany kružnice, na které leží hrana e_1, a rozebereme dva případy. V obou se budeme snažit sestrojit perfektní párování, jež neobsahuje ani jednu z hran e_1 a e_2, tedy bude vhodné i pro původní graf G. Můžeme nadále předpokládat, že $e_1 \in M_1$ a $e_2 \in M_2$; v opačném případě je M_1, resp. M_2 perfektní párování v G a ukončíme důkaz.

(1) Hrana $e_2 \notin H$.

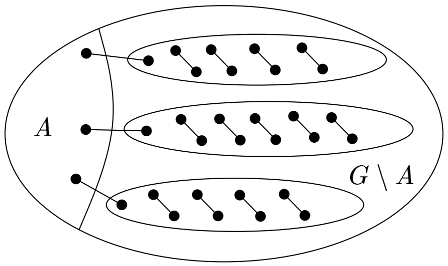


Do perfektního párování $\widetilde{M}$ mimo kružnici H můžeme zahrnout jak M_1, tak M_2, vybereme tedy třeba M_1, na kružnici H použijeme hrany M_2. Formálně

$$\widetilde{M} := (M_1 \setminus H) \cup (H \cap M_2).$$

(2) Hrana $e_2 \in H$. Bez újmy na obecnosti můžeme předpokládat, že vrcholy s, x, y, t leží na kružnici v pořadí, jaké je zakreslené na obrázku.

[3] Čili $M = (M_1 \cup M_2) \setminus (M_1 \cap M_2)$, „xor“ pro informatiky.



Druhá implikace je složitější. Nechť platí Tutteova podmínka. Důkaz provedeme matematickou indukcí podle počtu hran, netradičně však v obráceném pořadí, od $\binom{n}{2}$ k 0.

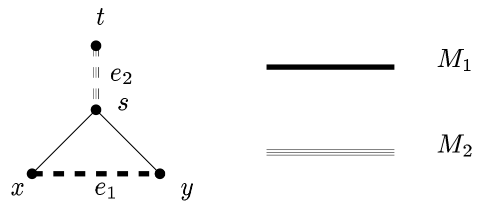
Pro případ úplného grafu K_n je situace jednoduchá. Pro $A = \emptyset$ dostaneme, že n musí být sudé, a tudíž lze jistě v K_n najít perfektní párování.

Nadále předpokládejme, že v grafu G chybí alespoň jedna hrana. Zvolíme množinu W vrcholů takových, že z nich vede hrana do všech ostatních vrcholů grafu G:

$$W := \{u;\ \deg_G u = n-1\}$$

Jestliže je graf $G \setminus W$ disjunktním sjednocením úplných grafů, dokážeme jednoduše najít perfektní párování. Komponenty souvislosti grafu $G \setminus W$ (což jsou úplné grafy) sudé velikosti jistě vypárujeme. V každé liché komponentě nelze spárovat pouze jeden vrchol, množinu všech takových vrcholů si označíme V_0. Z Tutteovy podmínky ovšem víme, že $|V_0| = C_\ell(G \setminus W) \leq |W|$, vrcholy V_0 tedy spárujeme s množinou W. Toto je mimochodem jediné místo v důkazu, kde se použije Tutteova podmínka.

Nechť nyní $G \setminus W$ disjunktním sjednocením úplných grafů není. Podle předchozího lemmatu v něm existuje $K_{1,2}$ jako indukovaný podgraf. Vrcholy tohoto $K_{1,2}$ si označíme x, s, y takto:



Ve vrcholu s nutně musí chybět hrana do nějakého vrcholu $t \in V(G)$; jinak by totiž s byl zařazen do množiny W. Označíme $e_1 = \{x, y\}$ jednou barvou a $e_2 = \{s, t\}$ druhou barvou. Ani e_1 ani e_2 nejsou hranami grafu G.

Uvědomíme si, že přidáním libovolné hrany do grafu se Tutteova podmínka neporuší. V $G + e \setminus A$ pro libovolnou hranu e a řez A zjevně počet lichých komponent nevzroste. Víme tedy, že oba grafy $G + e_1$ i $G + e_2$ taktéž splňují Tutteovu podmínku. Podle indukčního předpokladu proto v $G + e_1$ existuje

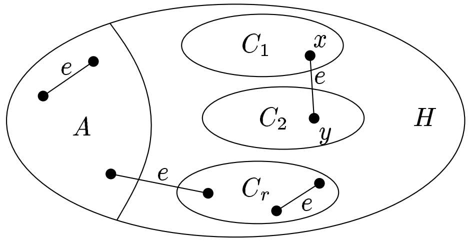
Lemma. *Pro každý graf G a libovolnou jeho hranu e platí*

$$k_v(G) - 1 \leq k_v(G - e) \leq k_v(G).$$

Důkaz. Úlohu si mírně přeformulujeme, budeme dokazovat pro graf $H = G - e$ nerovnost

$$k_v(H + e) \leq k_v(H) + 1.$$

V grafu H existuje vrcholový řez $A \subseteq V(H)$ takový, že $k_v(H) = |A|$. Při odebrání řezu A se nám graf H rozpadne na alespoň dvě komponenty $C_1, \ldots, C_r$. Rozebereme možnosti, odkud pochází hrana e:



(1) Alespoň jeden vrchol hrany e leží v řezu A. Přidáním hrany tedy nespojíme dohromady žádné komponenty, množina A bude stále řezem a tak

$$k_v(H + e) \leq |A| = k_v(H).$$

(2) Hrana e je celá uvnitř nějaké komponenty. Stejný argument.

(3) Hrana $e = \{x, y\}$ spojuje dvě komponenty, dejme tomu $x \in C_1$ a $y \in C_2$. Když jsou komponenty alespoň tři, je vše v pořádku, protože po přidání e se spojí jen dvě komponenty a množina A zůstane řezem. Ať jsou nadále komponenty pouze dvě a bez újmy na obecnosti $|C_1| \geq |C_2|$, $|C_1| \geq 2$.

Množina $A' = A \cup \{x\}$ bude vrcholovým řezem grafu $H + e$, protože C_2 a $C_1 \setminus \{x\}$ jsou komponenty souvislosti grafu $(H + e) \setminus A'$. Potom platí

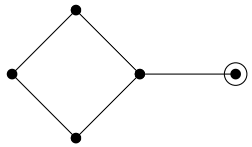
$$k_v(H + e) \leq |A'| \leq |A| + 1 = k_v(H) + 1.$$

Zbývá dořešit speciální případ, kdy jsou komponenty C_1 a C_2 jednoprvkové. Všech vrcholů je pak celkem $|V| = |A| + 2$ a spočítáme, že

$$k_v(H + e) \leq |V| - 1 = |V| - 2 + 1 = k_v(H) + 1.$$

□

Bylo by hezké, kdyby podobné lemma platilo i pro odebírání vrcholu, jenže ačkoli se vrcholová souvislost sníží nejvýše o jednu, může také vzrůst, třeba zde:



Věta. *Pro každý graf G platí*

$$k_v(G) \leq k_e(G).$$

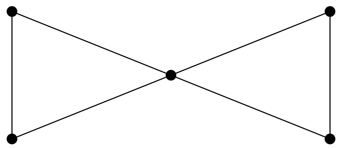
Důkaz. Budeme postupovat indukcí podle počtu hran. Je-li hran méně než $|V| - 1$ (tedy vrcholy nejsou pokryty ani stromem a graf je nesouvislý), bude zjevně $k_e(G) = k_v(G) = 0$. Nechť nadále $k_e(G) > 0$.

Vezmeme nejmenší hranový řez F a jeho jednu hranu e a na graf $G' = G - e$ použijeme indukční předpoklad. Dostaneme tak

$$k_v(G) - 1 \leq k_v(G - e) \leq k_e(G - e) = k_e(G) - 1.$$

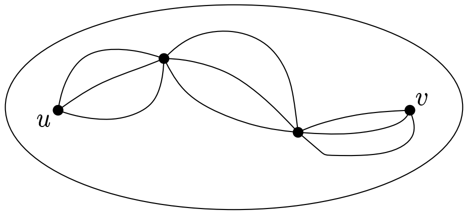
Odtud a z pomocných lemmat máme $k_v(G) \leq k_e(G)$. □

Máme-li potíže se zapamatováním, kterým směrem vlastně nerovnost platí, stačí si nakreslit motýlka:



Pro tento graf je $k_v = 1$ ale $k_e = 2$, druhý směr nerovnosti tak neplatí.

Věta (Ford, Fulkerson). *Pro každý graf a každé přirozené t platí $k_e(G) \geq t$, právě když mezi každými dvěma různými vrcholy u a v existuje alespoň t hranově disjunktních cest.*



5. Párování v obecných grafech

Hezkým příkladem na párování je problém hlídkujících strážníků. Strážníci hlídkují vždy po dvojicích a každý strážník je ochoten spolupracovat pouze s některými kolegy, s ostatními se nesnese. Každý strážník je reprezentován vrcholem grafu a hrana vede do každého snesitelného kolegy.

Zločin chceme samozřejmě potlačovat v co největší míře, a tedy nás zajímá, jaký největší počet strážníků může sloužit (a kdo s kým), a za jakých podmínek mohou být na obchůzce úplně všichni.

5.1 Perfektní párování a Tutteova věta.

Definice. Párování M v grafu G nazveme *perfektní*, pokud $|M| = \frac{|V(G)|}{2}$.

Věta (Tutte). *Graf G má perfektní párování, právě když pro každou $A \subseteq V(G)$ je počet komponent souvislosti grafu $G \setminus A$, které mají lichou velikost, nejvýše $|A|$; tato podmínka se nazývá Tutteova.*[2]

Než přistoupíme k důkazu, bude se nám hodit následující

Lemma. *Graf G je disjunktním sjednocením úplných grafů, právě když G neobsahuje $K_{1,2}$ (neboli „třešničku“) jako indukovaný podgraf.*

Důkaz. Pokud je G disjunktním sjednocením úplných grafů, triviálně nemůže obsahovat $K_{1,2}$. Naopak obměnou, nechť G není disjunktní sjednocením úplných grafů, chceme ukázat, že obsahuje $K_{1,2}$. Existuje tedy komponenta souvislosti C a v ní dva vrcholy x a y takové, že mezi nimi nevede hrana. Vrcholy x a y jsou v C propojené cestou P, na cestě P pak nutně musí existovat tři vrcholy tvořící $K_{1,2}$. Kdyby neexistovaly, každá trojice vrcholů by tvořila trojúhelník, jenže hrana chybí již mezi x a y. □

Důkaz Tutteovy věty. Počet lichých komponent souvislosti grafu G si označíme symbolem $C_\ell(G)$.

Nejprve nechť v grafu G existuje perfektní párování. Pro libovolnou množinu $A \subseteq V(G)$ tak v grafu $G \setminus A$ vzniknou nějaké komponenty souvislosti, zajímat nás ale budou jen komponenty liché velikosti. Ať bylo perfektní párování jaké chtělo, v každé liché komponentě přebývá alespoň jeden vrchol, který nutně musel být spárován s nějakým vrcholem z množiny A. Dvě liché komponenty nemohly sdílet stejný vrchol z množiny A, a tedy $C_\ell(G \setminus A) \leq |A|$.

---
[2] Vyslovujeme [*tat*] a [*tatova*].

Sestrojíme bipartitní graf $G = (S \cup H, E)$, kde $S = \{s_1, \ldots, s_n\}$ je množina sloupců latinského obdélníku a $H = \{1, \ldots, n\}$ je množina hodnot, ze kterých konstruujeme $(k+1)$-ní řádek. Hrana vede mezi s_i a x, pokud se hodnota x nevyskytuje v i-tém sloupci, tedy

$$E = \{\{s_i, x\};\ x \notin s_i\}.$$

Nyní chceme do nového řádku na j-tou pozici vybrat takový prvek x, že x se jednak nevyskytuje v sloupci s_j a jednak všechny takto vybrané prvky x jsou navzájem různé. To je ale ekvivalentní s existencí párování v grafu G, které plně uspokojuje část S.

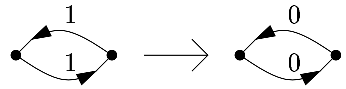
Jaké stupně vrcholů mají jednotlivé části grafu G? V latinském obdélníku o k řádcích chybí v každém sloupci jistě $n - k$ hodnot, proto pro každý $s_i \in S$ je $\deg_G s_i = n - k$.

Naopak, v kolika sloupcích může chybět prvek x? Prvek x se v každém řádku vyskytuje právě jednou, a protože řádků je k, je jím tedy obsazeno právě k různých sloupců; jinak by byly dva prvky x v jednom sloupci a nebyl by to latinský obdélník. Stejně jako v první části i pro každý $x \in H$ je $\deg_G x = n - k$ a podle předchozího tvrzení existuje párování uspokojující část S. □

Důkaz. Implikaci zprava doleva ukážeme sporem, nechť existuje hranový řez F menší než t. Po rozdělení grafu vezmeme dva vrcholy u a v ležící v různých komponentách souvislosti. Jenže mezi nimi vedlo t hranově disjunktních cest, řez F tedy na jejich rozpojení nemohl stačit.

V druhém směru použijeme skutečnosti dokázané o tocích v síti. Mějme tedy graf G takový, že $k_e(G) \geq t$, a pro dva vrcholy u a v hledáme hranově disjunktní cesty. Sestrojíme z grafu G síť $G' = (V(G), \{(x,y), (y,x);\ \{x,y\} \in E(G)\}, u, v, \mathbf{1})$ tak, že oboustranně zorientujeme původní hrany, za zdroj zvolíme vrchol u, za stok vrchol v a kapacity všech hran nastavíme na 1. Pak spustíme algoritmus hledání maximálního toku.

Nalezený tok bude celočíselný, tedy po hranách poteče 0 nebo 1. Provedeme ovšem ještě drobnou úpravu – když po obou hranách (a,b) i (b,a) teče jednička, nastavíme obě na nulu, velikost toku se tím nezmění:



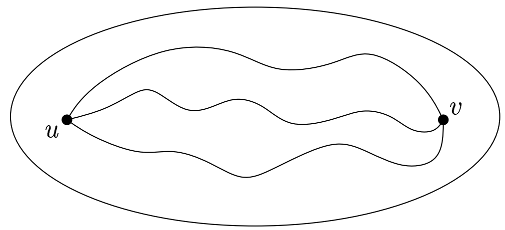
K našemu maximálnímu toku f potom existuje minimální hranový řez R.

Množina $F = \{\{x,y\};\ (x,y) \in R \vee (y,x) \in R\}$ je hranový řez v grafu G. Protože G je hranově t-souvislý, bude podle definice souvislosti i velikost řezu $|F| \geq t$. Jelikož za každou hranu z řezu R jsme do množiny F zařadili nejvýše jednu hranu, bude též $|R| \geq |F| \geq t$. Velikost toku f tak musí být alespoň t.

Zbývá zkonstruovat hranově disjunktní cesty z u do v. Existenci cest v G' ukážeme induktivně. Má-li tok velikost 1, jistě nalezneme cestu z u do v, kde na každé hraně teče jednička. Ve větším toku najdeme libovolnou jedničkovou cestu z u do v a hrany na ní vynulujeme. Velikost toku se tím sníží o 1 a existence zbývajících cest plyne z indukce. Na závěr už jen ke každé nalezené orientované cestě v G' vezmeme příslušnou neorientovanou cestu v G. □

Podobné tvrzení jako je Fordova–Fulkersonova věta platí i pro vrcholovou souvislost.

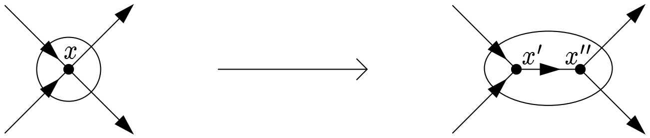
Věta (Menger). *Pro každý graf a každé přirozené t platí $k_v(G) \geq t$, právě když mezi každými dvěma různými vrcholy u a v existuje alespoň t cest, které jsou až na vrcholy u a v disjunktní.*



Důkaz. Nejprve implikaci zprava doleva, nechť mezi každými dvěma vrcholy existuje t disjunktních cest. Předpokládejme pro spor, že $k_v(G) < t$. V úplném grafu je $k_v(G) = n-1$ a jistě mezi dvěma vrcholy u a v nevede více než $n-1$ cest. Stejně tak v neúplném grafu, na rozpojení t cest je třeba řez velikosti alespoň t.

Druhým směrem, nechť je G vrcholově t-souvislý, zvolíme dva vrcholy u a v. Nejprve nechť $\{u,v\} \notin E(G)$.

Hrany grafu symetricky zorientujeme, tj. nahradíme každou původní hranu hranami tam a zpátky. Výjimku budou tvořit hrany obsahující u, ty povedou pouze z vrcholu u, a hrany obsahující v, ty povedou pouze do vrcholu v. Nyní všechny vrcholy vyjma u a v takto podrozdělíme:



Formálně zapsáno, utvoříme nový graf G' takový, že

$$V(G') = \{x', x'';\ x \in V(G),\ u \neq x \neq v\}$$
$$E(G') = \{(x'', y'), (y'', x');\ \{x,y\} \in E(G)\} \cup \underbrace{\{(x', x'');\ x \in V(G),\ u \neq x \neq v\}}_{F},$$

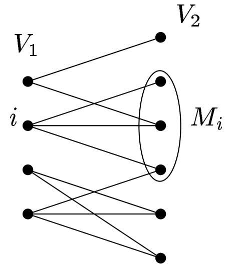
kde si množinu hran vedoucích uvnitř podrozdělených vrcholů označíme jako F. Hrany vzniklé rozdělením vrcholu tak vlastně „simulují“ vrchol grafu G, je tu korespondence mezi hranami z F a vrcholy grafu G.

Kapacity všech hran v grafu G' nastavíme na 1, za zdroj a stok vezmeme vrcholy u a v a nalezneme mezi nimi maximální tok. K němu bude existovat minimální (hranový) řez R. K řezu R ovšem vždy najdeme řez $R' \subseteq F$, $|R'| \leq |R|$, jednoduše tak, že každá hrana $e \in R \setminus F$ navazuje alespoň z jedné strany na nějakou hranu $e' \in F$; stačí uvažovat e' namísto e. Když bude R' dostatečně velký (tj. $|R| \geq t$), stejným postupem jako při důkazu Fordovy–Fulkersonovy věty v G' nalezneme t hranově disjunktních cest z u do v.

Pro spor předpokládejme, že existuje nějaký řez $R' \subseteq F$ takový, že $|R'| < t$. Uvážíme množinu $A = \{x;\ (x', x'') \in R\}$, která bude vrcholovým řezem grafu G, neboť v $G \setminus A$ nevede žádná cesta mezi u a v. Jenže G je vrcholově t-souvislý a není úplný, dostaneme tak $t \leq |A| = |R'| < t$.

A ještě zbývá zkonstruovat vrcholově disjunktní cesty v grafu G: pro cestu $(u, x_1', x_1'', x_2', x_2'', \ldots, v)$ v grafu G' vezmeme cestu $(u, x_1, x_2, \ldots, v)$. Kdyby se dvě takové cesty protínaly v nějakém x_i, musely by se protínat i původní cesty v hraně (x_i', x_i'').

Když $\{u,v\} \in E(G)$, zopakujeme celý postup pro graf $G - \{u,v\}$. Odebráním hrany se vrcholová souvislost sníží nejvýše o jednu, předchozí postup tedy nalezne $t-1$ cest, ke kterým dodáme ještě hranu $\{u,v\}$. □



Když o systému dokážeme, že v něm platí Hallova podmínka, bude v něm existovat systém různých reprezentantů. Tedy pro každý $i \in V_1$ bude existovat $x \in V_2$ tak, že mezi nimi povede hrana a všechna x budou navzájem různá. Ale to je přesně definice párování, které plně uspokojuje část V_1.

Zvolme $J \subseteq V_1$ a označíme si $k = \min_{j \in J}\{\deg_B j\}$. Z množiny J vede alespoň $k|J|$ hran do druhé části. Kdyby $|\bigcup_{j \in J} M_j| < |J|$, z přihrádkového principu by v části V_2 musel existovat vrchol x, do kterého by vedlo alespoň $k+1$ hran. Jeho stupeň by tak byl vyšší než stupeň vrcholu části V_1, což je spor s předpokladem tvrzení. □

Definice. *Latinský čtverec* A je čtvercová matice řádu n taková, že

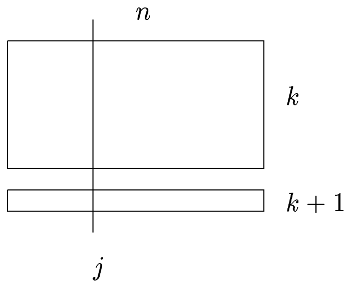
(1) $a_{ij} \in \{1, 2, \ldots, n\}$
(2) $a_{ij} \neq a_{i'j}$ pro každé j a $i \neq i'$
(3) $a_{ij} \neq a_{ij'}$ pro každé i a $j' \neq j$.

Latinský obdélník A je matice $k \times n$, splňující stejné podmínky jako latinský čtverec.

V každém řádku a sloupci latinského čtverce či obdélníku se tedy vyskytují jen navzájem různá čísla.

Tvrzení. *Každý latinský obdélník, kde počet řádků k je nižší než počet sloupců n, lze doplnit na latinský čtverec.*

Důkaz. K latinskému obdélníku budeme postupně konstruovat další a další řádky, dokud jich nebude n.



Jaká bude velikost řezu R?

$$c(R) = |A| + |B| = |I| - |J| + |B| \geq |I| - |J| + \Big|\bigcup_{j \in J} M_j\Big|.$$

Nyní použijeme Hallovu podmínku, která říká, že $|\bigcup_{j \in J} M_j| \geq |J|$, a dostaneme

$$c(R) \geq |I| - |J| + |J| = |I|,$$

tok g tak bude mít velikost alespoň $|I|$.

Zbývá říci, co je hledaným systémem různých reprezentantů. Jsou jím prostě prvky množiny X takové, že do nich tokem g teče jednička. Formálně, SRR bude funkce f, kde pro každé $i \in I$

$$f(i) = x \quad \text{takové, že} \quad g(i, x) = 1.$$

□

Hallova věta má celou řadu nejrůznějších aplikací, a proto stojí za to si ji zapamatovat. Předvedený důkaz samozřejmě není jediným možným, větu lze dokázat i přímo matematickou indukcí, bez použití toků. Nebo ji lze odvodit z Tutteovy věty, kterou vysvětlíme později.

Poznámka. Předpokládáme-li v teorii množin axiom výběru, lze dokázat variantu Hallovy věty, kde sice množiny M_i musí být konečné, ale množiny I a X mohou být nekonečné, dokonce nespočetné. Důkaz už je však poněkud složitější.

Když jsou X a I nekonečné (stačí spočetné) a povolíme nekonečné velikosti (taktéž stačí spočetné) množin M_i, analogie Hallovy věty neplatí. Stačí zvolit $X = \mathbb{N}$, $M_1 = X$ a ostatní M_i jsou všechny jednoprvkové podmnožiny množiny X. Zde systém různých reprezentantů nelze sestrojit.

4.1 Důsledky Hallovy věty.

Předvedeme si několik základních důsledků Hallovy věty, velké praktické použití má zejména následující

Tvrzení. *V bipartitním grafu $B = (V_1 \cup V_2, E)$ s neprázdnou množinou hran E takovém, že pro libovolné dva vrcholy $x \in V_1$ a $y \in V_2$ platí $\deg_B x \geq \deg_B y$, existuje párování, které má velikost $|V_1|$ („uspokojuje část V_1“).*

Důkaz. Zvolíme množinový systém $\mathcal{M} = (M_i, i \in I)$ nad V_2 takový, že $I = V_1$ a

$$M_i = \{x \in V_2;\ (i, x) \in E(B)\}.$$

Obrázkem:

Ještě si připomeneme

Lemma (ušaté). *Graf G je 2-souvislý, právě když je možné vytvořit ho z libovolné jeho kružnice přidáváním uší.*

4. Systémy různých reprezentantů

Definice. Nechť X a I jsou množiny. *Množinovým systémem nad X* nazveme množinu $\mathcal{M} = (M_i, i \in I)$, kde $M_i \subseteq X$. Obecně se uvažují nekonečné systémy, my budeme zkoumat pouze systémy konečné. Nechť jsou tedy navíc množiny X, $I \subset \mathbb{N}$ a M_i konečné.

Systém různých reprezentantů (SRR) je funkce $f\colon I \to X$ taková, že

(1) pro každé $i \in I$ je $f(i) \in M_i$
(2) f je prosté.

Jinými slovy, SRR je výběr jednoho prvku z každé množiny M_i takový, aby všechny vybrané prvky byly navzájem různé.

Definice. *Párování* v grafu G je množina hran $F \subseteq E(G)$ taková, že každý vrchol grafu G patří nejvýše do jedné hrany z F.

Incidenčním grafem systému $\mathcal{M}$ nazveme bipartitní graf

$$B_{\mathcal{M}} = (I \cup X, \{\{i, x\};\ x \in M_i\}).$$

Další ekvivalentní definicí párování je třeba to, že hrany v párování jsou disjunktní množiny nebo také že stupeň vrcholu v grafu obsahujícím pouze párovací hrany je nejvýše 1.

Pozorování. *Systém různých reprezentantů v $\mathcal{M}$ existuje, právě když v incidenčním grafu $B_{\mathcal{M}}$ existuje párování velikosti $|I|$.*

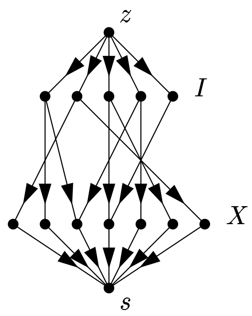
Následující věta je příkladem tvrzení, kde je na první pohled viditelná nutná podmínka zároveň i podmínkou postačující.

Hallova věta. *Systém různých reprezentantů v $\mathcal{M}$ existuje, právě když pro každou $J \subseteq I$ je $|\bigcup_{i\in J} M_i| \geq |J|$; tato podmínka se nazývá Hallova.*

Důkaz. Začneme jednodušší implikací, nechť v $\mathcal{M}$ existuje systém různých reprezentantů. Zvolíme libovolnou $J \subseteq I$. Pro každý index $j \in J$ existuje prvek p_j z množiny M_j tak, že všechny prvky p_j jsou navzájem různé. Proto

$$\Big|\bigcup_{j\in J} M_j\Big| \geq |\{p_j;\ j \in J\}| = |J|.$$

Naopak, mějme takový množinový systém $\mathcal{M}$, který splňuje Hallovu podmínku. Vezmeme bipartitní graf $B_{\mathcal{M}}$, který takto předěláme na síť:



Přesněji zapsáno, sestrojíme síť $S = (G, z, s, \mathbf{1})$, kde

$$V(G) = \{z, s\} \cup I \cup X$$
$$E(G) = \{(z, i);\ i \in I\} \cup \{(i, x);\ i \in I, x \in M_i\} \cup \{(x, s);\ x \in X\}.$$

Kapacity hran nastavíme na 1 a nalezneme Fordovým–Fulkersonovým algoritmem maximální tok g. V kapitole o tocích jsme si dokázali, že takový tok bude celočíselný. K toku g existuje minimální řez R'. Pro další průběh důkazu bude potřeba nějak se v řezu R' zbavit hran mezi I a X, aniž by se řez zvětšil. Vyrobíme tedy řez

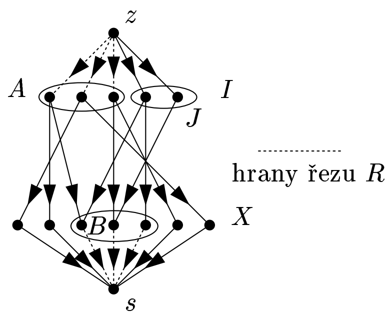
$$R = \{(x, s);\ (x, s) \in R'\} \cup \{(z, i);\ (z, i) \in R'\} \cup \{(z, i);\ (i, x) \in R'\}$$

tak, že když v R' byla hrana (i, x) mezi množinami I a X, zaměníme ji za odpovídající hranu (z, i); rozpojíme tak totiž stejnou cestu jako hranou (i, x).

Označíme si

$$A = \{i;\ (z, i) \in R\}, \quad B = \{x;\ (x, s) \in R\}, \quad J = I \setminus A.$$

Nakresleno



Hrany z J vedou do B, proto

$$\bigcup_{j\in J} M_j \subseteq B.$$