6 Vzdialenosť v grafe

Definition 6.1 *Nech $G = (V, E)$ je súvislý graf. Pre vrcholy v, $v' \in V$ definujme číslo $d_G(v, v')$ ako dĺžku najkratšej cesty z v do v' v grafe G. Číslo $d_G(v, v')$ sa nazýva vzdialenosť vrcholov v a v' v grafe G.*

Jednou zo základných algoritmických úloh v teórii grafov je nájsť najkratšiu cestu medzi dvoma zadanými vrcholmi v grafe. Takáto požiadavka vzniká v mnohých praktických aplikáciach. Známe algoritmy riešiace túto úlohu spravidla počítajú omnoho viacej – obvykle nájdu najkratšiu cestu z východzieho vrcholu do všetkých (alebo mnohých) vrcholov. Jedna z najjednoduchších a najdôležitejších takýchto metód je tvz. *Dijkstrov algoritmus*.

6.1 Hľadanie najlacnejšej cesty z daného vrcholu

Dijkstrov algortimus budeme aplikovať na graf, ktorý má hrany ohodnotené kladnými reálnymi číslami. To znamená, že spolu s každou hranou $e \in E(G)$ je dané kladné reálne číslo $w(e)$, alebo ináč povedané je daná funkcia $w\colon E(G) \to (0, \infty)$. Ohodnotenie $w(e)$ hrany e si môžeme predstavovať ako jej dĺžku, alebo cenu, ktorú musíme zaplatiť za jej prejdenie. Cena nejakej cesty v takto ohodnotenom grafe G je rovná súčtu cien jej hrán a cena $d_{G,w}(u, v)$ dvoch vrcholov $u, v \in V(G)$ je potom rovná najmenšej z cien všetkých ciest, spájajúcich u a v. Obyčajnú grafovú vzdialenosť dostaneme ako špeciálny prípad, totiž ak je $w(e) = 1$ pre každú hranu e.

Dijkstrov algoritmus.
Vstup: Graf $G = (V, E)$, ohodnotenie hrán $w(e)$, $e \in E$, štartovný vrchol $s \in V$.
Premenné: čísla $d(v)$, $v \in V$, číslo δ, množiny A, $N \subseteq V$.
Výstup: Po skončení algoritmu udáva $d(v)$ cenu najlacnejšej cesty do v z s, pre každé $v \in V$.

Krok 1: (Inicializácia)
$d(s) := 0$, $d(v) := \infty$ pre $v \in V \setminus \{s\}$; $A := V$.

Krok 2: (Test ukončenia)
Ak pre všetky $v \in A$ platí $d(v) = \infty$, algoritmus končí, inak sa pokračuje krokom 3.

Krok 3: (Voľba množiny N)
$\delta := \min\{d(y); y \in A\}$, $N := \{v \in A; d(v) = \delta\}$, $A := A \setminus N$

Krok 4: (Aktualizácia $d(y)$ pre všetkých susedov N)
Pre každú hranu $e = \{v, y\}$, kde $v \in N$ a $y \in A$, preveď príkaz

$$d(y) := \min\{d(y), d(v) + w(e)\}.$$

Po vyčerpaní všetkých takých hrán e pokračuj krokom 2.

Dôkaz správnosti algoritmu. Zrejme cyklus prebehne toľkokrát, koľko rôznych konečných cien z vrchola s do ostatných vrcholov v grafe G existuje (predpokladajme k). Stačí si uvedomiť, že po i-tom prechode cyklu ($1 \le i \le k$) v každom vrchole $v \in V$ hodnota $d(v)$ udáva najmenšiu možnú cenu z ciest obsahujúcich nanajvýš i hrán z vrchola s do vrchola v.

Poznámka. Predchádzajúci algoritmus spracuje graf o n vrcholoch v čase $O(n^2)$.

6.2 Reprezentácia grafov

Nech $G = (V, E)$ je obyčajný graf s n vrcholmi. Označme vrcholy grafu G v nejakom poradí $v_1, \ldots, v_n$. Matica susednosti grafu G je štvorcová $n \times n$ matica $A_G = (a_{ij})_{i,j=1}^n$ definovaná predpisom: $a_{ij} = 1$ pre $\{v_i, v_j\} \in E$ a 0 ináč.

A_G je vždy symetrická štvorcová matica, ktorej prvky sú 0 a 1 (v prípade orientovaných grafov matica samozrejme nemusí byť symetrická, v prípade ohodnotených grafov je na mieste (i, j) cena hrany $\{i, j\}$). Daný spôsob reprezentácie nie je najvhodnejší vzhľadom na pamäť počítača. Obzvlásť, keď má graf málo hrán býva výhodnejšie napr. pre každý vrchol zadať zoznam jeho susedov.

Vlastnosti matice susednosti.

Lemma 6.1 *Nech $G = (V, E)$ je graf, $V = \{v_1, \ldots, v_n\}$ a nech $A = A_G$ je jeho matica susednosti. Označme A^k k-tu mocninu matice susednosti. Nech $a_{ij}^{(k)}$ označuje prvok matice A^k v pozícii (i, j). Potom $a_{ij}^{(k)}$ je počet sledov dĺžky k z vrchola v_i do vrchola v_j v grafe G.*

Dôkaz. Indukciou podľa k. Pre $k = 1$ zrejmé.

Predpokladajme $k > 1$ a nech v_i, v_j sú nejaké vrcholy. Každý sled dĺžky k z v_i do v_j začína nejakou hranou idúcou z v_i do v_l a zvyšok je sled dĺžky $k - 1$ z v_l do v_j. Podľa IP je počet sledov dĺžky $k - 1$ z v_l do v_j rovný $a_{lj}^{(k-1)}$. Tým pádom je počet sledov dĺžky k z v_i do v_j rovný

$$\sum_{\{v_i, v_l\} \in E(G)} a_{lj}^{(k-1)} = \sum_{l=1}^{n} a_{il} a_{lj}^{(k-1)}.$$

To je ale podľa definície násobenia matíc presne prvok v pozícii (i, j) súčinu matíc A a A^{k-1}, alebo $a_{ij}^{(k)}$.
□

Poznámka. Pre vzdialenosť vrcholov platí

$$d_G(v_i, v_j) = \min\{k; a_{ij}^{(k)} \neq 0\}.$$

7 Stromy

Strom je jeden z najprirodzenejších útvarov ako v prírode, tak v matematike. Ako tento pojem vystihnúť? V teórii grafov existuje niekoľko rôznych charakterizácií pre stromy. Najčastejšie sa definujú nasledujúcu:

Definition 7.1 *Strom je súvislý graf, ktorý neobsahuje kružnicu.*

Z mnohých hľadísk je táto definícia nešikovná. Napríklad nie je jasné, ako jednoducho overiť, že daný graf je strom. Súvislosť je jednoduchá, horšie je to s neexistenciou kružníc. Preto sú dôležité alternatívne popisy stromov, ktoré budú nasledovať:

Lemma 7.1 *Každý strom s aspoň dvoma vrcholmi obsahuje aspoň dva vrcholy stupňa* 1*. Vrcholy stupňa* 1 *sa nazývajú koncové vrcholy alebo listy.*

Dôkaz. Nech $P = (v_0, e_1, v_1, \ldots, e_t, v_t)$ je cesta maximálnej možnej dĺžky v strome $T = (V, E)$. Zrejme dĺžka cesty P je aspoň 1. Tvrdíme, že v_0 a v_t sú koncové vrcholy T. To môžme nahliadnuť sporom: ak napríklad v_0 nie je koncový vrchol, potom existuje ešte iná hrana $e \neq e_1$ obsahujúca vrchol v_0; označme $e = \{v_0, v\}$. Potom buď je v niektorým z vrcholov cesty P, t.j. $v = v_i$, $i \geq 2$ (v tom prípade e spolu s úsekom cesty P od v_0 do v_i tvoria kružnicu), alebo $v \neq \{v_0, \ldots, v_t\}$ – čiže môžme cestu P predĺžiť o hranu e. V obidvoch prípadoch dostávame spor! □

Lemma 7.2 *(postupný „rast“ – „spiľovanie“ stromov) Pre daný graf G a jeho koncový vrchol v sú nasledujúce dve tvrdenia ekvivalentné:*
(i) G je strom,
(ii) $G \setminus v$ je strom ($G \setminus v$ je graf, ktorý dostaneme z G vyhodením vrcholu v a všetkých hrán, ktoré z neho vychádzajú).

Dôkaz.
(i) $\implies$ (ii): Ukážme, že $G \setminus v$ je súvislý graf. Nech x, y sú dva vrcholy grafu G rôzne od v a uvážme nejakú cestu z x do y. Táto cesta nemôže obsahovať vrchol stupňa 1 mimo x a y a teda neobsahuje v. Celá je teda obsiahnutá v $G \setminus v$, ktorý je teda nutne súvislý. Pretože G neobsahuje kružnicu, zrejme ani $G \setminus v$ neobsahuje kružnicu a je to teda strom.
(ii) $\implies$ (i): Nech $G \setminus v$ je strom. Pridaním koncového vrcholu v nemôžeme zjavne vytvoriť kružnicu. Súvislosť G je tiež zrejmá: medzi ľubovoľnými dvoma vrcholmi rôznymi od v existovala cesta už v grafe $G \setminus v$, a cesta z nejakého vrcholu x do vrcholu v sa dostane predĺžením cesty spájajúcej x s v' o hranu $\{v', v\}$, kde v' je jediný sused vrchola v v grafe G. □

Predchádzajúce lemma ukazuje, že graf je strom práve vtedy keď ho je možné previesť na jeden vrchol postupným odoberaním koncových vrcholov.
Nasledujúca veta zhrňuje základné znalosti o stromoch:

Veta 7.1 *(charakterizácia stromov) Pre graf G sú nasledujúce podmienky ekvivalentné:*

1. *G je strom,*
2. **jednoznačnosť cesty***: Pre každé dva vrcholy u a v z G existuje práve jediná cesta z u do v v G.*
3. **minimálna súvislosť***: Graf G je súvislý a vynechaním ľubovoľnej hrany vznikne nesúvislý graf (každá hrana je mostom).*
4. **maximálny graf bez kružníc***: Graf G neobsahuje kružnicu a každý graf, ktorý vznikne z G pridaním hrany už kružnicu obsahuje.*
5. **Eulerov vzorec***: G je súvislý a $|V| = |E| + 1$.*

Dôkaz. Dokážeme, že každé z tvrdení (ii) až (v) je ekvivalentné s (i). Dôkazy budú väčšinou robené indukciu podľa počtu vrcholov grafu G (pre jednovrcholové grafy tvrdenia zrejme platia) využitím postupného rastu stromov.

Najskôr ukážeme, že z (i) vyplýva (ii)-(v). Nech teda G je strom na aspoň dvoch vrcholoch, v jeho koncový vrchol a v' jediný sused v. Predpokladajme, že pre $G \setminus v$ platia (ii)–(v). Potom je zrejmá platnosť (ii), (iii) a (v) aj pre G. Pre (iv) nepotrebujeme ani indukciu: nakoľko G je súvislý, medzi ľubovoľnými dvoma vrcholmi $x, y \in V(G)$ vedie cesta a pokiaľ $\{x, y\} \neq E(G)$, potom hrana $\{x, y\}$ spolu so zmienenou cestou tvoria kružnicu.

Teraz dokážeme, že z každej z podmienok (ii)–(v) vyplýva (i). Pri (ii) a (iii) už súvislosť predpokladáme. A zrejme graf spĺňajúci (ii) alebo (iii) nemôže obsahovať kružnicu.

Pre platnosť implikácie (iv) $\implies$ (i) stačí overiť, že G je súvislý. Ak sú x, $y \in V(G)$ dva vrcholy, potom sú buď priamo spojené hranou, alebo graf $G + \{x, y\}$ obsahuje kružnicu a z tej po odobratí hrany $\{x, y\}$ vznikne ceta z x do y v G.

Implikácia (v) $\implies$ (i) sa dokáže opäť indukciou podľa počtu vrcholov. Nech teda G je súvislý graf na aspoň dvoch vrcholoch spĺňajúci $|V| = |E| + 1$. Súčet stupňov všetkých vrcholov je $2|V| - 2$ a pretože všetky stupne sú aspoň 1 (zo súvislosti), musí existovať vrchol v stupňa práve 1, alebo koncový vrchol grafu G. Graf $G' = G \setminus v$ je opäť súvislý a spĺňa $|V(G')| = |E(G')| + 1$, je to teda podľa indukčného predpokladu strom a teda i G je strom. □

Poznámka. Súvislý graf na n vrcholoch má aspoň $n - 1$ hrán, pričom ak má práve $n - 1$ hrán, tak je to strom.

Dôkaz. Ak je graf stromom, tak tvrdenie platí podľa prechádzajúcej vety, ináč graf obsahuje kružnicu. Vyhodením ľubovoľnej hrany z kružnice sa súvislosť grafu neporuší, len sa zníži počet jeho hrán. Postupne budeme vyhadzovať hrany z kružnice až raz skončíme pri strome. □