Dijkstrův algoritmus

Hledání nejkratší cesty v nezáporně hranově ohodnoceném grafu

Nechť je dán orientovaný graf $G = (V, H)$ a funkce, která každé hraně $h = (u, v) \in H$ přiřadí *nezáporné* reálné číslo označované $L(h)$ nebo $L(u, v)$.

Je-li $u_0, u_1, \ldots, u_k$ orientovaná cesta v G, pak její *délkou* nazveme číslo

$$\sum_{i=1}^{k} L(u_{i-1}, u_i).$$

Jsou-li dány dva vrcholy u, v grafu G, cílem je určit, zda existuje orientovaná cesta z u do v a především pokud existuje, tak nalézt nejkratší z nich. Známé algoritmy pro tento problém určují buď vzdálenosti z u do všech ostatních vrcholů (mezi nimi i do v) a nebo naopak vzdálenosti ze všech vrcholů do v. Problém nalezení nejkratší cesty v grafu je řešen Bellman-Fordovým algoritmem, který nevyžaduje předpoklad o nezápornosti délky hrany. Pokud ovšem jsou hrany nezáporné, což je v praxi velmi časté, pokud například délky hran souvisejí s fyzickou vzdáleností vrcholů, je vhodné používat algoritmus Dijkstrův, který je jednoduší a především rychlejší.

Počátek, ze kterého budeme hledat vzdálenosti do ostatních vrcholů si označíme jako v_0. Existuje řada variant implementace Dijkstrova algoritmu. Zde byla zvolena varianta, kde každý vrchol může být ve třech stavech: *nedosažený*, *dosažený* a *probraný*; dosažený přitom je chápán jako "dosažený, ale neprobraný"; pro každý vrchol w, který je dosažený nebo probraný, je dáno číslo $E(w)$ (E jako "Estimation"), což, jak později dokážeme, bude horní odhad vzdálenosti z počátku v_0 do w a na konci výpočtu přímo vzdálenost z v_0 do w. Pro nedosažené vrcholy bude $E(w)$ nedefinováno. Pokud w není počátek, pak jestliže $E(w)$ je definováno, bude pro takový vrchol dán vrchol $P(w)$; pomocí hodnot P bude možné pro každý vrchol w s definovaným $E(w)$ snadno nalézt cestu délky $E(w)$ z počátku do w způsobem, který bude popsán později.

Dijkstrův algoritmus

označ počátek v_0 jako dosažený a $E(v_0) := 0$;
for každý vrchol w různý od počátku v_0 **do**
    označ vrchol w jako nedosažený;
while existuje dosažený vrchol u **do begin**
    zvol jako u dosažený vrchol, který má minimální $E(u)$;
    označ u jako probraný;
    for všechny vrcholy w takové, že (u, w) je hrana **do**
        if w je nedosažený nebo $E(w) > E(u) + L(u, w)$ **then begin**
            if w je nedosažený **then** označ w jako dosažený;
            $E(w) := E(u) + L(u, w)$;
            $P(w) := u$;
            end;
      end.

Volba u znamená zvolit dosažený vrchol u takový, že $E(u) \leq E(w)$ pro každý dosažený vrchol w; takových vrcholů u může být více, pak je možno zvolit libovolný z nich.

Je snadné dokázat, že se výpočet algoritmu zastaví:

Lemma 1 *Stav vrcholu se může měnit jen z "nedosažený" na "dosažený" nebo z "dosažený' na "probraný".*

Důkaz: Jediné změny stavu proti počátečnímu nastavení mohou nastat v řádcích 6 a 9. ♣

Theorem 2 *Výpočet Dijkstrova algoritmu se zastaví ponejvýše N iteracích* **while** *cyklu, kde N je počet vrcholů grafu.*

Důkaz: Z popisu algoritmu a předchozího lemmatu plyne, že množina probraných vrcholů se každou iterací cyklu zvětší o jeden vrchol a její velikost je přitom mezi 0 a N. ♣

V dalším budeme dokazovat různé vztahy mezi hodnotami proměnných. Některé platí stále, některé mohou být dočasně narušeny a v zápětí obnoveny. Budeme se proto zajímat, zda platí v okamžicích, kdy se začíná testovat podmínka **while** cyklu. Tyto okamžiky budeme nazývat vstupy do cyklu.

V prvním řádku výpočtu se počátek v_0 označí jako dosažený a definuje se hodnota $E(v_0)$. Při prvním vstupu do **while** cyklu je počátek jediným dosaženým vrcholem, bude tedy vybrán jako vrchol u v řádku 5 a v zápětí v řádku 6 označen jako probraný. Na základě lemmatu 1 pak již zůstane probraným až do konce výpočtu. Status ostatních vrcholů popisuje následující

Lemma 3 *Pokud $w \neq v_0$, pak při každém vstupu do cyklu platí: $E(w)$ je definováno právě když je definováno $P(w)$ a to je právě když w je dosažený nebo probraný vrchol.*

Důkaz: Vrchol w jiný než počátek je po provedení inicializace v řádcích 2 a 3 nedosažený a $E(w)$ a $P(w)$ nejsou definovány. Uvažujme první okamžik, kdy se některé z těchto tvrzení změní. To se může stát pouze když je probírána hrana (u, w) pro nějaký vrchol u. V takovém případě se najednou změní klasifikace vrcholu w na dosažený a definují se hodnoty $E(w)$ a $P(w)$. Podle lemmatu 1 už vrchol w nikdy nestane nedosaženým a $E(w)$ a $P(w)$ zůstávají stále definovány. ♣

Velmi důležitým, i když na první pohled samozřejmým je následující

Lemma 4 *Je-li v probraný vrchol a w dosažený vrchol, pak $E(v) \leq E(w)$. Jestliže v je probraný vrchol, pak se hodnota $E(v)$ nemůže změnit.*

Důkaz: Podle předchozího lemmatu jsou pro probrané a dosažené vrcholy hodnoty E definovány. Tvrzení lemmatu dokážeme indukcí podle počtu kroků výpočtu. Po provedení inicializace v řádcích 1-3 triviálně platí, neexistují probrané vrcholy. Nechť platí do nějakého vstupu do **while** cyklu. Platnost tvrzení by se mohla porušit buď změnou stavu nějakého vrcholu nebo změnou hodnoty $E(w)$ nějakého vrcholu w při provádění těla cyklu.

Předpokládejme, že se provádí tělo cyklu a v řádku 6 algoritmu byl vybrán vrchol u. Ukážeme, že pro každý již probraný vrchol v a každý dosažený vrchol w platí $E(v) \leq E(u) \leq E(w)$. Při vstupu do cyklu $E(v) \leq E(u)$ představuje indukční předpoklad, nerovnost $E(u) \leq E(w)$ zase vyplývá z toho, jak byl vrchol u vybrán.

Změny, které mohou nastat nebo nastanou při provádění těla cyklu s vybraným vrcholem u, které by mohly narušit platnost tvrzení lemmatu, jsou následující:

Vrchol u je překlasifikován na probraný; jelikož $E(u) \leq E(w)$ pro každý dosažený w, lemma zůstává v platnosti.

Objeví se nový dosažený vrchol x; pro takový vrchol se nastaví $E(x) = E(u)+L(u,w)$, a v důsledku nezápornosti pro libovolné probrané v platí $E(v) \leq E(u) \leq E(x)$, což je v souladu s tvrzením lemmatu.

Pro některý dosažený vrchol x se hodnota $E(x)$ změní na $E(u) + L(u,w)$; podobně jako v předchozím odstavci se ukáže, že nová hodnota je stále alespoň tak velká, jako $E(v)$ pro libovolný probraný vrchol v.

Je-li vrchol x probraný a (u,x) je hrana, pak je $E(x) \leq E(u) \leq E(u)+L(u,x)$ (levá nerovnost plyne z indukčního předpokladu, druhá z nezápornosti délek hran) a tedy podmínka $E(x) > E(u) + L(u,x)$ v řádku 8 nebude nikdy splněna. U probraného vrcholu x proto již ke změně stavu nebo hodnoty $E(x)$ nebo $P(x)$ nemůže dojít. ♣

Lemma 5 *Hodnota $P(v_0)$ není nikdy definována.*

Důkaz: Po provedení inicializace $P(v_0)$ definováno není a pokud by bylo definováno později, stalo by se to současně se změnou $E(v_0)$ v okamžiku, kdy již v_0 je probraný. Tuto možnost ale předchozí lemma vyloučilo. ♣

Nyní ukážeme, že hodnoty $P(w)$ umožňují pro každý probraný nebo dosažený vrchol w najít cestu z počátku do w, která má délku $E(w)$.

Lemma 6 *Je-li $P(w)$ definováno a označíme-li $v = P(w)$, pak (v,w) je hrana v je probraný vrchol a $E(v) + L(v,w) = E(w)$.*

Důkaz: Uvažujme jistý okamžik τ výpočtu, kdy $P(w)$ je definováno. V okamžiku, kdy bylo $P(w)$ naposledy měněno a to na hodnotu v, byla probírána hrana (v,w), a provedlo se dosazení $P(w) := v$. $(P(w),w) = (v,w)$ je tedy hrana a vrchol v je probraný. Hodnota $E(w)$ byla tehdy změněna na $E(u) + L(v,w)$ a od té doby se do okamžiku τ nezměnila, protože změny $E(w)$ a $P(w)$ jsou spolu. Jelikož v bylo v okamžiku změny $P(w)$ probrané, předchozí lemma říká, že $E(v)$ se již také nezměnilo, proto rovnost $E(v) + L(v,w) = E(w)$ zůstává v platnosti. ♣

Lemma 7 *P je acyklická relace, neboli neexistuje posloupnost vrcholů $w_0, \ldots, w_k$ taková, že $w_{i-1} = P(w_i)$ pro $i = 1, \ldots, k$ a $w_k = P(w_0)$.*

Důkaz: Z lemmatu 1 a podmínky cyklu plyne, že každý vrchol w, pro který je v průběhu výpočtu $E(w)$ definováno, je na konci výpočtu probraný. Pro takový vrchol označme jako $T(w)$ takové číslo T, že vrchol se stal probraným v T-té iteraci **while** cyklu. Jelikož se v každé iteraci stane probraným jediný vrchol, je tato funkce prostá. Jestliže v některém okamžiku je v probraný vrchol a w je dosažený, pak zřejmě $T(v) < T(w)$. Kdykoli se nově definuje nebo mění $P(w)$ pro nějaký vrchol w, stane se tak při probírání hrany (u, w) v okamžiku, kdy u krátce předtím byl označen za probraný vrchol a w je ještě (nově nebo dříve) dosažený. Kdykoliv je tedy $u = P(w)$, je $T(u) < T(w)$, což ale vylučuje možnost cyklu relace P. ♣

Jestliže tedy $E(w)$ definováno, pak definujme posloupnost $w = z_0, z_1, \ldots$ předpisem $z_i = P(z_{i-1})$ pro $i = 1, \ldots$. Na základě předchozího lemmatu musí tato posloupnost být prostá a protože množina vrcholů je konečná, musí i právě definovaná posloupnost být konečná. Předpokládejme, že do ní už není možno přidat další prvek. Podle lemmat 3, 6 a 5 tedy tato posloupnost končí vrcholem v_0. Posloupnost má tedy tvar $w = z_0, \ldots, z_k = v_0$ a podle lemmatu 6 je obrácená posloupnost $v_0 = z_k, z_{k-1}, \ldots, z_0 = w$ cestou z v_0 do w v grafu G. Tuto cestu nazveme *vytčená* cesta do w. Vytčená cesta je jednoznačně dána svým cílovým vrcholem.

Lemma 8 *Je-li $E(w)$ definováno, pak existuje vytčená cesta do w a její délka je $E(w)$.*

Důkaz: Existence vytčené cesty $v_0 = w_0, w_1, \ldots, w_k = w$ z v_0 do w byla zdůvodněna v předchozím odstavci. Pro každou hranu (w_{i-1}, w_i) této cesty platí podle lemmatu 6 $E(w_i) - E(w_{i-1}) = L(w_{i-1}, w_i)$, z čehož ihned plyne, že délka celé cesty je

$$L(w_0, w_1) + \cdots + L(w_{k-1}, w_k) = [E(w_1) - E(w_0)] + \cdots + [E(w_k) - E(w_{k-1})] =$$

$$E(w_k) - E(w_0) = E(w) - E(v_0) = E(w).$$

♣

Řekneme,že cesta $v_0 = w_0, \ldots, w_k = w$ z počátku do vrcholu w je *definitivní*, jestliže její vrcholy $w_0, \ldots, w_{k-1}$ jsou probrané. Povšimněte si toho, že vrchol w_k nemusí být probraný, definitivní cesta tedy může vést i do vrcholu, který není probraný. Je ovšem zřejmé, že

Lemma 9 *Je-li při vstupu do* **while** *cyklu $v_0 = w_0, \ldots, w_k = w$ definitivní cesta, pak w je dosažený vrchol.*

Důkaz: Jakmile byl w_{k-1} označen za probraný, byla v zápětí probírána hrana (w_{k-1}, w_k) a tedy se vrchol w_k stal dosaženým, pokud již dosažený nebyl. ♣

Lemma 10 *Je-li $E(w)$ definováno, pak vytčená cesta do w je definitivní.*

Důkaz: Plyne ihned z lemmatu 6. ♣

Klíčové lemma důkazu správnosti Dijkstrova algoritmu je

Lemma 11 *Je-li při některém vstupu do* **while** *cyklu* $E(w)$ *definováno, pak* $E(w)$ *je délka nejkratší definitivní cesty z* v_0 *do* w.

Důkaz: Z předchozího lemmatu plyne, že $E(w)$ je délka vytčené cesty, která je definitivní. Je tedy třeba dokázat, že neexistuje definitivní cesta do w, kratší než $E(w)$.

Na začátku výpočtu existuje jediná definitivní cesta a to cesta v_0 (jediný vrchol, žádná hrana) do v_0 a ta má délku $0 = E(v_0)$, tvrzení tedy platí. Uvažujme okamžiky ukončení provádění některé iterace těla **while** cyklu a předpokládejme, že se platnost tvrzení v některém z nich poruší a uvažujme první okamžik, kdy se tak stane (před tím tedy tvrzení platí). Mohlo dojít buď k vytvoření nové definitivní cesty do w, která by byla kratší než $E(w)$ nebo ke změně $E(w)$.

Pokud vznikla nová definitivní cesta $v_0 = w_0, w_1, \ldots, w_k = w$ do w, stalo se tak v řádku 6 algoritmu, kdy byl některý vrchol u označen za probraný, takže tato cesta, která původně definitivní nebyla, se definitivní stala. V tom případě muselo být $u = w_i$ pro některé i takové, že $0 \leq i < k$.

Pokud bylo $i = k-1$, neboli nově probraný vrchol u byl předposlední vrchol cesty, pak cesta $v_0 = w_0, w_1, \ldots, w_{k-1}$ již definitivní byla, tedy dle indukčního předpokladu její délka byla alespoň $E(w_{k-1})$ a proto délka celé cesty $v_0 = w_0, w_1, \ldots, w_k = w$ je alespoň $E(w_{k-1})+L(w_{k-1}, w_k)$, což je hodnota, na kterou je vzápětí upraveno $E(w_k)$ při probírání hrany $(u, w) = (w_{k-1}, w_k)$; k porušení platnosti tvrzení lemmatu by tedy došlo jen na okamžik, ale do dokončení iterace **while** cyklu se zase platnost obnoví.

Pokud by bylo $i < k-1$ pak cesta $v_0 = w_0, w_1, \ldots, w_i$ již definitivní byla a tedy délka tohoto úseku byla alespoň $E(w_i)$ podle indukčního předpokladu. Jelikož délky hran jsou nezáporné, délka cesty $v_0 = w_0, w_1, \ldots, w_{k-1}$, rovná součtu délky cesty $v_0 = w_0, w_1, \ldots, w_i$ (alespoň $E(w_i)$) a délky cesty $w_i, w_1, \ldots, w_k$ (nezáporná),je alespoň $E(w_i)$. Jelikož ale vrchol u byl právě přeřazen z dosažených mezi probrané, plyne z lemmatu 4,že $E(w_i) \geq E(w_{k-1})$ a přitom $E(w_{k-1})$ je délka vytčené cesty do w_{k-1}. Délka cesty $w_0, w_1, \ldots, w_k$, což je délka cesty $w_0, w_1, \ldots, w_{k-1}$ plus délka hrany (w_{k-1}, w_k), tedy není nižší než délka cesty C tvořené vytčenou cestou do w_{k-1} prodlouženou o hranu (w_{k-1}, w_k). Cesta C ale již definitivní byla a má tedy délku alespoň $E(w_k)$ podle indukčního předpokladu, takže v tomto případě nemůže dojít k porušení platnosti lemmatu ani na krátký okamžik. ♣

Z toho již okamžitě plyne

Theorem 12 *V okamžiku ukončení výpočtu Dijkstrova algoritmu je pro každý vrchol* w, *do kterého existuje cesta z počátku* v_0, $E(w)$ *definováno a rovno délce nejkratší cesty z* v_0 *do* w *a vytčená cesta do* w *je nejkratší cesta do* w *(nebo jedna z nejkratších cest).*

Důkaz: Nakonci výpočtu nejsou v grafu žádné dosažené vrcholy a tedy každá cesta do w je dosažená. ♣

Dijkstrův algoritmus je velmi blízký Bellman-Fordovu algoritmu; vznikne z něho tím, že pro probírání nevybereme *první* vrchol ve frontě Q, ale vrchol u s nejnižším $E(u)$ (fronta je tvořena právě dosaženými vrcholy, které jsou identické s vrcholy označenými k probírání v Bellman-Fordově algoritmu). Jelikož v případě nezáporně označených hran má dosažený vrchol s nejnižším $E(u)$ již tuto hodnotu rovnou vzdálenosti od počátku, nebude již nikdy probírán a proto odpadá opakované probírání vrcholů, které způsobuje větší vypočetní složitost Bellman-Fordova algoritmu.

Jestliže ale se opakované probírání vrcholů zakáže, jak je tomu u výše podané varianty Dijkstrova algoritmu, pak v případě existence záporně ohodnocených hran můžeme dostat chybný výsledek.

Stejně tak se můžeme v řádku 7 algoritmu omezit (na rozdíl od obecného Bellman-Fordova algoritmu) na hrany vedoucí do jiných vrcholů než probraných, protože lemma 4 říká, že u hrany vedoucí do probraného vrcholu nemůže v grafu s nezáporně ohodnocenými hranami nikdy dojít ke zlepšení $E(w)$.