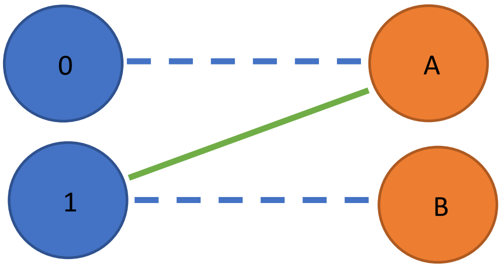
Daniel Skýpala

Sponzoři

Zamyšlení

Pojďme si problém převést na graf. Zvířata a sponzoři jsou vrcholy grafu, hrany jsou mezi zvířetem a jeho potencionálními sponzory.

Představme si že máme nějaké párování – podgraf grafu nahoře, kde zvířata a sponzoři mají nejvýše jednu hranu. Toto je možné řešení úlohy nemusí být však optimální. Ale lze ho vylepšit!



Jak na to? Představme si graf výše. Zde je spárovaný sponzor A se zvířetem 1, ale toto není optimální řešení. Pojďme najít vylepšovák – takovou cestu v grafu nahoře, když všechny její hrany „převrátíme“ (spárované na nespárované a naopak), tak dostaneme lepší párování (ale stále validní). (V příkladě je to 0-A-1-B). Jaké vlastnosti takový vylepšovák musí mít:

1. Musí se v něm střídat nespárovaná a spárovaná hrana. – Kdyby byly 2 nespárované hrany za sebou, vrchol mezi nimi po převrácení bude mít dvě spárované hrany – a to není validní. (někdo bude mít dva sponzory/zvířata). No a kdyby byly 2 spárované hrany za sebou, tak jsme nedostali párování.
2. Musí mít více nespárovaných hran než spárovaných. (Jinak bychom počet spárovaných dvojic po převrácení nezvýšili, a to by pak párování nevylepšilo.)
3. Podmínky 1 a 2 nám dávají dohromady, že první a poslední hrana musí být nespárovaná a to, že délka cesty je lichá.
4. Musí začínat a končit v nespárovaných vrcholech. (Jinak bychom vytvořili nevalidní párování.)

Vlastnosti vylepšováku nám dávají to, že párování určitě vylepší a zároveň vylepšené párování bude validní.

Lemma o optimálnosti: Vylepšovák neexistuje právě tehdy, když máme optimální řešení

Představme si, že máme optimální párování O a neoptimální párování N. Pro každou hranu, která se v obou párování liší (spárovaná / nespárovaná), ji obarvíme buď zeleně pokud patří O, jinak červeně (N). Všechny barevné hrany s příslušnými tvoří rozdílový graf.

Všimněme si, že vrcholy v rozdílovém grafu jsou nejvýše stupně 2. (Jinak by z něj vedly alespoň 2 stejně barevné hrany, což by znamenalo, že buď O nebo N není párování)

To nám dává, že v rozdílovém grafu mohou buď vznikat cesty sudé délky, nebo cesty liché délky. (Zároveň se v nich střídají hrany, jinak by O nebo N nebylo párování.) Pokud v grafu O přehodíme všechny hrany, které jsou v rozdílovém grafu v cestě sudé délky, počet spárovaných dvojic se nezmění (hrany v cestě se střídají spárovaná/nespárovaná a je jich stejný počet). Takže sudé cesty nemají vliv na to, zdali je řešení optimální. Pokud přehodíme lichou cestu, tak se řešení změní, nicméně lichá cesta se střídajícími hranami je vylepšovák. (dokázáno)

Zpátky: Takže stačí jenom hledat vylepšováky a převracet hrany na nich, dokud nám nezůstane optimální graf, který vylepšovák nemá.

Jak hledat vylepšováky?

Hopcroft-Karp

Protože nejsme troškaři, budeme hledat spoustu vylepšováků naráz. Jak to uděláme?

Najdeme si délku nejkratšího vylepšováku. Projdeme náš graf z nespárovaných zvířat do šířky postupně nespárovaná – spárovaná hrana, až dojdeme do nespárovaného sponzora. Necháme doběhnout BFS pro vrcholy, které mají stejnou vzdálenost od startu, jako nespárovaný sponzor. Do všech vrcholů, ke kterým jsme v BFS došli, si uložíme jejich vzdálenost od startovních vrcholů.

Nyní budeme hledat zlepšováky DFSkem. Začneme vždy ve nespárovaném zvířeti a prohledáme z něj do nespárovaného sponzora. Abychom hledali jen nejkratší zlepšováky, vždy budeme ze současného vrcholu prohledávat jen ty vrcholy, které mají o 1 vyšší vzdálenost od startovního vrcholu (zjištěnou v BFS). Poté co najdeme zlepšovák (skončíme v nespárovaném sponzorovi), přehodíme všechny hrany na něm. To budeme opakovat přes všechny volné vrcholy. (Mezi jednotlivými iteracemi DFS si budeme pamatovat prohledanost vrcholů, abychom neprohledávali slepé části grafu.)

Pokud budeme kombinaci BFS a DFS v cyklu opakovat, tak se náš graf bude postupně vylepšovat. S opakováním lze skončit, když BFS nedojde do nespárovaného sponzora – v tu chvíli žádný zlepšovák neexistuje.

Poté jen zjistíme, kteří zvířata a sponzoři jsou spárovaní a dvojice seřadíme a vypíšeme.

Časová složitost

Nejdřív, jak dlouho trvá kombinace BFS a DFS? Každý vrchol a hranu navštívíme jen jednou $O(N + M)$ Ale protože vrcholy bez hran jsou pro tuto úlohu nezajímavé a můžeme je vyřadit, tak po jejich vyřazení: $O(N) \leq O(M)$, tak celkem $O(M)$.

A kolikrát budeme tuto kombinaci opakovat? Můžeme si všimnout, že v každém kroku se délka zlepšováku zvýší o 1 (použijeme všechny nejlepší zlepšováky) a maximální délka zlepšováku je $O(N)$ (přes každý vrchol jen jednou). To dává: $O(NM)$.

No, ale ve skutečnosti to děláme rychleji. Představme si, že cyklus zastavíme po $O\left(\sqrt{N}\right)$ iteracích. Poté si představíme rozdílový graf našeho grafu N oproti optimálnímu grafu O. Všech lichých cest (zlepšováků) je v něm $O(N/\sqrt{N}) = O\left(\sqrt{N}\right)$ (cesty se nepřekrývají). To ale znamená, že optimální párování má o $O\left(\sqrt{N}\right)$ více spárovaných hran (každý zlepšovák zvedne počet spárovaných hran o 1). Zároveň v každé iteraci cyklu najdeme zlepšovák a zvýšíme počet spárovaných hran o 1, takže cyklus můžeme už provést jen $O\left(\sqrt{N}\right)$-krát. Zastavili jsme se po $O\left(\sqrt{N}\right)$ iteracích a zbývá nám $O\left(\sqrt{N}\right)$ iterací. Celkem provedeme $O\left(\sqrt{N}\right) + O\left(\sqrt{N}\right) = O\left(\sqrt{N}\right)$ iterací. Celková složitost bude $O\left(M\sqrt{N}\right)$.

(Závěrečné seřazení v $O(N \log \mathrm{N})$ je kvůli podmínce $O(N) \leq O(M)$ určitě rychlejší než celková složitost.)

Paměťová složitost

Pamatujeme si graf a také pár informací ke každému vrcholu. $O(N + M)$