Stromy

Kombinatorika pro bioinformatiky

25.11.2020

1 Stromy

Definice 1 *Graf se nazývá strom, pokud je souvislý a neobsahuje kružnici. Pokud graf neobsahuje kružnici, ale není souvislý, nazývá se les.*

Příklad:

Lemma 2 *Každý (konečný) strom s alespoň* 2 *vrcholy má alespoň* 2 *vrcholy stupně* 1 *(tzv. listy).*

Důkaz: Zvolme u, v takové, že $d(u, v)$ je maximální (přes všechny dvojice vrcholů). Potom u, v jsou listy.

□

Poznámka 3 *Lemma neplatí pro nekonečné stromy.*

Lemma 4 *Ať G je graf a v jeho vrchol stupně 1. Potom G je strom právě tehdy, když $G \setminus v$ je strom.*

Důkaz: Předpokládejme, že graf G je strom a vrchol v je list. Pak $G - v$ je zjevně bez kružnice, protože i G kružnici neobsahoval. Kdyby byl $G - v$ nesouvislý, musel by vrchol v mít stupeň alespoň 2, což nemá a tedy je $G - v$ souvislý.

Na druhou stranu, ať v je vrchol stupně 1 v G a $G - v$ je strom. Připojením vrcholu hranou k souvislému grafu dostáváme souvislý graf, tedy G je souvislý. Pokud by G obsahoval kružnici, ale $G - v$ ne, musí tato kružnice obsahovat vrchol v a ten má stupeň alespoň 2, spor. Takže G je bez kružnice a je to strom.

$\square$

Věta 5 (ekvivalentní definice stromů) *$G = (V, E)$, $|V| \geq 2$. Pak jsou následující podmínky ekvivalentní:*

1. *G je strom,*
2. *$\forall u, v \in V$ existuje právě jedna cesta z u do v,*
3. *G je minimální souvislý (tj. $\forall e \in E$ je $G \setminus e$ nesouvislý),*
4. *G je maximální bez kružnice (tj. $\forall e' \notin E$ má $G + e'$ kružnici),*
5. *G je souvislý a $|V| = |E| + 1$,*
6. *G je bez kružnice a $|V| = |E| + 1$.*

Důkaz: **1.⇒2.**: Mějme $u, v \in V$, $u \neq v$. G je strom, tedy je souvislý a existuje cesta z u do v. Pro spor předpokládejme, že existují dvě různé cesty P_1 a P_2. Procházejme obě ve směru od u k v. Označme x poslední vrchol společné části od u a y první další vrchol obsažený v obou cestách. Potom úseky obou cest mezi vrcholy x a y tvoří kružnici a G není strom, spor.
2.⇒3.: Jelikož jsou každé dva vrcholy spojeny (právě jednou) cestou, je G souvislý. Pro spor předpokládejme, že existuje $e = \{u, v\} \in E$ taková, že $G \setminus e$ zůstane souvislý. Potom v $G \setminus e$ existuje cesta z u do v a v G druhá cesta tvořená hranou e, což je spor.
3.⇒1.: Stačí dokázat, že G neobsahuje kružnici. Pokud by ji obsahoval, pak po odstranění libovolné libovolné hrany této kružnice zůstane graf souvislý. Spor.
1.⇒4.: Zjevně G nemá kružnici. Mějme libovolnou „nehranu“ $e' = \{u, v\} \notin E$. Potom v G existuje cesta z u do v která společně s e' vytvoří kružnici.
4.⇒1.: Zjevně je G bez kružnice. Kdyby nebyl souvislý, pak přidáním hrany spojující vrcholy z různých komponent souvislosti nevznikne kružnice. Spor.
1.⇒5. (resp. **6.**): Stačí ověřit vztah mezi počtem vrcholů a hran stromu. Postupujme matematickou indukcí podle počtu vrcholů. Pro $|V| = 1$ je $|E| = 0$

a tvrzení platí. Dále mějme $|V| \geq 2$, pak G má list v a $G \setminus v$ je strom s méně vrcholy a tedy tvrzení pro něj platí. Tedy $|V(G \setminus v)| = |E(G \setminus v)| + 1$. Také $|V| = |V(G \setminus v)| + 1$ a $|E| = |E(G \setminus v)| + 1$ a ověříme, že platí i $|V| = |E| + 1$.
5.⇒1.: G je souvislý, předpokládejme ale, že obsahuje kružnici. Potom lze vynechat hranu této kružnice a G zůstane souvislý. Vynechávejme hrany tak dlouho, dokud graf zůstává souvislý. Dostaneme graf G', který je podgrafem G a je souvislý a bez kružnice. Tedy je to strom se stejným počtem vrcholů a platí pro něj vztah $|V| = |E(G')| + 1 < |E| + 1$, spor.
6.⇒1.: G nemá kružnici, předpokládejme ale, že není souvislý Potom lze přidat hranu mezi vrcholy v různých komponentách souvislosti a G zůstane souvislý. Přidávejme hrany tak dlouho, dokud graf neobsahuje kružnici. Dostaneme graf G'' který je nadgrafem G a je souvislý a bez kružnice. Tedy je to strom se stejným počtem vrcholů a platí pro něj vztah $|V| = |E(G'')| + 1 > |E| + 1$, spor.

□

Důsledek 6 *Graf $G = (V, E)$ je les s k komponentami souvislosti právě tehdy, když G nemá kružnici a $|V| = |E| + k$.*

2 Kostry grafů

Definice 7 *Kostra souvislého grafu je jeho podgraf, který obsahuje všechny vrcholy grafu a je strom.*

Příklad:

Pozorování 8 *Každá kostra grafu s n vrcholy má $n - 1$ hran.*

Algoritmus pro nalezení kostry:

$G = (V, E), |V| = n, E = \{e_1, e_2, \dots, e_m\}$

1. $E_0 = \emptyset$
2. `for i=1 to m do if` $(V, E_{i-1} \cup \{e_i\})$ obsahuje kružnici
   `then` $E_i = E_{i-1}$
   `else` $E_i = E_{i-1} \cup \{e_i\}$
3. `if` $|E_m| = n - 1$ `return` $T = (V, E_m)$ `else return` G je nesouvislý

Poznámka 9 *Je-li* $|E_i| = n - 1$ *pro nějaké* i*, pak algoritmus může skončit a* $T = (V, E_i)$ *je kostra.*

Operace s grafy $G = (V, E)$

- odebrání vrcholu $G \setminus v = (V \setminus \{v\}, E \cap \binom{V \setminus \{v\}}{2})$,
- odebrání hrany $G \setminus e = (V, E \setminus \{e\})$,
- přidání hrany $G + e' = (V, E \cup \{e'\})$,
- dělení hrany $G\%e = (V \cup w, E \setminus \{e\} \cup \{\{u, w\}, \{v, w\}\})$, kde $e = \{u, v\}$,
- kontrakce hrany $G.e = (V \setminus \{u, v\} \cup \{w\}, E')$, pro $e = \{u, v\}$ a $E' = \{f \in E | e \cap f = \emptyset\} \cup \{\{x, w\} | x \in (N(u) \cup N(v)) \setminus \{u, v\}\}$.
- multigrafová kontrakce $G : e$, mohou vznikat násobné hrany, pokud x byl spojen s u i s v

Lemma 10 *Nechť* G *je graf,* e *jeho libovolná hrana a jako* $\kappa(G)$ *označujme počet koster grafu* G*. Pak platí* $\kappa(G) = \kappa(G - e) + \kappa(G : e)$*.*

Důkaz: Kostry grafu G rozdělíme do dvou skupin podle toho, zda obsahují nebo neobsahují hranu e.

Ty, které hranu e neobsahují, jsou vlastně kostrami grafu $G - e$ a je jich tedy $\kappa(G - e)$. Nyní si stačí uvědomit, že kostry grafu G obsahující hranu e vzájemně jednoznačně odpovídají kostrám multigrafu $G : e$ a je jich tedy $\kappa(G : e)$.

$\square$

Předchozí lemma je sice užitečné, ale spíše jako argument v důkazech či úvahách, pro praktické počítání se příliš nehodí.

Věta 11 (Cayleyho formule) *Pro* $n \geq 2$ *je* $\kappa(K_n) = n^{n-2}$.

Důkaz s obratlovci: Obratlovcem nazvěme kostru grafu K_n s jedním vrcholem v kroužku (hlavička) a jedním vrcholem ve čtverečku (zadeček) – může jít i dvakrát o tentýž vrchol. Potom počet obratlovců $o(n) = \kappa(K_n) \cdot n^2$.

Lemma 12 *Existuje bijekce mezi množinou všech obratlovců na* n *vrcholech a množinou zobrazení* $[n]$ *do* $[n]$.

Důkaz: Cestě spojující hlavičku a zadeček budeme říkat páteř. Uvažme posloupnost vrcholů páteře v pořadí podle velikosti (od nejmenšího k největšímu) a poté v pořadí od hlavičky k zadečku. napíšeme-li tato pořadí pod sebe, je možné to chápat jako permutaci vrcholů páteře. Vrchol mimo páteř zobrazíme vždy na jeho souseda směrem k páteři. Tím jsme získali zobrazení $[n]$ do $[n]$.

Obdobným způsoben dokážeme ze zobrazení $[n]$ do sebe sestrojit obratlovce tak, že tyto postupy jsou inverzní a jedná se tedy o bijekce.

$\square$

Počet zobrazení množiny $[n]$ do sebe je n^n a tedy i počet všech obratlovců na n vrcholech je $o(n) = n^n$. Z toho $\kappa(K_n) = o(n)/n^2 = n^{n-2}$.

$\square$