Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Matematički odsjek

Patricija Dovijanić

Dokazi Cayleyeve formule

Diplomski rad

Voditelj rada:
prof. dr. sc. Vedran Krčadinac

Zagreb, veljača 2024.

Ovaj diplomski rad obranjen je dana ____________________ pred ispitnim povjerenstvom u sastavu:

1. ______________________________ , predsjednik
2. ______________________________ , član
3. ______________________________ , član

Povjerenstvo je rad ocijenilo ocjenom ______________________ .

Potpisi članova povjerenstva:

1. ______________________________
2. ______________________________
3. ______________________________

Sadržaj

1 Uvod

Cayleyeva formula jedan je od osnovnih rezultata u teoriji grafova. Pripisuje se Arthuru Cayleyu koji ju je objavio 1889. nadovezavši se na Borchardtov rad iz 1860. Od tada je predmet proučavanja brojnih matematičara koji su je dokazali na još nebrojeno načina.

Iako se svaki matematički problem može riješiti na više načina, Cayleyeva formula posebna je po tome što su njeni dokazi zaista zadivljujuće brojni i raznoliki. Zbog toga je nastao ovaj rad u kojem su obrađeni odabrani dokazi - oni koji se ističu po svojoj poznatosti, slikovitosti, jedinstvenosti, ili nekom drugom, možda subjektivnom kriteriju.

Prvo, uvodno poglavlje ovog rada daje pregled osnova teorije grafova. Iskazana je Cayleyeva formula te je objašnjena njena pozadina i zanimljiva povijest uz kratko izlaganje Cayleyevog originalnog dokaza.

U drugom poglavlju izlažemo tri bijektivna dokaza: dokaz pomoću Prüferovih kodova, dokaz Joyalovom bijekcijom i dokaz pomoću funkcija parkiranja. Dokaz pomoću Prüferovih kodova možda je najpoznatiji dokaz Cayleyeve formule, a temelji se na uspostavljanju bijekcije između stabala i pripadnih Prüferovih kodova koristeći jednostavne algoritme. Joyal uvodi pojam grafa kralježnjaka, stabla s dva istaknuta vrha, te definira bijekciju grafova kralježnjaka i endofunkcija odakle Cayleyeva formula lako slijedi. Funkcije parkiranja, originalno korištene u računarstvu, imaju prirodnu korespondenciju sa stablima pa prebrojavanjem funkcija parkiranja dolazimo do Cayleyeve formule.

U trećem poglavlju nastavljamo s dva dokaza koji se temelje na matematičkoj indukciji: dokaz formulom uključivanja-isključivanja te Riordanov i Rényijev dokaz. Formula uključivanja-isključivanja dobro je poznat alat u kombinatorici, a jednom kad uočimo poveznicu s prebrojavanjem određenih surjekcija ona svoju primjenu nalazi i u ovom problemu. Ideje Riordana i Rényija daju dokaz Cayleyeve formule za šume na način da uočimo rekurzivnu relaciju među šumama koja nam omogućuje primjenu matematičke indukcije. Cayleyeva formula tada slijedi kao poseban slučaj.

Nadalje, u četvrtom poglavlju iznosimo tri dokaza dvostrukim prebrojavanjem: Pitmanov dokaz preko usmjerenih stabala, dokaz dvostrukim prebrojavanjem razapinjućih stabala i Clarkeov dokaz pomoću prespajanja. Pitman u svom dokazu na dva jednostavna načina prebrojava redoslijede dodavanja bridova u usmjerena korijenska stabla. Dokaz prebrojavanjem razapinjućih stabala grafa koji sadrže zadanu šumu kao poseban slučaj daje Cayleyevu formulu. Clarkeov dokaz temelji se na definiranju pojma prespajanja stabala koji pomaže u uspostavljanju veze među stablima koja se jedno iz drugog mogu dobiti izmjenom jednog brida.

Peto poglavlje posvećeno je Pólyinom dokazu pomoću funkcija izvodnica i Lagrangeove formule inverzije. Najprije dajemo kratak uvod u funkcije izvodnice i iskaz Lagrangeove formule inverzije. Dokaz slijedi uočavanjem veze među korijenskim stablima i korijenskim šumama te određivanja koeficijenata eksponencijalne funkcije izvodnice za broj korijenskih stabala.

U šestom poglavlju definiramo Laplacijan grafa te koristeći Binet-Cauchyjev teorem dokazujemo Kirchhoffov teorem, poznat i kao matrični teorem o stablima. Cayleyeva formula slijedi kao jednostavan korolar tog teorema.

Za kraj, u posljednjem poglavlju izlažemo vjerojatnosni dokaz preko Poissonovog procesa grananja. Najprije definiramo jednostavan proces grananja i načine označavanja stabla. Zatim računamo vjerojatnost da Poissonovim procesom grananja dobijemo zadano stablo, što kao posljedicu daje upravo Cayleyevu formulu.

Htjela bih zahvaliti svojoj obitelji, prijateljima i svima koji su vjerovali u mene. Hvala Borni Šimiću na upotpunjavanju ove kolekcije dokaza ustupanjem dokaza u odjeljku 5.2. Posebno zahvaljujem svom mentoru prof. dr. sc. Vedranu Krčadincu na izvrsnom vodstvu i velikom trudu kako bi ovaj rad bio što bolji.

2 Cayleyeva formula

Najprije uvedimo osnovne pojmove iz teorije grafova te uz nekoliko primjera iskažimo Cayleyevu formulu koju ćemo u ostatku rada proučavati.

Definicija 2.1. Graf *je uređen par* $G = (V, E)$, *pri čemu je* V *neprazan konačan skup* vrhova, *a* E *je skup dvočlanih podskupova od* V *koje zovemo* bridovima.

Za vrhove $x, y \in V$ kažemo da su *susjedni* ako su povezani bridom $e = \{x, y\} \in E$. Pišemo $x \sim y$ ako je $\{x, y\} \in E$, a u suprotnom pišemo $x \not\sim y$. Kažemo da je vrh x *incidentan* s bridom e ako je $x \in e$. Nadalje, *stupanj* vrha x jest broj bridova koji su incidentni s vrhom x te ga označavamo $d(x)$. *Izolirani vrh* je vrh stupnja nula, a *list* je vrh stupnja jedan. Za graf sa samo jednim vrhom kažemo da je *trivijalan.* Ako su svi vrhovi istog stupnja, kažemo da je graf *regularan.* Ako je svaki par vrhova brid, kažemo da se radi o *potpunom grafu.* Potpuni graf s n vrhova označavamo s K_n. Graf s n vrhova bez bridova nazivamo *nulgrafom* i označavamo s N_n.

Ako dopustimo *usmjerene bridove*, tj. bridove koji imaju orijentaciju od jednog vrha prema drugome, reprezentiramo ih uređenim parom (x, y) gdje je x *početni*, a y *krajnji vrh* tog brida. *Usmjereni graf* je graf kojem su svi bridovi usmjereni.

Podgraf grafa $G = (V, E)$ je graf kojem su skup vrhova i skup bridova podskupovi od redom V i E. *Razapinjući podgraf* je podgraf oblika $G' = (V, E')$.

Za niz vrhova $(v_0, v_1, \ldots, v_k)$ kažemo da je *šetnja* duljine k ako su v_{i-1} i v_i susjedni za svaki $i = 1, \ldots, k$. *Duljina šetnje* je broj bridova u nizu. *Put* je šetnja kojoj su vrhovi i bridovi medusobno različiti, osim možda v_0 i v_k. Ako je $v_0 = v_k$, kažemo da je put *zatvoren* i nazivamo ga *ciklusom* u grafu G. Graf $G = (V, E)$ je *povezan* ako za svaka dva vrha $x, y \in V$ postoji šetnja koja započinje u vrhu x i završava u vrhu y. *Komponenta povezanosti* grafa je podgraf u kojem su svaka dva vrha povezana nekim putem i koji nije povezan s preostalim vrhovima u grafu.

Definicija 2.2. Stablo *je graf koji je povezan i nema ciklusa.*

Razapinjuće stablo grafa je razapinjući podgraf koji je ujedno i stablo. *Korijensko stablo* je stablo s jednim istaknutim vrhom koji nazivamo *korijenom. Šuma* je graf kojem su komponente povezanosti stabla.

Lema 2.3. *Stablo s* $n \geq 2$ *vrhova ima bar dva lista.*

Dokaz. Neka je $P = (v_1, \ldots, v_m)$ najduži put u stablu. (Ako postoji više najdužih, odaberemo jedan od njih.) Tvrdimo da su v_1 i v_m listovi. Pretpostavimo da v_1 nije list, tj. da uz v_2 ima bar još jedan susjedan vrh w. On se ne pojavljuje u P jer bi inače postojao ciklus. Dakle, postoji $P' = (w, v_1, \ldots, v_n)$ koji je duži od P, što je kontradikcija. □

Propozicija 2.4. *Povezan graf s n vrhova ima točno $n-1$ bridova ako i samo ako je stablo.*

Dokaz. Dokazat ćemo matematičkom indukcijom da stablo s n vrhova ima $n-1$ brid. Za $n = 1$ očito imamo nula bridova. Pretpostavimo da tvrdnja vrijedi za neki k, tj. da stablo s k vrhova ima točno $k-1$ brid. Neka je v list u stablu S s $k+1$ vrhova koji postoji zbog leme 2.3. Stablo S bez vrha v je povezan graf jer je v bio list. Također, ne sadrži cikluse jer ih nije sadržavao ni S. Dakle, S bez v je stablo, pa po pretpostavci indukcije ima $k-1$ bridova. Slijedi da S ima k bridova.

Obratno, neka je G povezan graf s n vrhova i $n-1$ bridom. Pretpostavimo da G nije stablo, tj. da ima c ciklusa, $c \geq 1$. Uklonimo sve cikluse tako da uklonimo c bridova, po jedan iz svakog od ciklusa. Dobiveni graf je povezan jer uklanjanjem brida $\{v_1, v_2\}$ iz ciklusa, v_1 i v_2 i dalje su povezani preko preostalih vrhova iz ciklusa. Dakle, dobili smo povezan graf koji nema cikluse, tj. stablo, s n vrhova i $n-1-c$ bridova. Već smo dokazali da stablo s n vrhova ima $n-1$ brid. Kako je $n-1-c < n-1$, došli smo do kontradikcije. Dakle, G je stablo. □

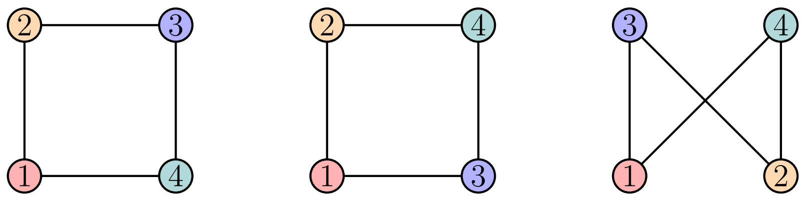
Propozicija 2.5. *Graf je stablo ako i samo ako za svaka dva vrha x i y postoji jedinstveni put od x do y.*

Dokaz. Pretpostavimo da u stablu postoje vrhovi x i y takvi da od x do y postoje dva različita puta $P_1 = (x, v_1, \ldots, v_n, y)$ i $P_2 = (x, w_1, \ldots, w_m, y)$. Tada bi postojao ciklus $(x, v_1, \ldots, v_n, y, w_m, \ldots, w_1, x)$ u stablu, što je kontradikcija.

Pretpostavimo da u grafu G za svaka dva vrha postoji jedinstveni put između njih. Očito je G povezan, pa još treba samo pokazati da nema cikluse. Pretpostavimo da u G postoji ciklus i označimo vrhove u tom ciklusu redom s $v_1, \ldots, v_m$. Tada npr. između vrhova v_1 i v_2 postoje dva puta: $P_1 = (v_1, v_2)$ i $P_2 = (v_1, v_m, v_{m-1}, \ldots, v_2)$, što je kontradikcija.

□

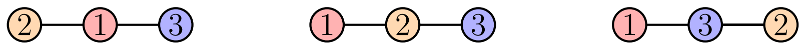
Grafovi $G_1 = (V_1, E_1)$ i $G_2 = (V_2, E_2)$ su *izomorfni* ako postoji bijekcija $\theta : V_1 \to V_2$ takva da je $E_2 = \theta(E_1) = \big\{\{\theta(x), \theta(y)\} \mid \{x, y\} \in E\big\}$. Izomorfni grafovi možda izgledaju različito, ali smatramo ih "jednakima", tj. izomorfnima jer među njima postoji bijekcija koja čuva bridove - svaki od vrhova zadržava iste susjede. Primjerice, sljedeća tri grafa su izomorfna:



U nastavku ćemo za skup vrhova koristiti skup prvih n prirodnih brojeva $\{1, 2, \ldots, n\}$. Odredimo broj stabala i broj klasa izomorfizma na manjim primjerima, konkretno za $n = 1, 2, 3, 4$. Kao u gornjem primjeru za grafove općenito, broj klasa izomorfizma će, naravno, biti manji ili jednak broju stabala.

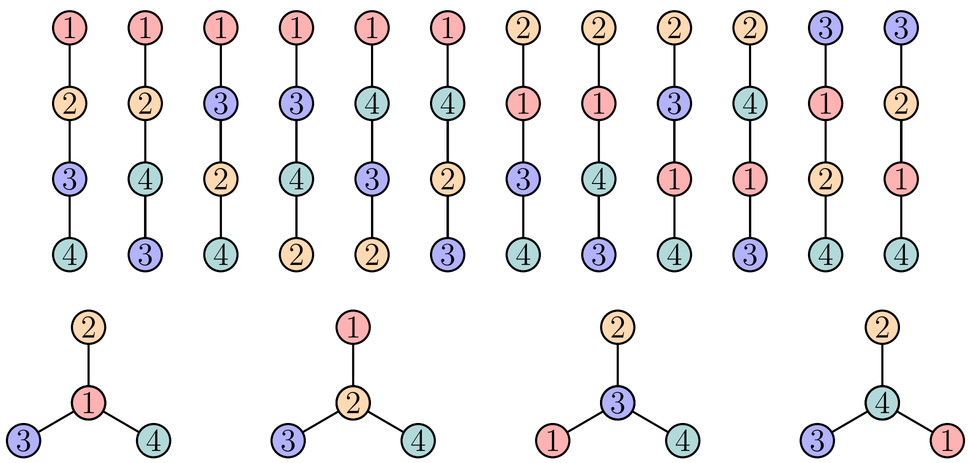
Promotrimo najprije slučaj $n = 1$. Tada očito postoji samo jedno stablo, i to trivijalno, pa je jasno da postoji i samo jedna klasa izomorfizma. Isto vrijedi i za $n = 2$.

Stvari postaju mrvicu kompliciranije za $n = 3$. Ispišimo sva stabla:



Stabala ima tri, no sva tri su međusobno izomorfna. Primjerice, bijekcija θ iz definicije izomorfizma koja slika bridove lijevog stabla u središnje stablo glasi $\theta(1) = 2$, $\theta(2) = 1$, $\theta(3) = 3$. Dakle, postoji samo jedna klasa izomorfizma.

Za kraj, proučimo što se događa za $n = 4$. Tu imamo nešto više stabala, njih 16, što nam daje naslutiti da broj stabala brzo raste u ovisnosti o n:



No, samo su dvije klase izomorfizma. Svi su grafovi u prvom retku izomorfni. Primjerice, bijekcija prvog i drugog grafa slijeva bila bi $\theta(1) = 1$, $\theta(2) = 2$, $\theta(3) = 4$, $\theta(4) = 3$. Isto vrijedi za sve grafove u drugom retku, a moguća bijekcija npr. prvog i drugog grafa slijeva bila bi $\theta(1) = 2$, $\theta(2) = 1$, $\theta(3) = 3$ i $\theta(4) = 4$. Da ne postoji bijekcija između ove dvije klase lako se vidi iz toga da ne možemo očuvati bridove koji izlaze iz vrha stupnja 3.

Prirodno se postavlja pitanje možemo li dobiti broj stabala i broj klasa izomorfizma u ovisnosti o broju vrhova n. Prebrojavanje stabala do na izomorfizam težak je problem koji ne proučavamo u ovom radu. No, problem određivanja broja različitih stabala s n vrhova ima relativno jednostavan odgovor koji je danas poznat kao Cayleyeva formula.

Teorem 2.6 (Cayleyeva formula). *Broj različitih stabala s n vrhova je n^{n-2}.*

Označimo broj različitih stabala s n vrhova s t_n. Cayleyeva formula tada glasi $t_n = n^{n-2}$. Formulu je prvi dokazao Borchardt 1860. godine pomoću determinante u članku [2], a sličan rezultat bio je poznat i Sylvestru nekoliko godina ranije, također u kontekstu determinanti. Cayleyev dokaz objavljen je 1889. u članku [4]. Iako je Cayleyev rezultat ekvivalentan Borchardtovom i Cayley se referira na njega u svom radu, Cayley ga je prvi izrazio u kontekstu teorije grafova, a njemu se pripisuje i da je prvi uveo pojam stabla. Od tada se ovaj rezultat asocira s Cayleyem.

Kako bismo demonstrirali kako izgleda Cayleyev dokaz, nastavno na već obrađene slučajeve $n = 1, 2, 3, 4$ prebrojimo stabla s 5 vrhova na način kako je Cayley opisao u svom dokazu. Promotrimo sljedeći polinom u 5 varijabli:

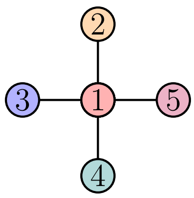
$$(a + b + c + d + e)^3 abcde.$$

Razvoj ovog polinoma ima $5^3 = 125$ monoma. Svaki od njih oblika je $\alpha^{\alpha'}\beta^{\beta'}\gamma^{\gamma'}\delta^{\delta'}\zeta^{\zeta'}$, gdje svaka od varijabli $\alpha, \beta, \gamma, \delta, \zeta$ odgovara točno jednoj od varijabli a, b, c, d, e, a $\alpha', \beta', \gamma', \delta', \zeta'$ su odgovarajuće potencije koje poprimaju vrijednosti od 1 do 4. Zapravo je multiskup eksponenata svakog pojedinog monoma jednak nekoj particiji broja 8 duljine 5: zbrojevi eksponenata u monomima iz razvoja $(a+b+c+d+e)^3$ su po multinomnom teoremu jednaki 3, a zbog množenja s $abcde$ rezultat je 8. U razvoju polinoma postoji $\binom{5}{1} = 5$ monoma oblika $\alpha^4\beta\gamma\delta\zeta$ od kojih se svaki pojavljuje točno jednom. Nadalje, postoji $2 \cdot \binom{5}{2} = 20$ monoma oblika $\alpha^2\beta^3\gamma\delta\zeta$ od kojih se svaki pojavljuje triput, te $\binom{5}{3} = 10$ monoma oblika $\alpha^2\beta^2\gamma^2\delta\zeta$ od kojih se svaki pojavljuje 6 puta:

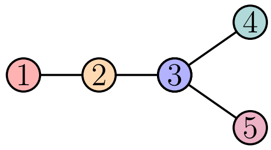
$$(a+b+c+d+e)^3 abcde = \left\{ \begin{array}{ll} \alpha^4\beta\gamma\delta\zeta & \ldots 5 \text{ puta}, \\ +\alpha^2\beta^3\gamma\delta\zeta & \ldots 60 \text{ puta}, \\ +\alpha^2\beta^2\gamma^2\delta\zeta & \ldots 60 \text{ puta} \end{array} \right\} \Rightarrow \text{ukupno 125 monoma.}$$

Postoje tri klase izomorfizma stabala s 5 vrhova. Cayley ih je pridružio gornjim trima oblicima tako da je skup potencija jednak skupu stupnjeva vrhova. Iako općenito nije slučaj da su stabla s istim multiskupom stupnjeva izomorfna, to vrijedi za $n = 5$ pa je pridruživanje jednostavno:

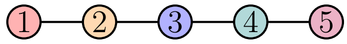
$\alpha^4\beta\gamma\delta\zeta$



$\alpha^2\beta^3\gamma\delta\zeta$



$\alpha^2\beta^2\gamma^2\delta\zeta$



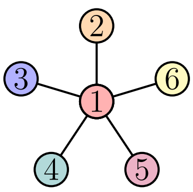
Primjerice, prvi graf slijeva ima jedan vrh stupnja 4 i četiri vrha stupnja 1, pa odgovara obliku $\alpha^4\beta\gamma\delta\zeta$. Zaista, postoji 5 stabala pridruženih obliku $\alpha^4\beta\gamma\delta\zeta$, ovisno koji vrh je stupnja 4. Nadalje, postoji $\binom{5}{2} \cdot 3! = 60$ stabala pridruženih obliku $\alpha^2\beta^2\gamma^2\delta\zeta$, jer prvo biramo dva lista povezana s vrhom stupnja 3, a zatim poredamo preostala tri vrha. Konačno, postoji $\frac{5!}{2} = 60$ stabala pridruženih obliku $\alpha^2\beta^2\gamma^2\delta\zeta$. Na taj način prebrojali smo 125 stabala s 5 vrhova.

Cayley je u svom dokazu gornji postupak proveo za $n = 6$. Polinom

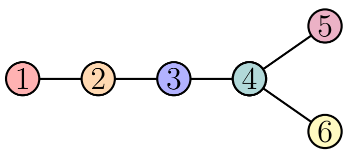
$$(a + b + c + d + e + f)^4 abcdef$$

ima $6^4 = 1296$ monoma u razvoju. Za $n = 6$ jedan od oblika monoma generira grafove iz dviju različitih klasa izomorfizma pa prebrojavanje zahtijeva nešto više opreza. (To možemo zaključiti i jer je broj klasa izomorfizma 6, dok je broj particija od 10 duljine 6 jednak 5.)

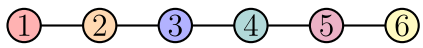
$\alpha^5\beta\gamma\delta\zeta\eta$



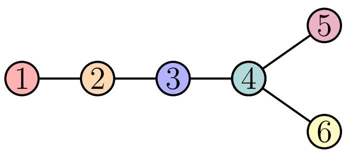
$\alpha^2\beta^2\gamma^3\delta\zeta\eta$



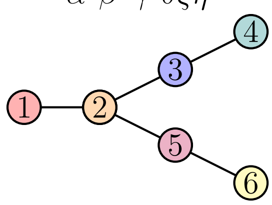
$\alpha^2\beta^2\gamma^2\delta^2\zeta\eta$



$\alpha^2\beta^2\gamma^3\delta\zeta\eta$



$\alpha^2\beta^2\gamma^3\delta\zeta\eta$



Konkretno, dvije klase izomorfizma u drugom retku reprezentirane su istim oblikom monoma. Pazeći na to, slično opisanom postupku za $n = 5$, Cayley je prebrojao 1296 stabala. Tvrdio je da analogna bijekcija postoji između stabala s n vrhova i n^{n-2} monoma razvoja $(x_1 + \ldots + x_n)^{n-2}x_1 \ldots x_n$, za sve $n \in \mathbb{N}$, no nije to dokazao. Proveo je dokaz samo za $n = 6$ i rekao da se dokaz može lagano poopćiti.

Iako sam Cayley nije dao dokaz u općem slučaju, Cayleyeva formula može se dokazati na brojne načine koristeći alate iz kombinatorike i drugih grana matematike. U sljedećim poglavljima krećemo na putovanje kroz razna područja matematike s jednim te istim ciljem: dokazivanje Cayleyeve formule. Sagledat ćemo ovaj problem s različitih točaka gledišta brojnih matematičara kroz povijest. Odabrani dokazi bit će sistematično obrađeni, a grupirani su u poglavlja po ključnim idejama.

3 Bijektivni dokazi

Funkcija $f : X \to Y$ je *bijekcija* ako za svaki $y \in Y$ postoji točno jedan $x \in X$ takav da $f(x) = y$, tj. ako je f i injekcija i surjekcija. Bijektivni dokazi su dokazi u kojima uspostavljamo bijekciju između dvaju skupova. Na taj način možemo odrediti broj elemenata skupa od interesa: definiramo bijekciju između tog skupa i skupa čije elemente znamo prebrojati. To je jedan od najčešćih načina rješavanja problema prebrojavanja u kombinatorici kada je teško direktno prebrojati elemente skupa koji nas zanima (npr. skup stabala s n vrhova), jer definiranjem prave bijekcije svodimo problem na prebrojavanje puno jednostavnijeg skupa.

Inverzna funkcija funkcije $f : X \to Y$ je funkcija $g : Y \to X$ za koju vrijedi $f \circ g = id_Y$ i $g \circ f = id_X$. U bijektivnim dokazima često je korisna sljedeća karakterizacija bijekcije pomoću inverzne funkcije.

Teorem 3.1. *Za funkciju $f : X \to Y$ postoji inverzna funkcija ako i samo ako je f bijekcija.*

Dokaz. Neka postoji inverzna funkcija $g : Y \to X$ od f. Tada za proizvoljni $y \in Y$ postoji $x = g(y)$ takav da je $f(x) = f(g(y)) = y$, pa je f surjekcija. Neka su x_1, $x_2 \in X$. Vrijedi $f(x_1) = f(x_2) \Rightarrow g(f(x_1)) = g(f(x_2)) \Rightarrow x_1 = x_2$, pa je f i injekcija.

Obratno, neka je f bijekcija. Definirajmo funkciju $g : Y \to X$ tako da svakom $y \in Y$ pridružimo jedinstveni $x \in X$ za koji vrijedi $f(x) = y$. Ona je dobro definirana jer je f bijekcija. Vrijedi $g(f(x)) = g(y) = x$, za sve $x \in X$, te $f(g(y)) = f(x) = y$, za sve $y \in Y$. Dakle, g je inverzna funkcija od f. $\square$

3.1 Prüferov dokaz

Prüferov kod (ili *Prüferov niz*), za prirodan broj $n \geq 2$, je uređena $(n-2)$–torka brojeva iz skupa $\{1, ..., n\}$ (uz dozvoljena ponavljanja). Jasno je da Prüferovih kodova duljine $n - 2$ ima n^{n-2}. Zbog toga Cayleyevu formulu možemo dokazati uspostavljanjem bijekcije između Prüferovih kodova duljine $n - 2$ i stabala s n vrhova. Na taj način Cayleyevu formulu dokazao je Heinz Prüfer 1918. u članku [14].

Za ovaj dokaz trebat će nam dva algoritma: Prüferov algoritam za kodiranje i za dekodiranje.

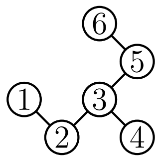
Algoritam 3.2 (Prüferov algoritam za kodiranje)**.** *Neka je zadano stablo S sa skupom vrhova $V = \{1, ..., n\}$.*

Postavimo $i = 1$.

Dok je $i \leq n - 2$, ponavljamo:

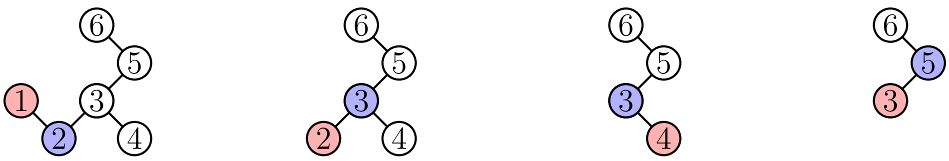
- *nađemo list v s najmanjom oznakom;*
- *stavimo oznaku jedinstvenog susjeda od v na i-to mjesto $(n-2)$-torke;*
- *izbacimo v iz grafa;*
- *povećamo i za 1.*

Primjer 3.3. *Odredimo Prüferov kod određen stablom sa slike.*



Prateći gore opisan algoritam, slijedi:

$i = 1$ $\quad$ $i = 2$ $\quad$ $i = 3$ $\quad$ $i = 4$



(2) $\quad$ $(2, 3)$ $\quad$ $(2, 3, 3)$ $\quad$ $(2, 3, 3, 5)$

Dakle, pripadni Prüferov kod je $(2, 3, 3, 5)$.

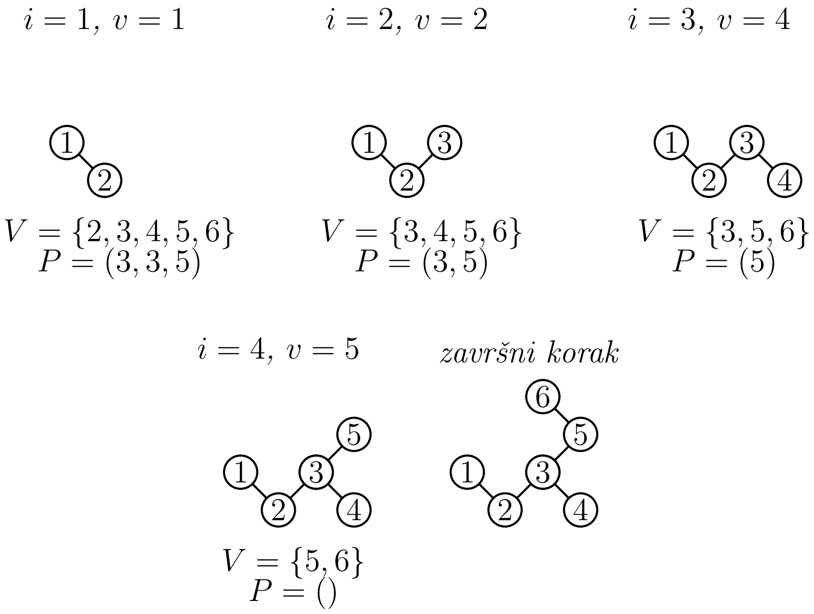
Algoritam 3.4 (Prüferov algoritam za dekodiranje)**.** *Neka je zadan graf s vrhovima* $V = \{1, ..., n\}$*, bez bridova, te neka je zadan Prüferov kod* $P = (p_1, p_2, \ldots, p_{n-2})$*. Postavimo* $i = 1$.

Dok je $i \leq n - 2$*, ponavljamo:*

- *nađemo najmanji element v skupa V koji nije u P;*
- *povežemo v i p_i;*
- *izbacimo v iz V i p_i iz P;*
- *povećamo i za 1.*

Na kraju povežemo preostala dva vrha iz V.

Primjer 3.5. *Nacrtajmo stablo određeno Prüferovim kodom* $P = (2, 3, 3, 5)$*. Na početku je* $V = \{1, 2, 3, 4, 5, 6\}$*. Pratimo gornji algoritam:*



Time smo dobili traženo stablo.

Uočimo da su se u prošla dva primjera algoritam za kodiranje i dekodiranje pokazali inverznima. To vrijedi i općenito, a zbog teorema 3.1 ekvivalentno je s time da se radi o bijekciji. Dokaz bijektivnosti zapravo je dokaz Cayleyeve formule.

Dokaz. Označimo preslikavanje sa skupa stabala na skup Prüferovih kodova iz algoritma za kodiranje s f. Očito je f dobro definirana funkcija. Dokažimo da je bijekcija, tj. da za svaku $(n-2)$-torku $(a_1, \ldots, a_{n-2})$ postoji jedinstveno stablo S takvo da je $f(S) = (a_1, \ldots, a_{n-2})$.

Koristimo jaku indukciju. Tvrdnja vrijedi za $n = 2$ jer postoji jedinstveno stablo s 2 vrha određeno praznim Prüferovim kodom. Neka je $n \geq 3$. Pretpostavimo da tvrdnja vrijedi za $1, \ldots, n-1$. Neka je $(a_1, \ldots, a_{n-2})$ $(n-2)$-torka s elementima iz skupa $\{1, \ldots, n\}$. Treba dokazati da određuje jedinstveno stablo.

Pretpostavimo da je $f(T) = (a_1, \ldots, a_{n-2})$ za neko stablo T s n vrhova. Tada su vrhovi $\{a_1, \ldots, a_{n-2}\}$ upravo oni vrhovi koji nisu listovi u T. Naime, da bi list v bio u $f(T)$, trebalo bi prvo obrisati njegovog susjeda, čime bi v postao izolirani vrh, što nije moguće. Također, ako vrh v nije list, tada će jedan od njegovih susjeda biti izbrisan (v ne može biti izbrisan prije toga) i v će biti dodan u Prüferov kod, pa je v u $f(T)$.

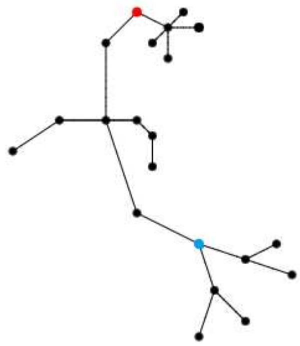
Slijedi da je oznaka prvog lista uklonjenog iz T minimalan element skupa $\{1, \ldots, n\} \setminus \{a_1, \ldots, a_{n-2}\}$. Označimo ga s m. Dakle, za svako stablo T

takvo da $f(T) = (a_1, \ldots, a_{n-2})$, m je list čiji je susjed a_1. Po pretpostavci indukcije postoji jedinstveno stablo U s vrhovima $\{1, \ldots, n\} \setminus \{m\}$ takvo da je $f(U) = (a_2, \ldots, a_{n-2})$. Dodavanjem vrha m i brida $\{a_1, m\}$ u U dobijemo jedinstveno stablo T takvo da je $f(T) = (a_1, \ldots, a_{n-2})$.

Dakle, definirali smo bijekciju između Prüferovih kodova duljine $n-2$ i označenih stabala s n vrhova, pa označenih stabala s n vrhova ima n^{n-2}. $\square$

3.2 Joyalov dokaz

André Joyal je 1981. godine u članku [9] dokazao Cayleyevu formulu uvodeći pojam *kralježnjaka* - stabla s dva istaknuta vrha (ne nužno različita). Istaknute vrhove nazivamo *glavom* i *repom* kralježnjaka, dok jedinstveni put od glave do repa nazivamo njegovom *kralježnicom*. U skladu s time, vrhove koji su elementi kralježnice nazivamo *kralješcima*. Preostale dijelove stabla možemo zamisliti kao udove kralježnjaka koji izlaze iz kralježnice.



Slika 1: Sredozemna medvjedica *(Monachus monachus)* u zaljevu Gökova u Turskoj. Možemo zamisliti grafa kralježnjaka s istaknutim vrhovima u glavi i repu medvjedice te njene udove koji izlaze iz kralježnice u obliku prsnih i repnih peraja - možda čak i brkova. Preuzeto s `https://www.alamy.com/stock-photo/mediterranean-monk-seal.html`.

Ovi slikoviti nazivi pomažu u vizualizaciji Joyalove bijekcije između kralježnjaka s n vrhova i funkcija $f : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$, tj. *endofunkcija*. Još jedan pojam koji ćemo koristiti je *periodična točka*. Periodična točka funkcije $f : X \to X$ je $x \in X$ za koju postoji $p \in \mathbb{N}$, tzv. *period*, takav da je $f^p(x) = x$. Potenciranje ovdje znači p uzastopnih kompozicija funkcije f sa samom sobom.

Cayleyevu formulu dokazat ćemo tako da ćemo prvo definirati funkciju J sa skupa kralježnjaka u skup endofunkcija, a zatim funkciju K sa skupa endofunkcija u skup kralježnjaka. Obje ćemo ilustrirati na primjerima. J je upravo Joyalova funkcija za koju ćemo dokazati da je bijekcija uz pomoć teorema 3.1: dokazat ćemo da je K njen inverz, tj. da vrijedi $K = J^{-1}$.

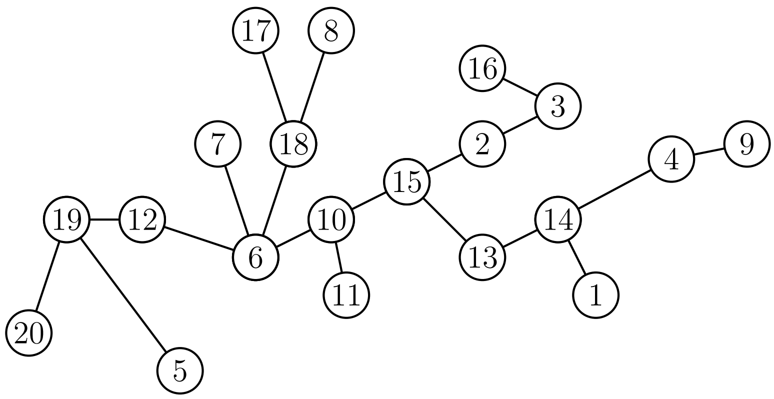
Funkcija J:

Definirajmo funkciju J sa skupa kralježnjaka u skup endofunkcija. Neka je zadan kralježnjak s n vrhova od kojih su njih k kralješci. Pridružujemo mu endofunkciju $f : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ tako da:

- označimo kralješke redom od glave do repa s $l_1, l_2, \ldots, l_k$, a zatim ih sortiramo uzlazno i označimo taj poredak s $m_1, m_2, \ldots, m_k$,
- definiramo f na kralješcima tako da $f(m_i) := l_i, i = 1, 2, \ldots, k$,
- preostale vrijednosti od f definiramo tako da svakom vrhu pridružimo prvi sljedeći vrh koji se nalazi na njegovom najkraćem putu do kralježnice.

Treći korak opravdan je jer preostali vrhovi zbog propozicije 2.5 imaju jedinstveni put do kralježnice. Dakle, J je dobro definirana.

Primjer 3.6. *Definirajmo endofunkciju određenu sljedećim grafom kralježnjakom. Neka je vrh* 12 *glava, a* 9 *rep.*



Kralježnicu čine redom vrhovi $12, 6, 10, 15, 13, 14, 4, 9$. *Uzlazni poredak je* $4, 6, 9, 10, 12, 13, 14, 15$ *pa definiramo* $f(4) = 12$, $f(6) = 6$, $f(9) = 10$, $f(10) = 12$, $f(12) = 13, f(13) = 14$, $f(14) = 4$, $f(15) = 10$. *Preostalim vrhovima pridružimo prvi sljedeći na putu do kralježnice: npr. vrhu* 17 *je put do kralježnice* $(17, 18, 6)$, *pa definiramo* $f(17) = 18$. *Ponovimo postupak za preostale vrhove i konačna funkcija je*

$$\begin{array}{cccccccccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 14 & 15 & 2 & 12 & 19 & 6 & 6 & 18 & 10 & 15 & 10 & 13 & 14 & 4 & 9 & 3 & 18 & 6 & 12 & 19 \end{array}$$

Kad bismo se htjeli iz zadane endofunkcije vratiti u polaznog kralježnjaka, najprije bismo trebali odrediti koje točke bi bile u kralježnici. Znamo da su sve točke osim onih u kralježnici nastale dodavanjem bridova $\{i, f(i)\}$ u graf. Kako na taj način nismo mogli dobiti ciklus za vrhove koji nisu u kralježnici, zaključujemo da su periodične točke endofunkcije upravo kralješci. Njihov poredak unutar kralježnice jednostavno utvrdimo promatranjem funkcijskih vrijednosti uzlazno sortiranih periodičnih točaka. Na taj način definiramo funkciju K.

Funkcija K:

Definirajmo sada funkciju K sa skupa endofunkcija u skup kralježnjaka. Neka je zadana endofunkcija $f : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$. Pridružujemo joj kralježnjaka s n vrhova tako da:

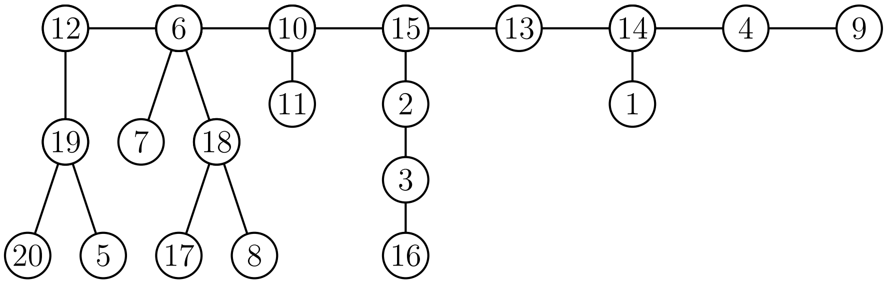
- označimo periodične točke od f u uzlaznom poretku s $m_1, \dots, m_k$,
- u nulgrafu N_n povežemo vrhove $m_1, \dots, m_k$ u kralježnicu tako da su vrhovi kralježice od glave do repa $f(m_1), f(m_2), \dots, f(m_k)$,
- u graf sa spojenom kralježnicom dodamo sve bridove određene s f, tj. bridove $\{i, f(i)\}$, za sve vrhove i koji nisu u kralježnici.

Opisanim povezivanjem vrhova u kralježnicu dobijemo povezan graf bez ciklusa, odnosno stablo. To svojstvo nije narušeno dodavanjem preostalih bridova jer preostale točke od f nisu periodične točke. Dakle, K je dobro definirana.

Primjer 3.7. *Nacrtajmo kralježnjaka određenog endofunkcijom*

$$\begin{array}{cccccccccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 14 & 15 & 2 & 12 & 19 & 6 & 6 & 18 & 10 & 15 & 10 & 13 & 14 & 4 & 9 & 3 & 18 & 6 & 12 & 19 \end{array}$$

Vrhovi u ciklusima sortirani uzlazno su $4, 6, 9, 10, 12, 13, 14, 15$*, a zbog* $f(4) = 12$*,* $f(6) = 6$*,* $f(9) = 10$*,* $f(10) = 12$*,* $f(12) = 13, f(13) = 14$*,* $f(14) = 4$*,* $f(15) = 10$*, kralježnicu redom od glave do repa čine vrhovi* $12, 6, 10, 15, 13, 14, 4, 9$*. Dodavanjem preostalih bridova rezultat je kralježnjak*



Kad bismo se htjeli iz kralježnjaka vratiti u polaznu endofunkciju, najprije bismo definirali periodične točke od f promatrajući kralješke. Preostale vrijednosti od f definiramo tako da svakoj pridružimo vrijednost susjednog vrha, jer je ostatak kralježnjaka nastao upravo dodavanjem bridova $\{\{i, f(i)\} \mid i = 1, 2, \dots, n\}$. Uočimo da smo ovime zapravo opisali prethodno definiranu funkciju J.

Dakle, funkcije su J i K inverzne ne samo na na gornjim primjerima, već i općenito, pa je time dobro definirana Joyalova bijekcija J grafova kralježnjaka i endofunkcija. Dokažimo to!

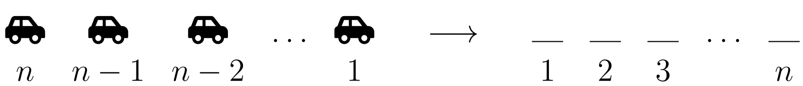
Dokaz. Zapišimo Cayleyevu formulu u obliku $n^2 \cdot t_n = n^n$ tako da desna strana bude broj endofunkcija $f : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$. Lijeva strana sada predstavlja broj stabala s n vrhova s dva istaknuta vrha, odnosno broj kralježnjaka s n vrhova. Definirajmo funkcije J i K kao gore. Iz definicija tih dviju funkcija jasno je da vrijedi $K = J^{-1}$. Dakle, J ima inverz pa je bijekcija po teoremu 3.1, odakle je $n^2 t_n = n^n$, tj. $t_n = n^{n-2}$.

□

Dokaz s jako sličnom idejom su 1986., pet godina nakon Joyalova, objavili Eğecioğlu i Remmel [7]. U njihovoj varijanti uspostavlja se bijekcija usmjerenih stabala s n vrhova i endofunkcija n-članog skupa s dvije fiksne točke. Fiksne točke endofunkcije uvelike odgovaraju ulozi dvaju istaknutih vrhova stabla i bijekcija se definira na jako sličan način kao u Joyalovom dokazu.

3.3 Funkcije parkiranja

Zamislimo n parkirnih mjesta s oznakama redom $1, \dots, n$ duž jednosmjerne ulice u koju ulazi kolona od n auta s oznakama redom od prvog do posljednjeg u koloni $1, \dots, n$.



Svaki od vozača tih auta redom bira parkirno mjesto te svaki ima svoje omiljeno parkirno mjesto. Ako je u trenutku dolaska vozača njegovo omiljeno mjesto slobodno, tu će se i parkirati, a ako nije, nastavlja voziti i parkira se na prvo sljedeće slobodno mjesto. Ako takvo ne postoji, vozač odustaje i odlazi. Neka je $A = (a_1, \dots, a_n)$ uređena n-torka koja svakom vozaču i pridužuje omiljeno parkirno mjesto a_i, $i = 1, \dots, n$. Ako se svi vozači uspiju parkirati uz gore navedene uvjete, kažemo da je A *funkcija parkiranja duljine* n.

Primjerice, $(1, 3, 2, 4, 4)$ je funkcija parkiranja duljine 5, jer iako vozač 5 ne može parkirati na svoje omiljeno mjesto, može parkirati na mjesto 5. Primjer uređene petorke koja nije funkcija parkiranja je $(1, 3, 5, 2, 5)$, jer se vozač 5 ne može parkirati (i mjesto 4 ostaje prazno). Jasno je da je n-torka funkcija parkiranja ako i samo ako sadržava bar jednu koordinatu manju ili jednaku 1, bar dvije manje ili jednake 2, ..., bar n manjih ili jednakih n. Zbog toga je jedan jednostavan način za provjeru sortirati koordinate uzlazno i provjeriti je li svaka od njih u tom poretku manja ili jednaka od svoje pozicije u tom poretku, tj. da za uzlazno sortirane koordinate u oznakama $b_1 \leq \ldots \leq b_n$ vrijedi $b_i \leq i$, za sve i.

Pojam funkcija parkiranja neovisno su uveli Pyke 1959. u članku [15] implicitno te Konheim i Weiss 1966. u članku [10] pri proučavanju hash-funkcija. Između ostalog, dokazali su da funkcija parkiranja duljine n ima $(n+1)^{n-1}$, odakle naslućujemo njihovu povezanost s brojem stabala. Od tada su funkcije parkiranja postale zanimljiv i dobro istražen matematički objekt, a mi ćemo proučiti kako pomoću njih dokazati Cayleyevu formulu. Sljedeći elegantni dokaz za broj funkcija parkiranja s n parkirnih mjesta dao je Pollak 1974., objavljen u [18]:

Teorem 3.8. *Funkcija parkiranja duljine n ima $(n+1)^{n-1}$.*

Dokaz. Uvedimo dodatno parkirno mjesto $n+1$ na kraj ulice i zatim početak i kraj ulice spojimo u krug tako da je mjesto 1 nakon mjesta $n+1$ u poretku. Dopuštamo i da je $n+1$ omiljeno mjesto. Kako je ulica kružna, svi će se uspjeti parkirati i točno jedno mjesto ostat će prazno. Funkcije parkiranja sada su one n-torke za koje parkirno mjesto $n+1$ ostaje prazno nakon što sve svi parkiraju.

Neka je zadana n-torka $A_0 = (a_1, \dots, a_n)$. Promotrimo n-torke $A_j = (a_1 + j, \dots, a_n + j) \pmod{n+1}$ za $j = 1, \dots, n$. One su dobivene rotacijom rasporeda parkiranja A_0 za j mjesta. Točno jedna od n-torki $A_0, A_1, \dots, A_n$ je funkcija parkiranja jer je u točno jednoj od mogućih rotacija prazno mjesto baš $n+1$.

Kako tih rotacija ima $n+1$, a mogućih polaznih n-torki ima $(n+1)^n$, broj funkcija parkiranja duljine n jednak je

$$\frac{(n+1)^n}{n+1} = (n+1)^{n-1}.$$

□

Sad možemo dokazati Cayleyevu formulu uspostavljanjem bijekcije između funkcija parkiranja i korijenskih stabala. Postoji više takvih bijekcija, a ovdje izlažemo onu koju su definirali Chassaing i Marckert 2001. u članku [5]. Temelji se poznatom algoritmu pretraživanja stabla po širini (eng. breadth-first search). U tom algoritmu počinjemo u korijenu i dodamo oznaku korijena u red (eng. queue). Zatim posjetimo redom slijeva na desno djecu korijena, tj. vrhove na razini 1, te i njih dodamo u red. Slijede djeca djece korijena, tj. vrhovi na razini 2, te nastavljamo dalje po razinama dok ne posjetimo sve vrhove. Općenito, u svakom koraku posjećujemo djecu prvog sljedećeg neposjećenog elementa u redu te ih dodajemo u red kako bismo i njih mogli posjetiti kasnije. Na taj način pretražili smo stablo po razinama, tj. po širini.

Pridružujemo korijensko stablo S s $n+1$ vrhova funkciji parkiranja $A = (a_1, \ldots, a_n)$ duljine n tako da konstruiramo stablo po širini algoritmom koji je jako sličan gore opisanom algoritmu za pretraživanje stabla po širini:

- definiramo skup A_i kao skup oznaka auta kojima je i omiljeno mjesto, za sve $i = 1, \ldots, n$,
- dodamo korijen 0 u S i dodamo ga u red $Q = (0)$,
- redom prolazimo skupove $A_1, \ldots, A_n$ na način da elemente skupa A_i sortiramo uzlazno i dodamo tim redom kao djecu vrha koji je prvi neposjećeni u Q, a zatim elemente od A_i dodamo na kraj reda Q.

Rezultat ovog postupka očito je stablo. Dobro je definirano jer koristimo algoritam pretraživanja tj. konstrukcije stabla po širini koji posjećuje svaki od vrhova i to točno jednom.

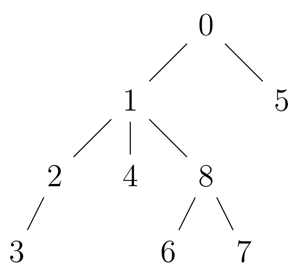
S druge strane, ako je zadano korijensko stablo s $n+1$ vrhom i korijenom 0, pridružujemo mu pripadnu funkciju parkiranja $(a_1, \ldots, a_n)$ duljine n tako da:

- promatramo stablo kao uređeno stablo i unutar svake od razina sortiramo djecu uzlazno,
- pretražujemo stablo po širini i redom definiramo skupove $A_1, \ldots, A_n$ kao redom skupove djece vrha kojeg smo posjetili i-tog po redu,
- definiramo svaku od koordinata a_j, $j = 1, \ldots, n$ u A kao indeks i onog skupa A_i koji sadrži vrijednost j.

Prvi korak potreban je kako bismo osigurali da svako stablo generira točno jednu funkciju parkiranja. Naime, promatramo li stablo kao uređeno bez da prvo sortiramo vrhove, mogli bismo iz jednog te istog stabla dobiti više različitih funkcija parkiranja ovisno o poretku vrhova. Posljednji korak dobro je definiran jer se svaka od oznaka nalazi točno u jednom A_i - svaki vrh u stablu ima točno jednog roditelja (osim korijena 0 koji svakako nije u konačnoj funkciji parkiranja). Da bi rezultat bila funkcija parkiranja trebamo $\sum_{j=0}^{i} |A_j| \geq i$, za sve i. To vrijedi jer u protivnom ne bismo imali stablo, tj. neki vrhovi ne bi imali roditelje jer bi neke razine ostale prazne.

Jasno je da su na ovaj način opisana dva inverzna postupka, što ćemo demonstrirati na primjeru.

Primjer 3.9. *Neka je zadana funkcija parkiranja* $A = (1, 2, 4, 2, 1, 6, 6, 2)$ *duljine* 8*. Pripadni skupovi iz gornjeg algoritma tada su jednaki* $A_1 = \{1, 5\}$, $A_2 = \{2, 4, 8\}$, $A_3 = \emptyset$, $A_4 = \{3\}$, $A_5 = \emptyset$, $A_6 = \{6, 7\}$, $A_7 = \emptyset$, $A_8 = \emptyset$ *i daju sljedeće korijensko stablo s* 9 *vrhova:*



S druge strane, ako nam je zadano gornje stablo, lako iščitavamo da su pripadni skupovi opet jednaki $A_1 = \{1, 5\}$, $A_2 = \{2, 4, 8\}$, $A_3 = \emptyset$, $A_4 = \{3\}$, $A_5 = \emptyset$, $A_6 = \{6, 7\}$, $A_7 = \emptyset$, $A_8 = \emptyset$, *odakle slijedi da je, kao i prije,*

$$A = (1, 2, 4, 2, 1, 6, 6, 2).$$

S obzirom da smo po teoremu 3.1 ovime definirali bijekciju između skupa stabala s n vrhova i funkcija parkiranja duljine $n - 1$ kojih po teoremu 3.8 ima $(n - 1 + 1)^{n-1-1} = n^{n-2}$, dokazana je Cayleyeva formula.

4 Induktivni dokazi

Princip matematičke indukcije jedan je od najčešćih alata pri dokazivanju tvrdnji koje vrijede za sve prirodne brojeve, pa nije neočekivano da se pojavljuje i u mnogim dokazima Cayleyeve formule, bilo kao glavna ideja dokaza ili kao međukorak u nekoj drugoj metodi dokazivanja. U ovom poglavlju iznosimo dokaze gdje je matematička indukcija temeljna ideja.

4.1 Formula uključivanja-isključivanja

Neka je dana konačna familija konačnih skupova $\{A_1, A_2, \ldots, A_n\}$. Tada broj elemenata njihove unije računamo pomoću tzv. *formule uključivanja - isključivanja*:

$$|A_1 \cup \cdots \cup A_n| = \sum_i |A_i| - \sum_{i<j} |A_i \cap A_j| + \sum_{i<j<k} |A_i \cap A_j \cap A_k| - \ldots + \\ + (-1)^{n-1} |A_1 \cap \cdots \cap A_n|.$$

Ekvivalentno, uz oznaku $A_I = \cap_{i \in I} A_i$,

$$\left| \bigcap_{i=1}^{n} A_i^c \right| = \sum_{I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|} |A_I|. \tag{1}$$

Cayleyevu formulu možemo dokazati induktivno tako da pomoću formule uključivanja-isključivanja povežemo t_n s brojem surjekcija s n-članog na k-člani skup, kako je obrađeno u skripti [11].

Lema 4.1. *Broj surjekcija s n-članog na k-člani skup jednak je*

$$\sum_{i=0}^{k} (-1)^i \binom{k}{i} (k-i)^n. \tag{2}$$

Dokaz. Neka je X skup funkcija s $\{1, \ldots, n\}$ na $\{1, \ldots, k\}$, A_i skup takvih funkcija koje nemaju element i u kodomeni, te A_I skup takvih funkcija koje nemaju elemente iz skupa I u kodomeni. Vrijedi $|X| = k^n$, $|A_i| = (k-1)^n$ i $|A_I| = (k - |I|)^n$. Uočimo da je funkcija s $\{1, \ldots, n\}$ na $\{1, \ldots, k\}$ surjekcija ako i samo ako nije sadržana ni u jednom A_i, pa zapravo tražimo kardinalnost skupa $A_1^c \cap \ldots \cap A_k^c$ i možemo iskoristiti formulu (1):

$$\left| \bigcap_{i=1}^{k} A_i^c \right| = \sum_{I \subseteq \{1,2,\ldots,k\}} (-1)^{|I|} (k - |I|)^n = \sum_{i=0}^{k} (-1)^i \binom{k}{i} (k-i)^n.$$

Posljednja jednakost vrijedi jer postoji $\binom{k}{i}$ skupova $I \subseteq \{1, \ldots, k\}$ kardinalnosti i za $i = 0, \ldots, k$. $\square$

Dokažimo sada Cayleyevu formulu.

Dokaz. Za $n = 1, 2$, tvrdnja je očita, pa je u nastavku $n \geq 3$. Neka je A_i skup stabala s n vrhova kojima je vrh i list. Vrijedi $|\bigcap_{i=1}^{n} A_i^c| = 0$. Naime, kako A_i^c predstavlja skup stabala kojima vrh i nije list, presjek skupova A_i^c za sve $i = 1, \ldots, n$ je skup stabala s n vrhova koji nemaju nijedan list, a taj skup je prazan zbog leme 2.3. Želimo prebrojati $t_n = |A_1 \cup \cdots \cup A_n|$, pa koristimo formulu uključivanja-isključivanja:

$$0 = \left| \bigcap_{i=1}^{n} A_i^c \right| = \sum_{I \subset \{1,2,\ldots,n\}} (-1)^{|I|} |A_I| = t_n + \sum_{\emptyset \neq I \subset \{1,2,\ldots,n\}} (-1)^{|I|} |A_I|$$

$$\Rightarrow t_n = \sum_{\emptyset \neq I \subset \{1,2,\ldots,n\}} (-1)^{|I|-1} |A_I|.$$

Kako bismo dobili stablo s n vrhova kojem je i list, možemo stablu bez vrha i dodati list i na jedan od $n - 1$ vrhova. Dakle, $|A_i| = (n-1)t_{n-1}$. Slično, $|A_i \cap A_j| = (n-2)^2 t_{n-2}$, jer vrhove i i j možemo dodati neovisno i to svakog na $n-2$ načina. Primijenimo li ovo na općeniti skup vrhova I, vrijedi $|A_I| = (n - |I|)^{|I|} t_{n-|I|}$, gdje je A_I skup stabala kojima su listovi elementi od I. Možemo ih odabrati na $\binom{n}{i}$ načina. Sada uvrštavanjem imamo rekurziju

$$t_n = \sum_{i=1}^{n} (-1)^{i-1} \binom{n}{i} (n-i)^i t_{n-1}. \tag{3}$$

Iz (2) znamo da je broj surjekcija s $(n-2)$-članog na n-člani skup jednak $\sum_{i=0}^{n} (-1)^i \binom{n}{i} (n-i)^{n-2}$, što je naravno jednako 0. Odatle je

$$0 = \sum_{i=0}^{n} (-1)^i \binom{n}{i} (n-i)^{n-2} = n^{n-2} + \sum_{i=1}^{n} (-1)^i \binom{n}{i} (n-i)^{n-2}$$

$$\Rightarrow n^{n-2} = \sum_{i=1}^{n} (-1)^{i-1} \binom{n}{i} (n-i)^i (n-i)^{n-i-2}. \tag{4}$$

Sada smo spremni matematičkom indukcijom dokazati da je $t_n = n^{n-2}$. Bazu indukcije za $n = 3$ provjerili smo u poglavlju 2. Pretpostavimo da vrijedi $t_k = k^{k-2}$ za $k = 1, 2, \ldots, n-1$. Zbog toga možemo izjednačiti (3) i (4) (uz zamjenu u zapisu n s k) i dobiti $t_{k-i} = (k-i)^{k-i-2}$, $k = 1, 2, \ldots, n-1$,

$i = 1, 2, \ldots, k$. Uvrstimo li to u rekurziju za t_n u (3), s desne strane imamo upravo sumu $\sum_{i=1}^{n}(-1)^{i-1}\binom{n}{i}(n-i)^i(n-i)^{n-i+2}$ koja je zbog (4) jednaka n^{n-2}.

□

4.2 Riordanov i Rényijev dokaz

Ideje mađarskog matematičara Alfréda Rényija i američkog matematičara Johna Francisa Riordana u njihovim radovima [16] i [17] o prebrojavanju stabala iz 1960-ih godina daju nam još jedan dokaz Cayleyeve formule. Ovaj dokaz zasniva se na tome da proučimo nešto općenitiji slučaj i dokažemo Cayleyevu formulu za slučaj šume.

Neka je $t_{n,k}$ broj šuma s n vrhova sastavljenih od k stabala. (Jasno je da je $k \leq n$.) Tada vrijedi:

Teorem 4.2 (Cayleyeva formula za šume)**.** $t_{n,k} = kn^{n-k-1}$.

Provodimo dokaz kako je izložen u [1]. Uočimo da je $t_{n,1} = t_n$ jer je stablo zapravo šuma sastavljena od jednog stabla, pa će Cayleyeva formula slijediti kao poseban slučaj ovog teorema. Da bismo ga dokazali, najprije definiramo rekurzivnu relaciju za $t_{n,k}$, a tvrdnja teorema iz nje će slijediti indukcijom.

Definirajmo najprije početne vrijednosti. Neka je $t_{0,0} = 1$ jer postoji jedna prazna šuma. Nadalje, neka je $t_{n,0} = 0$ za $n \geq 1$ jer ne postoji neprazna šuma bez stabala.

Neka su sada bez smanjenja općenitosti vrhovi $\{v_1, \ldots, v_k\}$ u različitim stablima. Neka vrh v_1 ima i susjeda. Vrijedi $i \in \{0, 1, \ldots, n-k\}$ jer v_1 nije povezan s $v_2, \ldots, v_k$ ni sa samim sobom. Prebrojat ćemo sve šume s n vrhova i k stabala sumiranjem po $i = 0, 1, \ldots, n-k$.

Ako iz šume F s n vrhova i k stabala uklonimo vrh v_1, dobit ćemo šumu s $n-1$ vrhom $\{v_2, \ldots, v_n\}$ i $k-1+i$ stabala, jer je v_1 imao i susjeda. Da bismo iz takve šume $F \backslash v_1$ rekonstruirali F, najprije odaberemo $i \in \{0, 1, \ldots, n-k\}$. Zatim odaberemo i susjeda od v_1 na $\binom{n-k}{i}$ načina. Za kraj, sjetimo se da smo imali $t_{n-1,k-1+i}$ polaznih šuma $F \setminus v_1$.

Slijedi rekurzivna formula

$$t_{n,k} = \sum_{i=0}^{n-k} \binom{n-k}{i} t_{n-1,k-1+i}. \tag{5}$$

Dokažimo sada teorem 4.2.

Dokaz. Koristimo matematičku indukciju. Neka je $n = 1$. Promotrimo sve $k \leq n$. Vrijedi $t_{1,0} = 0$ (po definiciji) i $t_{1,1} = 1$ (postoji jedna šuma s jednim vrhom i jednim stablom), što zadovoljava teorem: $t_{1,0} = 0 = 0 \cdot 1^{1-0-1}$ i $t_{1,1} = 1 = 1 \cdot 1^{1-1-1}$.

Pretpostavimo da vrijedi $t_{n-1,k} = k(n-1)^{n-1-k-i}$, za neki $n \in \mathbb{N}$ i sve $k = 0, \ldots, n-1$. Dokažimo da vrijedi $t_{n,k} = kn^{n-k-1}$.

Zbog (5) i pretpostavke indukcije, vrijedi

$$\begin{aligned} t_{n,k} &= \sum_{i=0}^{n-k} \binom{n-k}{i} t_{n-1,k-1+i} \\ &= \sum_{i=0}^{n-k} \binom{n-k}{i} (k-1+i)(n-1)^{n-1-k-i}. \end{aligned}$$

Zamijenimo sada indeks i s $j = n-k-i$ i pojednostavnimo izraz koristeći svojstvo simetrije binomnih koeficijenata $\binom{a}{b} = \binom{a}{a-b}$, $a, b \in \mathbb{N}_0$:

$$\begin{aligned} t_{n,k} &= \sum_{j=n-k}^{0} \binom{n-k}{n-k-j} (k-1+n-k-j)(n-1)^{n-1-k-(n-k-i)} \\ &= \sum_{j=0}^{n-k} \binom{n-k}{j} (n-1-j)(n-1)^{j-1}. \end{aligned}$$

Rastavimo faktor $(n-1-j)$ da bismo dobili razliku dviju suma.

$$\begin{aligned} t_{n,k} &= \sum_{j=0}^{n-k} \binom{n-k}{j} (n-1)(n-1)^{j-1} - \sum_{j=0}^{n-k} \binom{n-k}{j} j(n-1)^{j-1} \\ &= \sum_{j=0}^{n-k} \binom{n-k}{j} (n-1)^{j} - \sum_{j=1}^{n-k} \binom{n-k}{j} j(n-1)^{j-1}. \end{aligned} \tag{6}$$

Pojednostavnimo ove dvije sume. U prvoj sumi to ćemo napraviti tako da iskoristimo binomni teorem:

$$\begin{aligned} \sum_{j=0}^{n-k} \binom{n-k}{j} (n-1)^{j} &= \sum_{j=0}^{n-k} \binom{n-k}{j} (n-1)^{j} \cdot 1^{n-k-j} \\ &= (n-1+1)^{n-k} \\ &= n^{n-k}. \end{aligned} \tag{7}$$

U drugoj sumi iskoristimo identitet apsorpcije za binomne koeficijente $\binom{a}{b} = \frac{a}{b}\binom{a-1}{b-1}$, $a, b \in \mathbb{N}$, zamijenimo indekse, te na kraju primijenimo binomni teorem kao i u prvoj sumi:

$$\begin{aligned}
\sum_{j=1}^{n-k} \binom{n-k}{j} j(n-1)^{j-1} &= \sum_{j=1}^{n-k} \binom{n-k-1}{j-1} \frac{n-k}{j} j(n-1)^{j-1} \\
&= (n-k) \sum_{j=1}^{n-k} \binom{n-k-1}{j-1} (n-1)^{j-1} \\
&= (n-k) \sum_{l=0}^{n-k-1} \binom{n-k-1}{l} (n-1)^{l} \\
&= (n-k)(n-1+1)^{n-k-1} \\
&= n^{n-k} - kn^{n-k-1}. \qquad (8)
\end{aligned}$$

Uvrstimo sada (7) i (8) u (6). Rezultat je

$$t_{n,k} = n^{n-k} - n^{n-k} + kn^{n-k-1} = kn^{n-k-1},$$

što je i trebalo dokazati. □

Za kraj, uvrštavanjem $k = 1$ u teorem 4.2 i zbog $t_n = t_{n,1}$ imamo

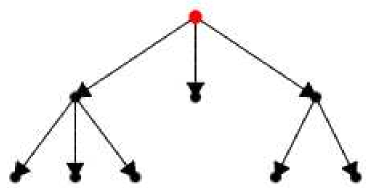
Korolar 4.3 (Cayleyeva formula). $t_n = n^{n-2}$.

5 Dokazi dvostrukim prebrojavanjem

Metoda dvostrukog prebrojavanja temelji se na tome da elemente jednog te istog skupa prebrojimo na dva načina. Dva dobivena izraza možemo izjednačiti i na taj način dobiti odgovor na razne kombinatorne probleme. U našem slučaju, na dva načina prebrojat ćemo neke veličine koje ovise o broju stabala i tako doći do Cayleyeve formule.

5.1 Pitmanov dokaz

Jim Pitman je 1999. u članku [12] objavio dokaz Cayleyeve formule dvostrukim prebrojavanjem. Za to je koristio usmjerena korijenska stabla u kojima su svi bridovi usmjereni u smjeru od korijena.



Slika 2: Primjer korijenskog usmjerenog stabla s bridovima usmjerenima u suprotnom smjeru od korijena

Prebrojavanjem redoslijeda dodavanja bridova u takvo stablo dolazimo do vjerojatno jednog od najjednostavnijih, ili bar najsažetijih, dokaza Cayleyeve formule.

Dokaz. Na dva ćemo načina prebrojiti broj nizova usmjerenih bridova koje dodajemo nulgrafu N_n da bismo dobili korijensko usmjereno stablo s n vrhova. Označimo taj broj s x. Počnimo od nulgrafa N_n.

Prvi način prebrojavanja je da počnemo od stabla s n vrhova na t_n načina. Korijen mu odaberemo na n načina, a time smo odredili i smjerove svih bridova jer svi moraju izlaziti iz korijena. Redom tih $n-1$ bridova možemo dodati na $(n-1)!$ načina. Dakle, $x = t_n \cdot n!$.

Drugi način je nulgrafu dodati $n-1$ bridova da dobijemo stablo. U svakom od $n-1$ koraka, neka je n_k broj usmjerenih bridova koje je moguće dodati u k-tom koraku. Svaki od njih određen je tako da nezavisno odaberemo početni i krajnji vrh. Očito je $n_1 = n(n-1)$. U k-tom koraku, početni vrh može biti bilo koji, ali krajnji mora biti neki od do tada neiskorištenih vrhova. Naime, ako bi pokazivao u neki iskorišteni vrh, tada bismo narušili svojstvo da svi bridovi pokazuju u smjeru suprotnom od korijena. Slijedi $n_k = n(n-k)$, pa je $x = n \cdot (n-1) \cdot n \cdot (n-2) \cdot \ldots 2 \cdot 1 = n^{n-2} \cdot n!$.

Izjednačavanjem ovih dvaju izraza za x i dijeljenjem s $n!$ slijedi $t_n = n^{n-2}$. $\square$

5.2 Razapinjuća stabla

Cameron i Kagan su u članku [3] dali dokaz Cayleyeve formule prebrojavanjem razapinjućih stabala grafa koji sadržavaju zadanu šumu. U svom dokazu koristili su Kirchhoffov teorem koji će biti obrađen u kasnijem poglavlju. Ovdje izlažemo elementarniji dokaz te tvrdnje pomoću dvostrukog prebrojavanja kao u [19].

Neka je T šuma s vrhovima $\{1, \ldots, n\}$ i komponentama $T_1, \ldots, T_k$. Dokazat ćemo da razapinjućih stabala na skupu vrhova $\{1, \ldots, n\}$ koja sadržavaju T ima

$$n^{k-2} \prod_{l=1}^{k} |T_l| \tag{9}$$

dvostrukim prebrojavanjem. Cayleyeva formula tada lagano slijedi za slučaj šume s n trivijalnih stabala: $n^{k-2} \prod_{l=1}^{k} |T_l| = n^{n-2} \prod_{l=1}^{n} 1 = n^{n-2}$.

Dokaz. Fiksirajmo n. Dokazat ćemo (9) indukcijom u čijem koraku koristimo dvostruko prebrojavanje. Za $k = 1$ tj. za stablo postoji samo jedno razapinjuće stablo koje ga sadržava. Kako je $n^{1-2} \prod_{l=1}^{1} |T_l| = n^{-1} \cdot n = 1$, tvrdnju smo dokazali za $k = 1$. Pretpostavimo da tvrdnja vrijedi za $k-1$ i dokažimo da vrijedi za k.

Dvaput prebrojimo parove razapinjućih stabala R koji sadržavaju šumu T i dvočlani skup indeksa $\{i, j\}$, $i < j$ za koje vrijedi da su komponente T_i i T_j spojene bridom u razapinjućem stablu, tj. parove $(R, \{i, j\})$ za takve R i takve $\{i, j\}$. Neka je x broj tih parova, a y traženi broj razapinjućih stabala koji sadrže T.

Za prvi način prebrojavanja, počnimo od R. Zamislimo svaku od komponenti $T_1, \ldots, T_k$ kao jednu točku $1, \ldots, k$. Tada je R zapravo stablo na skupu $1, \ldots, k$. Izaberemo jedan od $k-1$ bridova tog stabla čiji su vrhovi upravo jedan par indeksa. Dakle, $x = y(k-1)$.

U drugom načinu prebrojavanja prvo fiksirajmo indekse $\{i, j\}$. Između komponenti T_i i T_j povucimo brid na $|T_i||T_j|$ načina. Sada odaberimo stablo koje sadrži spoj T_i, T_j i preostale $k-2$ komponente. To je razapinjuće stablo šume s $k-1$ komponentom pa možemo iskoristiti pretpostavku indukcije. Sumiranjem po mogućim odabirima indeksa tada vrijedi

$$
\begin{aligned}
x &= \sum_{1\le i<j\le k} |T_i||T_j| \cdot n^{k-3}(|T_i|+|T_j|) \prod_{l\ne i,j} |T_l| \\
&= n^{k-3} \prod_{l=1}^{n} |T_l| \cdot \sum_{1\le i<j\le k} (|T_i|+|T_j|) \\
&= n^{k-3} \prod_{l=1}^{n} |T_l| \cdot (k-1)n \\
&= n^{k-2}(k-1) \prod_{l=1}^{n} |T_l|,
\end{aligned}
$$

gdje smo u prvoj jednakosti iskoristili pretpostavku indukcije, a u trećoj da je $\sum_{l=1}^{k} |T_l| = n$ i da se u ukupnoj sumi svaka od komponenti pojavljuje $k-1$ puta. Izjednačimo li dva dobivena izraza za x imamo

$$
y(k-1) = n^{k-2}(k-1) \prod_{l=1}^{n} |T_l|.
$$

Kako smo već provjerili slučaj $k = 1$, dijeljenjem s $k-1$ dobijemo da je $y = n^{k-2} \prod_{l=1}^{n} |T_l|$, što je i trebalo dokazati.

□

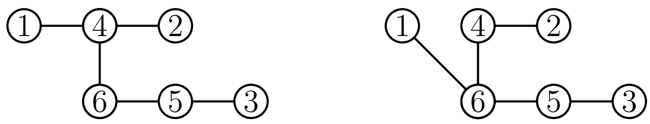
5.3 Clarkeov dokaz

Engleski matematičar L. E. Clarke je 1958. u članku [6] također dokazao Cayleyevu formulu dvostrukim prebrojavanjem, no nešto manje direktno. Broj stabala s n vrhova promatrao je kao zbroj brojeva stabala s n vrhova u kojima fiksan vrh ima stupanj $1, \ldots, n-1$, a posebno je izrazio vezu među stablima koja se jedno iz drugog mogu dobiti prespajanjem jednog brida koji sadrži taj vrh.

Neka je S stablo s vrhovima $\{v_1, \ldots, v_n\}$. Kažemo da je S *stablo tipa* l ako je $d(v_n) = l$. Neka su S_1 i S_2 stabla s n vrhova te neka je S_1 tipa $l-1$, a S_2 tipa l. Kažemo da se S_2 dobiva *prespajanjem* iz S_1 ako S_1 možemo transformirati u S_2 tako da odaberemo jedan vrh koji nije povezan s v_n, promotrimo jedinstveni put od tog vrha do v_n, uklonimo prvi brid na tom putu, te odabrani vrh spojimo s v_n.

Primjerice, sljedeća dva stabla su u tom odnosu jer S_2 možemo dobiti prespajanjem iz S_1 tako da uklonimo brid $\{1, 4\}$ i dodamo brid $\{1, 6\}$:

S_1 S_2



Sljedeća dva nisu u tom odnosu (unatoč tome što je stupanj vrha 6 za jedan veći u S_2 nego u S_1) jer ne možemo opisanom transformacijom prespajanja dobiti S_2 iz S_1:

S_1 S_2

Definiramo još i *prespajanje tipa l* kao uređen par stabala (S_1, S_2) gdje je S_1 tipa $l-1$, S_2 tipa l, i S_2 možemo dobiti iz S_1 prespajanjem. Dakle, u gornjem primjeru u kojem smo S_2 mogli dobiti iz S_1 prespajanjem vrha 1 na vrh 6, uređeni par (S_1, S_2) je upravo jedno prespajanje tipa 3 na stablima sa 6 vrhova.

Označimo s $N_{l,n}$ broj stabala s n vrhova tipa l i označimo s $M_{l,n}$ broj prespajanja tipa l na stablima s n vrhova, tj. broj opisanih uređenih parova stabala. Cayleyeva formula slijedit će iz nekih identiteta i rekurzivnih formula za $N_{l,n}$, a dokazat ćemo ih dvostrukim prebrojavanjem veličine $M_{l,n}$.

Dokaz. Prebrojimo $M_{l,n}$ na dva načina. Prvi način prebrojavanja slijedi direktno iz definicije. Neka je S_1 stablo tipa $l-1$ s n vrhova. Znamo da S_1 ima $(n-1)-(l-1) = n-l$ vrhova koji nisu povezani s v_n. Među njima biramo vrh koji možemo povezati s v_n da bismo prespajanjem dobili stablo S_2 tipa l. Dakle,

$$M_{l,n} = (n-l)N_{l-1,n}. \tag{10}$$

Za drugi način prebrojavanja $M_{l,n}$, neka je S_2 stablo tipa l s n vrhova te neka su bez smanjenja općenitosti vrhovi povezani s v_n upravo $v_1, \ldots, v_l$. Prebrojimo kako iz S_2 možemo dobiti stablo S_1 tipa $l-1$ s n vrhova. Uklonimo bridove koji sadrže v_n. Rezultat je šuma s $l+1$ stablom. Označimo brojeve vrhova u nastalim stablima p_i, $i = 1, \ldots, l$ tako da svaki p_i sadrži pripadni v_i. Vrijedi $\sum_{i=1}^{l} p_i = n-1$. Odaberimo jedan od vrhova, primjerice v_1, i spojimo ga s jednim od $n-1-p_1$ vrhova koji nisu jednaki v_n i koji nisu u stablu koje sadrži v_1. Zatim preostale vrhove $v_2, \ldots, v_l$ ponovno spojimo s v_n. Rezultat je traženi S_1. Kako isto možemo napraviti za sve $v_2, \dots, v_l$, slijedi

$$
\begin{aligned}
M_{l,n} &= N_{l,n} \sum_{i=1}^{l} (n-1-p_i) \\
&= N_{l,n} \left(l(n-1) - \sum_{i=1}^{l} p_i \right) \\
&= N_{l,n}(l(n-1)-(n-1)) \\
&= (l-1)(n-1)N_{l,n}. \qquad (11)
\end{aligned}
$$

Izjednačavanjem (10) i (11) slijedi rekurzivna formula za stabla tipa $l-1$ i l s n vrhova:

$$
N_{l-1,n} = \frac{(n-1)(l-1)}{n-l} N_{l,n}.
$$

Zamjenom indeksa $l-1 \to k$ i uočavanjem da za $k = n-1$ postoji samo jedno stablo tipa k (slučaj kada su svi vrhovi povezani s v_n, tzv. *zvjezdasti graf*), a za $k \geq n$ takva stabla ne postoje, vrijedi

$$
N_{k,n} = \begin{cases} \frac{(n-1)k}{n-k-1} N_{k+1,n}, & \text{ako } 1 \leq k \leq n-2, \\ 1, & \text{ako } k = n-1, \\ 0, & \text{ako } k \geq n. \end{cases} \qquad (12)
$$

Dokažimo da iz gornje rekurzivne relacije slijedi identitet

$$
N_{l,n} = \binom{n-2}{l-1} (n-1)^{n-l-1}. \qquad (13)
$$

Fiksirajmo n. Za proizvoljni $l \in \mathbb{N}$, znamo da formula vrijedi ako je $l \geq n-1$ jer

$$
\binom{n-2}{(n-1)-1} (n-1)^{n-(n-1)-1} = 1 = N_{n-1,n}
$$

i za $l \geq n$ jer

$$
\binom{n-2}{l-1} (n-1)^{n-l-1} = 0 = N_{l,n}.
$$

Preostaje provjeriti da vrijedi za $l = 1, \dots, n-2$. Koristimo silaznu indukciju po l. Bazu $l = n-1$ već smo provjerili. Pretpostavimo da (13) vrijedi za sve $l \geq k$, za $1 \leq k \leq n-2$ i dokažimo da tada vrijedi za k. Kako je $k+1 \geq k$, iz pretpostavke slijedi

$$N_{k+1,n} = \binom{n-2}{(k+1)-1}(n-1)^{n-(k+1)-1} = \binom{n-2}{k}(n-1)^{n-k-2}.$$

Uvrstimo li to u (12), dobijemo

$$\begin{aligned} N_{k,n} &= \frac{(n-1)k}{n-k-1}\binom{n-2}{k}(n-1)^{n-k-2} \\ &= \frac{(n-1)k}{n-k-1}\frac{(n-2)!}{k!(n-2-k)!}(n-1)^{n-k-2} \\ &= \frac{(n-2)!}{(k-1)!(n-k-1)}(n-1)^{n-k-1} \\ &= \binom{n-2}{k-1}(n-1)^{n-k-1}, \end{aligned}$$

čime je dokazano (13).

Konačno, kako je ukupan broj stabala s n vrhova jednak zbroju brojeva stabala tipa $1,\ldots,n$, primjenom redom (13) i binomnog teorema slijedi

$$\begin{aligned} t_n &= \sum_{l=1}^{n-1} N_{l,n} = \sum_{j=0}^{n-2} N_{j+1,n} \\ &= \sum_{j=0}^{n-2}\binom{n-2}{j}(n-1)^{n-j} \\ &= (n-1+1)^{n-2} \\ &= n^{n-2}. \end{aligned}$$

□

6 Dokaz pomoću funkcija izvodnica

Prsten formalnih redova potencija $\mathbb{C}[[z]]$ sadrži sve nizove $a : \mathbb{N}_0 \to \mathbb{C}$. Njegove elemente zapisujemo kao redove potencija i nazivamo *funkcijama izvodnicama*:

$$F(z) = \sum_{n \geq 0} a_n z^n.$$

Definiramo i operator $\langle z^n \rangle$ koji funkciji izvodnici pridružuje koeficijent uz z^n, tj. $\langle z^n \rangle F(z) = a_n$. Često je lakše raditi s nizom $(\frac{a_n}{n!})$ te u tom slučaju govorimo o *eksponencijalnim funkcijama izvodnicama*: $\hat{F}(z) = \sum_{n \geq 0} \frac{a_n}{n!} z^n$.

U kombinatorici se funkcije izvodnice koriste za rješavanje problema prebrojavanja. Koeficijenti funkcije izvodnice zapravo čine niz koji prebrojava traženu familiju konačnih skupova, a rad s tim nizom nam je olakšan jer koristimo formalne redove. To je još jedan način kako riješiti problem prebrojavanja označenih stabala.

Mađarski matematičar George Pólya dokazao je Cayleyevu formulu 1937. u članku [13] uspostavivši vezu između dviju funkcija izvodnica i računanjem njihovih koeficijenata. Da bismo proveli njegov dokaz trebamo još nekoliko pojmova i rezultata.

Definiramo *red* (eng. *order*) funkcije izvodnice $F(z) = \sum_{n \geq 0} a_n z^n$ za $F(z) \neq 0$ kao

$$\mathrm{ord} F(z) = \min\{n \in \mathbb{N}_0 \mid a_n \neq 0\}.$$

Kompozicija funkcija izvodnica $F(z) = \sum_{n \geq 0} a_n z^n$ i $G(z) = \sum_{n \geq 0} b_n z^n$, u oznaci $(F \circ G)(z)$, je beskonačna suma $\sum_{n \geq 0} a_i G(z)^i$. No, ona nije uvijek dobro definirana. U slučaju kada je F polinom dobro je definirana jer se radi o konačnoj linearnoj kombinaciji funkcija izvodnica. U općenitom slučaju, može se pokazati se da je definirana ako je $b_0 = 0$. Kada je kompozicija funkcija izvodnica definirana, ima ista svojstva kao kompozicija polinoma.

Bitan pojam u ovom poglavlju jest pojam *kompozicijskog inverza* funkcije izvodnice. Neutralni element za kompoziciju funkcija izvodnica jest $I(z) = z$: $F(z) \circ I(z) = I(z) \circ F(z) = F(z)$. Kompozicijski inverz od $F(z)$ je funkcija izvodnica $F^{(-1)}(z)$ takva da vrijedi $F^{(-1)}(z) \circ F(z) = F(z) \circ F^{(-1)}(z) = I(z) = z$. Za izračunavanje koeficijenata kompozicijskog inverza pomaže nam Lagrangeova formula inverzije.

Teorem 6.1 (Lagrangeova formula inverzije)**.** *Neka je* $F(z) = \sum_{i \geq 1} a_i z^i$*,* $a_1 \neq 0$*,* $n \in \mathbb{Z}$*. Tada je*

$$\langle z^n \rangle F^{(-1)}(z) = \frac{1}{n} \langle z^{n-1} \rangle \left(\frac{z}{F(z)} \right)^n.$$

Lagrangeovu formulu inverzije možemo iskoristiti i za rješavanje jednadžbi u $\mathbb{C}[[z]]$ oblika $f(z) = z \cdot G(f(z))$, gdje je $G \in \mathbb{C}[[z]]$ s netrivijalnim slobodnim članom zadana, a tražimo $f \in \mathbb{C}[[z]]$:

$$\langle z^n \rangle f(z) = \frac{1}{n} \langle z^{n-1} \rangle (G(z)^n).$$

Neka t'_n označava broj korijenskih stabala, tj. stabala s jednim istaknutim vrhom. Vrijedi $t'_n = n \cdot t_n$, pa Cayleyeva formula tada glasi $t'_n = n^{n-1}$. Šume kojima su komponente povezanosti korijenska stabla nazivamo *korijenskim šumama*. Označimo broj korijenskih šuma s f_n. Pólya je otkrio sljedeći odnos među ovim dvama nizovima:

Propozicija 6.2 (Pólya). *$t'_{n+1} = (n+1)f_n$, za $n \geq 0$.*

Dokaz. Neka je zadano korijensko stablo s $n+1$ vrhova. Odaberimo vrh $i \in \{1, \ldots, n+1\}$ na $n+1$ način i preimenujmo preostale vrhove redom u $\{1, \ldots, n\}$. Izbrišimo i te svakoj nastaloj komponenti povezanosti za korijen odaberimo onaj vrh koji je bio povezan s i.

Obratno, neka je zadana korijenska šuma s vrhovima $\{1, \ldots, n\}$. Možemo dobiti korijensko stablo tako da odaberemo $i \in \{1, \ldots, n+1\}$, preimenujemo vrhove šume u $\{1, \ldots, i-1, i+1, \ldots, n+1\}$, te vrh i povežemo s korijenima stabala koje su komponente povezanosti dane šume. Tako dobijemo korijensko stablo s korijenom i i $n+1$ vrhom.

Očito su ove dvije transformacije inverzne pa vrijedi tvrdnja propozicije. $\square$

Sada smo spremni za još jedan dokaz Cayleyeve formule.

Dokaz. Definirajmo eksponencijalnu funkciju izvodnicu za broj korijenskih stabala $T(z) = \sum_{k \geq 0} \frac{1}{k!} t'_k$. Dokazat ćemo najprije da vrijedi relacija

$$T(z) = ze^{T(z)}. \tag{14}$$

U tu svrhu, kombinatorno interpretirajmo izraz

$$e^{T(z)} = \sum_{k \geq 0} \frac{1}{k!} T(z)^k.$$

Promotrimo li odgovarajuće eksponente te uvrstimo koeficijente od $T(n)$, vidimo da za koeficijente uz z^n u toj sumi vrijedi

$$\begin{aligned}\langle z^n\rangle \frac{1}{k!}T(z)^k &= \frac{1}{k!}\sum_{i_1+\ldots+i_k=n}\frac{t'_{i_1}}{i_1!}\cdots\frac{t'_{i_k}}{i_k!}\\ &= \frac{1}{n!}\cdot\frac{1}{k!}\sum_{i_1+\ldots+i_k=n}\frac{n!}{i_1!\ldots i_k!}t'_{i_1}\ldots t'_{i_k}\\ &= \frac{1}{n!}\cdot\frac{1}{k!}\sum_{i_1+\ldots+i_k=n}\binom{n}{i_1\ldots i_k}t'_{i_1}\ldots t'_{i_k}.\end{aligned}$$

Uočimo da izraz $\frac{1}{k!}\sum_{i_1+\ldots+i_k=n}\binom{n}{i_1\ldots i_k}t'_{i_1}\ldots t'_{i_k}$ prebrojava korijenske šume s n vrhova koje imaju točno k komponenti. Multinomni koeficijent predstavlja biranje broja vrhova u svakoj komponenti kao $i_1,\ldots,i_k$. Njihova suma je n, a unutar j-te komponente imamo t'_{i_j} mogućih stabala. Na kraju dijelimo s $k!$ jer poredak komponenta nije bitan. Sumiramo li ovaj izraz za sve $k=0,\ldots,n$ dobit ćemo upravo f_n, broj korijenskih šuma s n vrhova. Označimo li odgovarajuću funkciju izvodnicu s $F(z)=\sum_{n\geq 0}\frac{f_n}{n!}z^n$, dokazali smo relaciju $e^{T(z)}=F(z)$. Primijenimo na to propoziciju 6.2:

$$e^{T(z)} = F(z) = \sum_{n\geq 0}\frac{f_n}{n!}z^n = \sum_{n\geq 0}\frac{t_{n+1}}{(n+1)!}z^n = \frac{1}{z}\sum_{n\geq 1}\frac{t_n}{n!}z^n = \frac{T(z)}{z}.$$

Dakle, vrijedi relacija (14). Iz nje lako slijedi dokaz Cayleyeve formule korištenjem Lagrangeove formule inverzije za slučaj implicitno zadanih funkcija izvodnica:

$$\langle z^n\rangle T(z) = \frac{1}{n}\left\langle z^{n-1}\right\rangle e^{nz} = \frac{1}{n}\left\langle z^{n-1}\right\rangle\sum_{k\geq 0}\frac{n^k}{k!}z^k = \frac{n^{n-1}}{n!} = \frac{t'_n}{n!}.$$

Slijedi $t'_n = n^{n-1}$, odnosno $t_n = n^{n-2}$. $\square$

7 Matrični dokaz

Linearna algebra koristan je alat u kombinatornim problemima, naročito u teoriji grafova. Mnoga svojstva grafova mogu se prikazati matricama, a one pak sa sobom nose mnoštvo korisnih rezultata iz linearne algebre koje možemo primijeniti. Na taj način može se dokazati i Cayleyeva formula. Gustav Kirchhoff je 1847. objavio teorem koji prebrojava razapinjuća stabla grafa koristeći determinante. U literaturi je njegov teorem poznat i kao matrični teorem o stablima (eng. matrix-tree theorem). Izlažemo verziju Kirchhoffovog dokaza iz knjige [20]. Jednom kad ga dokažemo, Cayleyeva formula direktno slijedi kao poseban slučaj Kirchhoffovog teorema.

Neka je G jednostavan graf s n vrhova $v_1, \dots, v_n$. Definiramo *matricu susjedstva* $A = [a_{ij}]$ grafa G kao $n \times n$ matricu s elementima

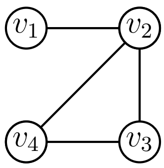
$$a_{ij} = \begin{cases} 1, & \text{ako } v_i \sim v_j \\ 0, & \text{inače.} \end{cases}$$

Matrica susjedstva je za jednostavne grafove simetrična jer nema usmjerenih bridova. *Matricu stupnjeva* $D = [d_{ij}]$ grafa G definiramo kao dijagonalnu $n \times n$ matricu u kojoj za elemente dijagonale vrijedi $d_{ii} = d(v_i)$. *Laplacijan* grafa G je matrica L definirana kao

$$L = D - A.$$

Laplacijan je također simetričan, jer su takve i D i A.

Primjer 7.1. *Odredimo Laplacijan grafa sa slike.*



Vrijedi $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$ *i* $D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$*, pa je*

$$L = D - A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 3 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ 0 & -1 & -1 & 2 \end{bmatrix}.$$

U ovom primjeru primjećujemo da je suma svakog retka i stupca Laplacijana jednaka 0. To vrijedi i općenito. U i-tom retku imamo $d(v_i)$ i točno $d(v_i)$ puta se ponavlja -1, za svakog od susjeda od v_i po jedan. Analogno vrijedi i za stupce. Ovakve matrice imaju brojna korisna svojstva, a jedno od njih vezano je za kofaktore matrice.

Kofaktor $A_{i,j}$ kvadratne matrice A je determinanta matrice A s uklonjenim i-tim retkom i j-tim stupcem pomnožena s $(-1)^{i+j}$. *Adjunkta* $\operatorname{adj}(A)$ kvadratne matrice A je matrica koja na poziciji (i,j) ima kofaktor $A_{j,i}$. Iz linearne algebre znamo da vrijedi $\operatorname{adj}(AB) = \operatorname{adj}(A)\operatorname{adj}(B)$ i $\operatorname{adj}(A)A = (\det A)I$.

Lema 7.2. *Ako svi retci i stupci $n \times n$ matrice A imaju sumu 0, tada su svi kofaktori matrice A jednaki.*

Dokaz. Dokažimo da su svi kofaktori na pozicijama unutar i-tog retka jednaki, za sve retke i. Kako je suma svakog retka 0, rang matrice je najviše $n-1$. Ako je manji od $n-1$, očito su svi kofaktori jednaki 0. Neka je rang točno $n-1$. Vektor jedinica $\mathbb{1} = (1,\dots,1)$ je u jezgri od A jer je suma svakog retka 0. Kako je $r(A) = n-1$, slijedi da je svaki element jezgre u potprostoru razapetom s $\mathbb{1}$, tj. da je jednak $\lambda\mathbb{1}$ za skalar $\lambda \in \mathbb{R}$. Iskoristimo sada jednakost $\operatorname{adj}(A)A = (\det A)I$. Zbog $\det A = 0$ je $\operatorname{adj}(A)A = 0$ pa su svi stupci od $\operatorname{adj}(A)$ u jezgri od A. Dakle, svi stupci od $\operatorname{adj}(A)$ su konstantni, što znači da su svi kofaktori istog retka jednaki. Analogno se dokaže da su svi kofaktori istog stupca jednaki. Promotrimo i-ti redak u $\operatorname{adj}(A)$. Svi njegovi elementi su jednaki, a kako isto vrijedi i za stupce koji svi imaju po jedan element u i-tom retku, vrijedi da su svi elementi od $\operatorname{adj}(A)$ jednaki. $\square$

Trebat će nam još i Binet-Cauchyjev teorem.

Teorem 7.3 (Binet-Cauchy)**.** *Neka je C produkt $n \times m$ matrice A i $m \times n$ matrice B. Tada je $\det C = \sum_{S \subseteq \{1,\dots,m\}, |S|=n} \det A_S \det B_S$, gdje je A_S (B_S) $n \times n$ podmatrica od A (B) s retcima (stupcima) određenima skupom S.*

Iskažimo i dokažimo Kirchhoffov teorem.

Teorem 7.4 (Kirchhoffov teorem)**.** *Neka je G jednostavan graf i L njegov Laplacijan. Tada je broj razapinjućih stabala od G jednak kofaktoru $L_{i,j}$.*

Dokaz. Zbog leme 7.2, svi kofaktori od L su jednaki pa je dovoljno tvrdnju dokazati za $i = j$. Dakle, dokazujemo da je broj razapinjućih stabala od G jednak jedinstvenoj vrijednosti kofaktora od L, tj. $\det L_i$, gdje je L_i matrica L s uklonjenim i-tim retkom i stupcem, za proizvoljan $i = 1,\dots,n$.

Bez smanjenja općenitosti, proizvoljno usmjerimo bridove od G i dobiveni usmjereni graf označimo s U. Označimo vrhove od U s $v_1,\dots,v_n$ i bridove s $e_1, \ldots, e_m$. Definirajmo matricu $M = [m_{ij}]$ u kojoj retci predstavljaju vrhove, a stupci bridove od U, te je

$$m_{ij} = \begin{cases} 1, & \text{ako je } v_i \text{ početni vrh od } e_j, \\ -1, & \text{ako je } v_i \text{ krajnji vrh od } e_j, \\ 0, & \text{inače.} \end{cases}$$

Promotrimo produkt matrica MM^T. Na poziciji (i, j) od MM^T u slučaju $i \neq j$ nalazi se -1 ako su vrhovi v_i i v_j susjedni jer je to umnožak brojeva 1 i -1 (jedan vrh je početni vrh brida, a drugi krajnji), a u slučaju da nisu susjedni imamo 0. Nadalje, ako $i = j$, na poziciji (i, i) od MM^T nalazi se stupanj vrha v_i jer zbrajamo jedinicu onoliko puta koliko ima bridova u kojima se nalazi v_i. Ovo je upravo definicija Laplacijana od G, pa vrijedi $L = MM^T$.

Zbog propozicije 2.4, sva razapinjuća stabla od G imaju $(n-1)$ bridova. Promotrimo zato proizvoljnu $(n-1) \times (n-1)$ podmatricu B od M. Ako pripadni graf matrice B nije povezan, B ima nulredak pa je $\det B = 0$. Ako pak ima ciklus, tada zbroj stupaca koji odgovaraju bridovima u ciklusu iznosi 0, pa je zbog linearne zavisnosti opet $\det B = 0$. Dakle, ako B nije stablo, vrijedi $\det B = 0$.

Tvrdimo da ako B jest stablo, tada je $\det B = \pm 1$. Dokažimo to indukcijom. Za $n = 1$ tvrdnja vrijedi jer 0×0 matrica po definiciji ima determinantu 1 što odgovara jedinom stablu s jednim vrhom, N_1. Pretpostavimo da je je T razapinjuće stablo čiji su bridovi stupci od B. Zbog propozicije 2.3, T ima bar dva lista. Zbog toga B sadrži redak koji odgovara listu od T i on ima samo jedan nenul element te je jednak ± 1. Računamo $\det B$ Laplaceovim razvojem po tom retku. Jedini nenul pribrojnik dobit ćemo iz umnoška tog elementa i determinante $(n-2) \times (n-2)$ podmatrice koja također odgovara stablu jer smo uklonili samo list pa nismo narušili ni povezanost ni nepostojanje ciklusa. Po pretpostavci indukcije taj umnožak je upravo ± 1.

Izračunajmo sada $\det(L_i)$ za proizvoljan $i = 1, \ldots, n$. Neka je M_i matrica M s uklonjenim i-tim retkom. Tada vrijedi $L_i = M_i M_i^T$. Iskoristimo Binet-Cauchyjev teorem: $\det(L_i) = \sum_{S \subseteq \{1, \ldots, m\}} \det(M_i) \det(M_i^T)$. Ako je $m < n-1$, sve determinante su jednake 0. Neka je $m \geq n-1$. Sumiranjem po svim podmatricama od M_i i M_i^T dobit ćemo da su determinante jednake 0 ako pripadni podgraf nije stablo, odnosno 1 ako jest. Dakle, konačna suma prebrojava stabla s $n-1$ bridom čiji su svi bridovi u G. To su upravo razapinjuća stabla od G.

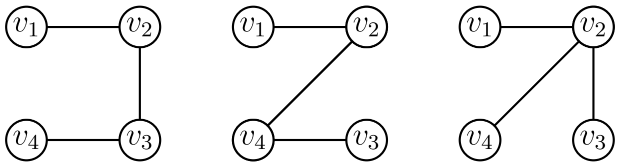
□

Ilustrirajmo tvrdnju Kirchhoffovog teorema na primjeru.

Primjer 7.5. *Odredimo broj razapinjućih stabala grafa iz prethodnog primjera. Neka je, primjerice, $i = 2$. Tada je*

$$\det(L_2) = \det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = 3.$$

Lako možemo provjeriti i da vrijedi $\det(L_1) = \det(L_3) = 3$, *tj. da su ti kofaktori jednaki. S obzirom da se radi o malenom* n, *možemo se u rezultat lako uvjeriti i ispisivanjem svih razapinjućih stabala:*



Cayleyeva formula zapravo je korolar Kirchhoffovog teorema. Treba prepoznati da je broj stabala s n vrhova zapravo broj razapinjućih stabala potpunog grafa K_n. Kako je u potpunom grafu svaki vrh susjedan svim vrhovima osim samom sebi, matrica susjedstva i matrica stupnjeva su

$$A = \begin{bmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \dots & 0 \end{bmatrix}, \quad D = \begin{bmatrix} n-1 & 0 & 0 & \dots & 0 \\ 0 & n-1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & n-1 \end{bmatrix}$$

pa je Laplacijan od K_n jednak

$$L = \begin{bmatrix} n-1 & -1 & -1 & \dots & -1 \\ -1 & n-1 & -1 & \dots & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & -1 & \dots & n-1 \end{bmatrix}.$$

Svi L_i su jednaki, pa odredimo primjerice $\det(L_1)$ svođenjem na gornjetrokutasti oblik Gaussovim transformacijama:

$$\det(L_1) = \det \begin{bmatrix} n-1 & -1 & \dots & -1 \\ -1 & n-1 & \dots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \dots & n-1 \end{bmatrix} = \det \begin{bmatrix} 1 & -1 & \dots & -1 \\ 1 & n-1 & \dots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & -1 & \dots & n-1 \end{bmatrix}$$

$$= \det \begin{bmatrix} 1 & -1 & \dots & -1 \\ 0 & n & \dots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & n \end{bmatrix} = n^{n-2},$$

gdje smo u drugoj jednakosti pribrojili sve retke prvom, a u trećoj oduzeli prvi redak od svih ostalih. Konačan rezultat slijedi množenjem vrijednosti na dijagonali jer se radi o trokutastoj matrici. Kako je matrica dimenzija $(n-1) \times (n-1)$ i sve vrijednosti na dijagonali osim jedne jedinice su n, dobili smo očekivanih n^{n-2} razapinjućih stabala. Time smo dokazali:

Korolar 7.6 (Cayleyeva formula)**.** *Broj razapinjućih stabala potpunog grafa K_n jednak je n^{n-2}.*

8 Vjerojatnosni dokaz

Vjerojatnost i kombinatorika su nerazdvojne, a u nastavku ćemo pokazati kako ni Cayleyeva formula nije iznimka. Iznosimo jedan nešto složeniji dokaz koji se temelji na Poissonovom procesu grananja. Ideja dokaza je jednostavna: cilj je izračunati vjerojatnost da rezultat slučajnog procesa bude zadano stablo, a recipročna vrijednost te vjerojatnosti je broj mogućih stabala. No, da bismo tu ideju proveli u djelo potrebno je provesti malo opširniju pripremu. Pratimo dokaz kako je izložen u [8].

Neka je $(Z_{n,i})_{n,i\geq 1}$ niz nezavisnih jednako distribuiranih nenegativnih cjelobrojnih slučajnih varijabli. *Jednostavan proces grananja* $(Z_n)_{n\geq 0}$, ili *Galton-Watsonov proces*, slučajan je proces u kojem Z_n predstavlja broj jedinki u n-toj generaciji. Proces počinje s jednom jedinkom, a svaka sljedeća generacija ima onoliko jedinki koliki je zbroj djece jedinki iz prethodne generacije, tj.

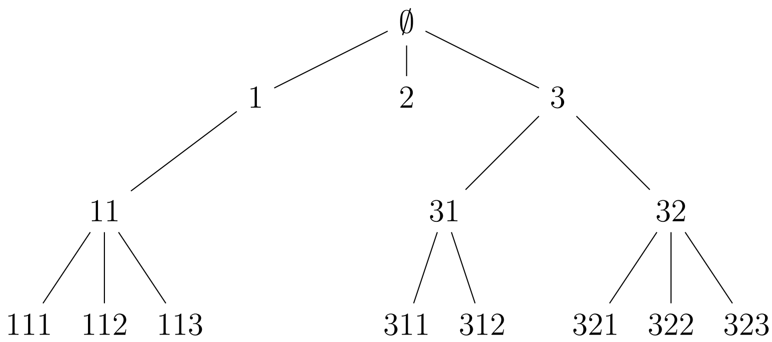
$$\begin{aligned}
Z_0 &= 1\\
Z_1 &= Z_{1,1}\\
Z_2 &= Z_{2,1} + Z_{2,2} + \ldots + Z_{2,Z_1}\\
&\vdots\\
Z_n &= Z_{n,1} + Z_{n,2} + \ldots + Z_{n,Z_{n-1}},
\end{aligned}$$

gdje nezavisne jednako distribuirane $Z_{n,i}$ predstavljaju broj djece i-te jedinke iz n-te generacije. Uočimo da $Z_n = 0$ povlači $Z_{n+k} = 0$, za $n \geq 1$ (*izumiranje*). Zbog pretpostavke jednake distribuiranosti niza $(Z_{n,i})_{n,i\geq 1}$, možemo definirati *distribuciju grananja* kao $Z_1 = Z_{1,1} \sim Z_{n,i}$.

Proces grananja svoje ime dobio je upravo iz vizualizacije preko obiteljskog stabla u kojem se pamti samo jedan roditelj. Jezikom teorije grafova, to je korijensko stablo. Korijen čini nultu generaciju, a n-tu generaciju čine svi vrhovi kojima je put do korijena duljine n. Za svaki vrh koji nije korijen nazivamo prvi vrh na njegovom putu do korijena njegovim *roditeljem* (posebno, svaki vrh osim korijena ima jedinstvenog roditelja). Preostale susjede svakog od vrhova nazivamo njegovom *djecom*. Dakle, ako je v_1 roditelj od v_2, tada je v_2 dijete od v_1. Ako su v_1 i v_2 djeca od v_3, tada kažemo da su v_1 i v_2 *braća*. Kao i u obiteljskom stablu u kojem su braća slijeva na desno sortirana po starosti, poredak među braćom je bitan. Kažemo da se radi o *uređenom stablu*.

Neka je ukupan broj vrhova (jedinki) stabla S jednak $|S|$. Neka je d_v broj djece vrha v (tj. $d_v = d(v) - 1$, osim za korijen za koji vrijedi $d(v) = d_v$). Dakle, d_v su realizacije nezavisnih jednako distribuiranih slučajnih varijabli

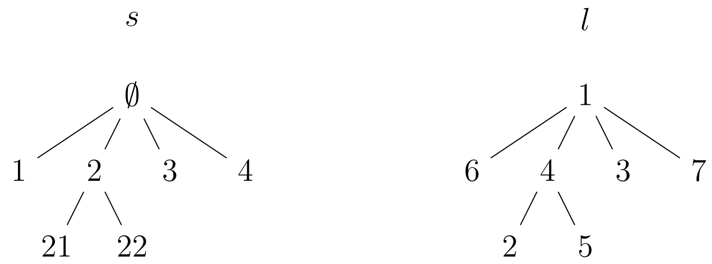
$Z_{n,i}$. Korijen stabla označimo s $\emptyset$, njegovu djecu s $1, 2, \ldots, d$, djecu vrha 1 s $11, 12, \ldots, 1d_1$, i tako dalje. Dakle, svaki vrh v u n-toj generaciji označavamo riječju s n znakova, gdje je prvih $n-1$ znakova jednako putu do korijena, a n-ti znak označava koji je po redu vrh v među braćom. Primjerice,



Ova reprezentacija vrhova poznata je kao *Ulam-Harrisova reprezentacija stabla.* U nastavku, naziv *obiteljsko stablo* označava uređeno korijensko stablo označeno na upravo ovaj način. Dva obiteljska stabla jednaka su ako su reprezentirana istim skupom riječi.

Uočimo da prebrojavanje obiteljskih stabala nije jednako prebrojavanju stabala u smislu Cayleyeve formule. Naime, cilj nam je prebrojati *neuređena stabla*, stabla u kojima poredak braće nije bitan (tj. dva stabla su jednaka ako imaju jednak skup bridova), dok u slučaju Ulam-Harrisove reprezentacije brojimo uređena stabla. U gornjem primjeru zamjenom vrhova 111 i 112 i dalje dobijemo isto neuređeno stablo. Štoviše, postoji $3! \cdot 3! \cdot 2! \cdot 3! = 432$ uređenih stabala nastalih procesom grananja koja čine isto neuređeno stablo, jer možemo zamijeniti poretke vrhova $1, 2, 3$, vrhova $111, 112, 113$, vrhova $311, 312$ i vrhova $321, 322, 323$. Dakle, kako bismo preko proučavanja procesa grananja došli do broja stabala, trebamo prvo uspostaviti vezu prebrojavanja uređenih i neuređenih stabala. U tu svrhu slijedi nekoliko definicija i lema.

Neka je s obiteljsko stablo s n vrhova. Označit ćemo ga na drugi način tako da korijenu pridružimo oznaku 1, a preostalim vrhovima nasumično (uniformno i bez ponavljanja) dodijelimo oznake $\{2, \ldots, n\}$. Rezultat je označeno i neuređeno korijensko stablo l. Radi jednostavnosti, u nastavku *označeno stablo* znači upravo takvo stablo, a *označavanje* obiteljskog stabla znači opisani proces dodjeljivanja oznaka. Za označeno stablo l i obiteljsko stablo s uvodimo oznaku $s \sim l$ ako se l može dobiti iz s promjenom oznaka. Zbog toga što je s definirano kao uređeno, a l kao neuređeno stablo, više različitih označavanja stabla s rezultirat će istim neuređenim stablom l. Zbog toga za zadano obiteljsko stablo s i označeno stablo l takve da $s \sim l$ definiramo $L(l)$ kao broj načina na koje možemo označiti s da bismo dobili stablo jednako stablu l. Neka su npr. zadana stabla



Označeno stablo sa zamijenjenim vrhovima $6, 3, 7$ ili vrhovima $2, 5$ jednako je stablu l jer l nije uređeno. Dakle, postoji $L(l) = 3! \cdot 2! = 12$ načina da iz s dobijemo l. Jasno je da je veličina $L(l)$ dobro definirana jer je bitno samo koje vrhove možemo permutirati u s da dobijemo isti l, što ne ovisi o izboru s dokle god vrijedi $s \sim l$.

Sada smo spremni za sljedeće dvije leme koje povezuju obiteljska i označena stabla.

Lema 8.1. *Za zadano označeno stablo l, broj obiteljskih stabala s za koje vrijedi $s \sim l$ jednak je*

$$|\{s : s \sim l\}| = \frac{\prod_{v \in l} d_v!}{L(l)}.$$

Dokaz. Neka je l označeno stablo i neka je s_0 obiteljsko stablo takvo da $s_0 \sim l$. Obiteljsko stablo s_0 može promjenom poredaka braće postati bilo koje drugo $s_i \in \{s : s \sim l\}$. Za svaki vrh v u s_0 možemo promijeniti poredak braće na $d_v!$ načina, što je ukupno $\prod_{v \in l} d_v!$ različitih stabala. No, kad ta stabla promatramo kao neuređena stabla, svako od njih jednako je kao još njih $L(l)$. Dakle, konačan rezultat je $\frac{\prod_{v \in l} d_v!}{L(l)}$. $\square$

Lema 8.2. *Za zadano obiteljsko stablo s i označeno stablo l takve da $s \sim l$, vjerojatnost događaja da označavanjem s dobijemo upravo l, u oznaci $s \to l$, jednaka je*

$$\mathbb{P}(s \to l) = \frac{L(l)}{(|l| - 1)!}.$$

Dokaz. Korijen uvijek ima oznaku 1, pa postoji $(|l|-1)!$ načina za uniformno i bez ponavljanja dodijeliti oznake $2, \dots, |l|$ preostalim vrhovima od s (očito $s \sim l \Rightarrow |s| = |l|$). Dakle, vjerojatnost da je proizvoljno označeno uređeno stablo s $|l|$ vrhova jednako l, ako l promatramo kao uređeno stablo, jednaka je $\frac{1}{(|l|-1)!}$. No, kako je l neuređeno stablo, to množimo s $L(l)$. Dakle, vrijedi $\mathbb{P}(s \to l) = \frac{L(l)}{(|l|-1)!}$. $\square$

Vrijeme je da dosadašnja razmatranja povežemo s procesima grananja. U našem slučaju, promatrat ćemo proces grananja s distribucijom grananja $Z_1 \sim \text{Pois}(1)$. Vjerojatnosna funkcija gustoće Poissonove distribucije za općeniti parametar $\lambda \in \langle 0, +\infty \rangle$ je $\mathbb{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$, $k \in \mathbb{N}_0$, pa u našem slučaju ona glasi

$$P(X = k) = \frac{e^{-1}}{k!}.$$

Lema 8.3. *Neka je zadano obiteljsko stablo s takvo da je $|s| = n$. Za obiteljsko stablo S nastalo procesom grananja s distribucijom grananja* Pois(1) *takvo da je $|S| = n$, vrijedi*

$$P(S = s) = \frac{e^{-n}}{\prod_{v \in s} d_v!}. \tag{15}$$

Dokaz. Kako je stablo S nastalo procesom grananja, svaki od n vrhova ima onoliko djece kolike su realizacije d_v od n nezavisnih jednako distribuiranih slučajnih varijabli s distribucijom Pois(1). Zbog nezavisnosti i jednake distribuiranosti, vjerojatnost da je S jednako upravo zadanom s jednaka je

$$P(S = s) = \prod_{v \in s} P(Z_1 = d_v) = \prod_{v \in s} \frac{e^{-1}}{d_v!} = \frac{e^{-n}}{\prod_{v \in s} d_v!}.$$

□

Uočimo da je ta vjerojatnost jednaka i za sva druga stabla s istim stupnjevima vrhova.

Lema 8.4. *Neka je obiteljsko stablo S nastalo procesom grananja s distribucijom grananja* Pois(1) *i neka je zadano označeno stablo l. Označimo S i neka je l' dobiveno označeno stablo. Tada je vjerojatnost da smo na taj način dobili upravo l jednaka*

$$P(l' = l) = \frac{e^{-|l|}}{(|l| - 1)!}.$$

Dokaz. Po definiciji, da bi bilo $l' = l$, najprije treba vrijediti $S = s_l$ za neki $s_l \sim l$, a zatim još i s_l treba označiti točno kao l. Dakle,

$$P(l' = l) = \sum_{s_l \sim l} \mathbb{P}(S = s_l)\mathbb{P}(s_l \to l).$$

Uvrstimo li rezultate iz lema 8.2 i 8.3, slijedi

$$P(l' = l) = \sum_{s_l \sim l} \frac{e^{-|l|}}{\prod_{v \in s} d_v!} \frac{L(l)}{(|l|-1)!}.$$

Možemo se riješiti sume jer za svako stablo $s_l \sim l$ imamo jednaku vjerojatnost, a zbog leme 8.1 znamo da takvih stabala ima $\frac{\prod_{v\in l} d_v!}{L(l)}$. Dakle,

$$P(l' = l) = \frac{\prod_{v\in l} d_v!}{L(l)} \frac{e^{-|l|}}{\prod_{v\in s} d_v!} \frac{L(l)}{(|l|-1)!} = \frac{e^{-|l|}}{(|l|-1)!}.$$

□

Za kraj, bez dokaza navodimo poseban slučaj zakona o ukupnom potomstvu (eng. law of total progeny) za Poissonovu distribuciju.

Teorem 8.5. *Vjerojatnost generiranja obiteljskog stabla S procesom grananja s distribucijom grananja* Pois(λ) *takvo da vrijedi $|S| = n$ jednaka je*

$$\mathbb{P}(|S| = n) = \frac{(\lambda n)^{n-1}}{n!} e^{-\lambda n}.$$

Pokažimo sada kako odavde slijedi Cayleyeva formula.

Dokaz. Neka je zadano označeno stablo l s n vrhova, tj. $|l| = n$. Neka je S obiteljsko stablo nastalo procesom grananja s distribucijom grananja Pois(1), i neka je l' označeno stablo nastalo označavanjem stabla S. Izračunajmo vjerojatnost da na taj način dobijemo l, tj. $\mathbb{P}(l' = l \mid |l'| = n)$. Vrijedi

$$\mathbb{P}(l' = l \mid |l'| = n) = \frac{\mathbb{P}(l' = l)}{\mathbb{P}(|l'| = n)}.$$

Zbog teorema 8.5 vrijedi $\mathbb{P}(|l'| = n) = \frac{(\lambda n)^{n-1}}{n!} e^{-\lambda n} = \frac{n^{n-2} e^{-n}}{(n-1)!}$, a u lemi 8.4 već smo izračunali vjerojatnost da dobijemo označeno stablo l s proizvoljnim brojem vrhova. Dakle,

$$\mathbb{P}(l' = l \mid |l'| = n) = \frac{e^{-|l|}}{(|l|-1)!} \left(\frac{n^{n-2} e^{-n}}{(n-1)!} \right)^{-1} = \frac{n^{n-2} e^{-n}}{(n-1)!} = \frac{1}{n^{n-2}}.$$

Kako je gornja vjerojatnost jednaka za sva označena stabla s n vrhova, slijedi da takvih stabala ima n^{n-2}.

□

Literatura

[1] M. Aigner, G. M. Ziegler, *Proofs from THE BOOK (4th ed.)*, Springer, 2010.

[2] C. W. Borchardt, *Ueber eine der Interpolation entsprechende Darstellung der Eliminations-Resultante*, Journal für die reine und angewandte Mathematik 57 (1860), 111–121.

[3] P. J. Cameron, M. Kagan, *Number of spanning trees containing a given forest*, dostupno na `https://arxiv.org/pdf/2210.09009.pdf` (siječanj 2024.).

[4] A. Cayley, *A Theorem on Trees*, The Quarterly Journal of Mathematics 23 (1889), 376-378.

[5] P. Chassaing, J. F. Marckert, *Parking Functions, Empirical Processes, and the Width of Rooted Labeled Trees*, The Electronic Journal of Combinatorics 8 (2001), R14

[6] L. E. Clarke, *On Cayley's Formula for Counting Trees*, The Journal of the London Mathematical Society 33 (1958), 471-474.

[7] Ö. Eğecioğlu, J. B. Remmel, *Bijections for Cayley trees, spanning trees, and their q-analogues*, Journal of Combinatorial Theory A 42 (1986), 15–30.

[8] R. van der Hofstad, *Random Graphs and Complex Networks Volume 1*, Cambridge University Press, 2016.

[9] A. Joyal, *Une Théorie Combinatorie des Séries Formelles*, Advances in Mathematics 42 (1981), 1-82.

[10] A. G. Konheim, B. Weiss, *An Occupancy Discipline and Applications*, SIAM Journal on Applied Mathematics 14 (1966), 1266-1274.

[11] I. Nakić, *Predavanja iz Diskretne matematike*, dostupno na `https://web.math.pmf.unizg.hr/nastava/komb/predavanja/predavanja.pdf` (studeni 2023.).

[12] J. Pitman, *Coalescent Random Forests*, Journal of Combinatorial Theory A 85 (1997), 165-193.

[13] G. Pólya, *Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen*, Acta Mathematica 68 (1937), 145-254.

[14] H. Prüfer, *Neuer Beweis eines Satzes über Permutationen*, Archiv der Mathematik Und Physik (3) 27 (1918), 142-144.

[15] R. Pyke, *The supremum and infimum of the Poisson process*, The Annals of Mathematical Statistics 30 (1959), 568-576.

[16] A. Rényi, *Some Remarks on the Theory of Trees*, Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei 4 (1956), 73-85.

[17] J. Riordan, *Forests of Labeled Trees*, Journal of Combinatorial Theory 5 (1968), 90-103.

[18] J. Riordan, *Ballots and trees*, Journal of Combinatorial Theory 6 (1969), 408–411.

[19] B. Sudakov, *Graph Theory Assignment 2*, dostupno na `https://www2.math.ethz.ch/education/bachelor/lectures/fs2016/math/graph_theory/homework2.pdf` (siječanj 2024.).

[20] D. B. West, *Combinatorial mathematics*, Cambridge University Press, 2021.

[21] J. Willis, *An Introduction to Combinatorics via Cayley's Theorem*, dostupno na `https://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=9757&context=etd` (studeni 2023.).

Sažetak

U ovom radu proučavamo Cayleyevu formulu i njene dokaze. Najprije uvodimo osnovne pojmove teorije grafova i iskaz Cayleyeve formule. Zatim prelazimo na njene odabrane dokaze, počevši s bijektivnim dokazima: dokaz pomoću Prüferovih kodova, dokaz Joyalovom bijekcijom te dokaz pomoću funkcija parkiranja. Potom provodimo dokaze koji se temelje na matematičkoj indukciji: dokaz formulom uključivanja-isključivanja te Riordanov i Rényijev dokaz. Slijede dokazi dvostrukim prebrojavanjem: Pitmanov dokaz preko usmjerenih stabala, dokaz preko razapinjućih stabala i Clarkeov dokaz pomoću prespajanja. Zatim obrađujemo Pólyin dokaz preko funkcija izvodnica i Lagrangeove formule inverzije. Kao primjer matričnog dokaza dajemo dokaz Kirchhoffovim matričnim teoremom o stablima. Završavamo vjerojatnosnim dokazom preko Poissonovog procesa grananja.

Summary

In this thesis, we study Cayley's formula and its various proofs. First, we introduce basic graph theory notation and the statement of Cayley's formula. We then move on to a selection of its proofs, starting with bijective proofs: a proof using Prüfer codes, a proof using Joyal's bijection, and a proof using parking functions. Then we carry out proofs based on mathematical induction: a proof using the principle of inclusion and exclusion, and Riordan and Rényi's proof. Proofs by double counting follow: Pitman's proof via directed trees, a proof via spanning trees, and Clarke's proof using linkages of trees. We continue by covering Pólya's proof via generating functions and Lagrange's inversion formula. We also present a proof using Kirchhoff's matrix-tree theorem. We conclude with a probabilistic proof using a Poisson branching process.

Životopis

Rođena sam 5.11.1999. u Dubrovniku. U Blatu na Korčuli završila sam Osnovnu školu Blato, Osnovnu glazbenu školu pri osnovnoj školi Blato i opću gimnaziju u Srednjoj školi "Ivo Padovan" Blato. Upisala sam Preddiplomski sveučilišni studij Matematika na Prirodoslovno-matematičkom fakultetu u Zagrebu 2018., gdje sam 2021. i nastavila obrazovanje na diplomskom studiju Matematička statistika. Uz studij volontiram u udruzi Mladi nadareni matematičari "Marin Getaldić" i u projektu RADDAR, sviram u orkestru "Narodna glazba Blato" te pjevam u zboru Prirodoslovno-matematičkog fakulteta "Cantus Naturae".