Yhtenäisten joukkojen määrä rajoitetun asteen verkoissa

Juho-Kustaa Kangas

Helsinki 21.5.2012

Pro gradu -tutkielma

HELSINGIN YLIOPISTO
Tietojenkäsittelytieteen laitos

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta — Fakultet — Faculty	Laitos — Institution — Department
Matemaattis-luonnontieteellinen	Tietojenkäsittelytieteen laitos

Tekijä — Författare — Author
Juho-Kustaa Kangas

Työn nimi — Arbetets titel — Title
Yhtenäisten joukkojen määrä rajoitetun asteen verkoissa

Oppiaine — Läroämne — Subject
Tietojenkäsittelytiede

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Pro gradu -tutkielma	21.5.2012	55 sivua

Tiivistelmä — Referat — Abstract

Suuntaamattoman verkon *yhtenäinen joukko* on solmujen osajoukko, jonka indusoima aliverkko on yhtenäisen. Tällaisten joukkojen lukumäärä eri verkkoperheissä on viime aikoina osoittautunut kiinnostavaksi suureeksi erityisesti algoritmien analyysin kannalta. Esimerkiksi kauppamatkustajan ongelma sekä niin sanottu Tutte-polynomi voidaan ratkaista ajassa, joka on polynomisen kertoimen päässä verkon yhtenäisten joukkojen määrästä. Koska täydellisen verkon kaikki osajoukot ovat yhtenäisiä, kysymykseksi on noussut erityisesti yhtenäisten joukkojen määrän asymptoottinen kasvu verkoissa, joiden suurin aste on rajoitettu. Tällä hetkellä paras tunnettu yläraja asymptoottiselle kasvulle tällaisissa verkoissa seuraa eräänlaiseen paikalliseen analyysiin perustuvasta entropiamenetelmästä. Ylärajan laatu riippuu olennaisesti siitä, kuinka monella tavalla mielivaltainen yhtenäinen joukko voi leikata kunkin solmun suljetun naapuruston. Tunnettu alaraja puolestaan seuraa yksinkertaisesta äärettömästä verkkoperheestä, jossa yhtenäisten joukkojen määrä osataan määrittää tarkasti. Kokeellisten arvioiden perusteella vaikuttaa kuitenkin selvältä, ettei kumpikaan nykyisistä rajoista ole tiukka ja lukumäärän asymptoottinen kasvu vaatii siten tarkempaa analyysiä.

Käymme läpi tunnetut analyyttiset ja algoritmiset tulokset liittyen yhtenäisten joukkojen lukumäärään sekä yleisissä että rajoitetun asteen verkoissa. Laajennamme entropiaan perustuvaa menetelmää tarkastelemalla tavallisten naapurustojen sijaan solmujen yleistettyjä naapurustoja, jotka sisältävät kaikki enintään kiinnitetyn säteen $r \geq 2$ päässä olevat solmut. Näytämme koneelliseen etsintään nojaten, että leikkausten suhteellinen lukumäärä on tällöin pienempi ja johtaa hieman parempaan ylärajaan, kun verkon suurin aste Δ on 3 tai 4. Näytämme myös kokeellisiin havaintoihin vedoten, että laajennus ei oletettavasti paranna aikaisempaa ylärajaa millään $\Delta > 5$ tai $r > 2$. Avoimeksi kysymykseksi jää, voidaanko ylärajaa parantaa kehittämällä entropiaan perustuvaa menetelmää edelleen vai vaatiko tarkempi analyysi uudenlaista lähestymistapaa. Tarkastelemme kysymystä lopuksi toisesta suunnasta osoittamalla asymptoottiselle kasvulle alarajan yksinkertaisella konstruktiolla. Luomme myös katsauksen pieniin verkkoihin, joissa yhtenäisen joukkojen määrä on suurin mahdollinen, ja tarkastelemme lähemmin tällaisten verkkojen muita ominaisuuksia.

ACM Computing Classification System (CCS):
F.2.2 [Nonnumerical Algorithms and Problems],
G.2.1 [Combinatorics],
G.2.2 [Graph Theory]

Avainsanat — Nyckelord — Keywords
yhtenäinen joukko, enumeratiivinen verkkoteoria, entropia, Shearerin epäyhtälö

Säilytyspaikka — Förvaringsställe — Where deposited

Muita tietoja — övriga uppgifter — Additional information

Sisältö

1 Johdanto

Monet verkkoteorian keskeisistä kysymyksistä ovat luonteeltaan kombinatorisia ongelmia, joissa tarkastelemme verkkojen tai niissä määriteltyjen alirakenteiden perheitä. Tyypillisimpiä ongelmia ovat muun muassa erilaiset päätös- ja optimointiongelmat, joissa etsimme tietyn ominaisuuden sisältäviä rakenteita. Esimerkiksi niin sanotussa Hamiltonin syklin ongelmassa haluamme selvittää, onko annetussa verkossa sykliä, joka kulkee täsmälleen kerran jokaisen solmun kautta [Bol98]. Kauppamatkustajan ongelma puolestaan on tunnettu optimointiongelma, jossa etsimme painotetun verkon keveintä Hamiltonin sykliä.

Enumeratiivinen kombinatoriikka on kombinatoriikan klassinen haara, joka tutkii objektien lukumäärään liittyviä kysymyksiä. Voidaan esimerkiksi näyttää, että n-solmuisella r-säännöllisellä verkolla on enintään $i_r^n = (2^{r+1} - 1)^{n/2r}$ niin sanottua *riippumatonta* solmujen osajoukkoa (engl. *independent set*), jossa mitkään solmut eivät ole toistensa naapureita [Zha10]. Tämä pitkään avoin kysymys osoitettiin hiljattain palauttamalla tarkastelu aiempaan todistukseen kaksijakoisille verkoille. Aiempi tulos puolestaan todistettiin käyttämällä hyväksi entropiaan perustuvia menetelmiä, jotka ovat yleisesti osoittautuneet hyödylliseksi kombinatorisessa analyysissä [Rad01]. Vastaavien ylärajojen tiukkuus osoitetaan tyypillisesti konstruoimalla ääretön joukko objekteja ja näyttämällä, että sen jäsenet saavuttavat ylärajan täsmälleen. Esimerkiksi yläraja i_r^n on tiukka ja saavutetaan täsmälleen kaksijakoisten verkkojen $K_{r,r}$ erillisessä yhdisteessä.

Lukumäärään liittyvät tulokset ovat kiinnostavia paitsi puhtaasti kombinatorisessa mielessä, myös algoritmien suunnittelun ja analyysin kannalta. Esimerkiksi verkon väritysongelmalle tunnetaan algoritmeja, joiden aikavaativuus riippuu olennaisesti verkon maksimaalisten riippumattomien joukkojen tai maksimaalisten kaksijakoisten aliverkkojen määrästä [FGPS05]. Tulos, jonka nojalla Hamiltonin syklien määrä on 3-säännöllisillä verkoissa luokkaa $\mathcal{O}(2^{3n/8})$, puolestaan liittyy suoraan vastaavaan ratkaisuun kauppamatkustajan ongelmalle [Epp03]. Tällaisissa ratkaisuissa tarvitsemme usein ongelman algoritmista puolta, jossa haluamme luetella tai laskea annetun ominaisuuden sisältävät objektit mahdollisimman tehokkaasti. Siinä missä luettelemisen vaatima aika riippuu vähintään lineaarisesti lueteltavien objektien määrästä, objektien laskeminen voidaan joissain tapauksissa suorittaa olennaisesti nopeammin.

Enumeratiiviset ongelmat voidaan monissa tapauksissa määritellä kysymällä tietyn ominaisuuden sisältävien aliverkkojen tai indusoitujen aliverkkojen lukumäärää. Indusoidun aliverkon määrää yksikäsitteisesti jokin solmujen osajoukko, jossa kaksi solmua on yhdistetty kaarella, jos ja vain jos ne on yhdistetty alkuperäisessä verkossa. Esimerkiksi edellä mainittu riippumaton joukko voidaan määritellä joukkona, jonka indusoima aliverkko on tyhjä. Muita tutkittuja joukkoperheisiin liittyviä kysymyksiä ovat muun muassa *klikkien* [Woo07] ja *dominoivien* joukkojen [FGPS05] lukumäärä. Klikillä (engl. *clique*) tarkoitamme solmujen osajoukkoa, jonka indusoima aliverkko on täydellinen. Dominoiva joukko (engl. *dominating set*) puolestaan sisältää verkon jokaista solmua kohti joko solmun itse tai jonkin sen naapurisolmun. Edellä mainitut joukkoperheet tunnetaan monista **NP**-vaikeista optimointiongelmista. Erityisesti tällaisten joukkojen laskeminen algoritmisesti on sekä yleisessä että monissa erityistapauksissa osoittautunut #**P**-täydelliseksi ongelmaksi [KOU11].

Tässä tutkielmassa tarkastelemme lukumäärään liittyviä kysymyksiä vähemmän tutkitulle mutta luonnolliselle indusoitujen aliverkkojen perheelle. Sanomme, että solmujen osajoukko on *yhtenäinen joukko* (engl. *connected set*), jos sen indusoima aliverkko on yhtenäinen: toisin sanoen joukon mitkä tahansa kaksi solmua voidaan yhdistää toisiinsa polulla, jonka kaikki solmut kuuluvat joukkoon. Yksinkertaisesta määritelmästä huolimatta kiinnostus yhtenäisten joukkojen lukumäärää kohtaan on varsin tuoretta ja ensimmäiset merkittävät tulokset ovat vain muutaman vuoden takaa. Tutkimuksen ensisijainen motivaatio on ollut lukumäärän hyödyllisyys algoritmianalyysissä: esimerkiksi kauppamatkustajan ongelma n-solmuiselle verkolle voidaan hieman klassista $\mathcal{O}(2^n n^2)$-algoritmia muokkaamalla ratkaista ajassa, joka on polynomisen kertoimen päässä verkon yhtenäisten joukkojen määrästä [BHKK08b]. Samassa ajassa voidaan laskea muun muassa myös verkon niin sanottu Tutte-polynomi, joka yleistää lukuisia verkkoteoreettisia käsitteitä ja on yksi merkittävimmistä verkkoinvarianteista [BHKK08a].

Koska täydellisen n-solmuisen verkon kaikki 2^n osajoukkoa ovat selvästi yhtenäisiä, luonnollinen kysymys on, voimmeko antaa yhtenäisten joukkojen määrälle eitriviaalin ylärajan jossain merkittävässä harvojen verkkojen perheessä. Björklund ym. [BHKK08b] näyttävät, että yhtenäisten joukkojen määrä on n-solmuisessa verkossa luokkaa $\mathcal{O}(\beta_\Delta^n)$, missä $\beta_\Delta = (2^{\Delta+1} - 1)^{1/(\Delta+1)} < 2$ riippuu ainoastaan verkon suurimmasta asteesta Δ. Myös tämän ylärajan todistus perustuu entropiamenetelmään: keskeinen työkalu on Shearerin entropiaepäyhtälöstä johdettu projektiolause, joka antaa ylärajan joukkoperheen koolle, jos sen koolle osataan antaa tietynlainen paikallinen yläraja perusjoukon riittävän monessa osajoukossa. Perrier ym. [PIM08] antavat yhtenäisten joukkojen määrälle kokeelliseen tutkimukseen perustuvan arvion, jonka perusteella ylärajassa on kuitenkin vielä parantamisen varaa. Myös tunnetut alarajat ovat suhteellisen kaukana kokeellisesti määritetystä arviosta, joten kysymys vaikuttaa kaipaavan tarkempaa analyysiä kummastakin suunnasta. Olemme erityisesti kiinnostuneita siitä, voidaanko entropiaan perustuvaa menetelmää jollain tavalla parantaa vai vaatiiko tiukan ylärajan osoittaminen toisenlaista lähestymistapaa.

Luomme tässä tutkielmassa katsauksen yhtenäisten joukkojen lukumäärää koskeviin kysymyksiin sekä analyyttisestä että algoritmisesta näkökulmasta. Määrittelemme tutkimuskysymyksen sekä olennaiset käsitteet formaalisti luvussa 2. Käymme myös tarkemmin läpi tunnettuja verkko-ongelmia, joiden vaativuus riippuu olennaisesti yhtenäisten joukkojen määrästä. Luvussa 3 tarkastelemme yhtenäisten joukkojen laskemista ja luettelemista algoritmisena ongelmana. Annamme yksinkertaisen algoritmin, joka luettelee yhtenäiset joukot yleiselle verkolle, sekä esitämme tehokkaamman rekursiivisen tavan laskea yhtenäisten joukkojen lukumäärä puissa. Luvussa 4 käymme läpi ylärajaan liittyvät tulokset rajoitetun asteen verkoissa. Esitämme erityisesti tarkemmin Björklundin ym. entropiaan perustuvan menetelmän sekä parannamme sitä, kun verkon suurin aste $\Delta \in \{3, 4\}$. Arvioimme myös mahdollisuutta parantaa tulosta edelleen samankaltaisilla menetelmillä. Luvussa 5 lähestymme ongelmaa toisesta suunnasta osoittamalla alarajan yhtenäisten joukkojen enimmäismäärälle rajoitetun asteen verkoissa. Luomme myös katsauksen pieniin verkkoihin, joissa yhtenäisten joukkojen määrä on suurin mahdollinen, ja pyrimme täten karakterisoimaan verkkoperheitä, joiden tarkemmalla analyysillä voitaisiin saavuttaa tiukka alaraja.

1.1 Yleisistä merkinnöistä

Käytämme merkintöjä $\mathbb{R}$ ja $\mathbb{N}$ reaalilukujen ja luonnollisten lukujen joukoista. Erityisesti sovimme, että $0 \in \mathbb{N}$. Joukko $\mathbb{R}^+$ sisältää kaikki positiiviset (nollaa suuremmat) reaaliluvut. Puhuessamme mielivaltaisesta luvusta tarkoitamme aina mielivaltaista reaalilukua (tai luonnollista lukua silloin, kun se on asiayhteydessä selvää). Merkinnällä $\log a$ tarkoitamme aina luvun a 2-kantaista logaritmia.

Osajoukkomerkintöjä $\subset$ ja $\supset$ käytetään kirjallisuudessa vaihtelevasti merkitsemään joko aitoa sisältyvyyttä tai sisältyvyyttä sekä mahdollista yhtäsuuruutta. Tässä tutkielmassa käytämme jälkimmäistä tulkintaa. Toisin sanoen joukoille A ja B pätee $A \subset B$, jos ja vain jos $x \in A \Rightarrow x \in B$ kaikilla x. Erityisesti siis $A = B \Rightarrow A \subset B$. Merkinnällä 2^A tarkoitamme joukon A potenssijoukkoa eli sen kaikkien osajoukkojen perhettä ja merkinnällä $A^{(k)}$ joukon A kaikkien k-alkioisten osajoukkojen perhettä. Formaalisti $2^A = \{S : S \subset A\}$ ja $A^{(k)} = \{S : S \subset A, |S| = k\}$. Erityisesti pätee $2^A = \bigcup_{k=0}^{|A|} A^{(k)}$.

Oletamme funktion asymptoottisen kasvun käsitteen sekä $\mathcal{O}$-notaation lukijalle tutuksi. Määrittelemme nämä kuitenkin tässä eksaktisti pohjustaaksemme määritelmää vähemmän tunnetulle $\mathcal{O}^*$-notaatiolle. Olkoot $f, g : \mathbb{N} \to \mathbb{R}^+$ funktioita. Sanomme, että f kasvaa asymptoottisesti *enintään* yhtä nopeasti kuin g, jos on olemassa $n_0 \in \mathbb{N}$ ja $c > 0$ siten, että kaikilla $n > n_0$ pätee $f(n) \leq c \cdot g(n)$. Toisin sanoen jostain n_0 alkaen funktio f ylittää funktion g enintään jollain vakiokertoimella c. Vastaavasti sanomme, että f kasvaa asymptoottisesti *vähintään* yhtä nopeasti kuin g, jos on olemassa $n_0 \in \mathbb{N}$ ja $c > 0$ siten, että kaikilla $n > n_0$ pätee $f(n) \geq c \cdot g(n)$. Jos molemmat edellisistä toteutuvat, sanomme, että f ja g kasvavat asymptoottisesti yhtä nopeasti.

Käytämme tavanomaista merkintää $\mathcal{O}(g)$ kaikkien sellaisten funktioiden joukosta, jotka kasvavat asymptoottisesti enintään yhtä nopeasti kuin g. Toisin sanoen kaikille f, g pätee $f \in \mathcal{O}(g)$, jos ja vain jos f kasvaa enintään yhtä nopeasti kuin g. Selvästi pätee esimerkiksi $\log \in \mathcal{O}(n^2)$. Tässä n^2 tarkoittaa implisiittisesti funktiota $n \mapsto n^2$. Jos lauseke sisältää useita muuttujia, kontekstista ilmenee, minkä muuttujan suhteen funktion kasvua tarkastellaan. Vastaavasti käytämme merkintää $\Omega(g)$ sellaisten funktioiden joukosta, jotka kasvavat asymptoottisesti vähintään yhtä nopeasti kuin g, ja merkintää $\Theta(g)$ sellaisten funktioiden joukosta, jotka kasvavat täsmälleen yhtä nopeasti. Erityisesti siis $\Theta(g) = \mathcal{O}(g) \cap \Omega(g)$ kaikilla g.

Sanomme, että funktio $f : \mathbb{N} \to \mathbb{R}^+$ kasvaa eksponentiaalisesti, jos $f \in \Omega(a^n)$ jollain vakiolla $a > 1$. Eksponentiaalisesti kasvavien funktioiden tapauksessa olemme yleensä kiinnostuneita nimenomaan eksponentiaalisesta tekijästä, emme niinkään funktion mahdollisista polynomisista kertoimista. Määrittelemme tähän tarkoitukseen $\mathcal{O}^*$-notaation, joka piilottaa vakiokertoimien lisäksi myös kaikki polynomiset kertoimet. Esimerkiksi kauppamatkustajan ongelma ratkeaa dynaamisella ohjelmoinnilla ajassa $\mathcal{O}(2^n n^2) \subset \mathcal{O}^*(2^n)$. Formaalisti $f \in \mathcal{O}^*(g)$, jos ja vain jos $f \in \mathcal{O}(g \cdot p)$, missä p on polynomifunktio, kaikilla f, g. Vastaavasti määrittelemme luokan Θ^* siten, että $f \in \Theta^*(g)$, jos ja vain jos $f \in \Theta(g \cdot p)$. Jatkossa tarkoitamme polynomifunktioilla aina funktioita, jotka kasvavat polynomisesti suhteessa verkon solmujen määrään.

Verkkoihin liittyvät merkinnät käsittelemme tarkemmin luvussa 2.1.

2 Yhtenäiset joukot

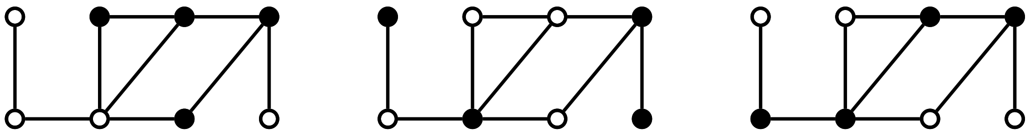
Tässä luvussa pohjustamme kysymyksiä, joita käsittelemme myöhemmissä luvuissa. Luvussa 2.1 käymme ensin läpi yleisiä verkkoihin liittyviä määritelmiä. Erityisesti määrittelemme täsmällisesti yhtenäiset joukot ja rajoitetun asteen verkot, jotka ovat tutkielman keskeisimmät käsitteet. Luvussa 2.2 esittelemme ongelmia, joiden aikavaativuus riippuu yhtenäisten joukkojen määrästä. Luvussa 2.3 määrittelemme formaalisti yhtenäisiin joukkoihin liittyvät ongelmat, joihin tulemme vastaamaan luvuissa 3–5.

2.1 Määritelmiä

Verkko on formaalisti järjestetty pari $G = (V, E)$, missä V on äärellinen joukko ja $E \subset V^{(2)}$. Kutsumme joukon V alkioita *solmuiksi* ja joukon E alkioita *kaariksi*. Käytämme kaaresta $\{u, v\}$ jatkossa lyhyempää merkintää uv tai määritelmän nojalla yhtäpitävästi vu. Sanomme, että kaari uv *yhdistää* solmut u ja v ja että nämä solmut ovat toistensa *naapureita*.

Sanomme, että solmujen jono $u_1 \ldots u_k$ on *polku* (engl. *path*), jos pätee $u_i \neq u_j$ kaikilla $i \neq j$ ja $u_i u_{i+1} \in E$ kaikilla $1 \leq i \leq k-1$. Tällöin polun *pituus* on $k-1$, solmut u_1 ja u_k ovat polun *päätepisteet* ja polku kulkee näiden solmujen *välillä*. Solmujen $u, v \in V$ välinen *etäisyys* $d(u, v)$ on lyhimmän polun pituus niiden välillä. Jos solmujen u, v välillä ei ole polkua, merkitsemme $d(u, v) = \infty$. Verkon *läpimitta* (engl. *diameter*) $\delta(G)$ on suurin etäisyys kahden solmun välillä: $\delta(G) = \max\{d(u, v) : u, v \in V\}$. Edelleen $\delta(G) = \infty$, jos $d(u, v) = \infty$ jollain $u, v \in V$. Sanomme, että G on *yhtenäinen* (engl. *connected*), jos kaikkien solmujen välillä on polku, toisin sanoen $\delta(G) \neq \infty$. Jos polulle $u_1 \ldots u_k$ pätee $k \geq 3$ ja $u_k u_1 \in E$, sanomme, että $u_1 \ldots u_k$ on *sykli* (engl. *cycle*). Toisin kuin poluilla, syklin pituudella tarkoitamme sen solmujen määrää k. Verkon *leveys* (engl. *girth*) $g(G)$ on sen lyhimmän syklin pituus. Jos verkko ei sisällä syklejä, merkitsemme $g(G) = \infty$. Nimitämme syklitöntä verkkoa *metsäksi* (engl. *forest*) ja syklitöntä yhtenäistä verkkoa *puuksi* (engl. *tree*).

Olkoon $u \in V$ ja $U \subset V$. Merkitsemme $N(u) = \{v \in V : uv \in E\}$ ja sanomme, että joukko $N(u)$ on solmun u (avoin) *naapurusto* (engl. *neighborhood*). Vastaavasti sanomme, että $N[u] = N(u) \cup \{u\}$, joka sisältää naapuruston lisäksi solmun u itse, on sen *suljettu naapurusto*. Yleistämme nämä määritelmät mielivaltaiselle solmujen osajoukolle $U \subset V$ merkitsemällä $N[U] = \bigcup_{u \in U} N[u]$ ja $N(U) = N[U] \setminus U$. Erityisesti siis $N(\{u\}) = N(u)$ ja $N[\{u\}] = N[u]$ kaikilla $u \in V$. Merkitsemme myös



Kuva 1: Yhtenäinen joukko, dominoiva joukko ja yhtenäinen dominoiva joukko.

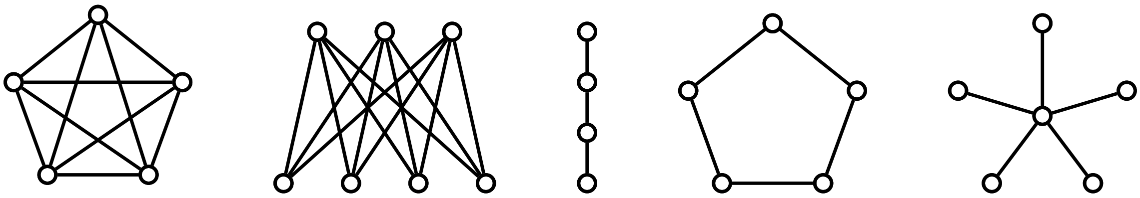
$N^0[u] = u$ ja rekursiivisesti $N^{k+1}[u] = N[N^k[u]]$ kaikilla $k \in \mathbb{N}$. Lisäksi merkitsemme $N^{k+1}(u) = N^{k+1}[u] \setminus N^k[u]$ kaikilla $k \in \mathbb{N}$. Toisin sanoen kaikilla $k \geq 0$ joukko $N^k(u)$ sisältää solmut, joiden etäisyys solmusta u on täsmälleen k, ja joukko $N^k[k]$ solmut, joiden etäisyys solmusta u on enintään k.

Solmun $v \in V$ *aste* (engl. *degree*) on sen naapurisolmujen lukumäärä $d(v) = |N(v)|$. Verkon *suurin aste* (engl. *maximum degree*) on $\Delta(G) = \max_{v \in V} d(v)$. Jos verkon kaikilla solmuilla on sama aste r, sanomme, että verkko on r-*säännöllinen* (engl. *regular*). Rajoitetun asteen verkoilla (engl. *bounded degree graph*) tarkoitamme verkkoperheitä $\mathcal{G}_\Delta = \{G \in \mathcal{G} : \Delta(G) \leq \Delta\}$, missä $\mathcal{G}$ on kaikkien verkkojen perhe ja $\Delta \in \mathbb{N}$. Perhe $\mathcal{G}_\Delta$ sisältää siis kaikki verkot, joiden suurin aste on enintään Δ, ja erityisesti kaikki Δ-säännölliset verkot. Merkinnällä $\mathcal{G}_{\Delta,n}$ tarkoitamme kaikkien n-solmuisten Δ-rajoitettujen verkkojen perhettä.

Sanomme, että verkko (V', E') on verkon G aliverkko, jos pätee $V' \subset V$ ja $E' \subset E$. Määrittelemme jokaiselle osajoukolle $U \subset V$ verkon G aliverkon $G[U] = (U, E[U])$, missä $E[U] = \{uv \in E : u, v \in U\}$, ja kutsumme tätä joukon U *indusoimaksi* aliverkoksi (engl. *induced subgraph*). Jos verkko $G[U]$ on yhtenäinen, sanomme, että joukko U on verkon G *yhtenäinen joukko* (engl. *connected set*). Toisin sanoen U on yhtenäinen, jos ja vain jos kaikilla $u, v \in U$ on olemassa polku $u \ldots v$, jonka kaikki solmut ovat joukossa U (kuva 1). Erityisesti tyhjä joukko on yhtenäinen. Sanomme lisäksi, että U on *dominoiva joukko* (engl. *dominating set*), jos se leikkaa verkon jokaisen solmun suljetun naapuruston: toisin sanoen jokaisella $v \in V \setminus U$ on naapurisolmu joukossa U. Käytämme merkintää $\mathcal{C}$ verkon yhtenäisten ja $\mathcal{D}$ dominoivien joukkojen perheistä. Olkoon lisäksi $\mathcal{E} = \mathcal{C} \cap \mathcal{D}$, siis sellaisten joukkojen perhe, jotka ovat sekä yhtenäisiä että dominoivia.

Merkitsemme edellä määriteltyjä ominaisuuksia $\delta(G), g(G), \mathcal{C}(G), \ldots$ lyhyemmin $\delta, g, \mathcal{C}, \ldots$, kun on selvää, mistä verkosta on kyse. Määrittelemme lopuksi kuusi yksinkertaista parametrisoitua verkkoperhettä (kuva 2), joihin viittaamme myöhemmin, sekä annamme lausekkeen niiden yhtenäisten joukkojen määrälle.

Tyhjä verkko (engl. *empty graph*) E_n on n-solmuinen verkko, joka ei sisällä lainkaan



Kuva 2: Esimerkkejä parametrisoiduista verkoista: K_5, $K_{4,3}$, P_4, C_5, S_6.

kaaria. Tyhjän verkon yhtenäiset joukot ovat täsmälleen tyhjä joukko sekä jokaisen solmun yksiö. Siten $\mathcal{C}(E_n) = n + 1$.

Täydellinen verkko tai *täysi verkko* (engl. *complete graph*) K_n on n-solmuinen verkko, jonka jokainen solmupari on yhdistetty kaarella. Täten jokainen solmujen osajoukko on yhtenäinen ja pätee $\mathcal{C}(K_n) = 2^n$.

Täydellinen kaksijakoinen verkko (engl. *complete bipartite graph*) $K_{k,m}$ sisältää solmut $u_1, \ldots u_k, u'_1, \ldots, u'_m$ sekä kaaren $u_i u'_j$ kaikilla $1 \leq i \leq k$ ja $1 \leq j \leq m$. Solmujen joukko on epäyhtenäinen täsmälleen silloin, kun se sisältää ainakin kaksi solmua ja kaikki solmut kuuluvat joko joukkoon $\{u_1, \ldots, u_k\}$ tai joukkoon $\{u'_1, \ldots, u'_m\}$. Täten $\mathcal{C}(K_{k,m}) = 2^{k+m} - 2^k - 2^m + k + m + 2$.

Polkuverkko (engl. *path graph*) P_n sisältää solmut $u_1, \ldots, u_n$ ja kaaren $u_i u_{i+1}$ kaikilla $1 \leq i \leq n - 1$. Tyhjää joukkoa lukuun ottamatta jokainen yhtenäinen joukko sisältää täsmälleen solmut $u_i, \ldots, u_j$ yksikäsitteisillä $1 \leq i \leq j \leq n$. Solmut u_i ja u_j voidaan selvästi valita $\binom{n+1}{2} = \frac{n(n+1)}{2}$ tavalla ja siten $\mathcal{C}(P_n) = \frac{n(n+1)}{2} + 1$.

Sykliverkko (engl. *cycle graph*) C_n saadaan lisäämällä polkuverkkoon P_n kaari $u_n u_1$. Tyhjää ja kaikkien solmujen joukkoa lukuun ottamatta jokainen yhtenäinen joukko indusoi polkuverkon: kaikilla $1 \leq i \leq n$ ja $1 \leq k \leq n - 1$ saadaan yksi k-solmuinen polku, joka ”alkaa” solmusta u_i. Siten $\mathcal{C}(C_n) = n(n - 1) + 2$.

Tähtiverkko (engl. *star graph*) S_n sisältää solmut $u_1, \ldots, u_n$ ja kaaren $u_1 u_i$ kaikilla $2 \leq i \leq n$. Tähtiverkko on isomorfinen kaksijakoisen verkon $K_{1,n-1}$ kanssa ja siten $\mathcal{C}(S_n) = 2^{n-1} + n$.

2.2 Ajassa $\mathcal{O}^*(|\mathcal{C}|)$ ratkeavia ongelmia

Ensisijainen mielenkiintomme yhtenäisten joukkojen määrää kohtaan liittyy analyyttisiin tuloksiin, joiden nojalla tietyt verkko-ongelmat voidaan ratkaista perheen $\mathcal{C}$ koosta riippuvassa ajassa. Kuten Björklund ym. [BHKK08b] näyttävät, tämä on helppo todeta esimerkiksi kauppamatkustajan ongelmalle (engl. *travelling salesman problem, TSP*), joka kysyy painotetun verkon keveintä Hamiltonin sykliä. Klassinen dynaamista ohjelmointia hyödyntävä algoritmi ratkaisee ongelman n-solmuiselle verkolle (V, E) etsimällä kaikille $U \subset V$ ja $u \in U$ optimaalisen polun, joka sisältää joukon U jokaisen solmun tasan kerran ja päättyy solmuun u. Muodostamalla jokaisen osapolun yhtä solmua lyhyemmästä jo lasketusta polusta algoritmi löytää optimaalisen silmukan ajassa ja tilassa $\mathcal{O}(n^2|2^V|)$, missä $|2^V| = 2^n$ seuraa suoraan osajoukkojen U lukumäärästä. Koska on selvää, että ainoastaan yhtenäinen osajoukko voi sisältää polun, joka kulkee sen jokaisen solmun kautta, riittää laskea optimaaliset polut kaikille $U \in \mathcal{C}$. Kuten näytämme luvussa 3, yhtenäiset joukot voidaan luetella ajassa $\mathcal{O}(n|\mathcal{C}|)$. Soveltamalla dynaamista ohjelmointia näiden yli TSP on siten mahdollista ratkaista ajassa ja tilassa $\mathcal{O}^*(|\mathcal{C}|)$.

Aikavaativuuden $\mathcal{O}^*(|\mathcal{C}|)$ toteaminen ei tässä tapauksessa ole vielä itsessään erityisen merkittävä havainto. Erityisesti se ei sano mitään ongelman yleisestä vaativuudesta, koska pahimmassa tapauksessa (esimerkiksi täysillä verkoilla) pätee $|\mathcal{C}| \in \Theta(2^n)$. Björklund ym. antavat kauppamatkustajan ongelmalle myös huomattavasti nopeamman algoritmin, jonka perusajatus on välttää epäyhtenäisten joukkojen lisäksi muun muassa sellaiset joukot, joissa kaikki solmut sisältävä polku ajautuu väistämättä ”umpikujaan”. Lisäksi jos kaarten painot ovat vakiolla rajoitettuja kokonaislukuja, he näyttävät, että ongelma voidaan ratkaista polynomisessa tilassa ja ajassa $\mathcal{O}^*(|\mathcal{C} \cap \mathcal{D}|)$. Täten olemme yhtenäisten joukkojen lisäksi kiinnostuneita sellaisten joukkojen määrästä, jotka ovat sekä yhtenäisiä että dominoivia.

Osoittautuu, että verkon Tutte-polynomi voidaan niin ikään laskea ajassa $\mathcal{O}^*(|\mathcal{C}|)$ [BHKK08a]. Tutte-polynomi on yleisin verkkoinvariantti, joka voidaan laskea niin sanotulla poisto–kontraktio-algoritmilla, ja sisältää siten erikoistapauksena kaikki samalla menetelmällä laskettavat invariantit, kuten kromaattisen polynomin, luotettavuuspolynomit ja vuopolynomin. Verkkoteorian ulkopuolella sillä on useita sovelluksia muun muassa kombinatoriikassa, solmuteoriassa ja tilastollisessa fysiikassa. Poisto–kontraktio-algoritmi laskee n-solmuisen verkon G Tutte-polynomin ajassa $\mathcal{O}^*(\tau(G))$, missä $\tau(G)$ on verkon G virittävien puiden määrä. Täysillä verkoilla pätee $\tau(G) = n^{n-2}$, joten $\mathcal{O}^*(|\mathcal{C}|)$-aika on huomattava parannus.

Viimeisin tuntemamme algoritmi, jonka aikavaativuus on $\mathcal{O}^*(|\mathcal{C}|)$, on Perrierin ym. algoritmi [PIM08], joka oppii optimaalisen Bayes-verkon annetusta aineistosta. Algoritmin keskeinen ajatus on etsiä ensin tietynlainen ”ylirakenne”, joka todistettavasti sisältää optimaalisen verkon aliverkkonaan. Sopiva ylirakenne on yleensä olennaisesti helpompi löytää kuin etsiä optimaalinen aliverkko suoraan kaikkien mahdollisuuksien joukosta. Kun ylirakenne tunnetaan, optimaalinen verkko löydetään suorittamalla etsintä sen yhtenäisten joukkojen yli. Siten algoritmin aikavaativuus on polynomisen kertoimen päässä ylirakenteen yhtenäisten joukkojen määrästä.

2.3 Tutkimuskysymys

Edellisten esimerkkien valossa yhtenäisten joukkojen määrän arvioiminen on algoritmianalyysin kannalta kiinnostava ongelma. Koska täydessä n-solmuisessa verkossa K_n kaikki 2^n osajoukkoa ovat yhtenäisiä, on luonnollista kysyä, miten yhtenäisten joukkojen määrä käyttäytyy, jos kaikkien verkkojen sijaan rajoitumme johonkin merkittävään verkkoperheeseen. On intuitiivisesti selvää, että verkkoperheen jäsenten on oltava jossain mielessä ”harvoja”, jotta $|\mathcal{C}|$ kasvaisi olennaisesti hitaammin kuin kaikkien osajoukkojen määrä. Yksi määritelmä verkkoperheen harvuudelle on, että kaarien määrä on enintään lineaarinen solmujen määrään nähden. Tämä ei kuitenkaan vielä riitä rajoittamaan yhtenäisten joukkojen määrää. Yksinkertainen esimerkki tästä on luvussa 2.1 esitelty tähtiverkko, jossa yksi solmu on kaikkien muiden solmujen naapuri ja muut solmut ainoastaan sen naapureita. Koska S_n sisältää täsmälleen $n-1$ kaarta, tähtiverkot ovat selvästi harvoja. Erityisesti verkon $S_n = (V, E)$ *keskimääräinen aste* on $(\sum_{v \in V} d(v))/|V| = 2|E|/|V| = 2(n-1)/n < 2$ kaikilla $n \in \mathbb{N}$. Tästä huolimatta yhtenäisten joukkojen määrä on luokkaa $\mathcal{C}(S_n) = 2^{n-1} + n \in \Theta(2^n)$.

Edellisen esimerkin nojalla harvuus tai edes matala keskimääräinen aste ei olennaisesti rajoita perheen $\mathcal{C}$ kokoa. Tähtiverkon tapauksessa yhtenäisten joukkojen suureen määrään riittää, että ainoastaan yhdellä solmulla on rajoittamaton aste. Sen sijaan osoittautuu, että yhtenäisten joukkojen määrää voidaan rajoittaa olennaisesti, jos suljemme tällaiset solmut pois kokonaan tarkastelemalla ainoastaan rajoitetun asteen verkkoperheitä $\mathcal{G}_\Delta$. Formaalisti määritellen olemme kiinnostuneita luvusta

$$C_{\Delta,n} = \max\{\mathcal{C}(G) : G \in \mathcal{G}_{\Delta,n}\}$$

kaikilla Δ ja n, siis yhtenäisten joukkojen suurimmasta mahdollisesta määrästä n-solmuisessa Δ-rajoitetussa verkossa.

Osoittautuu, että n-solmuiset verkot, joilla $|\mathcal{C}|$ on täsmälleen $C_{\Delta,n}$, ovat usein monella tavalla symmetrisiä ja siten teoreettisesti mielenkiintoisia. Käsittelemme tällaisia verkkoja pienillä Δ ja n tarkemmin luvussa 5. Algoritmianalyysin kannalta olemme kuitenkin ensisijaisesti kiinnostuneita siitä, miten $C_{\Delta,n}$ käyttäytyy asymptoottisesti, kun n kasvaa jollain kiinteällä Δ. Kuten osoitamme luvussa 5, $C_{\Delta,n}$ kasvaa eksponentiaalisesti kaikilla $\Delta \geq 3$. Ei ole kuitenkaan selvää, että välttämättä pätisi $C_{\Delta,n} \in \Theta(a_\Delta^n)$ tai edes $C_{\Delta,n} \in \Theta^*(a_\Delta^n)$ jollain vakiolla a_Δ. Koska olemme ensisijaisesti kiinnostuneita nimenomaan eksponentiaalisesta tekijästä, etsimme täten arviota luvulle

$$\sigma_\Delta = \inf\{a \in \mathbb{R} : C_{\Delta,n} \in \mathcal{O}(a^n)\}$$

kaikilla $\Delta \geq 3$. Toisin sanoen etsimme suurinta lukua, jolla ei ole pienempää lukua b siten, että $C_{\Delta,n} \in \mathcal{O}(b^n)$. Koska kantalukujen $a \in \mathbb{R}$ joukko on epätyhjä, infimum on reaalilukuaksioomien nojalla olemassa ja σ_Δ on siten hyvin määritelty.

Björklund ym. [BHKK08b] näyttävät, että $\sigma_\Delta \leq \beta_\Delta$, missä $\beta_\Delta = (2^{\Delta+1}-1)^{1/(\Delta+1)} < 2$ kaikilla $\Delta \geq 3$. Tämä on toistaiseksi paras todistettu ja julkaistu yläraja. Perrier ym. [PIM08] ovat tutkineet kysymystä kokeellisesti generoimalla satunnaisia verkkoja pienillä n ja Δ ja laskemalla niiden yhtenäiset joukot. Tämän katsauksen perusteella vaikuttaa uskottavalta, että $\sigma_3 \approx 1.81, \sigma_4 \approx 1.92$ ja $\sigma_5 \approx 1.96$. Vertailun vuoksi pätee $\beta_3 \approx 1.9680, \beta_4 \approx 1.9874$ ja $\beta_5 \approx 1.9948$. Tunnettu yläraja vaikuttaa siten olevan vielä huomattavan kaukana todellisesta σ_Δ. Alarajalle emme tunne lainkaan julkaistua tulosta, joskin on varsin helppo osoittaa, että pätee $\sigma_\Delta > 1$ kaikilla $\Delta \geq 3$.

Björklundin ym. todistuksen keskeinen työkalu on entropiaan perustuva projektiolause, joka antaa ylärajan joukkoperheen koolle, jos ylärajaa osataan arvioida paikallisesti perusjoukon osajoukoissa. Esitämme projektiolauseen sekä ylärajatodistuksen luvuissa 4.1 ja 4.2. Luvussa 4.3 laajennamme projektiolauseeseen perustuvaa analyysiä tarkastelemalla solmujoukon suurempia osajoukkoja. Luvussa 4.4 näytämme, että laajennus antaa paremman ylärajan, kun suurin aste $\Delta \in \{3, 4\}$, sekä arvioimme mahdollisuutta laajentaa analyysiä edelleen. Yhtenäisten joukkojen lisäksi käsittelemme lyhyesti myös yhtenäisten dominoivien joukkojen määrää ja parannamme niiden ylärajaa niin ikään tapauksessa $\Delta \in \{3, 4\}$.

Luvussa 5 esitämme alarajan luvulle σ_Δ kaikilla $\Delta \geq 3$ konstruoimalla rajoitetun asteen verkkoperheitä ja antamalla asymptoottisesti tiukan arvion niiden yhtenäisten joukkojen määrälle. Esitämme myös arvioita verkkoperheistä, joissa empiirisen tutkimuksen perusteella otaksumme yhtenäisten joukkojen määrän maksimoituvan, vaikkakaan emme osaa antaa niille tiukkaa arviota. Tätä ennen käsittelemme luvussa 3 ongelman algoritmista puolta, yhtenäisten joukkojen eksaktia luettelemista ja laskemista annetulle verkolle. Emme tunne käytännön sovelluksia, joissa yhtenäisten joukkojen tarkka laskeminen olisi tarpeellista, mutta hyödynnämme näitä menetelmiä muun muassa ylärajan koneellisessa analyysissä, jonka esitämme luvussa 4.

3 Luetteleminen ja laskeminen

Olkoon $\mathcal{A}$ äärellinen joukko. Sanomme, että perheen luetteluongelma (engl. *enumeration*) on tuottaa lista, joka sisältää joukon $\mathcal{A}$ jokaisen jäsenen tasan kerran, ja laskentaongelma (engl. *counting*) on laskea jäsenten lukumäärä $|\mathcal{A}|$. Luetteleminen vaatii määritelmällisesti vähintään $|\mathcal{A}|$ askelta ja ratkaisee samalla myös laskentaongelman. Toisaalta laskentaongelma on joissain tapauksissa mahdollista ratkaista olennaisesti nopeammin kuin luettelemalla.

Yhtenäisten joukkojen tapauksessa eksaktia luettelemista ja laskemista voidaan pitää eräänlaisina metaongelmia. Kombinatoriikan kannalta niillä on lähinnä teoreettista mielenkiintoa, mutta ne ovat hyödyllisiä työkaluja koneellista laskentaa hyödyntävässä analyysissä, kuten osoitamme luvussa 4. Tässä luvussa näytämme, että yhtenäiset joukot voidaan luetella mielivaltaiselle verkolle ajassa $\mathcal{O}^*(|\mathcal{C}|)$ ja laskea mielivaltaiselle puulle ajassa $\mathcal{O}(n)$.

3.1 Yleinen verkko

Mielivaltaisen n solmua ja m kaarta sisältävän verkon yhtenäisyys voidaan tarkistaa ajassa $\mathcal{O}(n+m)$ suorittamalla syvyyssuuntainen haku jostain solmusta. Siten yhtenäiset joukot voidaan luetella ajassa $\mathcal{O}((n+m)|2^V|) = \mathcal{O}^*(2^n)$ luettelemalla kaikki $U \subset V$ ja tarkistamalla kunkin $G[U]$ yhtenäisyys erikseen. Raa'an läpikäynnin lisäksi tunnemme kirjallisuudessa kaksi muuta menetelmää. Avis ja Fukuda [AF93] esittävät yleisen käänteishauksi (engl. *reverse search*) kutsutun tekniikan, joka soveltuu useisiin kombinatorisiin luetteluongelmiin. Sopivasti toteutettuna se toimii ajassa $\mathcal{O}(p(k)|\mathcal{A}|)$ ja tilassa $\mathcal{O}(p(k))$, missä $\mathcal{A}$ on luteltava perhe, k syötteen koko ja $p(k)$ sen polynomi. Erityisesti menetelmä luettelee yhtenäiset joukot ajassa $\mathcal{O}(nm|\mathcal{C}|)$. Tämä on myös ensimmäinen tuntemamme viittaus yhtenäisten joukkojen luetteluongelmaan.

Paras tunnettu algoritmi luettelee verkon yhtenäiset joukot ajassa $\mathcal{O}(n|\mathcal{C}|)$ ja tilassa $\mathcal{O}(n+m)$. Eksplisiittisesti sen ovat muotoilleet ensimmäisen kerran Gutin ym. [GJR$^+$07], jotka luettelevat erityisesti myös niin sanotut konveksit yhtenäiset joukot. Algoritmi perustuu siihen yksinkertaiseen havaintoon, että jokainen yhtenäinen joukko C voidaan konstruoida täydentämällä tyhjää joukkoa (tai mitä tahansa joukon C yhtenäistä osajoukkoa) solmu kerrallaan siten, että yhtenäisyys säilyy jokaisessa välivaiheessa. Myös Björklund ym. [BHKK08b] esittävät tämän periaatteen, mutta eivät muotoile ratkaisua formaalisti.

Lause 1. *Olkoon $G = (V, E)$ verkko. Perheen $\mathcal{C}(G)$ jäsenet voidaan luetella ajassa $\mathcal{O}(n|\mathcal{C}(G)|)$.*

Esitämme oman muotoilumme tälle algoritmille ja todistamme sen oikeellisuuden ja aikavaativuuden. Otamme ensin käyttöön uuden määritelmän: olkoot A ja B joukkoja, joilla $A \subset B$. Merkitsemme $[A, B] = \{U : A \subset U \subset B\}$ ja sanomme, että $[A, B]$ on joukkojen A ja B *väli*. Osoitamme nyt seuraavan lemman.

Lemma 2. *Olkoon $G = (V, E)$ verkko, $A \subset B \subset V$ ja $A \in \mathcal{C}(G)$. Yhtenäiset joukot välillä $[A, B]$ voidaan luetella ajassa $\mathcal{O}(n|[A, B] \cap \mathcal{C}(G)|)$.*

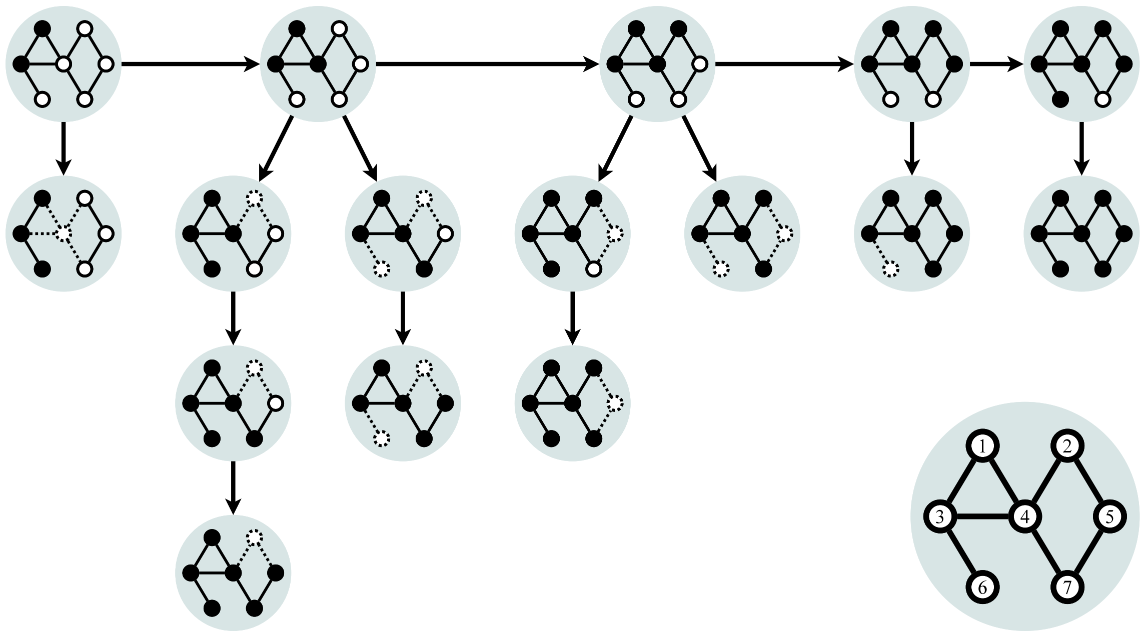
Ratkaisemme lemmassa 2 muotoillun osaongelman tulostamalla ensin joukon A. Tämän jälkeen käymme läpi kaikki solmut $u_1, \ldots, u_k$, joilla voimme laajentaa joukkoa A säilyttäen yhtenäisyyden, toisin sanoen $u_i \in B \setminus A$ ja $A \cup \{u_i\} \in \mathcal{C}$ kaikilla $1 \leq i \leq k$. Jos $A = \emptyset$, minkä tahansa solmun lisääminen säilyttää yhtenäisyyden. Muussa tapauksessa lisättäviä solmuja ovat täsmälleen ne solmut, joilla on naapurisolmu joukossa A, toisin sanoen kaikki $u \in N(A) \cap B$. Luetellaksemme loput joukot sovellamme nyt osaongelman ratkaisua rekursiivisesti: jokaisella $1 \leq i \leq k$ luettelemme kaikki $U \in \mathcal{C}$ välillä $[A \cup \{u_i\}, B \setminus \{u_j : j < i\}]$. Jokaisessa rekursiohaarassa laajennamme siis joukkoa A yhdellä solmulla. Lisäksi jätämme joukosta B pois edellisissä haaroissa lisätyt solmut. Täten varmistamme, ettei sama joukko tule lueteltua kahdessa eri haarassa. Esimerkki algoritmin toiminnasta on esitetty kuvassa 3.

Väitämme, että edellä kuvattu algoritmi toimii oikein, toisin sanoen luettelee täsmälleen perheen $[A, B] \cap \mathcal{C}$ joukot, kunkin täsmälleen yhden kerran.

Todistus. Induktio:

1) Tapaus $A = B$: Algoritmi luettelee selvästi vain joukon A, koska lisättäviä solmuja ei ole. Siten triviaalisti jokainen lueteltu joukko kuuluu perheeseen $[A, B] \cap \mathcal{C}$ ja luetellaan täsmälleen kerran.

2) Tapaus $A \subsetneq B$: Oletamme nyt, että rekursiiviset kutsut toimivat oikein. Koska jokainen rekursiivinen kutsu luettelee välin $[A, B]$ osajoukon, selvästi jokainen lueteltu joukko kuuluu välille $[A, B]$. Olkoon nyt $U \in [A, B] \cap \mathcal{C}$. Näytämme, että U luetellaan ainakin kerran. Jos $U = A$, väite pätee triviaalisti. Muuten koska U on yhtenäinen ja $A \subset U$, pätee $u_i \in U$ jollain $1 \leq i \leq k$. Olkoon m pienin luku, jolla $u_m \in U$. Tällöin $U \in [A \cup \{u_m\}, B \setminus \{u_j : j < m\}]$ ja siten U luetellaan haarassa m. Lopuksi näytämme, että jokainen joukko luetellaan korkeintaan yhdessä haarassa. Teemme vastaoletuksen, jonka mukaan $U \subset [A, B]$ luetellaan kahdessa eri rekursiohaarassa.



Kuva 3: Algoritmin suorituspuu eräällä 7 solmun verkolla. Joukkoon A kuuluvat solmut on merkitty mustalla ja joukkoon B kuulumattomat solmut katkoviivalla. Puun jokainen solmu vastaa yhtä rekursiokutsua ja siten yhtä yhtenäistä joukkoa. Kutsut suoritetaan samassa järjestyksessä kuin vastaavat solmut on numeroitu.

Olkoon u solmu, joka lisätään joukkoon A näistä haaroista ensimmäisessä. Siten $u \in U$. Toisaalta koska U luetellaan myös toisessa haarassa, jossa solmua u ei lisätä, pätee $u \notin U$, mikä on ristiriita. Siten kukin joukko luetellaan korkeintaan kerran. $\square$

Väitetyn aikavaativuuden näyttämiseksi toteamme ensin, että rekursiivisten kutsujen määrä on selvästi luokkaa $\mathcal{O}(|[A, B] \cap \mathcal{C}(G)|)$. Toteutuksessamme joukot A ja B talletetaan n-alkioisina boolean-taulukoina, jolloin solmun kuuluminen joukkoon voidaan tarkistaa vakioajassa. Taulukoista ei tehdä kopioita, vaan niitä muokataan ennen rekursiivista kutsua ja muokatut taulukot välitetään viitteinä. Ennen kutsua lisäämme yhden solmun joukkoon A ja kutsun jälkeen poistamme sen joukoista A ja B. Ennen kutsun päättymistä kaikki kutsun aikana joukosta B poistetut solmut lisätään takaisin. Yhteensä nämä operaatiot vaativat ajan $\mathcal{O}(n)$. Jokaisessa kutsussa joudumme käymään läpi joukon A naapurisolmut, mikä vaatii naiivisti toteutettuna ajan $\mathcal{O}(n^2)$ tai vieruslistaesitystä käyttäen $\mathcal{O}(\Delta n)$. Voimme kuitenkin ylläpitää joukon A lisäksi myös joukkoa A_N siten, että aina pätee $A_N = N(A)$. Lisätessämme solmun joukkoon A, poistamme sen joukosta A_N ja lisäämme sen naapurit joukkoon A_N. Pidämme kirjaa muutoksista, jotta voimme kumota ne kutsun päätyttyä. Tämä operaatio vaatii ajan $\mathcal{O}(n)$ ja suoritetaan kerran jokaista kutsua kohti. Vastaavasti joukon $N(A)$ alkioiden läpikäynti vaatii nyt ajan $\mathcal{O}(n)$. Siten lopullinen vaativuus on $\mathcal{O}(n|[A, B] \cap \mathcal{C}(G)|)$.

Lauseen 1 todistus seuraa nyt suoraan lemman 2 erityistapauksena sijoituksella $A = \emptyset, B = V$.

3.2 Puut

Osoittautuu, että n-solmuiselle puulle yhtenäiset joukot voidaan laskea rekursiivisesti ajassa $\mathcal{O}(n)$, jos oletamme tietyt peruslaskutoimitukset vakioaikaisiksi. Idea on laskea yhtenäisten joukkojen määrä alipuiden yhtenäisten joukkojen kombinaatioina aloittaen lehdistä ja edeten kohti juurta. Lisäksi joudumme pitämään erikseen kirjaa yhtenäisistä joukoista, jotka sisältävät alipuun juuren.

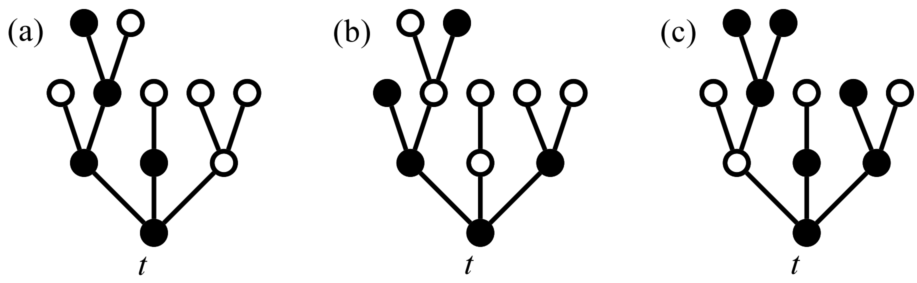
Olkoon $G = (V, E)$ puu eli syklitön yhtenäinen verkko. Haluamme antaa yhtenäisten joukkojen määrälle rekursiokaavan, joten kiinnitämme ensin mielivaltaisen juurisolmun $v \in V$, joka määrää verkolle G yksikäsitteisen alipuuhierarkian. Jatkossa tarkoitamme puun alipuulla sellaista alipuuta, jonka juuri on jokin puun lapsista ja joka sisältää kaikki juuren jälkeläiset. Tarkastellaan nyt mielivaltaista alipuuta T. Olkoon U sen solmujen joukko ja $t \in U$ sen juuri. Merkitsemme $\mathcal{C}(U) = \mathcal{C}(G[U])$ ja $\mathcal{C}_t(U) = \{C \in \mathcal{C}(U) : t \in C\}$. Toisin sanoen $\mathcal{C}(U)$ sisältää alipuun kaikki yhtenäiset joukot ja $\mathcal{C}_t(U)$ ne, jotka sisältävät alipuun juuren.

Jos $|U| = 1$, joukolla U on kaksi osajoukkoa, $\emptyset$ ja $\{t\}$. Molemmat ovat yhtenäisiä, joten tällöin $|\mathcal{C}(U)| = 2$ ja $|\mathcal{C}_t(U)| = 1$. Muussa tapauksessa olkoot $U_1, \ldots, U_k$ alipuun T alipuiden solmujoukot ja $t_1, \ldots, t_k$ niiden juuret. Laskemme nyt luvut $|\mathcal{C}(U)|$ ja $|\mathcal{C}_t(U)|$ alipuiden vastaavista lukumääristä.

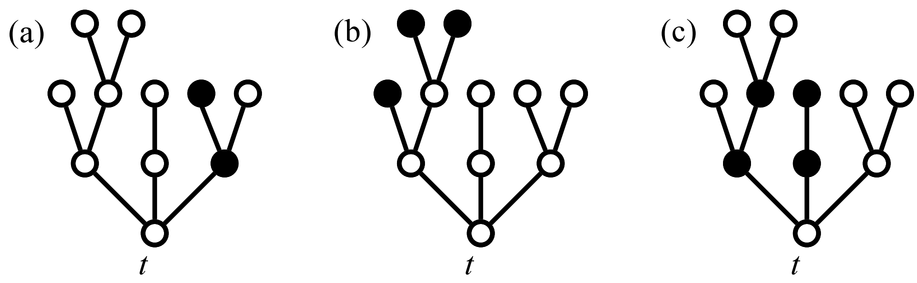
Laskemme ensin sellaiset yhtenäiset joukot, jotka sisältävät juuren t. Olkoon siis $C \subset U$ ja $t \in C$. Olkoon $C_i = C \cap U_i$ kaikilla $1 \leq i \leq k$, jolloin joukko C saadaan yhdisteenä $C = \{t\} \cup C_1 \cup \cdots \cup C_k$. Nyt joukko C on yhtenäinen, jos ja vain jos kaikilla $1 \leq i \leq k$ pätee joko $C_i = \emptyset$ tai $C_i \in \mathcal{C}_{t_i}(U_i)$ (kuva 4). Tällaisten kombinaatioiden määrä saadaan selvästi tulona

$$|\mathcal{C}_t(U)| = \prod_{i=1}^{k} (|\mathcal{C}_{t_i}(U_i)| + 1) \ .$$

Olkoon nyt $t \notin C$. Jos C on yhtenäinen, se voi selvästi sisältää korkeintaan yhden alipuun solmuja, koska ainoa polku kahden alipuun välillä kulkee solmun t kautta.



Kuva 4: Esimerkkejä joukoista (mustat solmut), jotka sisältävät juuren t. Joukko on yhtenäinen täsmälleen silloin, kun alipuiden kaikista solmuista on polku juureen t (a). Kääntäen joukko on epäyhtenäinen, jos projektio jonkin alipuun suhteen on epäyhtenäinen (b) tai epätyhjän alipuun juuri ei ole mukana (c).



Kuva 5: Esimerkkejä joukoista, jotka eivät sisällä juurta t. Tällöin epäyhtenäisiä joukkoja ovat täsmälleen ne tapaukset, joissa projektio jonkin alipuun suhteen on epäyhtenäinen (b) tai joukko sisältää solmuja kahdesta eri alipuusta (c).

Siispä C on yhtenäinen, jos ja vain jos $C_i \neq \emptyset$ korkeintaan yhdellä $1 \leq i \leq k$ ja $C_i \in \mathcal{C}(G)$ (kuva 5). Mahdollisten kombinaatioiden määrä saadaan siis olennaisesti summana alipuiden yhtenäisten joukkojen määristä. On kuitenkin huomioitava, että tyhjä joukko tulee tällöin laskettua kerran jokaista alipuuta kohti, joten se on lopuksi vähennettävä pois $k - 1$ kertaa. Tällaisten joukkojen määrä on siten

$$\sum_{i=1}^{k} |\mathcal{C}(U_i)| - k + 1 \ .$$

Yhdistämällä nämä tulokset saadaan rekursiokaava:

$$\begin{cases} |\mathcal{C}_t(U)| &= 1 \ , & \text{kun } |U| = 1, \\ |\mathcal{C}(U)| &= 2 \ , & \text{kun } |U| = 1, \\ |\mathcal{C}_t(U)| &= \prod_{i=1}^{k}(|\mathcal{C}_{t_i}(U_i)| + 1) \ , & \text{muulloin}, \\ |\mathcal{C}(U)| &= |\mathcal{C}_t(U)| + \sum_{i=1}^{k} |\mathcal{C}(U_i)| - k + 1 \ , & \text{muulloin}. \end{cases}$$

Triviaaliaskel suoritetaan kerran jokaista lehtisolmua kohti, joten sen suoritus vaatii yhteensä ajan $\mathcal{O}(n)$. Rekursiivinen askel suoritetaan kerran jokaiselle alipuulle, jonka juuri t on sisäsolmu. Olettaen, että summa ja tulo ovat vakioaikaisia operaatioita, askel vaatii selvästi ajan $\mathcal{O}(d(t))$. Yhteensä sen vaativuus on siten myös

$$\mathcal{O}\left(\sum_{v \in V} d(v)\right) = \mathcal{O}(2|E|) = \mathcal{O}(n) \ .$$

Jos jokaisella syvyydellä s olevilla solmuilla on k_s lasta, kaikki saman tason alipuut ovat samanlaisia ja jokainen lehtisolmu on samalla syvyydellä h. Tällöin yhtenäiset joukot voidaan laskea ajassa $O(h)$. Tämän havaitsemiseksi riittää todeta, että rekursiivinen askel pelkistyy muotoon:

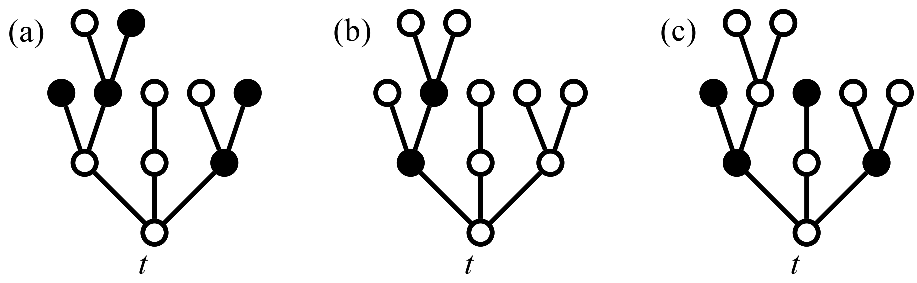
$$\begin{cases} |\mathcal{C}_t(U)| &= (|\mathcal{C}_{t_i}(U_i)| + 1)^k \ , \\ |\mathcal{C}(U)| &= |\mathcal{C}_t(U)| + k|\mathcal{C}(U_i)| - k + 1 \ , \end{cases}$$

missä U_i on mikä tahansa identtisistä alipuista. Olettaen, että myös potenssiinkorotus on vakioaikainen, rekursiivinen askel toimii ajassa $\mathcal{O}(1)$ ja se suoritetaan $h - 1$ kertaa. Triviaaliaskel suoritetaan ainoastaan kerran. Solmujen määrän suhteen kokonaisaikavaativuus on luokkaa $\mathcal{O}(\log_a n)$, missä $a = (\prod_{s=1}^{h} k_s)^{1/h}$.

3.3 Lehtiyhtenäiset joukot

Luvussa 4 annamme yhtenäisten joukkojen määrälle ylärajan rajoitetun asteen verkoissa käyttämällä tietynlaista paikallista analyysiä. Tunnemme tällöin verkosta vain pienen osajoukon A ja kysymme, kuinka moni joukon A osajoukoista voi olla leikkaus yhtenäisen joukon ja joukon A kanssa. Yhtäpitävästi haluamme tietää sellaisten joukkojen määrän, joissa jokaisesta solmusta on polku joukon A ulkopuoliseen solmuun. Esitämme seuraavaksi tavan laskea tällaiset joukot tapauksessa, jossa A on puu ja saman syvyyden solmuilla on yhtä monta lasta.

Määritelmä 3. Olkoon $T = (V, E)$ juurellinen puu, $U \subset V$ ja $u \in U$. Sanomme, että solmu u on *lehtiyhtenäinen* joukossa U, jos verkossa $T[U]$ on polku solmusta u johonkin verkon T lehtisolmuun. Vastaavasti u on *juuriyhtenäinen*, jos verkossa $T[U]$ on polku solmusta u verkon T juureen. Sanomme lisäksi, että joukko U on lehtiyhtenäinen alipuussa T, jos sen kaikki solmut ovat lehtiyhtenäisiä. Erityisesti tyhjä joukko on lehtiyhtenäinen. Määrittelemme nyt ominaisuudet A, …, E. Sanomme, että joukolle U pätee

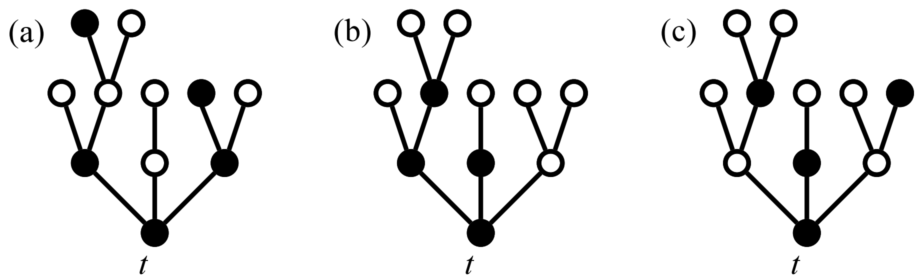


Kuva 6: Esimerkkejä joukoista, jotka eivät sisällä juurta. Selvästi millään joukolla ei ole ominaisuutta A eikä siten ominaisuutta D. Ainoastaan joukolla (a) on ominaisuus B ja siten myös C ja E. Joukot (b) ja (c) sisältävät solmuja, jotka eivät ole lehtiyhtenäisiä. Siten kummallakaan ei ole mitään laskettavista ominaisuuksista.

- A, jos verkon T juuri on joukossa U ja on lehtiyhtenäinen siinä.
- B, jos jokainen $u \in U$ on lehti- tai juuriyhtenäinen.
- C, jos se on lehtiyhtenäinen.
- D, jos sille pätee A ja B.
- E, jos sille ei päde A ja pätee B.

Olkoon nyt G juurellinen puu, jonka syvyys on h ja jossa kaikilla tason s solmuilla on k_s lasta. Laskemme rekursiivisesti sellaisten joukkojen määrät, joilla on ominaisuudet C, D ja E. Verkon G määritelmästä seuraa, että kaikki saman tason alipuut ovat samanlaisia, joten olkoot c_s, d_s ja e_s näiden joukkojen määrät tason s alipuussa kaikilla $0 \leq s \leq h$. Näytämme, miten nämä luvut riippuvat syvemmän tason vastaavista luvuista c_{s+1}, d_{s+1} ja e_{s+1}. Tarkastelemme nyt alipuuta $T = (V, E)$ jollain kiinteällä s. Olkoon t alipuun juuri. Jos $k_s = 0$, ainoat osajoukot ovat $\emptyset$ ja $\{t\}$. Tällöin pätee triviaalisti $c_s = 2$, $d_s = 1$ ja $e_s = 1$. Oletamme nyt, että $k_s > 0$. Olkoot $T_1, \ldots, T_{k_s}$ verkon T alipuut ja $U_1, \ldots, U_{k_s}$ näiden solmujoukot. Kuten yhtenäisten joukkojen kanssa, määritämme luvut c_s, d_s ja e_s erikseen tapauksille, joissa juuri on ja ei ole mukana osajoukossa.

Olkoot c'_s, d'_s, e'_s sellaisten joukkojen määrät, joissa juuri ei ole mukana (kuva 6). Tällöin $U \subset V$ on selvästi lehtiyhtenäinen, jos ja vain jos sen projektio jokaisen alipuun suhteen on lehtiyhtenäinen, toisin sanoen joukko $U \cap U_i$ on lehtiyhtenäinen alipuussa T_i kaikilla $1 \leq i \leq k_s$. Siten $c'_s = c_{s+1}^{k_s}$. Koska juuri ei ole mukana, pätee triviaalisti $d'_s = 0$. Edelleen pätee E $\iff$ B $\iff$ C, joten $e'_s = c'_s = c_{s+1}^{k_s}$.



Kuva 7: Esimerkkejä joukoista, jotka sisältävät juuren. Joukolla (a) on ominaisuus C (ja siten myös D), koska sen yhdellä alipuulla on ominaisuus D ja muilla E. Joukossa (b) kaikki solmut ovat yhteydessä vain juureen, joten sillä on ominaisuus E. Joukko (c) sisältää solmun, joka ei ole juuri- eikä lehtiyhtenäinen. Siten sillä ei ole mitään laskettavista ominaisuuksista.

Olkoot c_s'', d_s'', e_s'' sellaisten joukkojen määrät, joissa juuri on mukana (kuva 7). Jos joukolle $U \subset V$ pätee C, sen projektiolle ainakin yhdessä alipuussa täytyy päteä D. Muille alipuille riittävä ja välttämätön ehto on B eli joko D tai E. Jos D pätee i alipuulle, tällöin E pätee $k_s - i$ alipuulle. Alipuut, joille pätee D, voidaan valita $\binom{k_s}{i}$ tavalla ja kullakin valinnalla saadaan $d_{s+1}^i e_{s+1}^{k_s-i}$ kombinaatiota. Siten lehtiyhtenäisten joukkojen määrä on

$$c_s'' = \sum_{i=1}^{k_s} \binom{k_s}{i} d_{s+1}^i e_{s+1}^{k_s-i} \ .$$

Näemme, että tämä on binomikerroinkaava potenssille $(d_{s+1} + e_{s+1})^{k_s}$ ilman ensimmäistä termiä $e_{s+1}^{k_s}$. Siten $c_s'' = (d_{s+1} + e_{s+1})^{k_s} - e_{s+1}^{k_s}$. Koska juuri on mukana, pätee C $\iff$ D. Siten $d_s'' = c_s''$. Edelleen E pätee nyt täsmälleen silloin, kun projektiolle jokaisessa alipuussa pätee E. Siten $e_s'' = e_{s+1}^{k_s}$. Yhdistämällä nämä tapaukset saadaan

$$\begin{cases} d_s &= d_s' + d_s'' &= (d_{s+1} + e_{s+1})^{k_s} - e_{s+1}^{k_s} \ , \\ c_s &= c_s' + c_s'' &= c_{s+1}^{k_s} + d_s \ , \\ e_s &= e_s' + e_s'' &= c_{s+1}^{k_s} + e_{s+1}^{k_s} \ . \end{cases}$$

Tätä käyttäen lehtiyhtenäiset joukot voidaan laskea rekursiivisesti aloittaen syvimmästä tasosta ja edeten kohti juurta. Olemme erityisesti kiinnostuneita sellaisista puista, joissa juurella on d lasta ja muilla sisäsolmuilla $d-1$. Palaamme tähän luvussa 4.4.

4 Yläraja rajoitetun asteen verkoissa

Luvussa 2 totesimme, että rajoitetun asteen verkot ovat mahdollisesti yleisin merkittävä verkkoperhe, jossa yhtenäisten joukkojen määrä kasvaa olennaisesti hitaammin kuin kaikkien osajoukkojen määrä. Määrittelimme kaikilla $\Delta \geq 3$ ja n luvun $C_{\Delta,n}$, jolla tarkoitamme perheen $\mathcal{C}$ suurinta mahdollista kokoa n-solmuisessa Δ-rajoitetussa verkossa. Erityisesti määrittelimme kaikilla $\Delta \geq 3$ luvun σ_Δ, joka on olennaisesti paras arvio luvun $C_{\Delta,n}$ asymptoottista kasvua kuvaavalle eksponentiaaliselle tekijälle.

Björklund ym. [BHKK08b] osoittavat, että $\sigma_\Delta \leq \beta_\Delta$, missä $\beta_\Delta = (2^{\Delta+1} - 1)^{1/(\Delta+1)}$ kaikilla $\Delta \geq 3$, mutta jättävät tuloksen mahdollisen optimaalisuuden avoimeksi kysymykseksi. Keskeinen työkalu tämän ylärajan todistuksessa on Shearerin entropiaepäyhtälöstä johdettu projektiolause, joka intuitiivisesti antaa ylärajan joukkoperheen koolle, kun tunnetaan yläraja sen projektioiden koolle perusjoukon osajoukoissa.

Todistamme Shearerin epäyhtälön sekä sen kombinatorisen seurauksen luvussa 4.1 ja esitämme Björklundin ym. todistuksen luvussa 4.2. Luvussa 4.3 laajennamme Björklundin ym. analyysiä ja esitämme yleisen menetelmän asymptoottisen kasvun arvioimiseksi. Näytämme luvussa 4.4, että menetelmä parantaa aiempaa ylärajaa ainakin tapauksessa $\Delta \in \{3, 4\}$, mutta ei oletettavasti millään $\Delta > 5$.

4.1 Kombinatorisia työkaluja

Esitämme tässä luvussa kaksi keskeistä tulosta, Shearerin entropiaepäyhtälöstä seuraavaan projektiolauseen sekä Jensenin epäyhtälön, joita käytämme ylärajan analyysissä. Todistamme sekä Shearerin epäyhtälön että seurauslauseen lähtien entropian peruslauseista. Todistuksen seuraaminen ei vaadi syvällisiä esitietoja informaatioteoriasta, mutta lukijan oletetaan tuntevan todennäköisyyslaskennan ja erityisesti satunnaismuuttujien perusteet. Merkinnällä $\mathrm{P}(A)$ tarkoitamme tapahtuman A todennäköisyyttä ja merkinnällä $\mathrm{E}[X]$ satunnaismuuttujan X odotusarvoa.

Määritelmä 4. Olkoot X, Y diskreettejä satunnaismuuttujia, jotka saavat arvoja äärellisessä joukossa S. Tällöin muuttujan X (Shannon-)*entropia* on

$$\mathrm{H}[X] = -\sum_{s \in S} \mathrm{P}(X = s) \log \mathrm{P}(X = s) \ ,$$

missä sovimme, että $0 \cdot \log 0 = 0$. Muuttujan X entropia ehdolla $Y = y$ on

$$\mathrm{H}[X|Y = y] = -\sum_{s \in S} \mathrm{P}(X = s|Y = y) \log \mathrm{P}(X = s|Y = y)$$

kaikilla $y \in S$. Muuttujan X *ehdollinen entropia*, kun Y on annettu, on

$$\mathrm{H}[X|Y] = \mathrm{E}[\mathrm{H}[X|Y=y] : y \in S] = \sum_{y \in S} \mathrm{P}(Y=y)\mathrm{H}[X|Y=y] ,$$

siis entropian $\mathrm{H}[X|Y=y]$ odotusarvo muuttujan Y arvojen yli.

Entropia kuvaa intuitiivisesti satunnaismuuttujaan liittyvää epävarmuutta ja toisaalta saatavan informaation määrää, kun muuttujan arvo paljastetaan. Olkoot X, Y, Z diskreettejä satunnaismuuttujia, jotka ovat jakautuneet äärelliselle joukolle S. Entropialle pätevät seuraavat perustulokset [Rad01, Rom92].

1) Muuttujan X entropia on aina välillä

$$0 \leq \mathrm{H}[X] \leq \log |S| ,$$

missä yhtäsuuruus $\mathrm{H}[X] = \log |S|$ pätee täsmälleen silloin, kun X on tasaisesti jakautunut, ja $\mathrm{H}[X] = 0$ täsmälleen silloin, kun $\mathrm{P}(X = s) = 1$ jollain $s \in S$.

2) Ehdollinen entropia on *monotoninen*:

$$\mathrm{H}[X|Y, Z] \leq \mathrm{H}[X|Y] ,$$

toisin sanoen muuttujaan liittyvä epävarmuus ei voi olla suurempi tilanteessa, jossa tunnemme useamman muuttujan arvon.

3) Yhdistetyn muuttujan (X, Y) entropia palautuu ehdolliseen entropiaan (entropian Bayes-sääntö):

$$\mathrm{H}[X, Y] = \mathrm{H}[X|Y] + \mathrm{H}[Y] .$$

Yhdistetyn muuttujan $(X_1, X_2, \ldots, X_n)$ entropia saadaan ketjuttamalla Bayes-sääntöä:

$$\mathrm{H}[X_1, X_2, \ldots, X_n] = \mathrm{H}[X_1] + \mathrm{H}[X_2|X_1] + \ldots + \mathrm{H}[X_n|X_1, \ldots, X_{n-1}] .$$

Näistä tuloksista voidaan edelleen johtaa Shearerin epäyhtälö [CGFS86], joka antaa yleisemmän ylärajan muuttujan $(X_1, X_2, \ldots, X_n)$ entropialle, kun tunnetaan entropia muuttujien $X_1, X_2, \ldots, X_n$ riittävän monelle kombinaatiolle.

Lause 5. (Shearerin epäyhtälö) *Olkoon A äärellinen joukko ja olkoot $A_1, \ldots, A_k$ sen osajoukkoja siten, että jokainen $a \in A$ on ainakin δ osajoukossa. Olkoon X_a satunnaismuuttuja kaikilla $a \in A$ ja $X_S = (X_a : a \in S)$ kaikilla $S \subset A$. Tällöin*

$$\delta \cdot \mathrm{H}[X_A] \leq \sum_{i=1}^{k} \mathrm{H}[X_{A_i}] .$$

Todistus. Kiinnitämme joukon A alkioille jonkin mielivaltaisen järjestyksen. Tällöin Bayes-säännön nojalla kaikilla $S \subset A$ pätee

$$\mathrm{H}[X_S] = \sum_{a \in S} \mathrm{H}[X_a | X_b : b \in S, b < a] \ ,$$

missä monotonisuuden nojalla $\mathrm{H}[X_a|X_b : b \in S, b < a] \geq \mathrm{H}[X_a|X_b : b < a]$. Summa

$$\sum_{i=1}^{k} \mathrm{H}[X_{A_i}] \geq \sum_{i=1}^{k} \sum_{a \in A_i} \mathrm{H}[X_a|X_b : b < a]$$

sisältää nyt kaikilla $a \in A$ termin $\mathrm{H}[X_a|X_b : b < a]$ kerran jokaista A_i kohti, johon a kuuluu. Koska oletuksen nojalla tällaisia A_i on kaikilla $a \in A$ vähintään δ, pätee

$$\sum_{i=1}^{k} \mathrm{H}[X_{A_i}] \geq \sum_{a \in A} \delta \cdot \mathrm{H}[X_a|X_b : b < a] = \delta \cdot \mathrm{H}[X_A] \ .$$

□

Lauseella 5 on seuraava kombinatorinen seuraus.

Lause 6. (Projektiolause, Chung ym. [CGFS86]). *Olkoon A äärellinen joukko ja $A_1, \ldots, A_k$ sen osajoukkoja siten, että jokainen $a \in A$ esiintyy ainakin δ osajoukossa A_i. Olkoon nyt $\mathcal{F} \subset 2^A$ ja $\mathcal{F}_i = \{F \cap A_i : F \in \mathcal{F}\}$ kaikilla $1 \leq i \leq k$. Tällöin*

$$|\mathcal{F}|^{\delta} \leq \prod_{i=1}^{k} |\mathcal{F}_i| \ .$$

Todistus. Havaitsemme ensin, että voimme esittää jokaisen $F \in \mathcal{F}$ yksikäsitteisesti muodossa $(p_a : a \in A)$, missä $p_a = [a \in F]$ kaikilla $a \in A$: toisin sanoen $p_a = 1$, jos ja vain jos $a \in F$, ja muuten $p_a = 0$. Olkoon nyt X satunnaismuuttuja, joka on jakautunut tasaisesti joukolle $\mathcal{F}$. Käyttämällä edellistä esitystapaa joukoille $F \in \mathcal{F}$ voimme vastaavasti esittää muuttujan X yhdistettynä muuttujana $X = (X_a : a \in A)$, missä $X_a = [a \in X]$ kaikilla $a \in A$. Olkoon edelleen $X_{A_i} = (X_a : a \in A_i)$ kaikilla $1 \leq i \leq k$, jolloin jokainen X_{A_i} on jakautunut joukolle $\mathcal{F}_i$. Nyt kaikki lauseen 5 ehdot ovat voimassa. Lisäksi perustulosten nojalla pätee $\mathrm{H}[X] = \log |\mathcal{F}|$ ja $\mathrm{H}[X_{A_i}] \leq \log |\mathcal{F}_i|$ kaikilla $1 \leq i \leq k$. Yhdistämällä nämä saamme

$$\delta \log |\mathcal{F}| = \delta \cdot \mathrm{H}[X] \leq \sum_{i=1}^{k} \mathrm{H}[X_{A_i}] \leq \sum_{i=1}^{k} \log |\mathcal{F}_i| \ .$$

Väite seuraa ottamalla puolittain kaksikantainen eksponentti. □

Sanomme, että joukko $\mathcal{F}_i$ on osajoukkoperheen $\mathcal{F}$ *projektio* joukon A_i suhteen kaikilla $1 \leq i \leq k$. Intuitiivisesti lause 6 antaa ylärajan joukon $\mathcal{F}$ koolle, jos tunnetaan yläraja projektioiden koolle riittävän monen A_i suhteen ja joukot A_i ovat jakautuneet riittävän tasaisesti perusjoukolle A. Koska olemme tässä yhteydessä kiinnostuneita nimenomaan solmujen osajoukoista, tulemme käyttämään seuraavaa sovellusta.

Lemma 7. *Olkoon $G = (V, E)$ verkko ja $\mathcal{F} \subset 2^V$. Olkoon A_v joukon V osajoukko kaikilla $v \in V$ siten, että verkon jokainen solmu esiintyy ainakin δ osajoukossa. Määrittelemme kaikilla $v \in V$ projektiot $\mathcal{F}_v = \{F \cap A_v : F \in \mathcal{F}\}$. Tällöin*

$$|\mathcal{F}|^\delta \leq \prod_{v \in V} |\mathcal{F}_v| \ .$$

Lemma 7 seuraa suoraan lauseen 6 erityistapauksena. Perusjoukkona A toimii nyt solmujen joukko V, jonka osajoukkoperheen koolle haluamme antaa ylärajan. Lauseessa 6 osajoukkojen A_i määrä on jokin mielivaltainen k, mutta tässä käytämme implisiittisesti sijoitusta $k = |V|$ eli määrittelemme täsmälleen yhden osajoukon kutakin solmua kohti ja käytämme solmuja näiden joukkojen indekseinä.

Projektiolauseen lisäksi käytämme todistuksessa Jensenin epäyhtälön erityistapausta konkaaveille funktioille. Sanomme, että reaalivälillä I määritelty funktio $f : I \to \mathbb{R}$ on *konkaavi* (engl. *concave*) tai *kovera*, jos kaikilla $a, b \in I$ ja $t \in [0, 1]$ pätee

$$f(ta + (1-t)b) \geq tf(a) + (1-t)f(b) \ , \tag{1}$$

toisin sanoen $f(x) \geq y$ jokaisella (x, y), joka sijaitsee pisteiden $(a, f(a))$ ja $(b, f(b))$ välisellä janalla. Jos epäyhtälö 1 pätee toisin päin kaikilla $a, b \in I$ ja $t \in [0, 1]$, sanomme vastaavasti, että funktio f on *konveksi* (engl. *convex*) tai *kupera*.

Esimerkki 8. *Olkoon $c > 0$ vakio ja $f :]\log c, \infty[\to \mathbb{R}, f(x) = \log(2^x - c)$. Tällöin funktio f on konkaavi.*

Todistus. Oletamme tunnetuksi, että derivoituva funktio on konkaavi, jos sen derivaattafunktio on monotonisesti vähenevä. Seurauksena pätee, että funktio on konkaavi, jos sen toinen derivaatta on negatiivinen kaikkialla. Derivoimalla kerran saadaan

$$f'(x) = \log'(2^x - c) \cdot \mathrm{D}_x(2^x - c) = \frac{1}{(2^x - c)\ln 2} \cdot (2^x \ln 2) = \frac{2^x}{2^x - c}$$

ja toisen kerran

$$f''(x) = \frac{(\mathrm{D}_x 2^x)(2^x - c) - 2^x \mathrm{D}_x(2^x - c)}{(2^x - c)^2} = -\frac{c \cdot 2^x \ln 2}{(2^x - c)^2} \ .$$

Selvästi pätee $f''(x) < 0$ kaikilla $x > \log c$. $\square$

Konkaavin funktion f vastafunktio $-f$ on selvästi aina konveksi ja toisin päin, joten konkaaveja funktioita koskeville tuloksille saadaan yksinkertaisella sijoituksella analoginen käänteinen tulos konvekseille funktioille. Niculescu ja Persson [NP06] esittävät konveksin version Jensenin epäyhtälöstä.

Lause 9. Jensenin epäyhtälön diskreetti tapaus. *Olkoon $f : I \to \mathbb{R}$ konveksi funktio reaalivälillä I ja olkoon $x_1, \ldots, x_n \in I$, $\lambda_1, \ldots, \lambda_n \in [0, 1]$ ja $\sum_{i=1}^{n} \lambda_i = 1$. Tällöin*

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leq \sum_{i=1}^{n} \lambda_i f(x_i) \ .$$

Esimerkissä 11 tarvitsemme seuraavaa erityistapausta konkaaveille funktioille.

Korollaari 10. *Olkoon $\phi : \mathbb{R} \to \mathbb{R}$ konkaavi ja $x_1, \ldots, x_n \in \mathbb{R}$. Tällöin*

$$\frac{1}{n}\sum_{i=1}^{n} \phi(x_i) \leq \phi\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \ .$$

Korollaari 10 seuraa suoraviivaisesti lauseesta 9 sijoittamalla $f = -\phi$, $I = \mathbb{R}$ ja $\lambda_i = 1/n$ kaikilla $1 \leq i \leq n$.

Annamme nyt esimerkin projektiolauseen ja Jensenin epäyhtälön käytöstä antamalla ylärajan dominoivien joukkojen määrälle rajoitetun asteen verkossa.

Lause 11. *Olkoon $G = (V, E)$ verkko, jonka suurin aste on Δ ja $n = |V|$. Tällöin $|\mathcal{D}(G)| \leq \beta_\Delta^n$, missä $\beta_\Delta = (2^{\Delta+1} - 1)^{1/(\Delta+1)}$.*

Tässä $\mathcal{D}(G)$ on verkon G dominoivien joukkojen perhe. Määritelmän mukaan $U \subset V$ on dominoiva joukko, jos $N[v] \cap U \neq \emptyset$ kaikilla $v \in V$. Osoitamme lauseen 11 lemmaa 7 käyttäen. Todistus seuraa Björklundin ym. [BHKK10] esimerkkiä.

Todistus. Haluamme antaa ylärajan joukon $\mathcal{D}$ koolle, joten sovellamme lemmaa 7 sijoituksella $\mathcal{F} = \mathcal{D}$. Haluamme lisäksi käyttää sijoitusta $\delta = \Delta + 1$: toisin sanoen haluamme joukon V osajoukot siten, että jokainen $v \in V$ on täsmälleen $\Delta + 1$ osajoukossa. Saavuttaaksemme tämän määrittelemme alustavasti $A_v = N[v]$ kaikilla $v \in V$. Tämä määritelmä on selvästi riittävä, jos G on säännöllinen: tällöin kaikilla $v \in V$ pätee $|A_v| = \Delta + 1$ ja $v \in A_u$ kaikilla $u \in A_v$. Jotta vaatimus olisi voimassa myös ei-säännöllisille verkoille, teemme määritelmään seuraavan muutoksen: kaikilla $u \in V$, joille pätee $d(u) < \Delta$, valitsemme $\Delta - d(u)$ joukkoa A_v (mielivaltaisesti), joihin solmu u ei kuulu, ja lisäämme sen niihin. Tällöin jokainen $u \in V$ on täsmälleen $\Delta + 1$ joukossa A_v. Lisäksi tämän seurauksena pätee selvästi

$$\sum_{v \in V} |A_v| = n(\Delta + 1) \ . \tag{2}$$

Edelleen määrittelemme kaikilla $v \in V$ projektiot $\mathcal{D}_v = \{D \cap A_v : D \in \mathcal{D}_v\}$. Nyt kaikki oletukset ovat voimassa, joten lemman 7 nojalla pätee

$$|\mathcal{D}|^{\Delta+1} \leq \prod_{v \in V} |\mathcal{D}_v| \ .$$

Tarkastelemme nyt projektion $\mathcal{D}_v$ kokoa mielivaltaisella $v \in V$. Jokainen $U \in \mathcal{D}_v$ on selvästi jokin joukon A_v osajoukko. Koska $N[v] \subset A_v$, perheen $\mathcal{D}$ määritelmästä seuraa, että joukon U täytyy sisältää joko solmu v tai jokin sen naapureista. Siten ainakin yksi joukon A_v osajoukoista, tyhjä joukko $\emptyset$, ei kuulu projektioon $\mathcal{D}_v$. Täten pätee $|\mathcal{D}_v| \leq 2^{|A_v|} - 1$ ja edelleen

$$|\mathcal{D}|^{\Delta+1} \leq \prod_{v \in V} (2^{|A_v|} - 1) \ .$$

Käytämme nyt tulon arvioimiseen Jensenin epäyhtälöä. Olkoon $f(x) = \log(2^x - 1)$ kaikilla $x > 0$. Esimerkin 8 nojalla f on konkaavi funktio koko määrittelyalueellaan. Siten Jensenin epäyhtälön ja edelleen yhtälön 2 nojalla pätee

$$\frac{1}{n} \sum_{v \in V} f(|A_v|) \leq f\left(\frac{1}{n} \sum_{v \in V} |A_v|\right) = f(\Delta + 1) \ .$$

Ottamalla puolittain 2-kantainen eksponentti saadaan

$$\prod_{v \in V} 2^{f(|A_v|)} \leq (2^{f(\Delta+1)})^n$$

ja avaamalla funktion f määritelmä edelleen

$$\prod_{v \in V} \left(2^{|A_v|} - 1\right) \leq (2^{\Delta+1} - 1)^n \ .$$

Siispä

$$|\mathcal{D}|^{\Delta+1} \leq (2^{\Delta+1} - 1)^n \ .$$

Väite seuraa korottamalla puolittain potenssiin $\frac{1}{\Delta+1}$. $\square$

Taulukko 1: Luvun β_Δ arvoja pienillä Δ.

Δ	1	2	3	4	5	6	7	8
β_Δ	1.7321	1.9130	1.9680	1.9874	1.9948	1.9978	1.9991	1.9996

Osoittautuu, että edellä todistettu yläraja dominoivien joukkojen määrälle on tiukka. Olkoon $\Delta \in \mathbb{N}$. Tarkastellaan k täydellisen verkon $K_{\Delta+1}$ erillistä yhdistettä. Selvästi tällaisella verkolla on suurin aste Δ. Lisäksi solmujen osajoukko on dominoiva, jos ja vain jos se sisältää ainakin yhden solmun verkon jokaisesta yhtenäisestä komponentista. Siten mahdollisia kombinaatioita on $2^{\Delta+1}-1$ jokaisesta komponenttia kohti eli yhteensä $(2^{\Delta+1}-1)^k = (2^{\Delta+1}-1)^{n/(\Delta+1)}$.

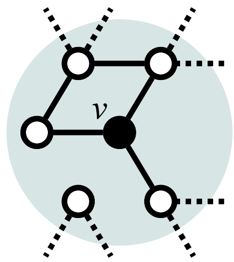
4.2 Yleinen yläraja yhtenäisille joukoille

Annamme nyt edellä esitettyjä tuloksia käyttäen ylärajan joukon $\mathcal{C}$ koolle rajoitetun asteen verkoissa. Toteamme ensin, että ongelma on triviaali, kun verkon suurin aste $\Delta < 3$. Olkoon n solmujen lukumäärä. Tapauksessa $\Delta = 0$ ainoa tarkasteltava verkko on tyhjä verkko. Sen yhtenäiset joukot ovat täsmälleen kaikki yksiöt sekä tyhjä joukko, joten pätee $|\mathcal{C}| = n + 1$. Tapauksessa $\Delta = 1$ suurin määrä saavutetaan verkolla, jossa jokaisella solmulla on täsmälleen yksi naapuri. Parittomalla n yksi solmu jää ilman paria. Tällöin $|\mathcal{C}| = n + 1 + \lfloor n/2 \rfloor$. Tapauksessa $\Delta = 2$ eniten yhtenäisiä joukkoja sisältävä verkko on sykli C_n, jolla $|\mathcal{C}| = n(n-1) + 2$. Esitämme nyt yhtenäisten joukkojen määrälle yleisen ylärajan tapauksessa $\Delta \geq 3$.

Lause 12. *Olkoon $\Delta \geq 3, G = (V, E)$ verkko, jonka suurin aste on Δ, ja $n = |V|$. Tällöin pätee $|\mathcal{C}(G)| \leq \beta_\Delta^n + n$, missä $\beta_\Delta = (2^{\Delta+1} - 1)^{1/(\Delta+1)}$.*

Todistus. Määrittelemme lemman 7 vaatimat osajoukot samoin kuin lauseen 11 todistuksessa. Olkoon siis alustavasti $A_v = N[v]$ kaikilla $v \in V$. Tämän jälkeen kaikilla $u \in V$, joille pätee $d(u) < \Delta$, valitsemme $\Delta - d(u)$ joukkoa A_v (mielivaltaisesti), joihin solmu u ei kuulu, ja lisäämme sen niihin. Tällöin jokainen $u \in V$ on täsmälleen $\Delta + 1$ joukossa A_v ja pätee $\sum_{v \in V} |A_v| = n(\Delta + 1)$.

Koska olemme nyt antamassa ylärajaa yhtenäisten joukkojen määrälle, ensimmäinen ajatus on käyttää suoraan sijoitusta $\mathcal{F} = \mathcal{C}$, jolloin saisimme vastaavasti projektiot $\mathcal{C}_v = \{C \cap A_v : C \in \mathcal{C}\}$ kaikilla $v \in V$. On kuitenkin helppoa konstruoida tapaus, jossa kaikilla $U \subset A_v$ on olemassa $C \in \mathcal{C}$ siten, että $C \cap A_v = U$. Toisin kuin esimerkissä



Kuva 8: Mahdoton leikkaus projektiossa $\mathcal{C}_v$. Ainoa yhtenäinen joukko, joka sisältää solmun v mutta ei sen naapureita, on poissuljettu joukko $\{v\}$.

11, erityisesti tyhjä joukko voi nyt olla mukana projektiossa. Siten pahimmassa tapauksessa pätee $|\mathcal{C}_v| = 2^{|A_v|}$ kaikilla $v \in V$, mikä yhdessä lemman 7 kanssa antaa vain triviaalin ylärajan $|\mathcal{C}| \leq 2^n$.

Osoittautuu kuitenkin, että voimme sulkea pois yhden leikkauksista $C \cap A_v$ poistamalla perheestä $\mathcal{C}$ vain muutaman joukon. Käytämme sijoitusta $\mathcal{F} = \mathcal{C}'$, missä $\mathcal{C}' = \mathcal{C} \setminus \{\{v\} : v \in V\}$. Perhe $\mathcal{C}'$ sisältää siis kaikki yhtenäiset joukot pois lukien ne, jotka sisältävät täsmälleen yhden solmun. Pois jääviä joukkoja on täsmälleen n kappaletta, joten ne on helppo laskea erikseen. Määrittelemme nyt kaikilla $v \in V$ projektiot $\mathcal{C}_v = \{C \cap A_v : C \in \mathcal{C}'\}$, jolloin lemman 7 nojalla pätee

$$|\mathcal{C}'|^{\Delta+1} \leq \prod_{v \in V} |\mathcal{C}_v| \ .$$

Tarkastelemme nyt projektion $\mathcal{C}_v$ kokoa mielivaltaisella $v \in V$. Jokainen $U \in \mathcal{C}_v$ on selvästi jokin joukon A_v osajoukko. Havaitsemme, että ainakin yksi näistä joukoista, yksiö $\{v\}$ (kuva 8), ei voi kuulua projektioon $\mathcal{C}_v$. Käänteisestä oletuksesta seuraisi, että $C \cap A_v = \{v\}$ jollain $C \in \mathcal{C}'$ ja siten $C \cap N(v) = \emptyset$. Jos C sisältää solmun v, mutta ei yhtäkään sen naapureista, C voi olla yhtenäinen vain jos $C = \{v\}$. Tämä taas on ristiriita, koska määritelmän nojalla $\{v\} \notin \mathcal{C}'$. Siten $\{v\} \notin \mathcal{C}_v$, joten $|\mathcal{C}_v| \leq 2^{|A_v|} - 1$ ja edelleen

$$|\mathcal{C}'|^{\Delta+1} \leq \prod_{v \in V} \left(2^{|A_v|} - 1\right) \ .$$

Nyt olemme saaneet perheen $\mathcal{C}'$ koolle täsmälleen saman ylärajan kuin perheen $\mathcal{D}$ koolle esimerkissä 11. Seuraamalla samaa päättelyä Jensenin epäyhtälön kanssa saamme vastaavasti edelleen

$$|\mathcal{C}'|^{\Delta+1} \leq (2^{\Delta+1} - 1)^n \ .$$

Lopullinen väite seuraa korottamalla puolittain potenssiin $\frac{1}{\Delta+1}$ ja huomioimalla, että $|\mathcal{C}| = |\mathcal{C}'| + n$. □

Tämän seurauksena pätee selvästi $\sigma_\Delta \leq (2^{\Delta+1} - 1)^{\Delta+1}$ kaikilla $\Delta \geq 3$.

Luonnollisena jatkona tälle todistukselle saadaan yläraja yhtenäisten dominoivien joukkojen määrälle. Olkoon $\mathcal{E}(G) = \mathcal{C}(G) \cap \mathcal{D}(G)$ eli verkon G yhtenäisten dominoivien joukkojen perhe. Todistamme seuraavan lauseen.

Lause 13. *Olkoon $\Delta \geq 3, G = (V, E)$ verkko, jonka suurin aste on Δ, ja $n = |V|$. Tällöin pätee $|\mathcal{E}(G)| \leq (2^{\Delta+1} - 2)^{n/(\Delta+1)} + n$.*

Todistus. Päättely on olennaisesti yhdistelmä lauseiden 11 ja 12 todistuksista. Määrittelemme joukot A_v jälleen samalla tavalla ja käytämme sijoitusta $\mathcal{F} = \mathcal{E}'$, missä $\mathcal{E}' = \mathcal{E} \setminus \{\{v\} : v \in V\}$. Toisin sanoen $\mathcal{E}'$ sisältää yksiöitä lukuun ottamatta kaikki joukot, jotka ovat sekä yhtenäisiä että dominoivia. Tarkastelemme nyt projektioita $\mathcal{E}_v = \{E \cap A_v : E \in \mathcal{E}'\}$ kaikilla $v \in V$. Lauseen 11 todistuksen tapaan toteamme, että pätee $\emptyset \notin \mathcal{E}_v$ kaikilla $v \in V$. Vastaavasti lauseen 12 todistuksen tapaan toteamme, että pätee $\{v\} \notin \mathcal{E}_v$ kaikilla $v \in V$. Siten $|\mathcal{E}_v| \leq 2^{A_v} - 2$ kaikilla $v \in V$, mistä edelleen lemman 7 nojalla seuraa

$$|\mathcal{E}'|^{\Delta+1} \leq \prod_{v \in V} \left(2^{|A_v|} - 2\right) .$$

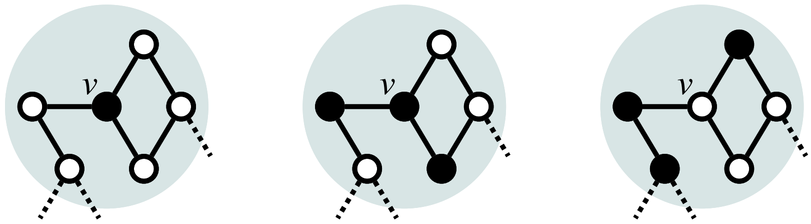
Esimerkin 8 nojalla funktio f on konkaavi, kun $f(x) = \log(2^x - 2)$ kaikilla $x > 1$. Käyttämällä jälleen vastaavaa päättelyä Jensenin epäyhtälön kanssa saadaan

$$|\mathcal{E}'|^{\Delta+1} \leq (2^{\Delta+1} - 2)^n .$$

Väite seuraa korottamalla potenssiin $\frac{1}{\Delta+1}$ ja huomioimalla, että $|\mathcal{E}| = |\mathcal{E}'| + n$. □

4.3 Analyysi yleistetyillä naapurustoilla

Olemme edellä osoittaneet ylärajan yhtenäisten ja dominoivien joukkojen määrälle rajoitetun asteen verkoissa tarkastelemalla näiden joukkoperheiden projektioita kunkin solmun naapuruston suhteen. Näytimme, että n-solmuisessa verkossa sekä yhtenäisten että dominoivien joukkojen määrä on luokkaa $\mathcal{O}(\beta_\Delta^n)$, kun $\Delta \geq 3$ on verkon suurin aste. Erityisesti näytimme, että saamamme yläraja on tiukka dominoiville joukoille konstruoimalla kaikilla $\Delta \geq 3$ verkkoperheen, jolle pätee



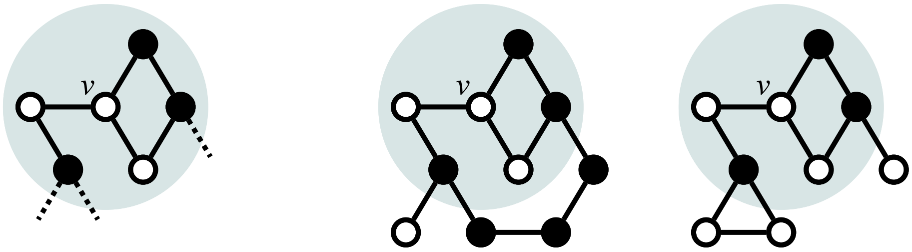
Kuva 9: Esimerkkejä mahdottomista leikkauksista tapauksessa $r = 2$.

$|\mathcal{D}| \in \Theta(\beta_\Delta^n)$. Sen sijaan vaikuttaa siltä, että yhtenäisille joukoille ylärajassa on huomattavasti parantamisen varaa. Esimerkiksi $\beta_3 = (2^{3+1} - 1)^{1/(3+1)} \approx 1.9680$, kun taas kokeellisten arvioiden perusteella näyttäisi pätevän $\sigma_3 \approx 1.81$ [PIM08].

Tässä luvussa esitämme yleisen menetelmän, jolla parannamme yleistä ylärajaa tapauksessa $\Delta \in \{3, 4\}$. Käytämme olennaisesti samaa todistusta lemman 7 kanssa kuin edellä, mutta laajennamme analyysiä käyttämällä paikallisina ympäristöinä solmujen *yleistettyjä naapurustoja*. Toisin sanoen tarkastelemme kullakin solmulla v kaikkia solmuja, jotka ovat enintään jonkin kiinnitetyn säteen r päässä solmusta v. Osoittautuu, että kaikilla $r \geq 2$ voidaan tällöin sulkea pois huomattavasti enemmän mahdollisia leikkauksia. Pohjustamme formaalia todistusta tarkastelemalla ensin esimerkkejä tällaisista leikkauksista.

Kuva 9 esittää solmun v yhtä mahdollista naapurustoa $N^2[v]$ sekä leikkauksia, jotka voimme tässä tapauksessa sulkea pois projektiosta. On helppo nähdä, että kahdessa ensimmäisessä tapauksessa joukko itse on ainoa yhtenäinen joukko, jonka leikkaus naapuruston $N^2[v]$ kanssa tuottaa kyseisen joukon. Luvun 4.2 tarkastelussa ainoa tällainen joukko oli yksiö $\{v\}$. Pääsimme siitä eroon jättämällä kaikki yksiöt tarkasteltavan perheen ulkopuolelle. Voimme nyt laajentaa tätä periaatetta sulkemalla pois kaikki sellaiset joukot, jotka sisältyvät kokonaan johonkin naapurustoon. Jos naapuruston koko on rajoitettu, tällaisten joukkojen määrä kasvaa enintään lineaarisesti suhteessa solmujen määrään. Tällöin voimme sulkea pois kuvan 9 kaksi ensimmäistä leikkausta. Viimeinen esimerkki on selvästi sisäisesti epäyhtenäinen ja siten olennaisesti uudenlainen tapaus. Suurin osa mahdottomista leikkauksista kaikilla $r \geq 2$ on nimenomaan viimeisen esimerkin kaltaisia.

Kuva 10 on esimerkki neljännestä leikkauksesta sekä kahdesta tapauksesta, joissa naapuruston ulkopuolinen rakenne vaikuttaa siihen, onko leikkaus mahdollinen. Ensimmäisessä tapauksessa on olemassa yhtenäinen joukko siten, että leikkaus on mahdollinen, jälkimmäisessä taas ei. Pahimmassa tapauksessa joudumme olettamaan,



Kuva 10: Esimerkki leikkauksesta, jonka joudumme olettamaan mahdolliseksi.

että tällainen joukko on olemassa emmekä siten voi sulkea leikkausta pois projektiosta. Mahdollisia leikkauksia ovat siten täsmälleen ne joukot, joissa jokaisesta solmusta on polku joukon ulkopuolelle.

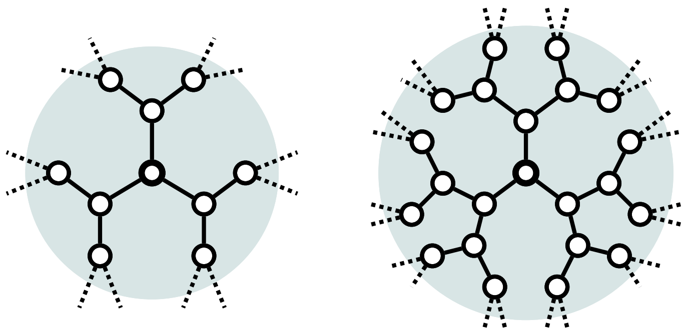
Yksinkertainen luetteleminen osoittaa, että kuvien 9 ja 10 naapurustolla $N^2[v]$ on 32 mahdollista leikkausta, mikä on täsmälleen puolet kaikista osajoukoista. Tämä tulos ei luonnollisesti päde yleisesti: on olemassa naapurustoja, joissa mahdollisten leikkausten määrä on sekä absoluuttisesti että suhteellisesti huomattavasti suurempi. Todistuksemme idea on siten etsiä ”pahin mahdollinen” naapurusto ja käyttää sen mahdollisten leikkausten määrää ylärajana jokaisen projektion koolle.

Tulemme käyttämään naapurustojen etsintään koneellista laskentaa, johon palaamme luvun lopussa. Tätä ennen esitämme todistuksemme rungon, jossa näytämme, miten yhtenäisten joukkojen määrä tarkalleen riippuu projektioiden koosta. Muistamme, että lemma 7 vaatii solmujen osajoukot siten, että jokainen solmu on vähintään δ osajoukossa. Koska käytämme solmujen ympäristöinä nyt yleistettyjä naapurustoja, haluamme valita luvuksi δ yleistetyn naapuruston suurimman mahdollisen koon.

Olkoon n_d niiden solmujen määrä, joiden etäisyys solmusta v on d kaikilla $0 \leq d \leq r$ naapurustossa $N^r[v]$. Selvästi $n_0 = 1$ ja $n_1 \leq \Delta$. Lisäksi kullakin etäisyydellä $d \geq 1$ olevalla solmulla voi olla enintään $\Delta - 1$ naapuria etäisyydellä $d + 1$. Siten $n_d \leq (\Delta - 1)n_{d-1}$ kaikilla $d \geq 2$. Yhdistämällä nämä saadaan

$$|N^r[v]| = \sum_{i=0}^{r} c_i \leq 1 + \Delta \sum_{i=0}^{r-1} (\Delta - 1)^i = 1 + \Delta \frac{1 - (\Delta - 1)^r}{1 - (\Delta - 1)} = \frac{\Delta(\Delta - 1)^r - 2}{\Delta - 2} .$$

Tämä on niin kutsuttu *Moore-raja* (engl. *Moore bound*) ja siitä käytetään merkintää $m_{\Delta,r}$ [MS05]. Moore-raja liittyy olennaisesti *Moore-verkkoihin*, joihin palaamme tarkemmin luvussa 5. On helppo nähdä, että naapuruston $N^r[v]$ koko saavuttaa Moore-rajan esimerkiksi silloin, kun $G[N^r[v]]$ on puu, jonka jokaisessa sisäsolmulla on Δ naapuria (kuva 11). Esimerkiksi 3-säännöllisessä verkossa kunkin solmun 2-säteinen



Kuva 11: Yksinkertaisin suurin mahdollinen naapurusto tapauksessa $\Delta = 3$ ja $r \in \{2, 3\}$. Muut suurimmat naapurustot voivat sisältää kaaria myös säteen r päässä olevien solmujen välillä.

naapurusto voi sisältää enintään $m_{3,2} = \frac{3(3-1)^2-2}{3-2} = 10$ solmua.

Olkoon nyt $\Delta \geq 3$, $r \geq 2$ ja $G = (V, E) \in \mathcal{G}_\Delta$. Haluamme soveltaa lemmaa 7 sijoituksella $\delta = m_{\Delta,r}$: toisin sanoen haluamme solmujen osajoukot A_v kaikilla $v \in V$ siten, että jokainen $u \in V$ on täsmälleen $m_{\Delta,r}$ joukossa A_v. Jos määrittelemme $A_v = N^r[v]$ kaikilla $v \in V$, tällöin kaikilla $u, v \in V$ pätee $u \in A_v \iff v \in N^r[u]$ ja siten jokainen $u \in V$ on täsmälleen $|N^r[u]|$ osajoukossa. Jos naapurusto $N^r[u]$ on kuvan 11 kaltainen maksimaalinen puu kaikilla $u \in V$, tällöin $|N^r[u]| = m_{\Delta,r}$ ja siten jokainen $u \in V$ on täsmälleen $m_{\Delta,r}$ joukossa A_v. Yleisessä tapauksessa kaikki naapurustot eivät luonnollisesti ole tällaisia. Saadaksemme ehdon voimaan kaikilla G joudumme siten jälleen lisäämään joukkoihin A_v ylimääräisiä solmuja. Nyt emme kuitenkaan voi lisätä solmuja mielivaltaisesti, koska haluamme pitää jokaisen A_v rajoitetun kokoisena. Tarkkaan ottaen haluamme seuraavanlaisen konstruktion.

Lemma 14. *Olkoon $\Delta \geq 3, r \geq 2, G = (V, E) \in \mathcal{G}_\Delta$ ja $\delta = m_{\Delta,r}$. Tällöin on olemassa joukon V osajoukot A_v kaikilla $v \in V$ siten, että*

1) Jokainen $u \in V$ esiintyy täsmälleen δ joukossa A_v.

2) $N^r[v] \subset A_v$ kaikilla $v \in V$.

3) $|A_v| \leq 2\delta$ kaikilla $v \in V$.

Ehdot (2) ja (3) toteutuvat selvästi sijoituksella $A_v = N^r[v]$ kaikilla $v \in V$, jonka jälkeen lisäämme joukkoihin solmuja solmuja saadaksemme myös ehdon (1) voimaan. Solmuja lisätessä joudumme pitämään huolta siitä, että ehto (3) säilyy: tämä ehto on olennainen, jotta voimme käsitellä erikseen sellaiset joukot, jotka sisältyvät johonkin A_v. Ei ole täysin selvää, että tällainen jako onnistuisi esimerkiksi lisäämällä puuttuvat

6	6	6	7	7	7	7
4	4	5	5	5	5	6
2	3	3	3	3	4	4
1	1	1	1	2	2	2
S_1	S_2	S_3	S_4	S_5	S_6	S_7

Kuva 12: Esimerkki lemman 15 mukaisesta jaosta joukolle $S = \{1, \ldots, 7\}$, kun $k = 4$.

solmut ahneesti ensimmäisiin joukkoihin, joissa niille on tilaa.

Todistamme lemman 14 käyttämällä konstruktiota, joka on intuitiivisesti yksinkertainen, mutta vaatii jonkin verran teknistä tarkastelua ollakseen formaalisti pätevä. Esitämme tätä varten ensin seuraavan apulauseen.

Lemma 15. *Olkoon S äärellinen joukko ja $|S| = p \geq k \in \mathbb{N}$. Tällöin on olemassa joukon S osajoukot $S_1, \ldots, S_p$ siten, että $|S_i| = k$ kaikilla $1 \leq i \leq p$ ja jokainen $s \in S$ esiintyy täsmälleen k osajoukossa.*

Sivuutamme formaalin todistuksen. Lemman 15 mukaiset joukot on kuitenkin helppo konstruoida esimerkiksi asettamalla joukon S alkiot jonoon, jossa jokainen $s \in S$ esiintyy k kertaa peräkkäin ja lisäämällä jonon jäsenet järjestyksessä joukkoihin $S_1, \ldots, S_p$ (kuva 12).

Annamme nyt todistuksen lemmalle 14.

Todistus. Lemman 15 nojalla kaikilla $v \in V$ on olemassa joukko $B_v \subset V$ siten, että $|B_v| = \delta$ ja jokainen $u \in V$ esiintyy täsmälleen δ joukossa B_v. Konstruoimme nyt joukot A_v seuraavasti: aluksi asetamme $A_v = N^r[v]$ kaikilla $v \in V$. Tämän jälkeen käymme läpi jokaisen $u \in V$, joka esiintyy $k < \delta$ joukossa. Edellisen konstruktion nojalla u esiintyy δ joukossa B_v, joten valitsemme mielivaltaisesti $\delta - k$ tällaista joukkoa B_v ja jokaisella niistä lisäämme solmun u joukkoon A_v.

Ehto (2) seuraa nyt suoraan alkuasetuksesta, koska sen jälkeen korkeintaan lisäämme joukkoihin solmuja. Myös ehto (1) pätee, koska jokainen solmu, joka kuuluu aluksi k joukkoon, lisätään $\delta - k$ joukkoon. Lisäksi koska kaikilla $v \in V$ joukkoon A_v valitaan solmuja vain joukoista $N^r[v]$ ja B_v, pätee $|A_v| \leq |N^r[v] \cup B_v| \leq |N^r[v]| + |B_v| \leq \delta + \delta = 2\delta$. Siten myös ehto (3) on voimassa. $\square$

Sovellamme nyt lemmaa 7 sijoittamalla $\delta = m_{\Delta,r}$ ja konstruoimalla joukot A_v lemman 14 mukaisesti. Kuten olemme edellä todenneet, emme halua käyttää suoraan sijoitusta $\mathcal{F} = \mathcal{C}$, koska haluamme sulkea pois tietynlaiset leikkaukset. Sen sijaan jaamme joukon $\mathcal{C}$ kahteen osaan ja annamme kummankin koolle ylärajan erikseen. Olkoon $\mathcal{C}'' = \{C \in \mathcal{C} : C \subset A_v \text{ jollain } v \in V\}$ ja $\mathcal{C}' = \mathcal{C} \setminus \mathcal{C}''$. Perhe $\mathcal{C}''$ sisältää siis jokaisen yhtenäisen joukon, joka on jonkin A_v osajoukko, ja perhe $\mathcal{C}'$ kaikki muut. Koska lemman 14 ehdon 3 nojalla jokaisen A_v koko on enintään 2δ, pätee

$$|\mathcal{C}''| \leq \sum_{v \in V} |2^{A_v}| = \sum_{v \in V} 2^{|A_v|} \leq n \cdot 2^{2\delta} ,$$

joka riippuu lineaarisesti solmujen määrästä. Selvästi suurin osa yhtenäisistä joukoista on perheessä $\mathcal{C}'$, joten käytämme sijoitusta $\mathcal{F} = \mathcal{C}'$. Määrittelemme nyt kaikilla $v \in V$ projektiot $\mathcal{C}_v = \{C \cap A_v : C \in \mathcal{C}'\}$, jolloin lemman 7 nojalla

$$|\mathcal{C}'|^{\delta} \leq \prod_{v \in V} |\mathcal{C}_v| . \tag{3}$$

Luvussa 4.2 osoitimme projektion $\mathcal{C}_v$ koolle ylärajan $|\mathcal{C}_v| \leq 2^{|A_v|} - 1$ kaikilla $v \in V$. Koska nyt pätee $N[v] \subset N^r[v] \subset A_v$ ja koska joukosta $\mathcal{C}'$ puuttuvat kaikki yhden solmut joukot, voisimme käyttää jälleen samaa argumenttia. Nyt haluamme kuitenkin osoittaa, että enintään jokin kiinteä osuus joukon A_v osajoukoista voi olla projektiossa $\mathcal{C}_v$. Olkoon siis c sellainen, että kaikilla $v \in V$ pätee $|\mathcal{C}_v| \leq c \cdot 2^{|A_v|}$. Selvästi esimerkiksi $c = 1$ on tällainen luku. Tällöin

$$\begin{aligned}
|\mathcal{C}'|^{\delta} &\leq \prod_{v \in V} \left(c \cdot 2^{|A_v|}\right) && \text{yhtälö 3} \\
&= c^n \prod_{v \in V} 2^{|A_v|} && c \text{ on tulossa vakio} \\
&= c^n \cdot 2^{\sum_{v \in V} |A_v|} && \text{potenssien tulosääntö} \\
&= c^n \cdot 2^{\delta n} . && \text{lemman 14 ehto (1)}
\end{aligned}$$

Korottamalla puolittain potenssiin $1/\delta$ saadaan $|\mathcal{C}'| \leq c^{n/\delta} \cdot 2^n$ ja edelleen

$$|\mathcal{C}| = |\mathcal{C}'| + |\mathcal{C}''| \leq c^{n/\delta} \cdot 2^n + n \cdot 2^{2\delta} \in \mathcal{O}((2\sqrt[\delta]{c})^n) .$$

Ylärajamme perheen $\mathcal{C}$ koolle riippuu täten siitä, kuinka hyvän arvion osaamme antaa luvulle c kullakin $\delta = m_{\Delta,r}$. Itse asiassa osoittautuu, että kaikilla Δ ja r on olemassa pienin luku $c_{\Delta,r}$ siten, että kaikilla mahdollisilla A_v pätee $|\mathcal{C}_v| \leq c_{\Delta,r} \cdot 2^{|A_v|}$. Tällöin saamme ylärajan $|\mathcal{C}| \in \mathcal{O}(\gamma_{\Delta,r}^n)$, missä $\gamma_{\Delta,r} = 2\sqrt[\delta]{c_{\Delta,r}}$ on täten yläraja luvulle

σ_Δ. Erityisesti $\gamma_{\Delta,1} = \beta_\Delta$ kaikilla $\Delta \geq 3$. Olennainen kysymys on, miten $\sqrt[\delta]{c_{\Delta,r}}$ käyttäytyy, kun r kasvaa rajatta. Se voi esimerkiksi saavuttaa globaalin optimin jollain r tai lähestyä jotain raja-arvoa, kun $r \to \infty$.

Näytämme seuraavaksi, että luku $c_{\Delta,r}$ seuraa käytännössä suoraan naapurustosta $N^r[v]$ eikä erityisesti riipu siitä, miten mahdolliset muut joukkoon A_v kuuluvat solmut on valittu. Siten luvun $c_{\Delta,r}$ määrittämiseksi riittää laskea projektion koko kaikille mahdollisille naapurustoille. Koska mahdollisia naapurustoja on äärellinen määrä kaikilla Δ ja r, tämä voidaan tehdä esimerkiksi koneellisella etsinnällä.

Tarkastelemme nyt kiinnittyä $v \in V$. Sanomme, että $U \subset A_v$ on *reunayhtenäinen* joukko, jos kaikilla $u \in U$ on polku $u_1 \ldots u_k x$, missä $u_i \in U$ kaikilla $1 \leq i \leq k, u_1 = u$ ja $x \in V \setminus A_v$. Erityisesti tyhjä joukko on reunayhtenäinen. Olkoon $\mathcal{C}_\partial$ tällaisten joukkojen perhe.

Lemma 16. *Pätee* $|\mathcal{C}_v| \leq |\mathcal{C}_\partial|$.

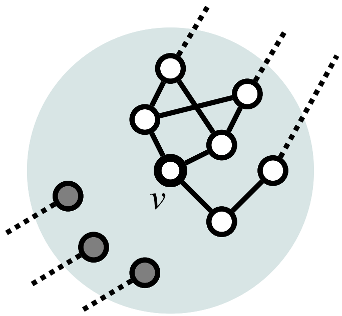
Todistus. Riittää osoittaa, että $U \in \mathcal{C}_v \Rightarrow U \in \mathcal{C}_\partial$ kaikilla $U \subset A_v$. Olkoon siis $U \in \mathcal{C}_v$. Joukon $\mathcal{C}_v$ määritelmän nojalla pätee $U = A_v \cap C$ jollain $C \in \mathcal{C}'$. Olkoon $X = C \setminus A_v$. Joukon $\mathcal{C}'$ määritelmän nojalla pätee $X \neq \emptyset$, joten olkoon $x \in X$. Nyt $U \cup X = (A_v \cap C) \cup (C \setminus A_v) = C$ on yhtenäinen, joten kaikilla $u \in U$ on polku $u \ldots x$. Siispä määritelmän nojalla $U \in \mathcal{C}_\partial$. $\square$

On helppo nähdä, että $|\mathcal{C}_v| \geq |\mathcal{C}_\partial|$ (ja siten edellisen perusteella $|\mathcal{C}_v| = |\mathcal{C}_\partial|$), jos $G[V \setminus A_v]$ on yhtenäinen. Täten luvun $c_{\Delta,r}$ määrittämiseksi annamme ylärajan joukon $\mathcal{C}_\partial$ koolle.

Näytämme nyt, että pahimmassa tapauksessa jokaisella $u \in A_v \setminus N^r[v]$ on naapureita ainoastaan joukon A_v ulkopuolella. Olkoot A'_v ja A''_v kaksi tapaa konstruoida joukko A_v. Olkoon A'_v mielivaltainen konstruktio, toisin sanoen $A'_v = N^r[v] \cup \{w'_1, \ldots, w'_k\}$, missä solmut $w'_i \in V \setminus N^r[v]$ ovat mielivaltaiset. Olkoon $A''_v = N^r[v] \cup \{w''_1, \ldots, w''_k\}$, missä jokaisella $w''_i \in V \setminus N^r[v]$ on naapureita joukossa $V \setminus A''_v$ muttei sen ulkopuolella. Jokaista $U' \subset A'_v$ vastaa nyt bijektiivisesti $U'' \subset A''_v$ siten, että $u \in U' \iff u \in U''$ kaikilla $u \in N^r[v]$ ja $w'_i \in U' \iff w''_i \in U''$ kaikilla $1 \leq i \leq k$. Näytämme, että U'' on reunayhtenäinen joukossa A''_v, jos U' on reunayhtenäinen joukossa A'_v.

Todistus. Olkoon U' reunayhtenäinen joukossa A'_v ja $u \in U''$. Riittää osoittaa, että on olemassa polku $u_1 \ldots u_m x$, missä $u_i \in U''$ kaikilla $1 \leq i \leq m, u_1 = u$ ja $x \in V \setminus A''_v$.

Tapaus $u \in N^r[v]$: tällöin pätee $u \in U'$ ja joukon U' reunayhtenäisyyden nojalla on olemassa polku $u_1 \ldots u_m x$, missä $u_i \in U'$ kaikilla $1 \leq i \leq m, u_1 = u$ ja $x \in V \setminus A'_v$.



Kuva 13: Pahimmassa tapauksessa A_v sisältää solmun v jonkin naapuruston sekä siitä erillisiä lisättyjä solmuja (tummennettu). Tällöin jokainen reunayhtenäinen joukko saadaan valitsemalla jokin naapuruston reunayhtenäinen joukko ja mielivaltainen kombinaatio lisättyjä solmuja. Olennaisesti naapuruston reunayhtenäisten joukkojen määrä kaksinkertaistuu jokaista lisättyä solmua kohti.

Olkoon y tämän polun ensimmäinen solmu, joka ei ole joukossa $N^r[v]$ ja u_l sitä edeltävä solmu. Nyt pätee $u_1, \ldots, u_l \in N^r[v]$ ja siten $u_1, \ldots, u_l \in U''$. Jos $y \notin A''_v$, tällöin $u_1 \ldots u_l y$ on haluttu polku. Jos $y \in A''_v$, tällöin $y = w''_i$ jollain $1 \leq i \leq k$. Solmulla w''_i on määritelmän nojalla naapuri $y' \in V \setminus A''_v$, joten $u_1 \ldots u_l w''_i y'$ on haluttu polku.

Tapaus $u \notin N^r[v]$: tällöin pätee $u = w''_i$ jollain $1 \leq i \leq k$. Kuten edellä, solmulla u on tällöin naapuri $y \in V \setminus A''_v$ ja uy on haluttu polku. □

Tästä seuraa, että naapurustolla A''_v on vähintään yhtä paljon reunayhtenäisiä joukkoja kuin naapurustolla A'_v.

Lemma 17. *Pätee* $|\mathcal{C}_\partial| \leq |\mathcal{C}_\partial \cap 2^{N^r[v]}| \cdot 2^{|A_v| - |N^r[v]|}$.

Todistus. Edellisen tarkastelun nojalla voimme nyt olettaa, että jokaisella $u \in M_v = A_v \setminus N^r[v]$ on naapureita ainoastaan joukon A_v ulkopuolella (kuva 13). Antamalla ylärajan tällaisille tapauksille annamme sen samalla kaikille muillekin.

Jokainen osajoukko $U \subset A_v$ voidaan esittää yhdisteenä $U_N \cup U_M$, missä $U_N \subset N^r[v]$ ja $U_M \subset M_v$. Koska oletuksen mukaan kaikilla $u \in N^r[v], w \in M_v$ pätee $uw \notin E$, joukkojen $N^r[v]$ ja M_v välillä ei ole polkuja joukossa A_v. Tästä seuraa, että pätee $U \in \mathcal{C}_\partial$, jos ja vain jos pätee $U_N \in \mathcal{C}_\partial$ ja $U_M \in \mathcal{C}_\partial$. Itse asiassa $U_M \in \mathcal{C}_\partial$ pätee nyt kaikilla $U_M \subset M_v$, koska oletuksen mukaan joukon M_v jokaisella solmulla on naapuri joukossa $V \setminus A_v$. Siten jokainen $U \in \mathcal{C}_\partial$ saadaan minkä tahansa $U_M \subset M_v$ ja $U_N \in \mathcal{C}_\partial$

kombinaationa. Siispä pahimmassa tapauksessa pätee

$$|\mathcal{C}_\partial| = |\mathcal{C}_\partial \cap 2^{N^r[v]}||2^{M_v}| = |\mathcal{C}_\partial \cap 2^{N^r[v]}| \cdot 2^{|M_v|} = |\mathcal{C}_\partial \cap 2^{N^r[v]}| \cdot 2^{|A_v|-|N^r[v]|} .$$

$\square$

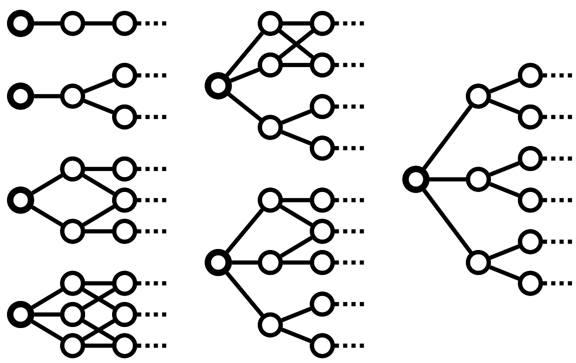
Olkoon $f_{\Delta,r}(a)$ joukon $\mathcal{C}_\partial \cap 2^{N^r[v]}$ suurin mahdollinen koko, kun pätee $|N^r[v]| = a$, missä $1 \leq a \leq \delta$. Nyt edellä todistettujen lemmojen nojalla pätee

$$|\mathcal{C}_v| \leq |\mathcal{C}_\partial \cap 2^{N^r[v]}| \cdot 2^{|A_v|-|N^r[v]|} \leq \max_{1\leq a\leq\delta} f_{\Delta,r}(a) \cdot 2^{|A_v|-a} = \left(\max_{1\leq a\leq\delta} 2^{-a} f_{\Delta,r}(a)\right) 2^{|A_v|} .$$

Täten $c_{\Delta,r} = \max\{2^{-a} f_{\Delta,r}(a) : 1 \leq a \leq \delta\}$. Koska mahdollisia naapurustoja on äärellinen määrä, $f_{\Delta,r}(a)$ ja siten luku $c_{\Delta,r}$ voidaan määrittää koneellisesti kaikilla $1 \leq a \leq \delta$. Esitämme seuraavaksi tämän laskennan tulokset sekä niistä seuraavan ylärajan.

4.4 Menetelmän tulokset ja rajoitukset

Osoittautuu, että edellä esitetty menetelmä antaa yleistä ylärajaa β_Δ paremman tuloksen tapauksessa $\Delta \in \{3, 4\}$ ja $r = 2$. Lisäksi otaksumme, että myös tapauksessa $\Delta = 5$ voidaan osoittaa hieman parempi yläraja. Koneellinen laskenta käy tässä tapauksessa kuitenkin työlääksi, joten otaksuma on toistaiseksi varmistamatta. Sen sijaan vaikuttaa siltä, että menetelmä ei paranna yleistä ylärajaa, kun $\Delta > 5$. Lisäksi säteellä $r > 2$ ei oletettavasti saada parannusta millään Δ.



Kuva 14: Pahimman tapauksen naapurustot kaikilla $3 \leq n \leq 10$, kun $\Delta = 3, r = 2$.

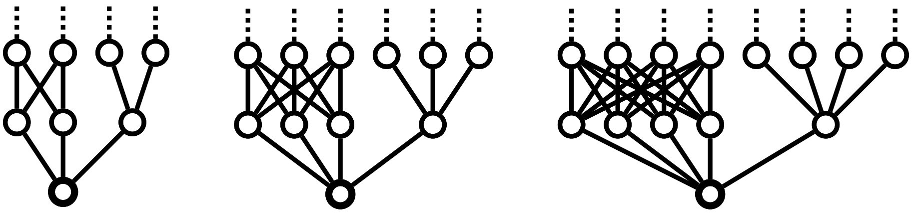
Taulukko 2: Reunayhtenäisten joukkojen enimmäismäärä naapurustoissa, kun $\Delta \in \{3, 4\}$ ja $r = 2$. Tapauksessa $a < 3$ ainoastaan tyhjä joukko on reunayhtenäinen.

a	$f_{3,2}(a)$	$f_{3,2}(a) \cdot 2^{-a}$	a	$f_{4,2}(a)$	$f_{4,2}(a) \cdot 2^{-a}$
3	4	0.5000	3	4	0.5000
4	10	0.6250	4	10	0.6250
5	22	0.6875	5	22	0.6875
6	44	0.6875	6	50	0.7813
7	91	0.7110	7	106	0.8282
8	184	**0.7188**	8	212	0.8282
9	366	0.7149	9	430	0.8399
10	730	0.7129	10	866	0.8458
			11	1 744	**0.8516**
			12	3 470	0.8472
			13	6 922	0.8450
			14	13 826	0.8439
			15	27 508	0.8395
			16	54 872	0.8373
			17	109 600	0.8362

Lause 18. *Pätee* $\sigma_3 < 1.9351$ *ja* $\sigma_4 < 1.9812$.

Todistus. Olemme laskeneet funktion $f_{\Delta,r}$ koneellisesti, kun $\Delta \in \{3, 4\}$ ja $r = 2$ (taulukko 2, kuva 14). Näemme, että tapauksessa $\Delta = 3$ tulo $f(a) \cdot 2^{-a}$ saa suurimman arvonsa, kun naapuruston koko $a = 8$. Siispä $c_{3,2} = f(8) \cdot 2^{-8} = 184 \cdot \frac{1}{256} = \frac{23}{32}$. Nyt aiemman todistuksen nojalla $\sigma_3 \leq \gamma_{3,2} = 2 \sqrt[10]{23/32} < 1.9351$. Vastaavasti tapauksessa $\Delta = 4$ tulo maksimoituu tapauksessa $a = 11$. Siten $c_{4,2} = f(11) \cdot 2^{-11} = \frac{1744}{2048} = \frac{109}{128}$ ja $\sigma_4 \leq \gamma_{4,2} = 2 \sqrt[17]{109/128} < 1.9812$. □

Tarkastellessamme pahimman tapauksen naapurustoja, kun $\Delta \in \{3, 4\}$, havaitsemme, että niillä on samankaltainen rakenne (kuva 15). Otaksumme, että tämä yleistyy muillekin asterajoille: toisin sanoen kaikilla $\Delta \geq 3$ pahimman tapauksen naapurusto voidaan määritellä solmun v sekä kaksijakoisten verkkojen $K_{\Delta-1,\Delta-1}$ ja $K_{1,\Delta-1}$ yhdisteenä, jossa v on yhdistetty Δ solmuun ja muut solmut naapuruston ulkopuolelle. Tällainen naapurusto sisältää $3\Delta - 1$ solmua ja sillä on $2^{3\Delta-1} - 5 \cdot 2^{2\Delta-2} + 2^{\Delta}$ reunayhtenäistä joukkoa, minkä näyttäminen on yksinkertainen laskenta-argumentti.



Kuva 15: Pahimman tapauksen naapurusto säteellä $r = 2$, kun $\Delta \in \{3, 4\}$, sekä otaksutusti, kun $\Delta = 5$.

Näitä tietoja käyttäen voimme laskea menetelmän antaman ylärajan kaikille Δ sillä oletuksella, että otaksuma pahimmasta tapauksesta pitää paikkansa. Tämä laskelma on esitetty pienille Δ taulukossa 3.

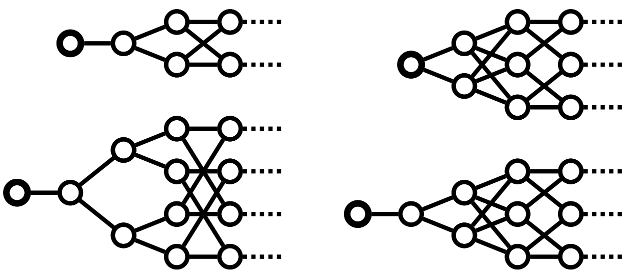
Konjektuuri 19. *Pätee* $c_{5,2} = 15136 \cdot 2^{-14}$, *jonka seurauksena* $\sigma_5 < 1.9940$.

Näemme, että tapauksessa $\Delta = 5$ otaksuttu pahin mahdollinen naapurusto sisältää 14 solmua ja 15136 reunayhtenäistä joukkoa. Suurin mahdollinen naapurusto sisältää $m_{5,2} = 26$ solmua, joten menetelmämme antaa konjektuurin 19 mukaisen ylärajan $\sqrt[26]{15136 \cdot 2^{-14}} < 1.9940$, jos otaksuma pätee. Tämä on hieman parempi kuin aiempi yläraja $\beta_5 < 1.9948$. Sen sijaan vaikuttaa selvältä, että uusi yläraja on aiempaa huonompi kaikilla $\Delta > 5$ siinäkin tapauksessa, että pahempia naapurustoja ei ole.

Niin ikään vaikuttaa siltä, että säteellä $r > 2$ ei saada parempaa ylärajaa millään Δ. Erityisesti tapauksessa $\Delta \in \{3, 4\}$ on todistetusti olemassa naapurustoja, jotka johtavat väistämättä huonompaan ylärajaan, kun $r \in \{3, 4\}$ (kuva 16). On syytä huomata, että nämä naapurustot eivät ole edes säännöllisiä ja siten todennäköisesti

Taulukko 3: Menetelmän antama yläraja pienillä Δ, jos kuvan 15 mukainen naapurusto on pahin tapaus.

Δ	$\lvert\mathcal{C}_\partial\rvert$	$\gamma_{\Delta,2}$ (?)	β_Δ
3	184	1.9351	1.9680
4	1 744	1.9812	1.9874
5	15 136	1.9940	1.9948
6	126 016	1.9979	1.9978
7	1 028 224	1.9993	1.9991
8	8 306 944	1.9997	1.9996



Kuva 16: Esimerkkejä naapurustoista, kun $\Delta, r \in \{3, 4\}$. Menetelmä antaa väistämättä huonomman ylärajan siinäkin (epäuskottavassa) tapauksessa, että nämä olisivat pahimpia mahdollisia naapurustoja.

kaukana pahimmasta tapauksesta.

Vaikka esittämämme menetelmä antaa paremman ylärajan tapauksessa $\Delta \in \{3, 4\}$, vaikuttaa selvältä, että tämäkin tulos on yhä kaukana todellisesta. Yksi ilmeinen syy on, että menetelmän todistus käyttää suurimman projektion kokoa ylärajana jokaiselle projektiolle. On kuitenkin helppo nähdä, että verkon kaikki naapurustot eivät voi olla kuvan 15 mukaisia otaksuttuja pahimman tapauksen naapurustoja. Kuten näytämme luvussa 5, vaikuttaa itse asiassa siltä, että eniten yhtenäisiä joukkoja sisältävien verkkojen naapurustot ovat kuvan 11 kaltaisia puita. Yläraja ei kuitenkaan parane merkittävästi edes siinä tapauksessa, että voimme olettaa kaikkien naapurustojen olevan tällaisia.

Taulukko 4: Menetelmän antama yläraja pienillä Δ ja r, kun kaikki naapurustot ovat puumaisia.

$r \backslash \Delta$	3	4	5	6	7	8
1	1.9680	1.9874	1.9948	1.9978	1.9991	1.9996
2	1.9335	1.9791	1.9924	1.9970	1.9988	1.9995
3	1.9206	1.9772	1.9920	1.9969	1.9987	1.9995
4	1.9158	1.9766	1.9919	1.9969	1.9987	1.9995
5	1.9136	1.9764	1.9918	1.9969	1.9987	1.9995
6	1.9125	1.9763	1.9918	1.9969	1.9987	1.9995
7	1.9120	1.9763	1.9918	1.9969	1.9987	1.9995
8	1.9117	1.9763	1.9918	1.9969	1.9987	1.9995
9	1.9116	1.9763	1.9918	1.9969	1.9987	1.9995

Koska reunayhtenäisyys on tässä yhteydessä olennaisesti sama asia kuin luvussa 3.3 esitetty lehtiyhtenäisyys, annettua rekursiokaavaa käyttäen voidaan laskea reunayhtenäisten joukkojen määrä jokaiselle Δ-säännölliselle r-syvyiselle puulle. Käyttämällä tätä ylärajana luvun 4.3 menetelmässä saamme edelleen ylärajan siinä tapauksessa, että kaikki naapurustot ovat puumaisia (taulukko 4). Näemme, että parannus on tässäkin optimistisessa tapauksessa enintään marginaalinen. Täten vaikuttaa selvältä, että optimaalisen ylärajan saavuttaminen vaatii merkittävämpiä muutoksia menetelmään tai toisenlaista lähestymistapaa.

Toteamme lopuksi, että menetelmä yleistyy lähes sellaisenaan yhtenäisille dominoiville joukoille. Tämän havaitseminen on sinänsä yksinkertainen mutta formaalisti työläs päättely, joten jätämme sen esittämättä. Olennaisesti laskemme koneellisesti osajoukot, jotka voivat olla leikkauksia yhtenäisten dominoivien joukkojen kanssa ja saamme siten funktiota $f_{\Delta,r}$ vastaan funktion $g_{\Delta,r}$. Osoittautuu, että tapauksessa $r = 2$ ja $\Delta \in \{3, 4\}$ maksimi saavutetaan samoissa naapurustoissa kuin yhtenäisillä joukoilla (taulukko 5). Soveltamalla aiempaa todistusta saadaan $|\mathcal{E}| \in \mathcal{O}(2\sqrt[10]{156 \cdot 2^{-8}}) \subset \mathcal{O}(1.9034^n)$, kun $\Delta = 3$ ja $|\mathcal{E}| \in \mathcal{O}(2\sqrt[17]{1638 \cdot 2^{-11}}) \subset \mathcal{O}(1.9739^n)$, kun $\Delta = 4$.

Taulukko 5: Leikkausten suurin mahdollinen määrä, kun $\Delta \in \{3, 4\}$ ja $r = 2$.

a	$g_{3,2}(a)$	$g_{3,2}(a) \cdot 2^{-a}$	a	$g_{4,2}(a)$	$g_{4,2}(a) \cdot 2^{-a}$
3	2	0.2500	3	2	0.2500
4	6	0.3750	4	6	0.3750
5	18	0.5625	5	18	0.5625
6	34	0.5313	6	42	0.6563
7	74	0.5782	7	98	0.7657
8	156	**0.6094**	8	192	0.7500
9	296	0.5782	9	388	0.7579
10	576	0.5625	10	794	0.7754
			11	1 638	**0.7999**
			12	3 182	0.7769
			13	6 270	0.7654
			14	12 446	0.7597
			15	24 238	0.7397
			16	47 822	0.7298
			17	94 990	0.7248

5 Alaraja rajoitetun asteen verkoissa

Edellisissä luvuissa olemme etsineet ylärajaa yhtenäisten joukkojen lukumäärälle rajoitetun asteen verkoissa. Erityisesti olemme näyttäneet, että n-solmuisella verkolla, jonka suurin aste on Δ, on enintään $(2^{\Delta+1} - 1)^{n/(\Delta+1)}$ yhtenäistä joukkoa, sekä parantaneet tätä rajaa jonkin verran tapauksessa $\Delta \in \{3, 4\}$. Vaikuttaa kuitenkin uskottavalta, että mikään näistä rajoista ei ole tiukka. Tässä luvussa tarkastelemme ongelman toista suuntaa. Kysymme: kuinka monta yhtenäistä joukkoa tällaisella verkolla voi *ainakin* olla? Jos voimme antaa yhtenäisten joukkojen enimmäismäärälle alarajan, saamme jonkinlaisen arvion sille, kuinka paljon nykyisissä ylärajoissa on mahdollisesti parantamisen varaa.

Luonteva tapa lähestyä alarajaongelmaa on konstruoida solmujen lukumäärällä n parametrisoitu perhe rajoitetun asteen verkkoja ja arvioida yhtenäisten joukkojen määrää näissä verkoissa. Esitämme tässä luvussa kaksi tällaista konstruktiota, timanttiketjut luvussa 5.1 sekä tikapuuverkot luvussa 5.2, ja annamme yhtenäisten joukkojen määrälle asymptoottisesti tiukan arvion (taulukko 6).

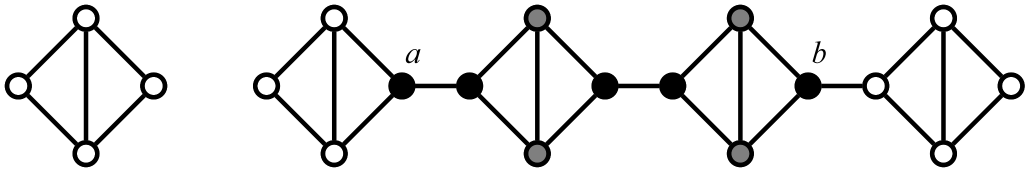
Luvussa 5.3 käsittelemme lyhyesti optimaalisia verkkoja pienillä n ja Δ. Osoittautuu, että verkot, joilla $|\mathcal{C}|$ saavuttaa ylärajan $C_{\Delta,n}$, ovat monessa mielessä kiinnostavia. Empiiristen havaintojen perusteella otaksumme erityisesti, että näiden verkkojen perhe sisältää kaikki yksikäsitteiset häkit ja siten muun muassa kaikki Moore-verkot.

5.1 Timanttiketjut

Timanttiverkko (engl. *diamond graph*) on 4 solmun verkko, joka saadaan esimerkiksi poistamalla kaari täydestä verkosta K_4 tai lisäämällä kaari sykliin C_4. Timanttiverkolla on selvästi suurin aste 3 ja kaksi solmua, joiden aste on 2. Kaksi timanttiverkkoa voidaan yhdistää valitsemalla kummastakin asteen 2 solmu ja lisäämällä niiden välille kaari. Selvästi näin saatavalla verkolla on edelleen suurin aste 3 ja kaksi asteen 2

Taulukko 6: Alaraja luvulle σ_Δ timanttiketjujen (δ_Δ) ja tikapuuverkkojen (λ_Δ) yleistyksillä pienille Δ.

Δ	3	4	5	6	7
δ_Δ	1.3160	1.4757	1.5704	1.6332	1.6784
λ_Δ	1.5537	1.6180	1.7320		

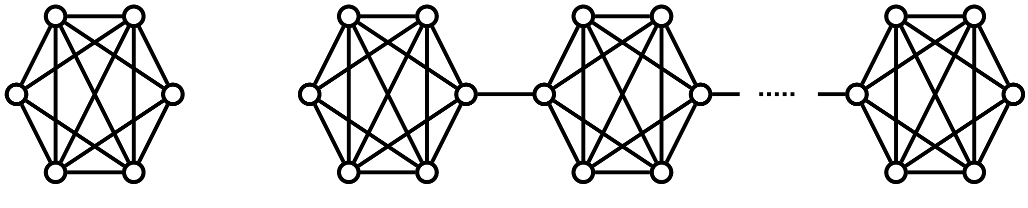


Kuva 17: Timanttiverkko ja neljän timantin ketju. Jos yhtenäinen joukko sisältää liitossolmut a, b, sen täytyy sisältää myös liitossolmut niiden väliltä sekä vähintään toinen solmu (harmaat) kunkin liitossolmun timantista.

solmua. Siten voimme yhdistää mielivaltaisen määrän timantteja toisiinsa ja kutsumme jokaista näin saatavaa verkkoa *timanttiketjuksi*. Sanomme timanttiketjun solmua *liitossolmuksi*, jos sillä on aste 2 tai naapuri ketjun toisessa timantissa.

Havaitsemme, että yhtenäisten joukkojen suhteen timanttiketju on olennaisesti polkumainen verkko: jos yhtenäinen joukko sisältää kaksi liitossolmua, sen täytyy sisältää myös kaikki liitossolmut niiden ”väliltä” (kuva 17). Koska kaksi solmua riittää määräämään liitossolmujen kombinaation, kombinaatioiden määrä on n-solmuisella timanttiketjulla luokkaa $\mathcal{O}(n^2)$. Olennaisempi tekijä ovat kunkin timantin kaksi muuta solmua, joista voimme valita jommankumman tai molemmat. Jos verkko sisältää k timanttia, mahdollisten kombinaatioiden määrä on siten luokkaa $\Theta^*(3^k) = \Theta^*(3^{n/4})$. Täten saamme $\sigma_3 \geq \sqrt[4]{3} > 1.3160$.

Voimme yleistää konstruktion myös muille asterajoille. Määrittelemme kaikilla $\Delta \geq 3$ verkon D_Δ, joka saadaan poistamalla kaari verkosta $K_{\Delta+1}$, ja kutsumme tätä asteen Δ timantiksi. Esimerkiksi D_4 saadaan poistamalla kaari verkosta K_5 ja D_3 on tavallinen timanttiverkko. Timantilla D_Δ on nyt analogisesti suurin aste Δ ja kaksi solmua, joiden aste on $\Delta - 1$. Yleistämme edelleen timanttiketjun tarkoittamaan verkkoa, joka saadaan yhdistämällä mielivaltainen määrä saman asteen timantteja edellä esitetyllä tavalla. Formaalisti $D_{\Delta,1}$ on verkko D_Δ ja $D_{\Delta,k}$ kaikilla $k > 1$ saadaan ottamalla verkkojen $D_{\Delta,k-1}$ ja $D_{\Delta,1}$ yhdiste, valitsemalla kummastakin asteen $\Delta - 1$ solmu ja



Kuva 18: Yleinen timantti D_5 ja sen ketju.

lisäämällä niiden välille kaari (kuva 18).

Liitossolmujen kombinaatioiden määrä yhtenäisissä joukoissa on edelleen luokkaa $\mathcal{O}(n^2)$, joten riittää laskea sellaiset joukot, jotka sisältävät kaikki liitossolmut. Jokainen timantti sisältää nyt $\Delta + 1$ solmua, joista $\Delta - 1$ voivat muodostaa minkä tahansa kombinaation tyhjää joukkoa lukuun ottamatta. Siten yhtenäisten joukkojen määrä k timantin verkossa on luokkaa $\Theta^*((2^{\Delta-1} - 1)^k) = \Theta^*((2^{\Delta-1} - 1)^{n/(\Delta+1)})$. Täten saamme $\sigma_\Delta \geq \delta_\Delta = (2^{\Delta-1} - 1)^{1/(\Delta+1)}$ kaikilla $\Delta \geq 3$.

5.2 Tikapuuverkot

Tikapuuverkko (engl. *ladder graph*) saadaan intuitiivisesti kahden saman pituisen polkuverkon yhdisteenä, joissa "vastakkaiset" solmut polkujen välillä on yhdistetty (kuva 19). Seuraavassa analyysissä käytämme eksplisiittistä määritelmää $L_k = (V, E)$, missä

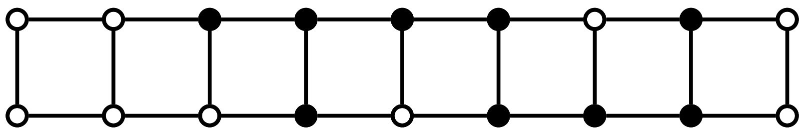
$$\begin{cases} V = \{u_1, \ldots, u_k, u'_1, \ldots u'_k\} \ , \\ E = \bigcup_{i=1}^{k-1} \{u_i u_{i+1}, u'_i u'_{i+1}\} \cup \bigcup_{i=1}^{k} \{u_i u'_i\} \ . \end{cases}$$

Tikapuuverkoilla on selvästi suurin aste $\Delta = 3$. Näytämme seuraavaksi, että yhtenäisten joukkojen määrä kasvaa tikapuuverkoissa asymptoottisesti nopeammin kuin timanttiketjuissa.

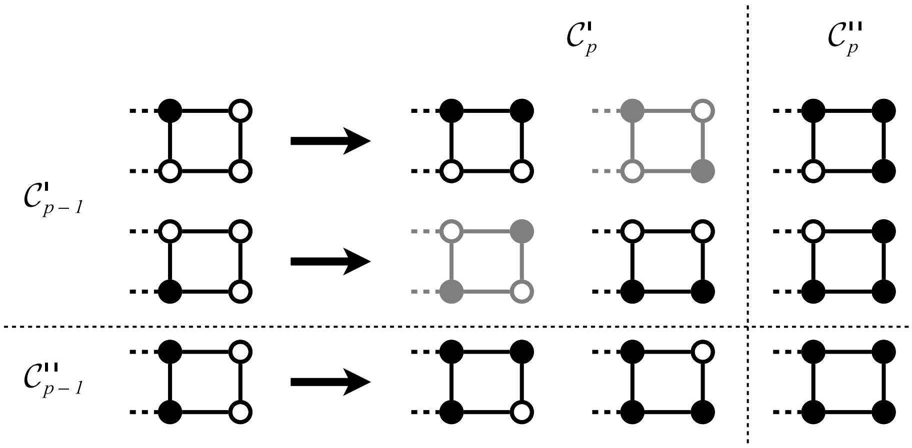
Lause 20. *Yhtenäisten joukkojen määrä verkossa L_k on luokkaa $\Theta^*\left((1 + \sqrt{2})^k\right)$.*

Todistus. Tarkastelemme verkkoa L_k mielivaltaisella k. Havaitsemme edellisen konstruktion tapaan, että riittää laskea sellaiset yhtenäiset joukot, jotka sisältävät vähintään yhden solmun kultakin "poikkipuolalta". Kaikki muut yhtenäiset joukot saadaan selvästi tällaisten joukkojen ja alitikapuiden leikkauksina. Koska alitikapuiden määrä on luokkaa $\mathcal{O}(k^2)$, tämä ei vaikuta eksponentiaaliseen tekijään.

Olkoon täten U_p ensimmäisten p puolan joukko $\bigcup_{i=1}^{p} \{u_i, u'_i\}$ kaikilla $1 \leq p \leq k$ ja olkoon $\mathcal{C}_p \subset 2^{U_p}$ sellaisten yhtenäisten joukkojen perhe, jotka sisältävät kaikilla $1 \leq i \leq p$ ainakin toisen solmuista u_i ja u'_i. Jaamme nyt perheen $\mathcal{C}_p$ kahteen osaan:



Kuva 19: Tikapuuverkko L_9 ja yksi sen yhtenäisistä joukoista.



Kuva 20: Jokainen yhtenäinen joukko perheessä $\mathcal{C}_p$ saadaan valitsemalla yhtenäinen joukko perheestä $\mathcal{C}_{p-1}$ ja lisäämällä jompikumpi tai molemmat solmuista u_p ja u'_p.

olkoon $\mathcal{C}'_p \subset \mathcal{C}_p$ niiden joukkojen perhe, jotka sisältävät täsmälleen toisen solmuista u_p ja u'_p, ja $\mathcal{C}''_p \subset \mathcal{C}_p$ niiden, jotka sisältävät molemmat. Täten $|\mathcal{C}_p| = |\mathcal{C}'_p| + |\mathcal{C}''_p|$ kaikilla $1 \leq p \leq k$.

Arvioimme nyt näiden kahden joukon kokoa kaikilla $1 \leq p \leq k$. Triviaalisti pätee $\mathcal{C}'_1 = \{\{u_1\}, \{u'_1\}\}$ ja $\mathcal{C}''_1 = \{\{u_1, u'_1\}\}$ ja siten $|\mathcal{C}'_1| = 2, |\mathcal{C}''_1| = 1$. Olkoon nyt $p > 1$. Näemme kuvaa 20 katsomalla, että tällöin jokainen $U \in \mathcal{C}'_p$ on jokin seuraavista kombinaatioista:

1) $U' \cup \{u_p\}$; missä $U' \in \mathcal{C}''_{p-1}$, tai $U' \in \mathcal{C}'_{p-1}$ ja $u_{p-1} \in U'$

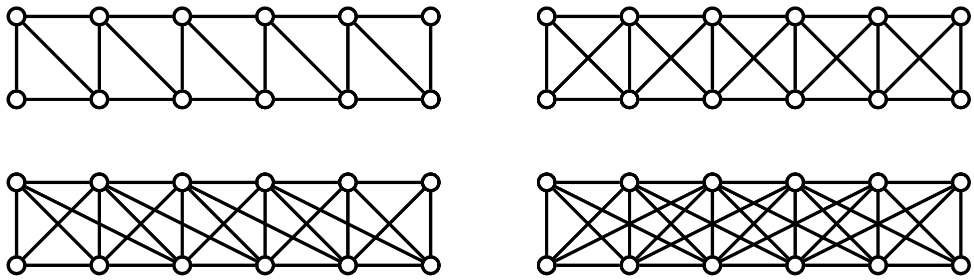
2) $U' \cup \{u'_p\}$; missä $U' \in \mathcal{C}''_{p-1}$, tai $U' \in \mathcal{C}'_{p-1}$ ja $u'_{p-1} \in U'$.

Siten perheessä $\mathcal{C}'_p$ on yksi joukko jokaista $U' \in \mathcal{C}'_{p-1}$ kohti ja kaksi joukkoa jokaista $U' \in \mathcal{C}''_{p-1}$ kohti. Vastaavasti jokainen $U \in \mathcal{C}''_p$ saadaan kombinaationa $U' \cup \{u_p, u'_p\}$, missä pätee joko $U' \in \mathcal{C}'_{p-1}$ tai $U' \in \mathcal{C}''_{p-1}$. Siten perheessä $\mathcal{C}''_p$ on yksi joukko kutakin $U \in \mathcal{C}'_{p-1}$ ja kutakin $U \in \mathcal{C}''_{p-1}$ kohti. Nämä havainnot yhdistämällä saadaan

$$\begin{cases} |\mathcal{C}'_p| = |\mathcal{C}'_{p-1}| + 2|\mathcal{C}''_{p-1}| \, , & (1) \\ |\mathcal{C}''_p| = |\mathcal{C}'_{p-1}| + |\mathcal{C}''_{p-1}| \, . & (2) \end{cases}$$

Näemme, että $\mathcal{C}'_p$ ja $\mathcal{C}''_p$ kasvavat asymptoottisesti yhtä nopeasti, joten riittää arvioida joukon $\mathcal{C}'_p$ kasvua. Merkitsemällä lyhyemmin $a_p = |\mathcal{C}'_p|$ ja $b_p = |\mathcal{C}''_p|$ voidaan kaikilla $p > 2$ kirjoittaa

$$\begin{aligned} a_p - a_{p-1} &= 2b_{p-1} \, , && \text{(yhtälö 1)} \\ &= 2a_{p-2} + 2b_{p-2} \, , && \text{(yhtälö 2)} \\ &= 2a_{p-2} + (a_{p-1} - a_{p-2}) && \text{(yhtälö 1)} \end{aligned}$$



Kuva 21: Tikapuuverkon L_6 yleistys asterajoilla $4 \leq \Delta \leq 7$.

ja sieventämällä edelleen

$$a_p = 2a_{p-1} + a_{p-2} \ .$$

Muotoa $h_n = c_1 h_{n-1} + c_2 h_{n-2} + \cdots + c_d h_{n-d}$ oleva yhtälö, jossa $c_1, c_2, \ldots, c_d$ ovat vakioita, on asteen d lineaarinen homogeeninen differenssiyhtälö. Sanomme tällöin, että $t^d - c_1 t^{d-1} - c_2 t^{d-2} - \cdots - c_d$ on differenssiyhtälön karakteristinen polynomi. Jos karakteristisella polynomilla on erisuuret juuret $r_1, r_2, \ldots, r_d$, tällöin yhtälöllä on suljettu muoto $h_n = q_1 r_1^n + q_2 r_2^n + \cdots + q_d r_d^n$ joillain vakioilla $q_1, \ldots, q_d$ [GK82].

Näemme, että $a_p = 2a_{p-1} + a_{p-2}$ on toisen asteen lineaarinen homogeeninen differenssiyhtälö kertoimilla $c_1 = 2, c_2 = 1$ ja sen karakteristinen polynomi on siten $t^2 - 2t - 1$. Tällä on erisuuret juuret $r_1 = 1 + \sqrt{2}$ ja $r_2 = 1 - \sqrt{2}$, joten yhtälöllä on ratkaisu $a_p = q_1(1+\sqrt{2})^p + q_2(1-\sqrt{2})^p$ joillain q_1, q_2. Määrittääksemme vakiot ratkaisemme ensin kaksi ensimmäistä termiä: $a_1 = 2$ ja $a_2 = a_1 + 2b_1 = 2 + 2 \cdot 1 = 4$. Tästä saamme lineaarisen yhtälöryhmän

$$\begin{cases} q_1(1+\sqrt{2})^1 + q_2(1-\sqrt{2})^1 = 2 \ , \\ q_1(1+\sqrt{2})^2 + q_2(1-\sqrt{2})^2 = 4 \ , \end{cases}$$

jolla on yksikäsitteinen ratkaisu $q_1 = \sqrt{2}/2, q_2 = -\sqrt{2}/2$. Täten

$$a_p = \frac{\sqrt{2}}{2}(1+\sqrt{2})^p - \frac{\sqrt{2}}{2}(1-\sqrt{2})^p \ .$$

Siispä $|\mathcal{C}(L_k)| \in \Theta^*(|\mathcal{C}_k|) = \Theta^*(|\mathcal{C}'_k|) = \Theta^*((1+\sqrt{2})^k)$. □

Solmujen määrä on nyt $n = 2k$, joten saamme, että yhtenäisten joukkojen määrä tikapuuverkoissa on luokkaa $\Theta^*(\lambda_3^n)$, missä $\lambda_3 = \sqrt{1+\sqrt{2}} > 1.5537$.

Tämäkin konstruktio on mahdollista yleistää muille asterajoille. Tapauksessa $\Delta = 4$ voidaan lisätä esimerkiksi kaari $u_i u'_{i+1}$ kaikilla $1 \leq i \leq k-1$ (kuva 21), jolloin yhtenäisten joukkojen määrä saadaan vastaavalla päättelyllä differenssiyhtälöstä

$$a_k = 3a_{k-1} - a_{k-2} \ .$$

Tälle saadaan edelleen $a_k \in \Theta^*((\frac{3+\sqrt{5}}{2})^k) = \Theta^*(\lambda_4^n)$, missä $\lambda_4 = \sqrt{\frac{3+\sqrt{5}}{2}} = \frac{1+\sqrt{5}}{2} > 1.6180$. Tapauksessa $\Delta = 5$ voidaan lisätä symmetrisesti myös kaari $u_i'u_{i+1}$ kaikilla $1 \leq i \leq k-1$. Tällöin $U \in \mathcal{C}_k$, jos ja vain jos vain kaikilla $1 \leq i \leq k$ ainakin toinen solmuista u_i, u_i' kuuluu joukkoon U. Siten yhtenäisten joukkojen määrä on $\Theta^*(3^k) = \Theta^*(\lambda_5^n)$, missä $\lambda_5 = \sqrt{3} > 1.7320$. Konstruktiota on mahdollista jatkaa lisäämällä edelleen toinen tai molemmat kaarista $u_iu_{i+j}', u_i'u_{i+j}$ kaikilla $1 \leq j \leq c$, missä c riippuu asterajasta Δ. Jätämme yleisen tapauksen analyysin esittämättä.

5.3 Optimaalisista verkoista

Olemme edellä antaneet alarajan luvulle σ_Δ kaikilla $\Delta \geq 3$ konstruoimalla solmujen määrällä n parametrisoituja verkkoperheitä ja analysoimalla yhtenäisten joukkojen määrää, kun n kasvaa rajatta. Vaikuttaa selvältä, että esittämämme konstruktiot eivät ole asymptoottisesti optimaalisia ja että alarajassa on vielä huomattavasti parantamisen varaa. Ei ole kuitenkaan selvää, millaisesta verkkoperheestä parempi alaraja voisi löytyä ja kuinka helposti se on osoitettavissa.

Tässä luvussa pyrimme vastaamaan tähän kysymykseen tarkastelemalla yksittäisiä optimaalisia verkkoja ja niiden ominaisuuksia. Sanomme, että n-solmuinen verkko G, jolla on suurin aste $\Delta \geq 3$, on yhtenäisten joukkojen suhteen optimaalinen, jos $|\mathcal{C}(G)| = C_{\Delta,n}$. Osoittautuu, että monet tällaisista verkoista ovat tunnettuja nimettyjä verkkoja, jotka maksimoivat tai minimoivat myös jonkin toisen kiinnostavan ominaisuuden, tai johdettavissa tällaisista verkoista yksinkertaisilla operaatioilla. Käsittelemme tarkemmin erityisesti tunnettuja optimointiongelmia, joissa pyrimme annetun asterajoituksen puitteissa minimoimaan verkon läpimitan tai maksimoimaan *leveyden* eli lyhimmän syklin pituuden. Otaksumme empiiristen havaintojen pohjalta, että näiden ongelmien ratkaisuihin lukeutuvat *Moore-verkot* sekä laajemmin kaikki *häkit* maksimoivat myös yhtenäisten joukkojen määrän.

Lähestymme kysymystä etsimällä optimaalisia verkkoja koneellisesti pienillä n ja Δ sekä vertaamalla niitä tunnettuihin verkkoihin ja verkkoperheisiin. Tämän katsauksen tulokset on esitetty yhteenvetona taulukoissa 7, 8 ja 9 kullakin $\Delta \in \{3,4,5\}$. Verkot on järjestetty yhtenäisten joukkojen määrän $|\mathcal{C}|$ mukaan ja kullakin n alin verkko (korostettu) on paras, jonka olemme onnistuneet löytämään. Verkot on nimetty joko

Taulukko 7: Joitain 3-säännöllisiä verkkoja ja niiden ominaisuuksia.

n	Verkko		δ	g	$\lvert\mathcal{C}\rvert$	
4	$\mathbf{K_4}$		**1**	**3**	16	häkki, Moore, $n_{3,1}$
6	$\mathbf{K_{3,3}}$		**2**	**4**	56	häkki, Moore
8	$GP_{4,1}, Q_3$	Kuutio	3	**4**	168	
	$\mathbf{M_8}$	**Wagner**	**2**	**4**	174	
10	$\mathbf{GP_{5,2}}$	**Petersen**	**2**	**5**	569	häkki, Moore, $n_{3,2}$
12	$GP_{6,2}$	Dürer	4	3	1 395	
	$GP_{6,1}$		4	4	1 441	
		Bidiakis-kuutio	**3**	4	1 531	
		Franklin	**3**	4	1 667	
	MAX		**3**	**5**	1 765	
14	$\mathbf{GH_{1,2}}$	**Heawood**	**3**	**6**	5 770	häkki, Moore
16	$\mathbf{GP_{8,3}}$	**Möbius-Kantor**	4	**6**	18 044	
18		Pappus	4	**6**	57 248	
	MAX		4	**6**	57 865	
20	$GP_{10,2}$	Dodekaedri	5	5	152 799	
	$GP_{10,3}$	Desargues	5	**6**	178 963	
	$GP_{10,4}$		4	5	179 379	
	J_5	Flower snark	4	5	179 973	
	$C_5 * F_4$		**3**	5	184 383	$n_{3,3}$
	MAX		4	**6**	185 363	
22	$GP_{11,3}$		**4**	**6**	584 962	
	MAX		**4**	**6**	596 029	
24	$GP_{12,5}$	Nauru	**4**	**6**	1 880 278	
		McGee	**4**	**7**	1 926 728	häkki
26		F26A	5	6	6 004 624	
	$\mathbf{GP_{13,5}}$		**4**	**7**	6 200 458	
28	$GP_{14,4}$		5	7	19 905 783	
		Coxeter	**4**	**7**	19 931 475	
	MAX		5	**7**	20 018 897	
30	$GP_{15,4}$		5	7	64 256 164	
		Tutte-Coxeter	**4**	**8**	64 904 615	häkki, Moore
32		Dyck	5	6	198 246 018	
	$GP_{16,6}$		5	7	208 773 469	
	MAX		5	7	209 348 074	

Taulukko 8: Joitain 4-säännöllisiä verkkoja ja niiden ominaisuuksia.

n	Verkko	δ	g	$\lvert\mathcal{C}\rvert$	
5	**$\mathbf{K_5}$**	**1**	**3**	32	häkki, Moore, $n_{4,1}$
6	**Oktaedri**	**2**	**3**	61	
7	**MAX**	**2**	**3**	115	
8	**$\mathbf{K_{4,4}}$**	**2**	**4**	234	häkki, Moore
9	$QC_{2,1}$	**2**	**3**	398	
	MAX	**2**	**3**	426	
10	**MAX**	**2**	**4**	839	
11	**4-Andrásfai**	**2**	**4**	1 564	
12	Chvátal	**2**	**4**	2 994	
	MAX	**2**	**4**	2 996	
13	**13-cyclotomic**	**2**	**4**	5 735	
14	**MAX**	**2**	**4**	10 955	
15	$K_3 * C_5$	**2**	3	20 712	$n_{4,2}$
	MAX	3	**4**	20 946	
16	Q_4, Terrasekti	4	**4**	37 294	
	Hoffman	4	**4**	37 878	
	MAX	**3**	**4**	40 461	
17	**MAX**	**3**	**4**	76 963	
18	**MAX**	**3**	**4**	148 368	
19	**Robertson**	**3**	**5**	285 205	häkki
20	Folkman	4	4	498 270	
	MAX	**3**	**5**	546 279	
21	Brinkmann	**3**	**5**	1 046 632	
	MAX	**3**	**5**	1 047 189	
22	**MAX**	**3**	**5**	2 009 404	
23	**MAX**	**3**	**5**	3 854 180	
24	**MAX**	4	**5**	7 404 890	
25	**MAX**	4	**5**	14 216 614	
26	**$\mathbf{GH_{1,3}}$**	**3**	**6**	27 327 419	häkki
27	Doyle, Holt	**3**	5	52 278 101	
27	**MAX**	4	5	52 499 729	

Taulukko 9: Joitain 5-säännöllisiä verkkoja ja niiden ominaisuuksia.

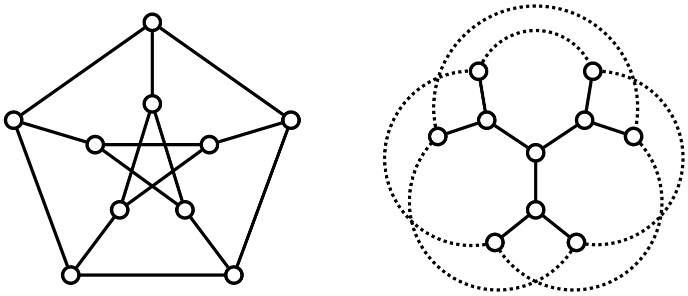
n	Verkko	δ	g	$\lvert\mathcal{C}\rvert$	
6	$\mathbf{K_6}$	**1**	**3**	64	häkki, Moore, $n_{5,1}$
8	**MAX**	**2**	**3**	242	
10	$\mathbf{K_{5,5}}$	**2**	**4**	972	häkki, Moore
12	Ikosaedri	3	3	3 211	
	MAX	3	3	3 594	
14	5-Andrásfai	**2**	**4**	13 864	
	MAX	3	**4**	13 929	
16	16-cyclotomic, Clebsch	**2**	**4**	53 082	
	MAX	3	**4**	53 386	
18	**MAX**	3	**4**	205 466	
20	**MAX**	3	**4**	789 785	
22	**MAX**	3	**4**	3 038 473	
24	$K_3 * X_8$	**2**	3	11 541 328	$n_{5,2}$
	MAX	3	**4**	11 708 194	
26	**MAX**	**3**	**4**	45 113 938	
28	**MAX**	**3**	**4**	174 049 489	
30	Robertson-Wegner	**3**	**5**	670 459 956	häkki
	Wong	**3**	**5**	670 631 650	häkki
	Foster	**3**	**5**	670 671 451	häkki
	Meringer	**3**	**5**	670 724 230	häkki

systemaattisia nimiä tai muita tunnettuja nimiä käyttäen. Parhaasta löydetystä verkosta käytämme merkintää **MAX** niissä tapauksissa, joissa emme ole löytäneet verkolle muuta nimeä. Otaksumme, että suurin osa parhaista löydetyistä verkoista on optimaalisia: todistetusti tämä pätee kuitenkin vain kaikilla $n \leq 12$, kun $\Delta \in \{3, 4\}$, ja kaikilla $n \leq 10$, kun $\Delta = 5$. Taulukoissa esitettyjen nimettyjen verkkojen osalta viittaamme Wolfram Mathworld -sivustoon [Wol] sekä Exoon ym. [EJ08] ja Millerin ym. [MS05] katsauksiin. Verkkojen niin sanottu tähtitulo (engl. *star product*) $G * H$ on esitetty jälkimmäisessä.

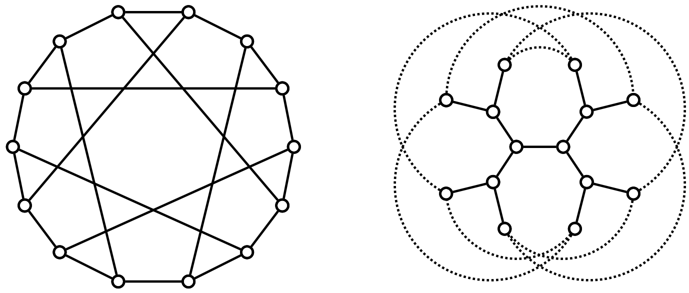
Ensimmäinen ilmeinen havainto on, että parhaat löytämämme verkot sisältävät enimmäismäärän kaaria. Koska tämä vaikuttaa intuitiivisesti selvältä, olemme yksinkertaisuuden vuoksi esittäneet taulukoissa ainoastaan säännölliset verkot. Syvällisempi havainto on, että parhaat verkot vaikuttaisivat olevan jokaisesta solmusta katsoen kuvan 11 kaltaisia puita, joissa lehtisolmut on yhdistetty toisiinsa. Myös Perrier ym. [PIM08] tekevät saman havainnon vastaavanlaisessa katsauksessa, jossa yhtenäisten joukkojen määrää maksimoidaan koneellisesti. Havainto voidaan ilmaista yhtäpitävästi toteamalla, että verkon lyhimmän syklin pituus g on suuri. Koska kaikki naapurustot ovat puumaista, tästä seuraa selvästi myös, että verkon läpimitta δ on suhteellisen pieni. Läpimitta on korostettu taulukoissa 7, 8 ja 9 niillä verkoilla, joilla olemme varmistaneet sen minimaaliseksi solmujen määrällä n. Vastaavasti leveys on korostettu, kun se on varmistetusti suurin mahdollinen.

Näiden havaintojen pohjalta on luonteva ajatus etsiä tiukkaa alarajaa verkoista, joilla on pieni läpimitta ja suuri leveys. Sekä läpimitan minimoiminen että leveyden maksimoiminen ovat kuitenkin ei-triviaaleja ongelmia säännöllisissä verkoissa. Näistä ensimmäinen on niin sanottu aste–läpimitta -ongelma (engl. *degree diameter problem*), joka on ollut tunnettu avoin ongelma verkkoteoriassa noin vuodesta 1960 alkaen [MS05]. Tavanomaisessa asettelussa kiinnitämme suurimman asteen Δ ja kysymme sen puitteissa suurinta mahdollista verkkoa, jolla on enintään annettu läpimitta δ. Olkoon $n_{\Delta,\delta}$ solmujen määrä tällaisessa verkossa. Jälkimmäisessä ongelmassa taas etsimme pientä Δ-säännöllistä verkkoa, jolla on annettu leveys g. Kutsumme pienintä tällaista verkkoa (Δ, g)-*häkiksi* (engl. *cage*) ja käytämme sen solmujen määrästä merkintää $\mu_{\Delta,g}$ [EJ08]. Voidaan näyttää, että tällainen verkko on olemassa kaikilla Δ ja g [Bol78].

Luvulle $n_{\Delta,\delta}$ saadaan yksinkertainen yläraja toteamalla, että läpimittaehdon nojalla kaikki solmut ovat enintään etäisyydellä δ jostain kiinnitetystä solmusta. Kuten näytimme luvussa 4.3, r-säteisen naapuruston suurin mahdollinen koko saadaan



Kuva 22: Petersenin verkon yleinen esitys ja esitys Moore-verkkona.



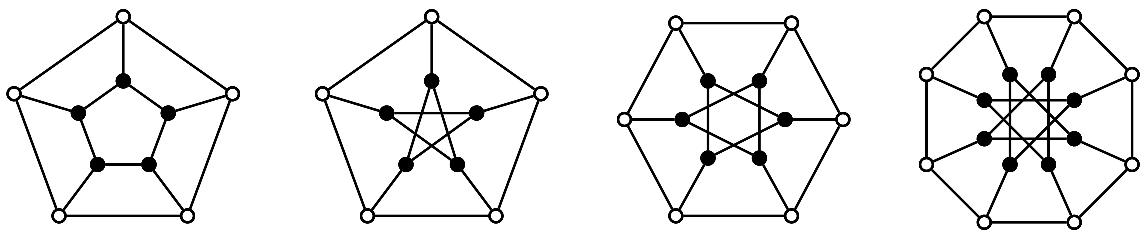
Kuva 23: Heawoodin verkon yleinen esitys ja esitys Moore-verkkona.

Moore-rajasta $m_{\Delta,r}$ ja siten pätee $n_{\Delta,\delta} \leq m_{\Delta,\delta}$ kaikilla Δ ja δ. Tekemällä vastaava tarkastelu leveydelle on helppo nähdä, että $m_{\Delta,\delta}$ on myös alaraja luvulle $\mu_{\Delta,g}$, kun $g = 2\delta + 1$ [Bol78, Bol98]. Parillisille g saadaan samankaltainen konstruktio ja nämä yhdistämällä yleinen alaraja $\mu_{\Delta,g} \geq M_{\Delta,g}$, missä

$$M_{\Delta,g} = \begin{cases} 1 + \Delta \displaystyle\sum_{i=0}^{(g-3)/2} (\Delta - 1)^i \underset{(\Delta>2)}{=} \dfrac{\Delta(\Delta-1)^{(g-1)/2} - 2}{\Delta - 2}, & \text{kun } g \text{ on pariton,} \\ 2 \displaystyle\sum_{i=0}^{g/2-1} (\Delta - 1)^i \underset{(\Delta>2)}{=} \dfrac{(\Delta-1)^{g/2} - 2}{\Delta - 2}, & \text{kun } g \text{ on parillinen} \end{cases}$$

kaikilla $\Delta \geq 2$ ja $g \geq 3$. Tässä $M_{\Delta,g}$ on Moore-rajan yleistys, joka on parametrisoitu läpimitan sijaan leveydellä. Parittomilla $g = 2\delta + 1$ pätee nyt erityisesti $M_{\Delta,g} = m_{\Delta,\delta}$.

Häkkejä, joissa solmujen määrä saavuttaa alarajan $M_{\Delta,g}$, kutsutaan *Moore-verkoiksi* (engl. *Moore graph*). Esimerkiksi Petersenin verkko (kuva 22) ja Heawoodin verkko (kuva 23) ovat yksikäsitteiset $(3, 5)$- ja $(3, 6)$-häkit ja molemmat Moore-verkkoja. Edellisen tarkastelun perusteella Moore-verkoilla on suurin mahdollinen leveys ja parittomilla g myös pienin mahdollinen läpimitta. Lisäksi kaikki tarkastelemamme Moore-verkot vaikuttaisivat olevan optimaalisia yhtenäisten joukkojen suhteen ja siten lupaava kandidaatti optimaaliseksi perheeksi.



Kuva 24: Yleistetyt Petersenin verkot $GP_{5,1}$, $GP_{5,2}$, $GP_{6,2}$ ja $GP_{8,3}$. Verkkoperheen jokainen jäsen muodostetaan yhdistämällä sykli (valkoiset solmut) tähtimonikulmioon (mustat solmut).

Moore-verkot ovat kuitenkin osoittautuneet suhteellisen harvinaisiksi. On helppo nähdä, että tapauksessa $\Delta = 2$ Moore-verkkoja ovat täsmälleen kaikki syklit, tapauksessa $g = 3$ kaikki täydelliset verkot ja tapauksessa $g = 4$ kaikki täydelliset kaksijakoiset verkot. Näiden triviaalitapausten lisäksi lisäksi Moore-verkkoja on olemassa vain, kun $g \in \{5, 6, 8, 12\}$. Tapauksessa $g = 5$ ainoat Moore-verkot saadaan, kun $\Delta \in \{2, 3, 7\}$, sekä mahdollisesti, kun $\Delta = 57$. Muilla g Moore-verkko on olemassa, jos ja vain jos on olemassa symmetrinen yleistetty n-kulmio, joka sisältää täsmälleen $\Delta - 1$ solmua [EJ08].

Vaikka Moore-verkot ovat harvinaisia, katsauksen perusteella vaikuttaa uskottavalta, että yhtenäisten joukkojen määrä maksimoituu myös häkeissä yleisesti. Häkit eivät kuitenkaan ole yksikäsitteisiä kaikilla Δ ja g: esimerkiksi $(5, 5)$-häkki voi viitata neljään eri verkkoon, jotka on esitetty taulukon 9 lopussa. Yleisesti otaksumme siten, että ainakin yksi (Δ, g)-häkeistä on optimaalinen kaikilla $\Delta \geq 2$ ja $g \geq 3$.

Tapauksessa $\Delta = 3$ olemme häkkien lisäksi kiinnostuneita erityisesti niin sanotuista yleistetyistä Petersenin verkoista, jotka saadaan yhdistämällä n-solmuinen sykliverkko n-solmuiseen tähtimonikulmioon kuvan 24 mukaisella tavalla. Formaalisti kaikilla $k \geq 3$ ja $1 \leq m \leq \lfloor k/2 \rfloor$ *yleistetty Petersenin verkko* (engl. *generalized Petersen graph*) $GP_{k,m}$ sisältää syklin $u_1 \dots u_k$ sekä solmut $u'_1, \dots, u'_k$, kaaren $u_i u'_i$ kaikilla $1 \leq i \leq k$ ja kaaren $u'_i u'_j$ kaikilla $i \equiv_m j$ [Bol78]. Katsauksen perusteella vaikuttaa siltä, että monilla k on ainakin yksi m siten, että $\mathcal{C}(GP_{k,m})$ on hyvin lähellä optimia. Erityisesti $GP_{5,2}$ eli Petersenin verkko itse on todistetusti optimaalinen ja $GP_{8,3}$ ja $GP_{13,5}$ ovat parhaiden löydettyjen joukossa. Otaksumme, että nykyistä alarajaa voisi parantaa huomattavasti yleistettyjen Petersenin verkkojen tarkemmalla analyysillä.

6 Jatkotutkimus

Entropiaan perustuva menetelmä on toistaiseksi ainoa tunnettu tekniikka, joka antaa yhtenäisten joukkojen määrälle ei-triviaalin ylärajan rajoitetun asteen verkoissa. Ylärajan todistus on luonteeltaan paikallinen tarkastelu, jossa arvioimme, kuinka monella tavalla yhtenäiset joukot voivat leikata kunkin solmun suljetun naapuruston. Olemme laajentaneet tätä menetelmää tarkastelemalla suljettujen naapurustojen sijaan solmujen yleistettyjä naapurustoja, jotka sisältävät kaikki säteen $r > 1$ sisällä olevat solmut. Käymällä tällaiset naapurustot läpi koneellisesti tapauksessa $r = 2$ olemme edelleen osoittaneet, että laajennus parantaa aiempaa ylärajaa, kun verkon suurin aste $\Delta \in \{3, 4\}$, sekä mahdollisesti, kun $\Delta = 5$. Vastaavat tulokset saadaan menetelmää soveltamalla myös yhtenäisille dominoiville joukoille. Sen sijaan kokeellisen katsauksen perusteella vaikuttaa selvältä, että pahimman tapauksen naapurustot johtavat väistämättä huonompaan ylärajaan, kun $\Delta > 5$ tai $r > 2$.

Koska sekä yleinen että tapauksen $\Delta \in \{3, 4\}$ yläraja ovat edelleen suhteellisen kaukana kokeellisesta arviosta, on luontevaa kysyä, missä määrin entropiaan perustuvaa menetelmää voidaan parantaa edelleen. Ensimmäinen ilmeinen havainto on, että menetelmä käyttää nykyisellään pahimman tapauksen naapurustoa ylärajana kaikille naapurustoille. Koska kaikki naapurustot eivät voi olla pahimman tapauksen kaltaisia, on selvää, että menetelmän antama yläraja on ainakin jossain määrin yliarvioitu. Toisaalta kaikilla Δ ja r on mahdollista, että jokaisen solmun naapurusto on Δ-säännöllinen r-säteinen puu. Kuten olemme näyttäneet, menetelmän antama parannus ylärajaan on tässä tapauksessa enintään marginaalinen pahimpaan tapaukseen nähden ja edelleen kaukana kokeellisesta arviosta.

Toinen havainto on, että naapurustojen keskimääräinen koko δ on menetelmässä aina täsmälleen Moore-raja $m_{\Delta,r}$. Vaikka säteen kasvattaminen pienentää leikkausten suhteellista määrää, δ vaikuttaa kasvavan suhteessa liian nopeasti, kun $r > 2$. Avoimeksi kysymykseksi jää, miten yläraja käyttäytyy, jos joukkojen keskimääräiseksi kooksi valitaan Moore-rajan sijaan mielivaltainen luku esimerkiksi välillä $]m_{\Delta,1}, m_{\Delta,3}[$. Alustavan katsauksen perusteella arvelemme, että yläraja ei tällöin parane ainakaan millään $\delta < m_{\Delta,2}$, mutta tämä on toistaiseksi varmistamatta.

Kolmanneksi toteamme, että esitetty projektiolause ei yleisessä tapauksessa anna tiukkaa ylärajaa edes silloin, kun osaamme määrittää mahdollisen leikkausten lukumäärän tarkasti. Täten kysymme, voitaisiinko mahdollisten leikkausten määrää vähentää muullakin tavalla kuin naapurustojen sopivalla valinnalla, esimerkiksi sulkemalla yhtenäisten joukkojen perheestä pois muitakin joukkoja kuin naapurustojen osajoukot. Mikäli tämä ei onnistu, tarkempi analyysi saattaa vaatia perustavampia muutoksia entropiamenetelmään tai kokonaan toisenlaista lähestymistapaa.

Myös paras tunnettu alaraja vaikuttaa olevan huomattavan kaukana asymptoottisen kasvun kokeellisesta arviosta. Siinä missä r-säännöllisen verkon riippumattomille ja dominoiville joukoille saadaan tiukka alaraja verkkojen K_{r+1} ja $K_{r,r}$ erillisillä yhdisteillä, yhtenäisille joukoille emme tunne vastaavaa yksinkertaista konstruktiota. Tunnetut alarajat seuraavat timanttiketjujen ja tikapuuverkkojen yleistyksistä, kun taas katsauksemme pieniin verkkoihin antaa uskoa, että optimaalisella verkkoperheellä tulisi olla pieni läpimitta ja suuri leveys. Tässä valossa otaksumme, että tiukka tai ainakin parempi alaraja voitaisiin saavuttaa tällaisten verkkoperheiden tarkemmalla analyysillä.

Lähteet

AF93 Avis, D. ja Fukuda, K., Reverse search for enumeration. *Discrete Applied Mathematics*, 65(1993), sivut 21–46.

BHKK08a Björklund, A., Husfeldt, T., Kaski, P. ja Koivisto, M., Computing the tutte polynomial in vertex-exponential time. *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science, FOCS'08*. IEEE Computer Society, 2008, sivut 677–686.

BHKK08b Björklund, A., Husfeldt, T., Kaski, P. ja Koivisto, M., The travelling salesman problem in bounded degree graphs. *Proceedings of the 35th International Colloquium on Automata, Languages and Programming, ICALP'08, Part I*, osa 5125 sarjasta *Lecture Notes in Computer Science*. Springer, 2008, sivut 198–209.

BHKK10 Björklund, A., Husfeldt, T., Kaski, P. ja Koivisto, M., Trimmed moebius inversion and graphs of bounded degree. *Theory of Computing Systems*, 47,3(2010), sivut 637–654.

Bol78 Bollobás, B., *Extremal graph theory*. Numero nid. 11 sarjassa L.M.S. monographs. Academic Press, 1978.

Bol98 Bollobás, B., *Modern graph theory*. Graduate texts in mathematics. Springer, 1998.

CGFS86 Chung, F. R., Graham, R. L., Frankl, P. ja Shearer, J. B., Some intersection theorems for ordered sets and graphs. *Journal of Combinatorial Theory, Series A*, 43,1(1986), sivut 23–37.

EJ08 Exoo, G. ja Jajcay, R., Dynamic Cage Survey. *The Electronic Journal of Combinatorics*, 15(2008).

Epp03 Eppstein, D., The traveling salesman problem for cubic graphs. *Proceedings of the 8th International Workshop on Algorithms and Data Structures, WADS'03*, osa 2748 sarjasta *Lecture Notes in Computer Science*. Springer, 2003, sivut 307–318.

FGPS05 Fomin, F. V., Grandoni, F., Pyatkin, A. V. ja Stepanov, A. A., On maximum number of minimal dominating sets in graphs. *Electronic Notes in Discrete Mathematics*, 22(2005), sivut 157–162.

GJR+07 Gutin, G., Johnstone, A., Reddington, J., Scott, E., Soleimanfallah, A. ja Yeo, A., An algorithm for finding connected convex subgraphs of an acyclic digraph. *Algorithms and Complexity in Durham ACID*, 7,D(2007), sivut 1–15.

GK82 Greene, D. ja Knuth, D., *Mathematics for the analysis of algorithms*. Progress in computer science. Birkhäuser, 1982.

KOU11 Kijima, S., Okamoto, Y. ja Uno, T., Dominating set counting in graph classes. *Proceedings of the 17th Annual International Conference on Computing and Combinatorics, COCOON'11*, osa 6842 sarjasta *Lecture Notes in Computer Science*. Springer, 2011, sivut 13–24.

MS05 Miller, M. ja Siran, J., Moore graphs and beyond: A survey of the degree–diameter problem. *Electronic Journal of Combinatorics*, Dynamic Survey DS14(2005).

NP06 Niculescu, C. ja Persson, L., *Convex functions and their applications: a contemporary approach*. Numero nid. 13 sarjassa CMS books in mathematics. Springer, 2006.

PIM08 Perrier, E., Imoto, S. ja Miyano, S., Finding optimal Bayesian network given a super-structure. *Journal of Machine Learning Research*, 9(2008), sivut 2251–2286.

Rad01 Radhakrishnan, J., Entropy and counting. *IIT Kharagpur, Golden Jubilee Volume, on Computational Mathematics, Modelling and Algorithms*, (2001).

Rom92 Roman, S., *Coding and information theory*. Graduate texts in mathematics. Springer, 1992.

Wol Wolfram Mathworld, `http://mathworld.wolfram.com`. [26.4.2012].

Woo07 Wood, D. R., On the maximum number of cliques in a graph. *Graphs and Combinatorics*, 23(2007), sivut 337–352.

Zha10 Zhao, Y., The number of independent sets in a regular graph. *Combinatorics, Probability and Computing*, 19,2(2010), sivut 315–320.