Tampereen yliopisto

Mikael Nyström

Suppenemislaki satunnaisverkoille

Informaatioteknologian ja viestinnän tiedekunta
Pro gradu -tutkielma
Matematiikka
Toukokuu 2020

TIIVISTELMÄ

Mikael Nyström: Suppenemislaki satunnaisverkoille
Pro gradu -tutkielma
Tampereen yliopisto
Matematiikan ja tilastotieteen tutkinto-ohjelma
Toukokuu 2020

Tässä tutkielmassa todistetaan, että satunnaisesti muodostetun verkon kasvaessa rajatta voidaan aina löytää raja-arvo todennäköisyydelle, että verkkoa koskeva ensimmäisen kertaluvun logiikan väittämä pitää paikkansa. Tätä varten tehdään joitakin tarpeellisia satunnaisverkon astejonoa koskevia oletuksia. Tutkielma perustuu James F. Lynchin artikkelissaan Convergence Law for Random Graphs With Specified Degree Sequence tekemään tutkimukseen.

Tullaan huomaamaan, että raja-arvon olemassaoloa koskevan väitteen todistuksessa ei tarvitse perehtyä satunnaisten skaalautumattomien verkkojen tutkimiseen, mikä olisi vaivalloista. Sen sijaan näytetään, että riittää tarkastella helpommin satunnaisesti muodostettavaa rakennetta, jota kutsutaan kokoonpanoksi. Samalla perehdytään verkkojen ominaisuuksia kuvaaviin rakenteisiin, kuten puihin, metsiin sekä ympäristöihin.

Verkkojen ominaisuuksien vertailemiseen käytetään Ehrenfeuchtin peliä, joka esitellään pinnallisesti. Tämän ohella suuri osa todistuksesta pohjautuu verkkojen kombinatoriseen tarkasteluun samaistamalla verkkojen käyttäytyminen niin kutsuttuihin haarautumisprosesseihin. Lisäksi esitellään tarvittavia todennäköisyyslaskennan ja kombinatoriikan tuloksia, jotka helpottavat päälauseen todistamista.

Avainsanat: satunnaisverkko, astejono, kokoonpano
Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -ohjelmalla.

Sisältö

1 Johdanto

Satunnaisverkot ovat laajalti käytettyjä työkaluja eri tieteenaloilla, kuten teknillisissä tieteissä tai sosiologiassa. Niillä voidaan havainnollistaa muun muassa tautien leviämistä, ravintoketjuja sekä Internetin rakennetta. Näille esimerkeille yhteinen piirre on, että niiden käyttäytymistä kuvastavat ennalta arvaamattomat prosessit, missä lisätään ja poistetaan pisteiden välisiä yhteyksiä. Matematiikassa satunnaisverkkojen teoria yhdistää verkkoteoriaa ja todennäköisyyslaskentaa.

Tämän tutkielman tavoitteena on todistaa suppenemislaki satunnaisverkoille, joiden astejono on kiinnitetty. Osoitetaan siis, että jos G on satunnaisesti muodostettu n-solmuinen verkko ja φ on verkon ominaisuutta kuvaava ensimmäisen kertaluvun logiikan lause, niin todennäköisyys, että verkolla G on ominaisuus φ, suppenee, kun n kasvaa rajatta. Tuloksesta voi olla käytännön hyötyä esimerkiksi suurten tietokantojen kyselyalgoritmien kehittämisessä. Tutkielmassa seurataan läheisesti menetelmiä, jotka James F. Lynch esittelee artikkelissaan Convergence Law for Random Graphs With Specified Degree Sequence [8].

Tutkielman alussa perehdytään verkkoteorian perusteisiin, jotka esitellään teoksen [2] mukaisesti. Verkko on rakenne, joka koostuu joukosta solmuja ja joitakin solmupareja yhdistävistä särmistä. Solmusta lähtevien särmien lukumäärää kutsutaan solmun asteeksi ja verkon astejonolla tarkoitetaan luonnollisten lukujen jonoa, jonka i:s jäsen kuvaa niiden solmujen lukumäärää, joiden aste on i.

Tämän jälkeen siirrytään määrittelemään joitakin tarvittavia käsitteitä logiikasta, joista tärkeimmät ovat lauseen kvanttorisyvyys sekä Ehrenfeuchtin peli. Näistä jälkimmäinen on pelaajien A ja B välinen peli, joka osoittautuu tehokkaaksi työkaluksi verkkojen ja muiden samankaltaisten rakenteiden vertailussa. Kaksi rakennetta määritellään k-ekvivalenteiksi, mikäli rakenteiden ominaisuuksille pätevät täsmälleen samat k-syvyiset ensimmäisen kertaluvun logiikan väittämät. Tässä tutkielmassa käytetään myös alunperin Ehrenfeuchtin todistamaa tulosta, jonka mukaan kaksi rakennetta ovat k-ekvivalentit, mikäli pelaajalla B on voittostrategia k-kierroksisessa Ehrenfeuchtin pelissä, joka pelataan näillä rakenteilla. Logiikkaan liittyvät määritelmät mukailevat teoksen [12] lähestymistapaa.

Jotta äärettömän suuriksi kasvavien verkkojen tutkiminen olisi mielekästä, tarkasteltaville astejonoille joudutaan asettamaan joitakin rajoituksia, jotka eivät kuitenkaan vaikuta tutkielman päälauseen kattavuuteen. Näistä konkreettisimpana esimerkkinä voidaan pitää sitä, että tarkasteltavissa verkoissa ei sallita olevan eristettyjä solmuja.

Satunnaisverkkojen muodostaminen ja perinpohjainen tutkiminen on vaivalloista. Tästä syystä määritellään helpommin satunnaisesti muodostettavissa oleva rakenne, jota kutsutaan kokoonpanoksi. Tullaan huomaamaan, että kokoonpanoja ja kokoonpanoista muodostettavia tekijäverkkoja kuvaavat ominaisuudet ja erityisesti niiden todennäköisyydet voidaan helposti kääntää verkkoja kuvaaviksi. Tästä syystä suurin osa loppututkielmasta käsittelee kokoonpanoja ja niiden rakennetta, eli keskitytään todistamaan suppenemislaki satunnaiskokoonpanoille. Keskeisimpänä käsitteenä on solmun r-säteinen ympäristö, jonka ominaisuuksia tarkastelemalla saadaan runsaasti tietoa kokoonpanon käyttäytymisestä. Tällaisen ympäristön ominaisuuksia kutsutaan kokoonpanon paikallisiksi ominaisuuksiksi. Erityisesti äärellisten ja erillisten ympäristöjen lukumäärä ja käyttäytyminen määrittävät, onko jokin rakennetta koskeva lause totta. Ollaan myös kiinnostuneita ympäristöjen indusoimien alikokoonpanojen rakenteesta ja erityisesti niiden syklisyydestä. Syklillä tarkoitetaan solmujen yhdistymistä toisiinsa kehän kaltaisesti.

Kokoonpanojen vertaileminen tapahtuu määrittelemällä ekvivalenssirelaatio $\sim_k$ ja näyttämällä, että mikäli kaksi kokoonpanoa ovat relaatiossa $\sim_k$, ovat ne myös k-ekvivalentit.

Loput päätuloksen todistuksesta tehdään kombinatorisesti. Aloitetaan määrittelemällä puuprosessi ja esittelemällä niiden tarkasteluun tarvittavia merkintöjä ja keskeisiä tuloksia. Viimeisessä luvussa esitellään myös joitain todennäköisyysteorian tuloksia, joita tarvitaan päälauseen todistamiseen.

Tämän jälkeen näytetään, että melkein kaikki kokoonpanot koostuvat syklittömistä tai yksisyklisistä ympäristöistä, minkä jälkeen todetaan, että melkein kaikilla kokoonpanoilla on ääretön määrä ympäristöjä jokaisessa syklittömässä $\sim_k$-ekvivalenssiluokassa, ja että jokaiselle yksisykliselle ympäristön $\sim_k$-ekvivalenssiluokalle sekä luonnolliselle luvulle j, todennäköisyys, että kokoonpanolla on täsmälleen j ympäristöä kyseisessä luokassa, suppenee. Osoitetaan myös, että solmujen lukumäärän kasvaessa rajatta lähes jokainen kokoonpano on k-rikas sekä k-yksinkertainen.

Näiden tarkastelujen nojalla saadaan yleistettyä suppenemislaki kaikille rajatta kasvaville kokoonpanoille, josta seuraa tutkielman päätulos. Solmujen lukumäärän kasvaessa rajatta todennäköisyys, että ominaisuus pätee verkolle suppenee, mikäli todennäköisyys, että ominaisuus pätee kokoonpanolle, suppenee.

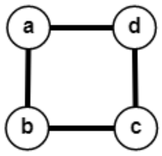
2 Verkkoteorian peruskäsitteitä

Tässä tutkielmassa tullaan tarkastelemaan rakenteita, jotka muodostuvat joukosta alkioita ja joitakin alkiopareja yhdistävästä relaatiosta. Tällaista rakennetta kutsutaan verkoksi. Verkkojen tutkimukseen liittyy runsaasti käsitteitä, jotka esitellään perinpohjaisesti teoksessa [2]. Tässä luvussa esitellään vain tämän tutkielman kannalta oleellisimmat verkkoteorian käsitteet.

Määritelmä 2.1. *Verkko* on pari (V, E), missä E on symmetrinen ja antirefleksiivinen kaksipaikkainen relaatio joukossa V. Sanotaan, että alkio $v \in V$ on *solmu* ja jos alkiot $v, u \in V$ ovat relaatiossa E, niin sanotaan, että solmujen välillä on *särmä*. Lisäksi sanotaan, että (V, E) on *n-verkko*, mikäli $V = \{1, 2, \ldots, n\}$.

Jos verkon (V, E) solmujen v ja u välillä on särmä, merkitään $v\ E\ u$. Tällöin voidaan myös sanoa, että *solmusta v lähtee särmä solmuun u*. Sanotaan, että solmusta $v \in V$ lähtevien särmien lukumäärä on solmun v *aste*, ja siitä käytetään merkintää $\deg(v)$.

Esimerkki 2.2. Olkoon $V = \{a, b, c, d\}$ ja määritellään relaatio E siten, että verkon kaikki särmät ovat $a\ E\ b$, $b\ E\ c$, $c\ E\ d$ ja $d\ E\ a$. Nyt (V, E) on verkko, jonka kuva on nelikulmio.

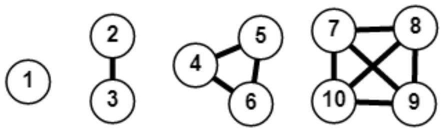


Määritelmä 2.3. Verkon (V, E) *astejono* on jono $(d_0, d_1, \ldots)$, missä jokaiselle $i = 0, 1, \ldots$ pätee $d_i = |\{v \in V \mid \deg(v) = i\}|$. Toisin sanoen d_i on niiden solmujen lukumäärä, joista lähtee täsmälleen i särmää. Sanotaan, että mikä tahansa lukujono $(d_0, d_1, \ldots)$ on *kelpaava*, mikäli se on jonkin n-verkon astejono.

Huomautus. Äärellisen n-verkon astejonolle $D = (d_0, d_1, \ldots)$ pätee $d_i = 0$, kun $i \geq n$. Tästä seuraa, että astejonon merkintä voidaan aina katkaista n:nen jäsenen jälkeen, eli voidaan merkitä $D = (d_0, d_1, \ldots, d_{n-1})$. Lisäksi huomataan, että

$$\sum_{i=0}^{\infty} d_i = n.$$

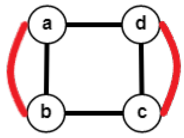
Esimerkki 2.4. Jono $(1, 2, 3, \ldots, i)$ on kelpaava, sillä kaikilla $i \in \mathbb{N}$ voidaan muodostaa sellainen i-verkko, että sen kaikki solmut ovat astetta $i - 1$. Esimerkiksi kuvan 10-verkolla on astejono $(1, 2, 3, 4)$.



Määritelmä 2.5. Verkko (V, E) on *paritus*, mikäli kaikkien solmujen $v \in V$ aste on 1. Huomataan, että paritukselle pätee $|V| = 2|E|$.

Määritelmä 2.6. Olkoon (V,E) verkko ja $V' \subseteq V$ sekä $E' \subseteq E$. Sanotaan, että pari (V',E') on verkon (V,E) *aliverkko*, mikäli sen solmujoukko V' sisältää kaikki solmut, jotka esiintyvät särmäjoukossa E'. Tällöin sanotaan myös, että verkko (V,E) *sisältää* verkon (V',E').

Esimerkki 2.7. Tarkastellaan esimerkin 2.2 verkkoa (V,E). Tämä verkko ei ole paritus, sillä sen jokaisen solmun aste on 2. Määritellään solmujoukolle $V = \{a,b,c,d\}$ relaatio E' siten, että verkolla (V,E') on vain särmät $a\ E'\ b$ ja $c\ E'\ d$. Nyt verkko (V,E') on paritus. Verkko (V,E') on myös verkon (V,E) aliverkko.



Kuvassa esiintyy nyt kaksi verkkoa, joilla on sama solmujoukko. Verkon (V,E) solmujen välistä relaatiota merkitään mustalla särmällä ja parituksen (V,E') solmujen välistä relaatiota punaisella särmällä.

Määritelmä 2.8. Olkoon (V,E) verkko ja $x_i \in V$, $i = 1,2,\ldots$ sen solmuja. Sanotaan, että jono $(x_1,x_2,\ldots,x_n)$ on *kulku solmusta* x_1 *solmuun* x_n, mikäli kaikille $1 \leq i < n$ pätee $x_i\ E\ x_{i+1}$.

Määritelmä 2.9. Sanotaan, että verkko (V,E) on *yhtenäinen*, mikäli kaikille $x,y \in V$ on olemassa kulku solmusta x solmuun y.

Määritelmä 2.10. Olkoon $G = (V,E)$ verkko ja $G' = (V',E')$ sen aliverkko. Sanotaan, että G' on verkon G *komponentti*, mikäli se on yhtenäinen ja se ei sisälly mihinkään yhtenäiseen verkon G aliverkkoon, jossa on enemmän särmiä kuin verkossa G'.

3 Logiikkaa

Tässä tutkielmassa tarkastellaan verkkoihin liittyviä väitteitä ensimmäisen kertaluvun logiikan näkökulmasta. Sen avulla voidaan kuvailla joitakin, mutta ei läheskään kaikkia verkon ominaisuuksia. Tarkasteltavissa kaavoissa voi siis esiintyä atomikaavojen lisäksi muuttuja- sekä relaatiosymboleita, universaali- ja eksistenssikvanttoreita sekä tavanomaisia loogisia konnektiiveja. Muodollinen lähestymistapa logiikkaan pohjautuvassa tarkastelussa ohitetaan ja tyydytään käyttämään esimerkiksi myöhemmin esiteltävän Ehrenfeuchtin pelin kaltaisia menetelmiä tunnettuina tuloksina. Määritelmät mukailevat teoksen [12] esitystapaa.

Esimerkki 3.1. Oletetaan, että (V, E) on verkko ja $v, u, w \in V$. Ominaisuus ”On olemassa kolmio” on esitettävissä ensimmäisen kertaluvun logiikan lauseena

$$\varphi \equiv \exists v \exists u \exists w((v\ E\ u) \wedge (v\ E\ w) \wedge (u\ E\ w)).$$

3.1 Lauseen syvyys

Jotta verkkojen ominaisuuksia voidaan vertailla, määritellään ensimmäisen kertaluvun logiikan kaavoille kvanttorisyvyys.

Määritelmä 3.2. Olkoot φ, ψ ja θ ensimmäisen kertaluvun logiikan kaavoja. Lauseen φ *kvanttorisyvyys* tai lyhyemmin vain *syvyys* $d(\varphi)$ on

- 0, jos φ on atomikaava,
- $d(\psi)$, jos $\varphi \equiv \neg\psi$,
- $\max(d(\psi), d(\theta))$, jos $\varphi \equiv \psi \vee \theta$ tai $\varphi \equiv \psi \wedge \theta$,
- $d(\psi) + 1$, jos $\varphi \equiv \exists x(\psi)$ tai $\varphi \equiv \forall x(\psi)$.

Esimerkki 3.3. Olkoon (V, E) verkko. Tällöin kaava ”Solmusta u lähtee särmä ainakin yhteen solmuun, mutta ei kaikkiin verkon solmuihin” voidaan esittää muodossa

$$\theta \equiv (\exists v(u\ E\ v)) \wedge (\exists w(\neg(u\ E\ w))),$$

ja sen syvyys on $d(\theta) = 1$.

Määritelmä 3.4. Sanotaan, että verkot G ja H ovat *k-ekvivalentit*, mikäli kaikille k-syvyisille ensimmäisen kertaluvun logiikan lauseille φ pätee $G \models \varphi$, jos ja vain jos $H \models \varphi$. Tällöin merkitään $G \equiv_k H$.

3.2 Ehrenfeuchtin peli

Ehrenfeuchtin peli on hyödyllinen menetelmä, kun halutaan näyttää kahden rakenteen olevan k-ekvivalentit. Se on täydellisen informaation peli, jota pelataan k kierrosta kahdella samankaltaisella rakenteella $\mathcal{U}^1$ ja $\mathcal{U}^2$. Tässä tutkielmassa tarkasteltavat rakenteet ovat joko verkkoja tai osittaisia kokoonpanoja, joista jälkimmäinen määritellään myöhemmin.

Jokaisella kierroksella j valitaan molemmista rakenteista $\mathcal{U}^i$, $i = 1, 2$, yksi solmu ja merkitään se symbolilla a^i_j. Ensin pelaaja A valitsee luvun i sekä haluamansa solmun lukua i vastaavasta rakenteesta, minkä jälkeen valitulle solmulle annetaan merkintä a^i_j. Pelaajan B tehtävänä on näyttää toisesta rakenteesta sellainen solmu a_j^{3-i}, joka vastaa pelaajan A valitsemaa solmua. Kun k kierrosta on pelattu, on saatu muodostettua solmujoukot $\{a_1^1, a_2^1, \ldots, a_k^1\}$ ja $\{a_1^2, a_2^2, \ldots, a_k^2\}$. Tällöin sanotaan, että pelaaja B voittaa pelin, jos ja vain jos muodostettujen joukkojen välillä on isomorfismi, joka kuvaa alkion a_j^1 alkiolle a_j^2. Toisin sanoen pelaaja B voittaa verkoilla $G = (V, E)$ ja $G' = (V', E')$ pelatun k-kierroksisen Ehrenfeuchtin pelin, jos ja vain jos kaikille $h, j \in \{1, 2, \ldots, k\}$ on voimassa

$$a_h^1 = a_j^1, \quad \text{jos ja vain jos} \quad a_h^2 = a_j^2 \tag{1}$$

$$a_h^1 \, E \, a_j^1, \quad \text{jos ja vain jos} \quad a_h^2 \, E' \, a_j^2. \tag{2}$$

Määritelmä 3.5. Oletetaan, että $k > j$ ja k-kierroksista Ehrenfeuchtin peliä on pelattu j kierrosta ja lisäksi pelaaja A on tehnyt siirron a_{j+1}^i. Käytetään tehtyjä siirtoja vastaavista solmuista muodostetusta jonosta merkintää $\pi = (a_1^i, a_1^{3-i}, a_2^i, a_2^{3-i}, \ldots, a_{j+1}^i)$. Tällöin sanotaan, että pelaajan B *strategia* on kuvaus σ, joka liittää jokaiseen jonoon π siirron a_{j+1}^{3-i}.

Määritelmä 3.6. Sanotaan, että pelaajan B *voittostrategia* k-kierroksisessa Ehrenfeuchtin pelissä on sellainen strategia σ, joka johtaa varmasti pelin voittoon.

Toisin sanoen pelaajalla B on voittostrategia, mikäli pelaaja B voi vastata mihin tahansa pelaajan A siirtoon siten, että se johtaa pelin voittoon.

Verkkojen vertailemista varten tarvitaan vielä tärkeä Ehrenfeuchtin [3] todistama lause:

Lause 3.7. *Olkoot G ja G' verkkoja. Tällöin $G \equiv_k G'$, jos ja vain jos pelaajalla B on voittostrategia k-kierroksisessa Ehrenfeuchtin pelissä verkoilla G ja G'.* □

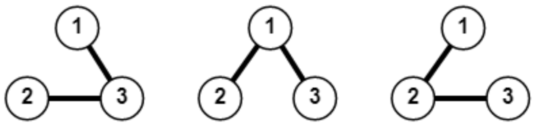
Luvussa 5 kuvaillaan Ehrenfeuchtin peli toisenlaisille rakenteille, nimittäin kokoonpanoille. Näiden rakenteiden välisen Ehrenfeuchtin pelin tarkastelu mahdollistaa tutkielman päätuloksen todistamisen.

4 Satunnaisverkko

Määritelmä 4.1. *Satunnaisverkko* on satunnaismuuttuja, joka saa arvoikseen verkkoja. Lisäksi sanotaan, että jos $n \in \mathbb{N}$ ja $\mathbb{G}_n$ on n-verkkojen joukko, niin *satunnainen n-verkko* on verkko, joka on muodostettu valitsemalla umpimähkään jokin verkko joukosta $\mathbb{G}_n$.

Satunnaisverkkoja on useita erilaisia. Yksi tapa muodostaa satunnaisverkko on kiinnittää *särmätodennäköisyys p*, jolla jokaisen solmuparin välille muodostetaan toisistaan riippumattomasti särmä. Tässä tutkielmassa kuitenkin tarkastellaan satunnaisverkkoja samasta näkökulmasta kuin artikkelissa [8], eli oletetaan verkon astejonon olevan kiinnitetty. Toisin sanoen jos $\mathbb{G}_D$ on astejonoa D noudattavien verkkojen joukko, niin todennäköisyys, että satunnaisverkko saa arvokseen verkon G on sama kaikille $G \in \mathbb{G}_D$.

Esimerkki 4.2. Tarkastellaan satunnaista 3-verkkoa, joka noudattaa astejonoa $(0, 2, 1)$. Tällaisia verkkoja on täsmälleen 3 kappaletta, joten todennäköisyys, että tällaista verkkoa muodostaessa tullaan valinneeksi juuri tietty kuvassa alla näkyvistä verkoista, on $\frac{1}{3}$.



Kun verkon astejono kiinnitetään, tullaan samalla kiinnittäneeksi verkon solmujen lukumäärä. Tutkielman tarkoituksena on kuitenkin tutkia verkon käyttäytymistä solmujen lukumäärän kasvaessa rajatta, joten seuraava määritelmä on tarpeellinen:

Määritelmä 4.3. Olkoot $D_0, D_1, \ldots$ sellaisia lukujonoja, että jokaisella $n \in \mathbb{N}$, jono $D_n = (d_0(n), d_1(n), \ldots, d_{n-1}(n))$ on kelpaava ja $\sum_{i=0}^{n-1} d_i(n) = n$. Tällöin sanotaan, että perhe $(D_n)_{n\in\mathbb{N}}$ on *asymptoottinen astejono*. Jos n-verkko noudattaa astejonoa D_n jollakin $n \in \mathbb{N}$, niin sanotaan, että *verkko noudattaa asymptoottista astejonoa $(D_n)_{n\in\mathbb{N}}$*.

Tarkastellaan verkon ominaisuutta P ja satunnaista n-verkkoa, joka noudattaa asymptoottista astejonoa $(D_n)_{n\in\mathbb{N}}$. Todennäköisyyttä, että verkolla on ominaisuus P, merkitään $\mathbb{P}(P, n, D_n)$. Jos P on lausuttavissa ensimmäisen kertaluvun logiikan lauseena φ ja jos astejono on kontekstista selvä, merkitään vain $\mathbb{P}(\varphi, n)$.

4.1 Asymptoottisten astejonojen ominaisuuksia

Jotta asymptoottisten todennäköisyyksien tutkiminen suurissa satunnaisverkoissa olisi mielekästä, on yhdenmukaisuuden nimissä tarpeen kiinnittää asymptoottisille astejonoille joitakin ehtoja. Näitä ehtoja pidetään standardioletuksina kyseistä aihetta tutkittaessa. Erityisesti artikkelissa [11] käytetään seuraavia määritelmiä.

Olkoon tässä aliluvussa $(D_n)_{n\in\mathbb{N}} = ((d_0(n), d_1(n), \ldots, d_{n-1}(n)) \mid n \in \mathbb{N})$ satunnaisen n-verkon asymptoottinen astejono.

Määritelmä 4.4. Jos jokaiselle $i \in \mathbb{N}$ on olemassa sellainen $\lambda_i \in [0, 1]$, että

$$\lim_{n\to\infty} \frac{d_i(n)}{n} = \lambda_i,$$

niin sanotaan, että asymptoottinen astejono $(D_n)_{n\in\mathbb{N}}$ on *tasainen*. Jos $(D_n)_{n\in\mathbb{N}}$ on tasainen ja $\lambda_i = 0$, $i \in \mathbb{N}$, niin oletetaan myös, että $d_i(n) = 0$ kaikilla $n \in \mathbb{N}$.

Määritelmä 4.5. Olkoon $(D_n)_{n\in\mathbb{N}}$ tasainen ja oletetaan, että verkon i-asteisten solmujen suhteellisille lukumäärille λ_i pätee $\sum_{i=0}^{\infty} \lambda_i = 1$. Sanotaan, että $(D_n)_{n\in\mathbb{N}}$ on *harva*, mikäli

$$\sum_{i=0}^{\infty} i\lambda_i = \Lambda$$

on äärellinen.

Määritelmä 4.6. Oletetaan, että asymptoottinen astejono $(D_n)_{n\in\mathbb{N}}$ on harva ja Λ kuten edellisessä määritelmässä. Sanotaan, että $(D_n)_{n\in\mathbb{N}}$ on *hyvin käyttäytyvä*, mikäli pätee

$$\lim_{n\to\infty} \sum_{i=1}^{n-1} \frac{i d_i(n)}{n} = \Lambda.$$

Edellisten määritelmien lisäksi tehdään rajoitus, joka mahdollistaa luvussa 6 esiteltävät apulauseet, joita tarvitaan tutkielman päälauseen todistamiseen.

(4.1) On olemassa sellainen vakio $\alpha < \frac{1}{4}$, että kaikille n ja $i > n^{\alpha}$ pätee $d_i(n) = 0$.

Viimeisenä tehdään oletus, että $\lambda_0 = 0$. Toisin sanoen oletetaan, että tarkasteltavissa verkoissa ei ole merkittävää määrää eristettyjä solmuja. Tämä oletus yksinkertaistaa verkkojen ominaisuuksien tutkimista. Seuraava apulause osoittaa, että tutkielman päätuloksen yleisyys ei kärsi tämän oletuksen myötä.

Apulause 4.7. Olkoon $(D_n)_{n\in\mathbb{N}} = ((d_0(n), d_1(n), \ldots, d_{n-1}(n)) \mid n \in \mathbb{N})$ tasainen asymptoottinen astejono, jolle $\lambda_0 < 1$. Merkitään $n' = n - d_0(n)$, kun $n \in \mathbb{N}$, ja lisäksi $I = \{n' \mid n \in \mathbb{N}\}$. Olkoon J mielivaltainen luonnollisten lukujen joukko, joka sisältää jokaista $s \in I$ kohti täsmälleen yhden luvun $n \in J$, jolle $n' = s$.

Tällöin on olemassa sellainen tasainen asymptoottinen astejono $(D'_{n'})_{n'\in I} = ((d'_0(n'), d'_1(n'), \ldots, d'_{n'-1}(n')) \mid n' \in I)$, että jokaisella $n' \in I$ pätee $d'_0(n') = 0$ ja että jokaista ensimmäisen kertaluvun logiikan lausetta φ vastaa ensimmäisen kertaluvun logiikan lause φ', jolle

$$\mathbb{P}(\varphi, n, D_n) = \mathbb{P}(\varphi', n', D'_{n'}),$$

kun $n \in J$ on tarpeeksi suuri.

Todistus. Koska $d_0(n) = n - n'$, huomataan, että joukon I on oltava ääretön, sillä muutoin asymptoottisen astejonon tasaisuuden nojalla pätee

$$\lambda_0 = \lim_{n\to\infty} \frac{d_0(n)}{n} = \lim_{n\to\infty} \frac{n - n'}{n} = 1,$$

mikä on ristiriidassa oletuksen $\lambda_0 < 1$ kanssa. Määritellään nyt tasainen asymptoottinen astejono $(D'_{n'})_{n'\in I} = ((d'_0(n'), d'_1(n'), \ldots, d'_{n'-1}(n')) \mid n' \in I)$, missä

$$d'_0(n') = 0,$$

sekä

$$d'_i(n') = d_i(n), \text{ kun } i > 0 \text{ ja } n \in J.$$

Jos $\lambda_0 = 0$, niin tasaisuuden määritelmän mukaan $d_0(n) = 0$, kun n on tarpeeksi suuri. Tästä seuraa, että $n' = n$ ja edelleen $d_i'(n) = d_i(n)$, kun n on tarpeeksi suuri, joten väite on selvä.

Oletetaan sitten, että $0 < \lambda_0 < 1$. Tarkastellaan ensimmäisen kertaluvun logiikan lausetta φ ja merkitään sen syvyyttä $d(\varphi) = k$. Nyt koska $\lambda_0 > 0$, niin asymptoottisen astejonon tasaisuuden määritelmän nojalla on olemassa sellainen $n_0 \in \mathbb{N}$, että jos $n > n_0$, niin $d_0(n) \geq k$.

Olkoon $G = (V, E)$ mielivaltainen n-verkko ja $G' = (V', E \upharpoonright V')$ verkko, missä $V' = \{u \in V \mid deg(u) \neq 0\}$. Tällöin kuvaus verkolta G verkolle G' säilyttää todennäköisyydet seuraavassa mielessä: Todennäköisyys, että G on tiettyä isomorfiatyyppiä asymptoottista astejonoa $(D_n)_{n\in\mathbb{N}}$ noudattavien n-verkkojen luokassa, on sama kuin todennäköisyys, että G' on vastaavaa isomorfiatyyppiä asymptoottista astejonoa $(D'_{n'})_{n'\in I}$ noudattavien n-verkkojen luokassa.

Olkoot G sellainen n-verkko, ja H sellainen m-verkko, jotka noudattavat asymptoottista astejonoa $(D_n)_{n\in\mathbb{N}}$. Oletetaan, että $m, n > n_0$. Tällöin verkkojen G ja H eristettyjen solmujen lukumäärä on $d_0(n) \geq k$.

Osoitetaan, että $G \equiv_k H$, jos ja vain jos $G' \equiv_k H'$. Tämä tehdään havaitsemalla, että tarkasteltaessa verkoilla G ja H sekä verkoilla G' ja H' pelattavia k-kierroksisia Ehrenfeuchtin pelejä, pelaajan B voittostrategia siirtyy pelistä toiseen. Jos pelaajalla B on voittostrategia σ verkoilla G' ja H' pelatussa pelissä, niin verkoilla G ja H pelatussa pelissä pelaajan B voittostrategia on seuraava:

(1) Jos pelaaja A merkitsee eristetyn solmun jommassakummassa verkossa, niin pelaaja B voi aina merkitä toisesta verkosta eristetyn solmun, sillä molemmissa verkoissa on vähintään k eristettyä solmua.

(2) Jos pelaaja A merkitsee solmun, joka ei ole eristetty, niin pelaaja B voi tehdä siirron strategian σ mukaisesti. Toisin sanoen jos tähän mennessä verkosta G on valittu s solmua, jotka eivät ole eristettyjä, niin pelaaja B voi tehdä saman siirron, kuin olisi tehnyt verkoilla G' ja H' pelatussa pelissä kierroksella $s + 1$.

Toisaalta jos pelaajalla B on voittostrategia σ verkoilla G ja H pelatussa pelissä, niin verkoilla G' ja H' pelatussa pelissä pelaaja B voi aina tehdä strategian σ mukaisen siirron. Toisin sanoen jos verkoilla G ja H pelatussa pelissä pelaaja A merkitsee kierroksella j solmun, joka ei ole eristetty ja jos tähän mennessä verkosta G on valittu s solmua, jotka ovat eristettyjä, niin pelaaja B voi tehdä verkoilla G' ja H' pelattavassa pelissä kierroksella $j - s$ saman siirron, kuin toisessa pelissä kierroksella j.

Pelaajalla B on siis voittostrategia verkoilla G ja H pelatussa Ehrenfeuchtin pelissä, jos ja vain jos pelaajalla B on voittostrategia verkoilla G' ja H' pelatussa Ehrenfeuchtin pelissä, joten lauseen 3.7 nojalla $G \equiv_k H$, jos ja vain jos $G' \equiv_k H'$.

Viimeiseksi muodostetaan apulauseen väitteen mukainen ensimmäisen kertaluvun logiikan lause φ'. Olkoon K' kaikkien sellaisten verkkojen luokka, jotka ovat isomorfisia jonkin verkon G' kanssa jollakin G, joka toteuttaa lauseen φ. Tällöin luokka K' on suljettu ekvivalenssin $\equiv_k$ suhteen asymptoottista astejonoa $(D'_{n'})_{n'\in I}$ noudattavien verkkojen luokassa, kunhan tarkasteltavat verkot ovat tarpeeksi suuria. Tästä seuraa, että luokka K' pystytään karakterisoimaan ensimmäisen kertaluvun logiikan lauseella φ', mikä tarkoittaa, että verkko kuuluu luokkaan K', jos ja vain jos se toteuttaa lauseen φ'.

Näiden tarkastelujen nojalla väite on todistettu. □

Tämän tutkielman tavoitteena on näyttää, että jokaiselle hyvin käyttäytyvää asymptoottista astejonoa $(D_n)_{n\in\mathbb{N}}$ noudattavalle n-verkolle sekä ensimmäisen kertaluvun logiikan lauseelle φ on olemassa raja-arvo $\lim_{n\to\infty} \mathbb{P}(\varphi, n, D_n)$. Tulos voidaan kuitenkin helposti laajentaa koskemaan mitä tahansa ääretöntä luonnollisten lukujen osajonoa, mikä tarkoittaa, että edellisen apulauseen todistuksen merkintöjä käyttäen on myös olemassa raja-arvo

$$\lim_{\substack{n' \to \infty \\ n' \in I}} \mathbb{P}(\varphi', n', D'_{n'}).$$

5 Rakenteiden vertailemisesta

Monet tunnetuista verkoista, kuten Internet ja jotkin sosiaaliset verkostot, ovat sellaisia, että niiden solmujen välisten särmien lukumäärä noudattaa potenssilakeja. Tällaisista verkoista sanotaan, että ne ovat *skaalautumattomia*. Skaalautumattomassa verkossa solmut ja särmät ovat toisistaan riippumattomia, joten verkon asymptoottista rakennetta ja käyttäytymistä on mahdotonta päätellä. Näitä verkkoja on myös haastavaa muodostaa satunnaisesti. Reaalimaailmassa esiintyvien verkkojen skaalautumattomuuteen liittyvä tutkimus on vielä kehitysvaiheessa, ja siitä on tarkempia tietoja teoksessa [1]. Tässä tutkielmassa määritellään helpommin satunnaisesti muodostettavissa oleva rakenne, jota kutsutaan kokoonpanoksi. Tullaan huomaamaan, että satunnaiskokoonpanoihin liittyvät tulokset voidaan kääntää kiinnitettyä astejonoa noudattavia satunnaisverkkoja kuvaaviksi.

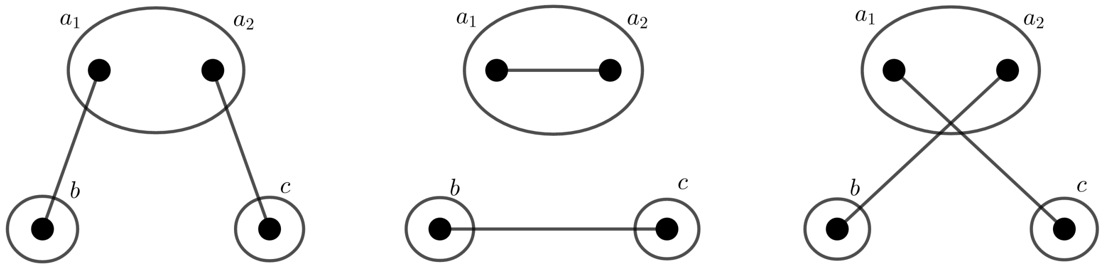
5.1 Kokoonpano

Määritelmä 5.1. *Kokoonpano* on rakenne $(U, \equiv, M)$, missä (U, M) on paritus ja $\equiv$ on ekvivalenssirelaatio joukossa U. Mikäli ekvivalenssiluokkien lukumäärä on n, sanotaan, että $(U, \equiv, M)$ on *n-kokoonpano*.

Lisäksi sanotaan, että $(U, \equiv, M)$ on *verkkoa* (V, E) *vastaava kokoonpano*, mikäli paritus (U, M) on määritelty siten, että $x_1 \; E \; x_2$, jos ja vain jos on olemassa sellaiset $u \in [x_1]$ ja $v \in [x_2]$, että $u \; M \; v$.

Voidaan ajatella, että verkon (V, E) jokainen solmu on kokoonpanon $(U, \equiv, M)$ solmujoukon ekvivalenssiluokan edustaja ja solmujoukko U muodostetaan ottamalla $\deg(v)$ kappaletta erillisiä kopioita jokaisesta solmusta $v \in V$. Tästä johtuen kokoonpanolla on yhteensä $m = \sum_{i=1}^{n-1} i d_i(n)$ solmua.

Esimerkki 5.2. Olkoon (V, E) verkko, jonka solmujoukko on $V = \{a, b, c\}$ ja jonka särmät ovat $a \; E \; b$ sekä $a \; E \; c$. Verkolla on siis astejono $(d_0, d_1, d_2) = (0, 2, 1)$. Muodostetaan 3-kokoonpanon $(U, \equiv, M)$ solmujoukko määrittelemällä $U = \{a_1, a_2, b, c\}$. Lisäksi määritellään ekvivalenssirelaatio $\equiv$ siten, että $[a_1] = \{a_1, a_2\}$, $[b] = \{b\}$ ja $[c] = \{c\}$. Nyt kokoonpano voidaan muodostaa kolmella eri tavalla:



Kuvan kokoonpanoista reunimmaiset ovat verkkoa (V, E) vastaavia kokoonpanoja, mutta keskimmäinen ei ole.

Määritelmä 5.3. Kokoonpanon $(U, \equiv, M)$ (joukon $W \subseteq U$ indusoima) *alikokoonpano* on $(W, \equiv\restriction W, M \restriction W)$, missä $\equiv\restriction W$ on ekvivalenssirelaation $\equiv$ rajoittuma solmujoukkoon W ja relaatio $M \restriction W$ on määritelty kuten relaatio M, mutta siitä on poistettu sellaiset särmät $u\ M\ v$, joille $u \notin W$ tai $v \notin W$. Alikokoonpanoa $(W, \equiv\restriction W, M \restriction W)$ merkitään lyhyemmin $(W, \equiv, M)$.

Huomautus. Alikokoonpano ei välttämättä ole kokoonpano, sillä siihen saattaa kuulua parittamattomia solmuja.

Määritelmä 5.4. Sanotaan, että rakenne $(U, \equiv, M)$ on *osittainen kokoonpano*, mikäli se on jonkin kokoonpanon alikokoonpano.

Huomautus. Kokoonpano $(U, \equiv, M)$ on itsensä alikokoonpano, eli kaikki kokoonpanot ovat osittaisia kokoonpanoja. Osittaisessa kokoonpanossa relaation M ei kuitenkaan tarvitse olla paritus, eli osittaisessa kokoonpanossa saa esiintyä parittamattomia solmuja.

Esimerkki 5.5. Tarkastellaan esimerkin 5.2 keskimmäistä kokoonpanoa, jossa relaatio M on määritelty siten, että kokoonpanon solmuja yhdistää särmät $b\ M\ c$ ja $a_1\ M\ a_2$. Nyt joukon $[a_1] \subset U$ indusoima alikokoonpano $([a_1], \equiv, M)$ koostuu yksinkertaisesti solmuista a_1 ja a_2 sekä niitä yhdistävästä särmästä $a_1\ M\ a_2$.

Olkoot $\mathcal{U}^1 = (U^1, \equiv^1, M^1)$ ja $\mathcal{U}^2 = (U^2, \equiv^2, M^2)$ osittaisia kokoonpanoja ja a_n^i, $i = 1, 2$, $n \in \{1, 2, \ldots, k\}$, kuten kappaleessa 3.2. Nyt pelaaja B voittaa osittaisilla kokoonpanoilla $\mathcal{U}^1$ ja $\mathcal{U}^2$ pelatun k-kierroksisen Ehrenfeuchtin pelin, jos ja vain jos kaikille $h, j \in \{1, 2, \ldots, k\}$ on voimassa

$$a_h^1 = a_j^1, \quad \text{jos ja vain jos} \quad a_h^2 = a_j^2 \tag{1}$$

$$a_h^1 \equiv^1 a_j^1, \quad \text{jos ja vain jos} \quad a_h^2 \equiv^2 a_j^2 \tag{2}$$

$$a_h^1\ M^1\ a_j^1, \quad \text{jos ja vain jos} \quad a_h^2\ M^2\ a_j^2. \tag{3}$$

Samaan tapaan kuin luvussa 3 tehtiin verkoille, esitellään Ehrenfeuchtin artikkelissaan [3] todistama tulos osittaisille kokoonpanoille:

Lause 5.6. *Olkoot $\mathcal{U}^1$ ja $\mathcal{U}^2$ osittaisia kokoonpanoja. $\mathcal{U}^1 \equiv_k \mathcal{U}^2$, jos ja vain jos pelaajalla B on voittostrategia k-kierroksisessa Ehrenfeuchtin pelissä rakenteilla $\mathcal{U}^1$ ja $\mathcal{U}^2$.* □

5.2 Tekijäverkko

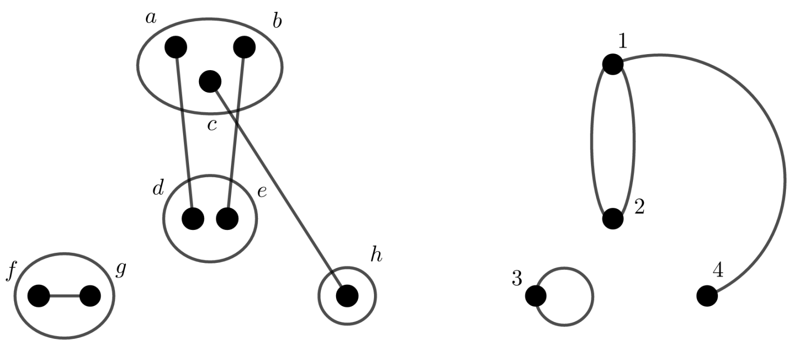
Määritelmä 5.7. *Moniverkko* on pari (V, M), missä V on solmujoukko kuten määritelmässä 2.1 ja M on monijoukko, jonka alkiot ovat joukkoja $\{v, u\}$, missä $v, u \in V$. Huomautetaan, että sallitaan $v = u$. Näitä pareja tai yksiöitä merkitään $v\ M\ u$ ja niitä sanotaan moniverkon särmiksi.

Huomautus. Moniverkot poikkeavat verkoista vain siten, että solmun sallitaan yhdistyä itseensä särmällä ja saman solmuparin välille voidaan muodostaa useita särmiä. Moniverkon särmille ja solmuille käytetään myös samoja nimityksiä kuin verkkojen tapauksessa, mistä syystä monijoukkoa M kutsutaan usein moniverkon *relaatioksi*.

Määritelmä 5.8. Olkoon $\mathcal{U} = (U, \equiv, M)$ kokoonpano ja $U/\equiv$ sen ekvivalenssiluokkien joukko. Käytetään alkion $x \in U$ ekvivalenssiluokasta merkintää $[x]$. Kokoonpanon $(U, \equiv, M)$ *tekijäverkko* on moniverkko $(U/\equiv, M/\equiv)$, missä relaatio $M/\equiv$ on määritelty siten, että ekvivalenssiluokkien $[x]$ ja $[y]$ välillä on yksikäsitteinen särmä jokaista sellaista paria $\{x', y'\}$ kohti, joille pätee $x' \equiv x$ ja $y' \equiv y$ sekä $x'\ M\ y'$.

Esimerkki 5.9. Tarkastellaan jälleen 3-kokoonpanoa $(U, \equiv, M)$, missä $U = \{a_1, a_2, b, c\}$, $[a_1] = \{a_1, a_2\}$, $[b] = \{b\}$ ja $[c] = \{c\}$, sekä relaatio M on määritelty siten, että $b\ M\ c$ ja $a_1\ M\ a_2$. Tämän kokoonpanon tekijäverkko on $(U/\equiv, M/\equiv)$, missä $U/\equiv\ = \{[a_1], [b], [c]\}$ ja $M/\equiv\ = \{\{[a_1], [a_1]\}, \{[b], [c]\}\}$. Solmu $[a_1]$ yhdistyy siis itseensä särmällä, eli tekijäverkot eivät välttämättä ole verkkoja.

Esimerkki 5.10. Kuvassa on kokoonpano ja sen tekijäverkko, kun ympyröidyt joukot kuvastavat solmujoukon $U = \{a, b, c, d, e, f, g, h\}$ ekvivalenssiluokkia. Huomataan, että tämän kokoonpanon tekijäverkko ei ole verkko.



Määritelmä 5.11. Sanotaan, että verkon ominaisuus P pätee kokoonpanolle, mikäli se pätee kokoonpanon tekijäverkolle.

Tarkastellaan nyt n-verkkoja G ja H, joilla on sama astejono $(d_0, d_1, \ldots, d_{n-1})$. Verkot G ja H ovat sellaisten kokoonpanojen tekijäverkkoja, joilla on n ekvivalenssiluokkaa, ja kokoa i olevien ekvivalenssiluokkien lukumäärä on d_i. Tällaisten kokoonpanojen lukumäärä riippuu vain jonosta $(d_0, d_1, \ldots, d_{n-1})$, joten verkot G ja H ovat yhtä monen kokoonpanon tekijäverkkoja. Tämä tulos pätee kaikille samaa astejonoa noudattaville verkoille.

Luvussa 4 annettiin määritelmä satunnaiselle verkolle. Satunnainen kokoonpano määritellään vastaavasti. Jos φ on kokoonpanon ominaisuus, merkitään todennäköisyyttä, että satunnaisella n-kokoonpanolla on kyseinen ominaisuus, $\mathbb{P}_c(\varphi, n)$. Olkoon nyt P jokin verkon ominaisuus, ja Q se kokoonpanon ominaisuus, että kokoonpanon tekijäverkko on verkko. Tällöin edellisen pohdinnan nojalla

$$\mathbb{P}(P, n) = \frac{\mathbb{P}_c(P \wedge Q, n)}{\mathbb{P}_c(Q, n)}.$$

Ominaisuus Q on määriteltävissä ensimmäisen kertaluvun logiikan avulla. Lisäksi $\mathbb{P}_c(Q, n)$ suppenee kohti jotain positiivista vakiota, kun $n \to \infty$. Tämän tuloksen todistus pohjautuu artikkelissa [10] tehtyyn tarkasteluun, jossa tutkitaan sellaisten $m \times n$-matriisien lukumääriä, jotka saavat arvoja joukosta $\{0, 1\}$ ja joiden rivien ja sarakkeiden summat ovat kiinnitetty. Artikkelissa kuvaillaan tarkemmin, miten näiden matriisien lukumääriä koskevat tulokset voidaan kääntää kaksijakoisten verkkojen, ja sitä kautta tekijäverkkojen lukumäärää kuvaaviksi tuloksiksi. Näiden tulosten vuoksi voidaan verkkojen tutkimisessa käyttää hyväksi tietoja kokoonpanojen käyttäytymisestä. Erityisesti satunnaisverkon suppenemislaki seuraa suoraan kokoonpanojen suppenemislaista. Näin saadaan seuraava lause:

Lause 5.12. *Jos on olemassa raja-arvo* $\lim_{n\to\infty} \mathbb{P}_c(P \wedge Q, n)$*, niin on myös olemassa raja-arvo* $\lim_{n\to\infty} \mathbb{P}(P, n)$. □

5.3 Puu ja ympäristö

Kokoonpanoille voidaan määritellä useita samoja käsitteitä kuin verkoillekin. Seuraavat määritelmät mahdollistavat kokoonpanon sisällä liikkumisen ja solmujen etäisyyksien tarkastelemisen. Samalla tehdään mahdolliseksi kokoonpanon paikallisten ominaisuuksien tutkiminen. Määritelmät on muotoiltu artikkelien [8] ja [4] mukaisesti.

Määritelmä 5.13. Olkoon $(U, \equiv, M)$ kokoonpano ja $x, y \in U$. Sanotaan, että *d-mittainen kulku solmusta x solmuun y* on jono $(x_0, x_1, \ldots, x_d)$, missä $x_0 = x$, $x_d = y$ sekä $x_i \equiv x_{i+1}$ tai $x_i \; M \; x_{i+1}$ kaikille $0 \leq i < d$. Sanotaan lisäksi, että kulku on *suljettu*, mikäli $x = y$.

Merkintä. Lyhimmän kulun solmusta x solmuun y pituutta merkitään $\delta(x, y)$. Lisäksi joukon U osajoukoille X ja Y määritellään $\delta(X, Y) = \min\{\delta(x, y) \in \mathbb{N} \mid x \in X, y \in Y\}$, sekä $\delta(X, x) = \delta(X, \{x\})$.

Määritelmä 5.14. Sanotaan, että kokoonpano $(U, \equiv, M)$ on *yhtenäinen*, mikäli kaikille $x, y \in U$ on olemassa kulku solmusta x solmuun y.

Esimerkki 5.15. Tarkastellaan esimerkin 5.9 kokoonpanoa. Ekvivalenssiluokasta $[a_1]$ ei lähde särmää mihinkään muuhun ekvivalenssiluokkaan, joten ei ole olemassa kulkua ekvivalenssiluokan $[a_1]$ solmusta ekvivalenssiluokan $[b]$ solmuun. Kokoonpano ei siis ole yhtenäinen.

Määritelmä 5.16. Olkoon $(x_0, \ldots, x_d)$ kulku, jolle $x_i \neq x_j$ kun $i \neq j$. Sanotaan, että $(x_0, \ldots, x_d)$ on *polku*, mikäli joukon $\{x_0, \ldots, x_d\}$ indusoimalle alikokoonpanolle pätee $|[x_i]| \leq 2$, kun $0 \leq i \leq d$.

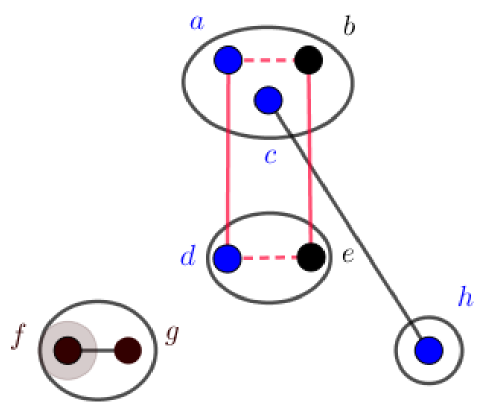
Huomautus. Toisin sanoen polku on sellainen kulku, jossa yhden ekvivalenssiluokan sisällä liikutaan korkeintaan yhden askeleen verran.

Määritelmä 5.17. Oletetaan, että $(x_0, \ldots, x_d)$ on vähintään 4-mittainen suljettu kulku, missä $x_0 \neq x_1 \neq \cdots \neq x_{d-1}$. Sanotaan, että $(x_0, \ldots, x_d)$ on *sykli*, mikäli joukon $\{x_0, \ldots, x_{d-1}\}$ indusoimalle alikokoonpanolle pätee $|[x_i]| = 2$, kun $0 \leq i < d$.

Kokoonpano on *syklitön*, mikäli se ei sisällä yhtäkään sykliä. Jos kokoonpano sisältää täsmälleen yhden syklin, sanotaan, että se on *yksisyklinen*.

Huomautus. Sykli on suljettu polku, jossa yhtä särmää pitkin kuljetaan korkeintaan kerran. Lisäksi tässä tutkielmassa samaistetaan samat solmut sisältävät syklit, jotka eroavat toisistaan vain syklin suuntaa tai alkupistettä vaihtamalla.

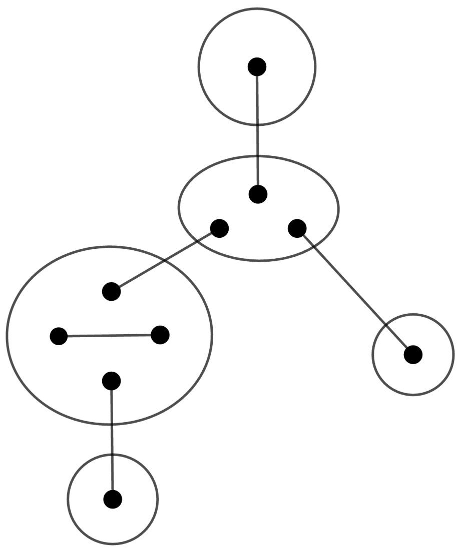
Esimerkki 5.18. Esimerkin 5.10 kokoonpano on yksisyklinen, sillä sen ainoa sykli on (a, b, e, d, a). Kuvassa on merkitty punaisella kokoonpanon sykliä, missä katkoviiva tarkoittaa ekvivalenssiluokan sisällä siirtymistä. Sinisellä on merkitty toista polkua (h, c, a, d). Tästä huomataan, että ekvivalenssiluokkaan $[h]$ päästään polkua pitkin, mutta solmu h ei kuulu mihinkään sykliin tässä kokoonpanossa.



Esimerkki 5.19. Esimerkin 5.9 kokoonpano on syklitön, sillä mistään solmusta ei pääse suljettua polkua pitkin takaisin itseensä kulkematta samaa särmää pitkin kahdesti.

Määritelmä 5.20. Osittainen kokoonpano on *puu*, mikäli se on yhtenäinen ja syklitön.

Esimerkki 5.21. Kuvan kokoonpano on puu. Myös esimerkin 5.2 kokoonpanoista kaksi ovat puita. Ne ovat kaikki syklittömiä, mutta keskimmäinen ei ole yhtenäinen, joten se ei ole puu.

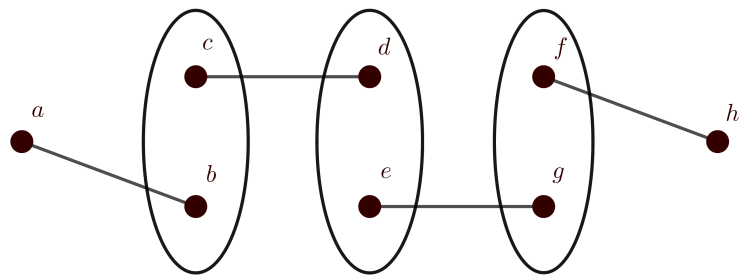


Määritelmä 5.22. Olkoon $R \subseteq U$ ja $r \geq 0$. Sanotaan, että *joukon R r-säteinen (pallon)kuori* on $S(R, r) := \{y \in U \mid \delta(R, y) = r\}$. Solmun v r-säteinen kuori on siis niiden solmujen joukko, jotka ovat täsmälleen r askeleen etäisyydellä solmusta v.

Määritelmä 5.23. *Joukon R r-säteinen ympäristö* on $N(R, r) := \bigcup_{t=0}^{r} S(R, t)$.

r-säteisen ympäristön voidaan siis ajatella muodostuvan r-säteisestä kuoresta sekä kaikista sen sisäpuolelle jäävistä solmuista.

Esimerkki 5.24. Tarkastellaan kuvan kokoonpanoa. Olkoon $R = \{d, e\}$. Nyt joukolla R on 2-säteinen kuori $S(R, 2) = \{b, f\}$ ja 2-säteinen ympäristö $N(R, 2) = \{b, c, d, e, f, g\}$.



5.4 Juurrosmetsä

Tehdään kokoonpanon rakenteeseen lisäys, joka helpottaa lauseen 5.6 mukaisen voittostrategian olemassaolon tutkimista. Laajennetaan siis tarkasteltavia rakenteita yksipaikkaisella relaatiolla R, jota kutsutaan juurijoukoksi:

Määritelmä 5.25. Olkoon $(U, \equiv, M)$ osittainen kokoonpano ja $R \subseteq U$. Sanotaan, että rakenne $(U, \equiv, M, R)$ on osittainen *juurroskokoonpano* ja alkiot $x \in R$ ovat sen *juuria*.

Lisäksi, ympäristön $N(R, r)$ indusoimasta alikokoonpanosta käytetään merkintää $\mathcal{N}(R, r)$. Alikokoonpanon juurijoukko on sama kuin alkuperäisen osittaisen kokoonpanon, ellei toisin mainita.

Määritelmä 5.26. Olkoon $(U, \equiv, M, R)$ juurroskokoonpano. Solmun $u \in U$ *syvyys* on $d(u) = \delta(R, u)$, eli lyhimmän polun pituus juuresta R solmuun u. Sanotaan, että solmu v on solmun u *jälkeläinen*, mikäli $d(v) > d(u)$ ja $d(v) = d(u) + \delta(u, v)$.

Huomataan, että solmun v syvyys $d(v) = k$, jos ja vain jos $v \in S(R, k)$.

Esimerkki 5.27. Tarkastellaan esimerkin 5.24 kokoonpanoa ja laajennetaan se juurroskokoonpanoksi määrittelemällä juurijoukko $R = \{d, e\}$. Nyt esimerkiksi solmu a on solmun b jälkeläinen, ja solmujen d ja e jälkeläisiä ovat kaikki solmut, jotka eivät kuulu juurijoukkoon R.

Määritelmä 5.28. Sanotaan, että juurroskokoonpano $(U, \equiv, M, R)$ on *juurrospuu*, mikäli kokoonpano $(U, \equiv, M)$ on puu ja R on yksiö. Lisäksi sanotaan, että juurrospuun yhtenäinen alikokoonpano on *alipuu*.

Olkoot $\mathcal{U}_1 = (U_1, \equiv_1, M_1)$ ja $\mathcal{U}_2 = (U_2, \equiv_2, M_2)$ osittaisia kokoonpanoja, joiden solmujoukoille pätee $U_1 \cap U_2 = \emptyset$. Tällöin myös rakenteiden särmäjoukot ovat erillisiä. Tällaisia osittaisia kokoonpanoja kutsutaan erillisiksi. Näille rakenteille määritellään yhdiste asettamalla $\mathcal{U}_1 \cup \mathcal{U}_2 = (U_1 \cup U_2, \equiv_1 \cup \equiv_2, M_1 \cup M_2)$. Erillisyys ja yhdiste määritellään juurroskokoonpanoille vastaavasti.

Määritelmä 5.29. Olkoot $\mathcal{T}_1, \ldots, \mathcal{T}_n$ erillisiä juurrospuita. Sanotaan, että $\mathcal{F} = \bigcup_{i=1}^{n} \mathcal{T}_i$ on *juurrosmetsä*. Lisäksi sanotaan, että juurrosmetsän $\mathcal{F}$ syvyys on sen puiden syvimmän solmun syvyys.

Jos $(U, \equiv, M, \{x\})$ on juurrospuu, niin $\mathcal{N}(x, r)$ on juurrospuu. Merkitään niiden solmujen, joiden syvyys on korkeintaan $r \geq 0$, indusoimaa juurrosmetsän $\mathcal{F}$ alikokoonpanoa $\mathcal{F} \upharpoonright r$. Huomataan, että myös $\mathcal{F} \upharpoonright r$ on juurrosmetsä.

5.5 Osittaisten kokoonpanojen vertailemisesta

Seuraavaksi mukaillaan artikkelissa [8] käytettyjä menetelmiä ja määritellään osittaisten kokoonpanojen välinen ekvivalenssirelaatio $\sim_k$, ja osoitetaan, että kaksi relaatiossa $\sim_k$ olevaa rakennetta ovat myös k-ekvivalentit. Tämän jälkeen luvussa 6 näytetään, että todennäköisyys, että satunnainen kokoonpano kuuluu tiettyyn $\sim_k$-ekvivalenssiluokkaan, suppenee. Tarkastellaan kokoonpanon solmua ja sen jonkin kokoista ympäristöä. Tämän ympäristön toteuttamia ominaisuuksia sanotaan kokoonpanon *paikalliseksi ominaisuudeksi*. Relaatio $\sim_k$ määräytyy kokoonpanon paikallisten ominaisuuksien mukaan. Artikkelissa [4] todistetaan, että paikalliset ominaisuudet määrittelevät kaikki kokoonpanon ensimmäisen kertaluvun logiikalla ilmaistavat ominaisuudet. Erityisesti äärellisten ja erillisten ympäristöjen lukumäärä ja käyttäytyminen määrittävät, onko jokin kokoonpanoa koskeva lause totta.

Relaatio $\sim_k$ määritellään ensin ympäristöille, joilla on enintään yksi sykli. Tarkemmin ottaen relaatio $\sim_k$ määritellään ensin juurrospuille ja -metsille (määritelmä 5.31) ja tämän jälkeen määritelmässä 5.36 tehdään sama niin kutsutuille keskitetyille kokoonpanoille. Tällöin voidaan luokitella tällaiset ympäristöt syklittömiksi tai yksisyklisiksi. Hyödynnetään Ehrenfeuchtin peliä ja näytetään, että kaksi tällaista ympäristöä kuuluvat samaan $\sim_k$-ekvivalenssiluokkaan, jos ja vain jos ne ovat k-ekvivalentit. Tämän jälkeen relaatio $\sim_k$ laajennetaan kokoonpanoihin, jotka koostuvat pelkästään syklittömistä tai yksisyklisistä, pienistä ympäristöistä. Tällaisia kokoonpanoja kutsutaan *k-rikkaiksi* ja *-yksinkertaisiksi*. Määritellään, että nämä kokoonpanot ovat relaatiossa $\sim_k$, mikäli niillä on sama määrä tällaisia ympäristöjä jokaista $\sim_k$-ekvivalenssiluokkaa kohden. Yksityiskohdat löytyvät määritelmästä 5.41.

Lopuksi lauseessa 5.45 todistetaan, että jos k-rikkaat ja k-yksinkertaiset kokoonpanot kuuluvat samaan $\sim_k$-ekvivalenssiluokkaan, niin ne ovat k-ekvivalentit. Luvussa 6 todistetaam, että määritelty relaatio $\sim_k$ todella kattaa lähes kaikki tarpeeksi suuret kokoonpanot. Toisin sanoen näytetään, että melkein kaikilla kokoonpanoilla on jokaista $\sim_k$-ekvivalenssiluokkaa kohden ääretön määrä ympäristöjä, ja että jokaista $\sim_k$-ekvivalenssiluokkaa ja luonnollista lukua j kohden todennäköisyys, että kokoonpanolla on täsmälleen j ympäristöä kyseisessä luokassa, suppenee.

Loppututkielman ajan oletetaan, että kokoonpanoista puhuttaessa juurijoukko R on määritelty eli että kaikki tarkasteltavat kokoonpanot ovat juurroskokoonpanoja. Juurijoukko voidaan kuitenkin jättää kokoonpanon merkinnästä pois, mikäli sitä ei tarvita kyseisessä tarkastelussa.

Määritelmä 5.30. Metsää, jossa ei ole yhtään puuta, kutsutaan *nollametsäksi*, ja sen syvyyden määritellään olevan -1.

Oletetaan, että k on kiinnitetty. Seuraavaksi määritellään ekvivalenssirelaatio $\sim_k$ juurrosmetsille induktiivisesti juurrosmetsän syvyyden suhteen. Induktion vaiheessa r määritellään relaatio $\sim_k$ r-syvyisille juurrospuille relaation $\sim_k$ suhteen, joka on määritelty korkeintaan $r-1$-syvyisille juurrosmetsille. Sen jälkeen relaatio $\sim_k$ määritellään r-syvyisille juurrosmetsille relaation $\sim_k$ suhteen, joka on määritelty niille juurrospuille, joiden syvyys on korkeintaan r.

Määritelmä 5.31. Määritellään relaatio $\sim_k$ induktiolla juurrosmetsän syvyyden suhteen. Nollametsiä on vain yksi, joten selvästi kaikki -1-syvyiset juurrosmetsät ovat relaatiossa $\sim_k$. Oletetaan nyt relaation $\sim_k$ olevan määritelty juurrosmetsille, joiden syvyys on korkeintaan $r-1$, missä $r \geq 0$. Oletetaan lisäksi, että $\mathcal{U}^i = (U^i, \equiv^i, M^i, \{x_i\})$, missä $i = 1, 2$, on juurrospuu, jonka syvyys on r. Olkoon $y \in S(x_i, 1)$. Määritellään, että $\mathcal{V}_y^i$ on solmun y ja sen jälkeläisten indusoima alipuu, jonka juuri on y. Olkoon lisäksi $\mathcal{F}^i$ metsä, joka koostuu niistä puista $\mathcal{V}_y^i$, joille pätee $x_i \equiv^i y$. Jos on olemassa sellainen yksikäsitteinen $y \in U^i$, että $x_i \; M^i \; y$, niin määritellään $\mathcal{T}^i = \mathcal{V}_y^i$. Muulloin $\mathcal{T}^i$ on nollapuu. Nyt $\mathcal{U}^1 \sim_k \mathcal{U}^2$, mikäli

$$\mathcal{T}^1 \sim_k \mathcal{T}^2 \quad \text{ja} \quad \mathcal{F}^1 \sim_k \mathcal{F}^2.$$

Oletetaan nyt, että $\mathcal{M}^1$ ja $\mathcal{M}^2$ ovat korkeintaan r-syvyisiä juurrosmetsiä. Oletetaan, että jokaista korkeintaan r-syvyisten juurrospuiden $\sim_k$-ekvivalenssiluokkaa τ kohden $t_{i,\tau}$ on niiden metsän $\mathcal{M}^i$, $i = 1, 2$, juurrospuiden lukumäärä, jotka kuuluvat ekvivalenssiluokkaan τ. Tällöin $\mathcal{M}^1 \sim_k \mathcal{M}^2$, jos ja vain jos jokaiselle τ pätee

$$t_{1,\tau} = t_{2,\tau} \quad \text{tai} \quad t_{1,\tau}, t_{2,\tau} \geq k.$$

Seuraava apulause osoittautuu hyödylliseksi relaatiota $\sim_k$ tutkittaessa ja sen todistus voi auttaa havainnollistamaan yllä esiteltyä määritelmää:

Apulause 5.32. Jos $\mathcal{M}^1$ ja $\mathcal{M}^2$ ovat juurrosmetsiä, joille pätee $\mathcal{M}^1 \sim_k \mathcal{M}^2$, niin niiden syvyydet ovat samat sekä kaikille $s, j \leq k$ pätee $\mathcal{M}^1 \upharpoonright s \sim_j \mathcal{M}^2 \upharpoonright s$.

Todistus. Todistetaan induktiolla juurrosmetsien maksimisyvyyden suhteen. Mikäli $\mathcal{M}^1$ ja $\mathcal{M}^2$ ovat nollametsiä, väite on selvä.

Tehdään nyt induktio-oletus, että jos $\mathcal{M}^1$ ja $\mathcal{M}^2$ ovat korkeintaan $r-1$-syvyisiä ja $\mathcal{M}^1 \sim_k \mathcal{M}^2$, niin $\mathcal{M}^1$ ja $\mathcal{M}^2$ ovat samansyvyisiä sekä kaikille $s \geq 0$ ja $j \leq k$ pätee $\mathcal{M}^1 \upharpoonright s \sim_j \mathcal{M}^2 \upharpoonright s$.

Tarkastellaan nyt tapausta, jossa $\mathcal{M}^1$ ja $\mathcal{M}^2$ ovat korkeintaan r-syvyisiä ja $\mathcal{T}^i$ ja $\mathcal{F}^i$, $i = 1, 2$, ovat kuten määritelmässä 5.31. Voidaan olettaa, että metsän $\mathcal{M}^1$ syvyys on r. Nyt koska $\mathcal{U}^1 \sim_k \mathcal{U}^2$, niin $\mathcal{T}^1 \sim_k \mathcal{T}^2$ ja $\mathcal{F}^1 \sim_k \mathcal{F}^2$. Toisaalta tiedetään, että puut $\mathcal{T}^i$ ja metsät $\mathcal{F}^i$, $i = 1, 2$, ovat korkeintaan $r - 1$-syvyisiä ja erityisesti joko puun $\mathcal{T}^1$ tai metsän $\mathcal{F}^1$ syvyys on $r - 1$. Tällöin induktio-oletuksesta seuraa, että joko puut $\mathcal{T}^1$ ja $\mathcal{T}^2$ ovat syvyydeltään $r - 1$ tai metsät $\mathcal{F}^1$ ja $\mathcal{F}^2$ ovat syvyydeltään $r - 1$. Tästä seuraa, että $\mathcal{M}^1$ ja $\mathcal{M}^2$ ovat syvyydeltään r.

Oletetaan sitten, että $\mathcal{M}^1$ ja $\mathcal{M}^2$ ovat r-syvyisiä ja $\mathcal{M}^1 \sim_k \mathcal{M}^2$. Oletetaan lisäksi $s < r$, sillä jos $s \geq r$, niin $\mathcal{M}^i \upharpoonright s = \mathcal{M}^i$, $i = 1, 2$, ja väite seuraa suoraan oletuksesta $\mathcal{M}^1 \sim_k \mathcal{M}^2$. Nyt puut $\mathcal{T}^1$ ja $\mathcal{T}^2$ sekä metsät $\mathcal{F}^1$ ja $\mathcal{F}^2$ ovat korkeintaan $r - 1$-syvyisiä sekä $\mathcal{T}^1 \sim_k \mathcal{T}^2$ ja $\mathcal{F}^1 \sim_k \mathcal{F}^2$. Tällöin induktio-oletuksesta seuraa, että kaikille $0 < s < r$ ja $j \leq k$ pätee $\mathcal{T}^1 \upharpoonright s \sim_j \mathcal{T}^2 \upharpoonright s$ sekä $\mathcal{F}^1 \upharpoonright s \sim_j \mathcal{F}^2 \upharpoonright s$. Tällöin määritelmän 5.31 nojalla väite pätee myös metsille $\mathcal{M}^1 \upharpoonright s$ ja $\mathcal{M}^2 \upharpoonright s$.

Induktioperiaatteen nojalla väite pätee mielivaltaisille juurrosmetsille $\mathcal{M}^1$ ja $\mathcal{M}^2$, joille $\mathcal{M}^1 \sim_k \mathcal{M}^2$. □

5.6 Keskitetty kokoonpano

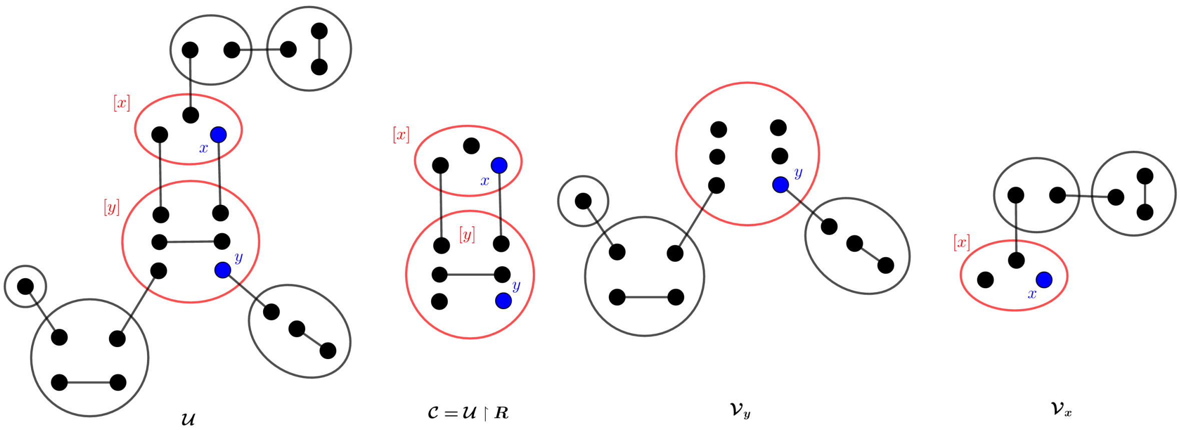
Seuraavaksi laajennetaan relaatio $\sim_k$ suurempaan osittaisten kokoonpanojen luokkaan, joka sisältää kaikki äärelliset ympäristöt, jotka löytyvät miltei kaikista kokoonpanoista.

Määritelmä 5.33. Sanotaan, että osittainen kokoonpano, jolla on juurijoukko R, on *keskitetty*, mikäli jokaisella solmulla on täsmälleen yksi polku joukkoon $\bigcup_{x \in R} [x]$.

Huomautus. Keskitetyssä kokoonpanossa jokainen syklin solmujen muodostama joukko sisältyy joukkoon $\bigcup_{x \in R} [x]$, sillä syklin sisällä on aina kaksi tapaa päästä haluamastaan solmusta toiseen.

Määritelmä 5.34. Olkoon $\mathcal{U} = (U, \equiv, M, R)$ keskitetty kokoonpano ja $C = \mathcal{U} \upharpoonright R$. Oletetaan, että $x \in R$. Olkoon V_x niiden solmujen joukko, joiden polku joukkoon R päättyy solmuun, joka on luokassa $[x]$, ja olkoon $\mathcal{V}_x$ joukon V_x indusoima puu, jolla on juuri x, ja jonka kaikki särmät $m \in M \upharpoonright [x]$ on poistettu. Sanotaan, että joukko $\{C, \mathcal{V}_x \mid x \in R\}$ on kokoonpanon $\mathcal{U}$ *kanoninen hajotelma*.

Esimerkki 5.35. Kuvassa on keskitetty kokoonpano $\mathcal{U}$, jonka juurijoukko on $R = [x] \cup [y]$ ja kokoonpanon $\mathcal{U}$ kanonisen hajotelman alkiot $\mathcal{C} = \mathcal{U} \upharpoonright R$, $\mathcal{V}_y$ ja $\mathcal{V}_x$. Huomataan, että alikokoonpanojen $\mathcal{V}_x$ ja $\mathcal{V}_y$ solmujoukoille V_x ja V_y pätee $V_x \cap R = [x] \cap R$ sekä $V_x \cap V_y = \emptyset$. Havaitaan, että sama pätee yleisesti kaikille keskitettyjen kokoonpanojen kanonisille hajotelmille.



Määritelmä 5.36. Olkoot $\mathcal{U}^1$ ja $\mathcal{U}^2$ yhtenäisiä, keskitettyjä kokoonpanoja, joilla on kanoniset hajotelmat $\{\mathcal{C}^1, \mathcal{V}^1_x \mid x \in R^1\}$ ja $\{\mathcal{C}^2, \mathcal{V}^2_x \mid x \in R^2\}$. Tällöin $\mathcal{U}^1 \sim_k \mathcal{U}^2$, jos on olemassa sellainen isomorfismi $f\colon \mathcal{C}^1 \to \mathcal{C}^2$, että kaikille $x \in R^1$ pätee $\mathcal{V}^1_x \sim_k \mathcal{V}^2_{f(x)}$.

Jos $\mathcal{U}^1$ ja $\mathcal{U}^2$ ovat sellaisten kahden kokoonpanon komponentteja, joilla pelataan Ehrenfeuchtin peliä, vaaditaan isomorfismilta f lisäksi, että se säilyttää "merkityt" solmut. Toisin sanoen a^1_j on merkintä rakenteessa $\mathcal{U}^1$, jos ja vain jos a^2_j on merkintä rakenteessa $\mathcal{U}^2$, sekä $f(a^1_j) = a^2_j$.

Huomautus. Laajennetaan tarkasteltavia kokoonpanoja lisäämällä niihin rakennetta merkitsemällä kokoonpanon juurijoukon solmuja Ehrenfeuchtin pelin mukaisilla vakioilla. Kun kokoonpanon $(U, \equiv, M, R)$ juurijoukkoon R on tehty merkinnät $a_1, \ldots, a_j$, niin merkitään kyseistä kokoonpanoa $(U, \equiv, M, R, a_1, \ldots, a_j)$

Laajennetaan nyt relaatio $\sim_k$ mielivaltaisiin keskitettyihin kokoonpanoihin $\mathcal{U}^1$ ja $\mathcal{U}^2$.

Määritelmä 5.37. Jos $\mathcal{U}^1$ ja $\mathcal{U}^2$ ovat keskitettyjä kokoonpanoja, niin $\mathcal{U}^1 \sim_k \mathcal{U}^2$, jos ja vain jos kaikille yhtenäisten ja keskitettyjen kokoonpanojen $\sim_k$-ekvivalenssiluokille τ pätee

$$h_{1,\tau} = h_{2,\tau} \quad \text{tai} \quad h_{1,\tau}, h_{2,\tau} \geq k,$$

missä $h_{i,\tau}$ on niiden kokoonpanon $\mathcal{U}^i$, $i = 1, 2$, komponenttien lukumäärä, jotka kuuluvat ekvivalenssiluokkaan τ.

5.7 Rikas ja yksinkertainen kokoonpano

Viimeiseksi relaatio $\sim_k$ laajennetaan sellaiseen kokoonpanojen luokkaan, joka sisältää lähes kaikki kokoonpanot.

Määritelmä 5.38. Olkoon $\mathcal{U} = (U, \equiv, M)$ kokoonpano. Sanotaan, että kokoonpano $\mathcal{U}$ on *k-rikas*, mikäli kaikilla $r \leq 3^{k-1}$, kaikilla juurrospuilla $\mathcal{T}$ sekä kaikilla korkeintaan $k-1$-alkioisilla joukoilla $S \subseteq U$ on olemassa sellainen solmu $x \in U$, että

$$\begin{aligned}
&\delta(S, x) > 2 \cdot 3^{k-1} \\
&\delta(C, x) > 2 \cdot 3^{k-1} \text{ jokaiselle syklille } C \subseteq U, \text{ jolle pätee } |C| \leq 2 \cdot 3^{k-1}, \\
&\mathcal{N}(x, r) \sim_k \mathcal{T}.
\end{aligned}$$

Huomautus. Vaikka erilaisten juurrospuiden lukumäärä on ääretön, on $\sim_k$ ekvivalenssiluokkia vain äärellinen määrä, mikä helpottaa rikkauden kolmannen ehdon tarkastelemista.

Määritelmä 5.39. Sanotaan, että kokoonpano $\mathcal{U}$ on k*-yksinkertainen*, mikäli sillä ei ole kahta korkeintaan 3^k-pituista sykliä, joiden etäisyys on korkeintaan $2 \cdot 3^{k-1}$.

Apulause 5.40. Jos $\mathcal{U} = (U, \equiv, M)$ on k-rikas ja k-yksinkertainen, niin $\mathcal{U}$ on myös j-rikas ja j-yksinkertainen kaikilla $j < k$.

Todistus. Olkoon $\mathcal{U}$ k-rikas ja k-yksinkertainen kokoonpano ja $j < k$. Olkoon lisäksi $r \leq 3^{j-1}$, $\mathcal{T}$ juurrospuu ja $S \subseteq U$ korkeintaan $j - 1$-alkioinen joukko. Koska $\mathcal{U}$ on k-rikas, on olemassa solmu $x \in U$, joka toteuttaa määritelmän 5.38 mukaiset ehdot. Nyt koska $3^{k-1} > 3^{j-1}$, niin solmu x toteuttaa myös j-rikkauden ehdot:

Koska kaikille korkeintaan $k-1$-alkioisille joukoille $A \subseteq U$ pätee $\delta(A, x) > 2 \cdot 3^{k-1} > 2 \cdot 3^{j-1}$, niin erityisesti joukolle S pätee $\delta(S, x) > 2 \cdot 3^{j-1}$. Solmu x toteuttaa j-rikkauden toisen ehdon vastaavasti. Koska k-rikkauden kolmannen ehdon mukaan $\mathcal{N}(x, s) \sim_k \mathcal{T}$ kaikilla $s \leq 3^{k-1}$, niin apulauseen 5.32 nojalla pätee erityisesti $\mathcal{N}(x, r) \sim_j \mathcal{T}$. Kokoonpano $\mathcal{U}$ on siis j-rikas.

Kokoonpanon j-yksinkertaisuus seuraa suoraan havainnosta $3^j < 3^k$. □

Suurempia satunnaisia kokoonpanoja muodostettaessa myös kokoonpanoon muodostuvien syklien pituudet kasvavat. Toisaalta on myös hahmotettavissa, että kokoonpanon ”ulkoreunalta” voidaan löytää kaukana sykleistä sijaitsevia solmuja. Luvussa 6 todistetaan, että lähes kaikki suuret kokoonpanot todella ovat k-rikkaita ja k-yksinkertaisia.

Määritelmä 5.41. Olkoot $\mathcal{U}^1$ ja $\mathcal{U}^2$ k-rikkaita ja k-yksinkertaisia kokoonpanoja. Oletetaan, että jokaista korkeintaan 3^{k-1}-syvyisten keskitettyjen yksisyklisten kokoonpanojen, joiden syklin pituus on korkeintaan $2 \cdot 3^{k-1}$, $\sim_k$-ekvivalenssiluokkaa τ kohden $c_{i,\tau}$ on niiden kokoonpanon $\mathcal{U}^i$, $i = 1, 2$, syklien C lukumäärä, joille pätee $\mathcal{N}(C, 3^{k-1}) \in \tau$. Tällöin $\mathcal{U}^1 \sim_k \mathcal{U}^2$, jos ja vain jos kaikille τ pätee

$$c_{1,\tau} = c_{2,\tau} \quad \text{tai} \quad c_{1,\tau}, c_{2,\tau} \geq k.$$

Seuraavaksi osoitetaan, että jos kaksi rakennetta ovat relaatiossa $\sim_k$, niin ne ovat myös k-ekvivalentit: Todistus on sovellus lauseesta 5.6. Samaan tapaan kuin relaatio $\sim_k$ laajennettiin ensin juurrosmetsistä keskitettyihin kokoonpanoihin ja lopulta k-rikkaisiin ja k-yksinkertaisiin kokoonpanoihin, käytetään samaa kaavaa ja kuvaillaan voittostrategia pelaajalle B.

Olkoot $\mathcal{U}^1 = (U^1, \equiv^1, M^1, R^1)$ ja $\mathcal{U}^2 = (U^2, \equiv^2, M^2, R^2)$ keskitettyjä kokoonpanoja. Kun $j = 0, \ldots, k$ ja $i = 1, 2$, olkoon

$$\mathcal{U}^i_j = (U^i, \equiv^i, M^i, R^i_j, a^i_1, \ldots, a^i_j),$$

missä

$$R^i_j = R^i \cup \left\{ x \in U_i \mid x \text{ kuuluu polkuun jostain solmusta } a^i_h \text{ joukkoon } \bigcup_{x \in R^i} [x] \right\}.$$

Tällöin $\mathcal{U}^1_j$ ja $\mathcal{U}^2_j$ ovat keskitettyjä kokoonpanoja, joille on tehty Ehrenfeuchtin pelin mukaiset merkinnät $a^i_1, \ldots, a^i_j$. Pelaajan B tehtävänä on säilyttää ehto

$$\mathcal{U}^1_j \sim_{k-j} \mathcal{U}^2_j \tag{5.1}$$

jokaisella kierroksella j. Ehtoa 5.1 kutsutaan *kanoniseksi invariantiksi*. Jos $j = 0$, niin kanoninen invariantti vastaa ehtoa $\mathcal{U}^1 \sim_k \mathcal{U}^2$. Jos taas kanoninen invariantti on voimassa kun $j = k$, on pelaaja B voittanut. Riittää siis näyttää, että jos kanoninen invariantti pätee kierroksella j, voi pelaaja B aina tehdä sellaisen siirron, että kanoninen invariantti pätee myös kierroksella $j + 1$.

Apulause 5.42. Olkoot $\mathcal{U}^1$ ja $\mathcal{U}^2$ juurrosmetsiä ja $0 \leq j < k$. Jos kanoninen invariantti pätee Ehrenfeuchtin pelin (rakenteilla $\mathcal{U}^1$ ja $\mathcal{U}^2$) kierroksella j, niin pelaaja B voi aina tehdä sellaisen siirron, että kanoninen invariantti pätee myös kierroksella $j + 1$.

Todistus. Määritelmän 5.31 tapaan käytetään induktiota metsien $\mathcal{U}^1$ ja $\mathcal{U}^2$ syvyyden suhteen. Jos metsän syvyys on -1, tulos on triviaali. Oletetaan nyt, että $r \geq 0$ sekä että väite pätee sellaisille pareille juurrosmetsiä, joiden syvyys on pienempi kuin r. On tarkasteltava erikseen tapaus, jossa $\mathcal{U}^1$ ja $\mathcal{U}^2$ ovat puita: Olkoot siis $\mathcal{U}^1$ ja $\mathcal{U}^2$ juurrospuita, joiden syvyys on r. Käytetään määritelmän 5.31 merkintöjä. Nyt pelaaja A voi valita solmun a^i_{j+1} kahdella eri tavalla:

(1) Oletetaan, että a^i_{j+1} on puussa T^i. Määritelmän 5.31 mukaan $\mathcal{T}^1 \sim_k \mathcal{T}^2$ ja koska kanoninen invariantti pätee, niin $\mathcal{T}^1_j \sim_{k-j} \mathcal{T}^2_j$. Koska puiden $\mathcal{T}^1$ ja $\mathcal{T}^2$ syvyydet ovat korkeintaan $r - 1$, induktio-oletuksen nojalla pelaaja B voi valita solmun a^{3-i}_{j+1} puusta $\mathcal{T}^{3-i}$ siten, että $\mathcal{T}^1_{j+1} \sim_{k-j-1} \mathcal{T}^2_{j+1}$. Apulauseen 5.32 nojalla tästä seuraa, että kanoninen invariantti pätee kierroksella $j + 1$.

(2) Oletetaan, että a^i_{j+1} on metsässä $\mathcal{F}^i$. Sama strategia kuin kohdassa (1) toimii, kun puun $\mathcal{T}^i$ sijaan tarkastellaan metsää $\mathcal{F}^i$.

Oletetaan nyt, että $\mathcal{U}^1$ ja $\mathcal{U}^2$ ovat r-syvyisiä juurrosmetsiä. Sanotaan, että metsään $\mathcal{U}^i$ kuuluva puu on *aktiivinen*, mikäli sen solmulle on tehty Ehrenfeuchtin pelin mukainen merkintä a^i_j. Muulloin sanotaan, että puu on *epäaktiivinen*. On jälleen tarkasteltava kahta eri tapausta:

(1) Oletetaan, että a^i_{j+1} on epäaktiivisessa puussa $\mathcal{T}^i$, joka kuuluu ekvivalenssiluokkaan τ. Nyt koska $\mathcal{U}^1 \sim_k \mathcal{U}^2$, niin $t_{1,\tau} = t_{2,\tau}$ tai $t_{1,\tau}, t_{2,\tau} \geq k$, mistä seuraa että on oltava olemassa metsän $\mathcal{U}^{3-i}$ epäaktiivinen puu $\mathcal{T}^{3-i}$, joka kuuluu luokkaan τ. Tästä seuraa, että $\mathcal{T}^1 \sim_k \mathcal{T}^2$ ja voidaan käyttää samaa strategiaa, kuin juurrospuita tarkasteltaessa. Siis pelaaja B voi aina tehdä sellaisen siirron, että $\mathcal{T}^1_1 \sim_{k-1} \mathcal{T}^2_1$. Tällöin apulauseen 5.32 nojalla kanoninen invariantti pätee kierroksella $j + 1$.

(2) Oletetaan, että a^1_{j+1} kuuluu aktiiviseen puuhun $\mathcal{T}^i$. Olkoon $\mathcal{T}^{3-i}$ vastaava puu toisessa rakenteessa $\mathcal{U}^{3-i}$. Nyt koska kanoninen invariantti pätee kierroksella j, on voimassa $\mathcal{T}^1_j \sim_{k-j} \mathcal{T}^2_j$, ja käyttämällä samaa strategiaa, kuin juurrospuita tarkasteltaessa, pelaaja B pystyy aina tekemään sellaisen siirron, että $\mathcal{T}^1_{j+1} \sim_{k-j-1} \mathcal{T}^2_{j+1}$, ja jälleen kanoninen invariantti pätee kierroksella $j + 1$.

Pelaajan A valinnasta riippumatta pelaaja B voi siis aina tehdä sellaisen siirron, että kanoninen invariantti pätee kierroksella $j + 1$. □

Apulause 5.43. Olkoot $\mathcal{U}^1$ ja $\mathcal{U}^2$ keskitettyjä kokoonpanoja, joilla on kanoniset hajotelmat $\{C^1, \mathcal{V}_x^1 \mid x \in R^1\}$ ja $\{C^2, \mathcal{V}_x^2 \mid x \in R^2\}$. Jos $\mathcal{U}^1 \sim_k \mathcal{U}^2$ ja kanoninen invariantti pätee Ehrenfeuchtin pelin kierroksella j, niin pelaaja B voi aina tehdä sellaisen siirron, että kanoninen invariantti pätee myös kierroksella $j+1$.

Todistus. Oletetaan ensin, että $\mathcal{U}^1$ ja $\mathcal{U}^2$ ovat yhtenäisiä. Olkoon $f\colon R^1 \to R^2$ määritelmän 5.36 mukainen isomorfismi. Jos pelaaja A valitsee solmun, joka kuuluu joukkoon $\mathcal{V}_x^i$, niin pelaaja B voi vastata valitsemalla solmun joukosta $\mathcal{V}_{f(x)}^{3-i}$ käyttäen samaa strategiaa, kuin apulauseessa 5.42.

Mielivaltaisille keskitetyille kokoonpanoille pelaajan B strategia on samankaltainen kuin juurrosmetsiä tarkasteltaessa. Huomataan, että kun käydään erikseen läpi tapaus, joissa kokoonpanon $\mathcal{U}^i$ komponentti on aktiivinen ja tapaus, jossa se on epäaktiivinen, saadaan johdettua haluttu tulos. □

Viimeiseksi kuvaillaan voittostrategia, kun $\mathcal{U}^1$ ja $\mathcal{U}^2$ ovat k-rikkaita sekä k-yksinkertaisia. Olkoot $i = 1, 2$ ja

$$A^i = \bigcup \left\{ C \subseteq U^i \mid C \text{ on sykli, jolle pätee } |C| \leq 2 \cdot 3^{k-1} \right\}.$$

Olkoot lisäksi $j = 0, \ldots, k$ ja

$$\begin{aligned}
U_j^i &= N\left(\{a_1^i, \ldots, a_j^i\} \cup A^i, 3^{k-j}\right), \\
R_j^i &= \{x \in U_j^i \mid x \text{ on joko polulla solmusta } a_g^i \text{ solmuun } a_h^i \text{ tai solmusta } a_g^i \text{ solmuun } y \in A^i\}, \\
\mathcal{U}_j^i &= (U_j^i, \equiv^i, M^i, R_j^i, a_1^i, \ldots, a_j^i).
\end{aligned}$$

Koska $\mathcal{U}_j^i$ on k-yksinkertainen, se on myös keskitetty. Nyt pelaajan B tehtävänä on jälleen säilyttää kanoninen invariantti tarkasteltaessa uudelleen määriteltyjä kokoonpanoja $\mathcal{U}_j^1$ ja $\mathcal{U}_j^2$.

Apulause 5.44. Olkoot $\mathcal{U}^1 \sim_k \mathcal{U}^2$ k-rikkaita ja k-yksinkertaisia kokoonpanoja. Olkoon lisäksi $0 \leq j < k$. Jos kanoninen invariantti pätee Ehrenfeuchtin pelin kierroksella j, niin pelaaja B voi aina tehdä sellaisen siirron, että kanoninen invariantti pätee myös kierroksella $j+1$.

Todistus. Pelaajan A siirto jakaa tarkastelun jälleen useaan eri tapaukseen:

(1) Oletetaan, että $N\left(a_{j+1}^i, 3^{k-j-1}\right) \nsubseteq U_j^i$. Tällöin on olemassa solmu $u \in N\left(a_{j+1}^i, 3^{k-j-1}\right)$, jonka etäisyys joukosta $\{a_1^i, \ldots, a_j^i\} \cup A^i$ on vähintään $3^{k-j} + 1$. Jokaisen solmun $v \in N\left(a_{j+1}^i, 3^{k-j-1}\right)$ etäisyys solmusta u on korkeintaan $2 \cdot 3^{k-j-1}$, eli

$$N\left(a_{j+1}^i, 3^{k-j-1}\right) \cap N\left(\{a_1^i, \ldots, a_j^i\} \cup A^i, 3^{k-j-1}\right) = \emptyset,$$

ja $\mathcal{N}\left(a_{j+1}^i, 3^{k-j-1}\right)$ on syklitön. Oletetaan, että $\mathcal{N}\left(a_{j+1}^i, 3^{k-j-1}\right)$ kuuluu $\sim_k$-ekvivalenssiluokkaan τ. Koska $\mathcal{U}^{3-i}$ on k-rikas, on olemassa sellainen $a_{j+1}^{3-i} \in U^{3-i}$, että myös $\mathcal{N}\left(a_{j+1}^{3-i}, 3^{k-j-1}\right)$ kuuluu luokkaan τ ja

$$N\left(a_{j+1}^{3-i}, 3^{k-j-1}\right) \cap N\left(\{a_1^{3-i}, \ldots, a_j^{3-i}\} \cup A^{3-i}, 3^{k-j-1}\right) = \emptyset.$$

Tällöin apulauseen 5.32 nojalla kanoninen invariantti pätee kierroksella $j+1$.

(2) Oletetaan, että $N\left(a^i_{j+1}, 3^{k-j-1}\right) \subseteq N(C^i, 3^{k-j})$, missä C^i on sykli, jolle $|C| \leq 2 \cdot 3^{k-1}$, joka kuuluu johonkin epäaktiiviseen kokoonpanon $\mathcal{U}^i_j$ komponenttiin. Tällöin

$$N\left(\{a^i_1, \ldots, a^i_j\}, 3^{k-j-1}\right) \cap N\left(C^i, 3^{k-j-1}\right) = \emptyset.$$

Koska $\mathcal{U}^1$ ja $\mathcal{U}^2$ ovat k-yksinkertaisia, $\mathcal{U}^1 \sim_k \mathcal{U}^2$ ja $\mathcal{U}^1_j \sim_{k-j} \mathcal{U}^2_j$, niin kokoonpanossa $\mathcal{U}^{3-i}$ on olemassa sykli C^{3-i}, jolle pätee

$$\mathcal{N}(C^1, 3^{k-j-1}) \sim_k \mathcal{N}(C^2, 3^{k-j-1})$$

ja

$$N\left(\{a^{3-i}_i, \ldots, a^{3-i}_j\}, 3^{k-j-1}\right) \cap N\left(C^{3-i}, 3^{k-j-1}\right) = \emptyset.$$

Nyt apulauseen 5.43 nojalla on olemassa sellainen $a^{3-i}_{j+1} \in N(C^{3-i}, 3^{k-j-1})$, että

$$\left(N(C^1, 3^{k-j-1}), \equiv^1, M^1, R^1\right) \sim_{k-1} \left(N(C^2, 3^{k-j-1}), \equiv^2, M^2, R^2\right),$$

missä $R^i = C^i \cup P^i$ ja P^i on polku solmusta a^i_{j+1} joukkoon $\bigcup_{x \in C^i}[x]$. Nyt apulauseen 5.32 nojalla kanoninen invariantti pätee kierroksella $j + 1$.

(3) Oletetaan, että $N(a^i_{j+1}, 3^{k-j-1}) \subseteq W^i$, missä W^i on jokin kokoonpanon $\mathcal{U}^i$ aktiivinen komponentti. Olkoon W^{3-i} vastaava aktiivinen komponentti kokoonpanossa $\mathcal{U}^{3-i}$. Olkoot $i = 1, 2$ ja $\overline{b^i}$ se jonon $(a^i_1, \ldots, a^i_j)$ alijono, joka koostuu solmuista $a^i_h \in W^i$.
Olkoon lisäksi $\mathcal{W}^i = (W^i, \equiv^i, M^i, R^i_j, \overline{b^i})$. Nyt koska kanoninen invariantti pätee, $\mathcal{W}^1 \sim_{k-j} \mathcal{W}^2$, joten apulauseen 5.43 nojalla pelaaja B voi valita solmun $a^{3-i}_{j+1} \in W^{3-i}$ siten, että

$$\left(W^1, \equiv^1, M^1, R^1 \cup P^1, \overline{b^1}, a^1_{j+1}\right) \sim_{k-j-1} \left(W^2, \equiv^2, M^2, R^2 \cup P^2, \overline{b^2}, a^2_{j+1}\right),$$

missä P^i on polku solmusta a^i_{j+1} joukkoon $\bigcup_{x \in R^i}[x]$. Tällöin apulauseen 5.32 nojalla kanoninen invariantti pätee kierroksella $j + 1$.

□

Seuraava lause kokoaa yhteen tämän luvun tulokset ja mahdollistaa luvussa 6 esiteltävän tutkielman päälauseen todistuksen:

Lause 5.45. *Olkoot $\mathcal{U}^1$ ja $\mathcal{U}^2$ k-rikkaita ja k-yksinkertaisia kokoonpanoja. Jos $\mathcal{U}^1 \sim_k \mathcal{U}^2$, niin $\mathcal{U}^1 \equiv_k \mathcal{U}^2$.*

Todistus. Jos $\mathcal{U}^1 \sim_k \mathcal{U}^2$, niin apulauseiden 5.42, 5.43 ja 5.44 nojalla pelaajalla B on voittostrategia k-kierroksisessa Ehrenfeuchtin pelissä kokoonpanoilla $\mathcal{U}^1$ ja $\mathcal{U}^2$. Nyt lauseen 5.6 nojalla $\mathcal{U}^1 \equiv_k \mathcal{U}^2$. □

6 Puuprosessit

Edellisessä luvussa näytettiin, että k-rikkaan ja k-yksinkertaisen kokoonpanon $\sim_k$-ekvivalenssiluokka määrittyy sen mukaan, kuinka monta pienikokoista $\sim_k$-tyyppistä yksisyklistä ympäristöä sillä on kussakin $\sim_k$-ekvivalenssiluokassa. Tutkitaan nyt todennäköisyyttä, että satunnaisella kokoonpanolla on jokin tietty määrä pieniä syklittömiä tai yksisyklisiä ympäristöjä. Aloitetaan tarkastelemalla todennäköisyyttä, että annetun solmujoukon ympäristö mielivaltaisessa satunnaisessa kokoonpanossa on isomorfinen jonkin keskitetyn kokoonpanon kanssa. Tarkastelu pohjautuu pitkälti artikkeliin [8].

6.1 Peruskäsitteitä

Haarautumisprosessi on populaation kasvamista kuvaava malli, jossa jokaisella populaation edustajalla on satunnainen määrä jälkeläisiä. Vaikka tavallisesti tällaiset prosessit voivat jatkua äärettömästi, ollaan tässä tutkielmassa kiinnostuneita vain äärellisen mittaisista prosesseista. Tässä luvussa esitellään lyhyesti tärkeimmät haarautumisprosesseihin liittyvät tulokset, joita tarvitaan tutkielman päälauseen todistamiseen. Lisätietoja haarautumisprosesseista löytyy teoksesta [6].

Tapaa, jolla satunnainen juurrospuu tai -metsä muodostetaan, sanotaan *puuprosessiksi* ja sen tuottaman juurrospuun tai -metsän alkioita sanotaan *yksilöiksi*. Lisäksi sanotaan, että *sukupolvi* r koostuu kaikista yksilöistä, joiden syvyys on r.

Yksinkertaisuuden vuoksi samaistetaan puuprosessi ja sen tuottama rakenne. Näin voidaan merkitä puuprosessia symbolilla $\mathcal{P}$. Yksilöiden jälkeläisten lukumäärät ovat riippumattomia toisistaan, mutta todennäköisyys, että yksilöllä on tietty määrä jälkeläisiä, on sama kaikille yksilöille. Tästä syystä puuprosessin todennäköisyysavaruuden määrittelee kuvaus $p\colon \mathbb{N} \to [0,1]$, missä yksilön *haarautumistodennäköisyys* $p(j)$ on todennäköisyys, että yksilöllä on j jälkeläistä. Jos $\mathcal{P}$ on puuprosessi ja $r \geq 0$, niin merkitään lisäksi, että $\mathcal{P} \restriction r$ on sellainen prosessin $\mathcal{P}$ tuottama äärellinen puu tai metsä, joka on $\mathcal{P}$ rajoitettuna ensimmäiseen r sukupolveen.

Määritelmä 6.1. *Puuprosessin määrittelemä haarautumisprosessi* on jono $(Z_0, Z_1, \dots)$, missä Z_r on sukupolven r koko.

Apulause 6.2. Oletetaan, että Z_0 on kiinnitetty. Jos yksilön jälkeläisten lukumäärän odotusarvo on $\mu < \infty$, niin kaikille $r \in \mathbb{N}$ pätee $\mathbb{E}(Z_r) = Z_0\mu^r$.

Todistus. Todistetaan induktiolla luonnollisen luvun r suhteen. Yksilön jälkeläisten lukumäärän odotusarvo on $\mu = \sum_{j=1}^{\infty} jp(j)$ ja sukupolvessa 0 on Z_0 yksilöä. Koska yksilöiden jälkeläisten lukumäärät ovat riippumattomia toisistaan, sukupolven 1 koon odotusarvo on odotusarvon lineaarisuuden nojalla

$$\mathbb{E}(Z_1) = \mathbb{E}\left(\sum_{i=1}^{Z_0} \mu\right) = \mu\mathbb{E}\left(\sum_{i=1}^{Z_0} 1\right) = \mu Z_0.$$

Tehdään nyt induktio-oletus, että $\mathbb{E}(Z_r) = \mu^r Z_0$. Tarkasteltaessa sukupolvea $r+1$ voidaan olettaa, että sukupolven r koko on kiinnitetty. Tällöin $\mathbb{E}(Z_{r+1} \mid Z_r) = \mu Z_r$. Nyt pätee

$$\mathbb{E}(Z_{r+1}) = \mathbb{E}(\mathbb{E}(Z_{r+1} \mid Z_r)) = \mathbb{E}(\mu Z_r) = \mu\mathbb{E}(Z_r) = \mu\mu^r Z_0 = \mu^{r+1} Z_0$$

Induktioperiaatteen nojalla väite pätee kaikille $r \in \mathbb{N}$. $\square$

Määritelmä 6.3. Olkoot $(a_n)_{n\in\mathbb{N}}$ ja $(b_n)_{n\in\mathbb{N}}$ positiivisia lukujonoja. Jos $\frac{a_n}{b_n} \to 0$, kun $n \to \infty$, niin merkitään $a_n = o(b_n)$.

Määritelmä 6.4. Olkoot $(a_n)_{n\in\mathbb{N}}$ ja $(b_n)_{n\in\mathbb{N}}$ positiivisia lukujonoja. Jos on olemassa vakio $C > 0$, jolle $\frac{a_n}{b_n} \leq C$, kaikilla riittävän suurilla $n \in \mathbb{N}$, niin sanotaan, että $\frac{a_n}{b_n}$ on *oleellisesti rajoitettu*. Tällöin merkitään $a_n = O(b_n)$.

Huomautus. Merkinnät $a_n = o(b_n)$ ja $a_n = O(b_n)$ voivat olla hämääviä, jos niitä käytetään yhtälönratkaisussa molempiin suuntiin. Esimerkiksi pätee $a_n = \frac{1}{n} = o(1)$ ja $b_n = \frac{1}{n^2} = o(1)$, mutta selvästi $a_n \neq b_n$. Tästä syystä merkinnän $a_n = o(b_n)$ sijaan olisi järkevämpää käyttää merkintää $a_n \in o(b_n)$. Yhtäsuuruusmerkin käyttäminen on kuitenkin hyvin yleinen merkintätapa ja se helpottaa monimutkaisten lausekkeiden sieventämistä huomattavasti.

Esimerkki 6.5. Olkoot $a_n = \frac{1}{n}$ ja $b_n = 2n$. Tällöin esimerkiksi $a_n = o(1)$ ja $b_n = o(n^2)$. Intuitiivisesti merkintä $x_n = o(y_n)$ tarkoittaa siis, että y_n kasvaa paljon nopeammin kuin x_n.

Huomataan, että kaikille lukujonoille $(a_n)_{n\in\mathbb{N}}$, joille $a_n = o(b_n)$ pätee myös $a_n = O(b_n)$. Tämä ei kuitenkaan päde toiseen suuntaan, sillä esimerkiksi $x_n = 2n^2 = O(n^2)$, mutta $x_n \neq o(n^2)$.

Määritelmä 6.6. Olkoot $(a_n)_{n\in\mathbb{N}}$ ja $(b_n)_{n\in\mathbb{N}}$ lukujonoja. Merkitään $a_n = \Omega(b_n)$, jos ja vain jos $b_n = O(a_n)$. Lisäksi jos $a_n = \Omega(b_n)$ ja $a_n = O(b_n)$, niin merkitään $a_n = \Theta(b_n)$.

6.2 Ympäristöjen vertailemisesta

Seuravaksi tutkitaan ympäristöjen ja puuprosessien välistä suhdetta. Tämän luvun apulauseiden todistuksissa käytetään menetelmää, jossa satunnainen rakenne ”paljastetaan” osa kerrallaan, jolloin todennäköisyyksiä tarkasteltaessa voidaan olettaa jo paljastettujen osien olevan kiinnitettyjä. Tällaista menetelmää käytetään laajalti satunnaisrakenteiden tutkimuksessa.

Määritelmien 4.5 ja 4.6 mukaan hyvin käyttäytyvälle astejonolle pätee

$$m = \sum_{i=1}^{n-1} i d_i(n) = n(\Lambda + o(1)).$$

Olkoon $\mathcal{W} = (W, \equiv', M', S)$ keskitetty kokoonpano, jolla on kanoninen hajotelma $\{\mathcal{S}, \mathcal{W}_x \mid x \in S\}$. Tarkasteltavissa tapauksissa $\mathcal{W}$ on joko juurrosmetsä tai yksisyklinen kokoonpano, jonka sykli on S. Olkoon lisäksi $\mathcal{U} = (U, \equiv, M)$ satunnainen n-kokoonpano, $r \geq 0$, $R \subseteq U$, $\mathcal{R} = (R, \equiv, M)$ sekä $\mathcal{N} = \mathcal{N}(R, r)$.

Olkoon $f\colon S \to R$ bijektio. Merkitään $\mathcal{S} \overset{f}{\cong} \mathcal{R}$, mikäli f on rakenteiden $\mathcal{S}$ ja $\mathcal{R}$ välinen isomorfismi ja merkitään $\mathcal{W} \xrightarrow{f} \mathcal{N}$, mikäli $\mathcal{N}$ on keskitetty kokoonpano, jolla on kanoninen hajotelma $\{\mathcal{R}, \mathcal{V}_x \mid x \in R\}$ ja f voidaan laajentaa joukkoon W siten, että $f \restriction W_x$ on rakenteiden $\mathcal{W}_x$ ja $\mathcal{V}_{f(x)}$ välinen isomorfismi kaikilla $x \in S$. Edellisessä W_x on kokoonpanon $\mathcal{W}$ kanonisen hajotelman komponentin $\mathcal{W}_x$ solmujoukko. Tällöin

$$\mathcal{W} \cong \mathcal{N}, \text{ mikäli on olemassa bijektio } f\colon S \to R, \text{ jolle } \mathcal{S} \overset{f}{\cong} \mathcal{R} \text{ ja } \mathcal{W} \xrightarrow{f} \mathcal{N}.$$

Kiinnittämällä bijektion $f\colon S \to R$ löydetään lauseke todennäköisyydelle, että $\mathcal{S} \overset{f}{\cong} \mathcal{R}$. Sen jälkeen tutkitaan todennäköisyyttä, että $\mathcal{W} \xrightarrow{f} \mathcal{N}$ ehdolla $\mathcal{S} \overset{f}{\cong} \mathcal{R}$. Tämä tehdään tarkastelemalla kuvauksen f laajennusta paljastamalla yksi $\equiv$-ekvivalenssiluokka kerrallaan. Luokan $[f(x)]$

tarkastelu tapahtuu kahdessa osassa: Samaistetaan $[f(x)]$ jonkin joukon $\{1,\ldots,m\}$ ekvivalenssiluokan kanssa ja sitten merkitään solmuja kuuluvaksi luokkaan $[f(x)]$ siten, että luokat $[f(x)]$ ja $[x]$ sisältävät yhtä monta alkiota. Käytetään järjestystä $x_1, x_2, \ldots, x_w \in W$, missä $x_1, \ldots, x_s \in S$ ovat joukon S ekvivalenssiluokkien edustajat, eli $[x_i] \neq [x_j]$ kaikilla $i, j = 1, \ldots, w$, $W = \bigcup_{i=1}^{w}[x_i]$, $S \subseteq \bigcup_{i=1}^{s}[x_i]$ ja jos $[x_j]$ on luokan $[x_i]$ jälkeläinen, niin $i \leq j$. Tätä järjestystä noudattamalla kuvauksen f kasvamista voidaan arvioida puuprosessilla, missä prosessin yksilöt vastaavat $\equiv$-ekvivalenssiluokkia.

Olkoon $1 \leq i \leq s$ ja

$$\begin{aligned}
b_i &= \big|[x_i] \cap S\big| \\
c_i &= \big|\{y \in [x_i] \cap S \mid \text{on olemassa } z \in S, \text{ jolle } y\ M'\ z\}\big| \\
C &= \sum_{j=1}^{s} \frac{c_j}{2}.
\end{aligned}$$

Toisin sanoen C on niiden joukon M' särmien lukumäärä, jotka yhdistävät joukon S solmuja. Olkoon $1 \leq i \leq w$ ja

$$\begin{aligned}
a_i &= |[x_i]|, \text{ sekä} \\
A_i &= \sum_{j=1}^{i} a_j.
\end{aligned}$$

Oletetaan, että seuraavaksi tarkasteltavat astejonot ovat hyvin käyttäytyviä ja tasaisia. Nyt on olemassa sellainen kuvaus $\rho\colon \mathbb{N} \to \mathbb{R}^+$, että kaikille $i, n > 0$ pätee

$$\begin{aligned}
&\lim_{n\to\infty} \rho(n) = 0, \\
&\left|\frac{d_i(n)}{n} - \lambda_i\right| \leq \rho(n), \\
&\left|\sum_{i=1}^{n-1} \frac{i d_i(n)}{n} - \Lambda\right| \leq \rho(n),
\end{aligned}$$

sekä

$$\rho(n) \geq n^{-\frac{1}{2}}.$$

Viimeistä epäyhtälöä käytetään seuraavissa apulauseissa rajoittamaan solmujoukon kokoa. Tässä luvussa noudatetaan edellä mainittuja ehtoja sekä merkintöjä.

Apulause 6.7. Jos $|W| \leq \dfrac{1}{\sqrt{\rho(n)}}$, niin

$$\mathbb{P}_c(\mathcal{S} \overset{f}{\cong} \mathcal{R}, n) = \left(\prod_{i=1}^{s} \sum_{j=b_i}^{\infty} \frac{(j-1)!\lambda_j}{(j-b_i)!}\right) \cdot \frac{1 + O\left(\sqrt{\rho(n)}\right)}{m^{|S|+C-s}}.$$

Todistus. Paljastetaan yksitellen ekvivalenssiluokat $[f(x_i)]$, missä $1 \leq i \leq s$. Olkoon

$$j_i = |[f(x_i)]|$$

ja

$$J_i = \sum_{h=1}^{i} j_h + s - i.$$

Tällöin tulee olla $j_i \geq b_i$ ja rajoituksen 4.1 nojalla $J_i = o\left(\frac{n^{\frac{1}{4}}}{\sqrt{\rho(n)}}\right)$. Palautetaan mieleen, että astejonon jäsen $d_i(n)$ tarkoittaa myös niiden verkkoa vastaavan kokoonpanon ekvivalenssiluokkien lukumäärää, joiden koko on i. Olkoon $h < i$ sellainen luku, että $|[f(x_h)]| = j_i$. Tällaisten lukujen lukumäärä on välillä $[0, i-1]$, joten todennäköisyys, että $|[f(x_i)]| = j_i$, on lukujen

$$\frac{d_{j_i}(n) - i + 1}{n - i + 1} \quad \text{ja} \quad \frac{d_{j_i}(n)}{n - i + 1}$$

välillä. Siis tämä todennäköisyys on

$$\lambda_{j_i}\left(1 + O\left(\rho(n) - \frac{1}{n\sqrt{\rho(n)}}\right)\right) = \lambda_{j_i}\left(1 + O(\rho(n))\right).$$

Nyt koska $\sum_{h=1}^{i-1} j_h + 1$ solmua on paljastettu kuuluvan ekvivalenssiluokkiin $[f(x_1)], \ldots, [f(x_{i-1})]$ ja solmut $f(x_{i+1}), \ldots, f(x_s)$ tullaan paljastamaan muista ekvivalenssiluokista, erilaisia tapoja sijoittaa jäljellä olevat solmut luokkaan $[f(x_i)]$ on

$$\binom{m - J_{i-1}}{j_i - 1}$$

kappaletta. Huomataan kuitenkin, että b_i kappaletta joukon $f([x_i] \cap S)$ solmuja täytyy sijoittaa luokkaan $[f(x_i)]$, joten jäljelle jää

$$\binom{m - J_{i-1} - b_i + 1}{j_i - b_i}$$

tapaa valita jäljellä olevat solmut. Tämän vuoksi todennäköisyys, että $j_i - b_i$ muuta solmua on paljastettu kuuluvan luokkaan $[f(x_i)]$, on rajoituksen 4.1 ollessa voimassa

$$\frac{\binom{m - J_{i-1} - b_i + 1}{j_i - b_i}}{\binom{m - J_{i-1}}{j_i - 1}} = \frac{(j_i - 1)!}{(j_i - b_i)!}\frac{(m - J_{i-1} - b_i + 1)_{j_i - b_i}}{(m - J_{i-1})_{j_i - 1}} = \frac{(j_i - 1)!}{(j_i - b_i)!}\frac{1}{(m - J_{i-1})_{b_i - 1}}$$

$$= \frac{(j_i - 1)!}{m^{b_i - 1}(j_i - b_i)!\left(1 - O\left(\frac{J_i}{m}\right)\right)^{b_i - 1}} = \frac{(j_i - 1)! \cdot \left(1 + o\left(\frac{1}{n^{\frac{3}{4}}\sqrt{\rho(n)}}\right)\right)^{b_i - 1}}{m^{b_i - 1}(j_i - b_i)!}.$$

Yllä on käytetty merkintää $\dfrac{n!}{(n-k)!} = (n)_k$.

Todennäköisyys, että f säilyttää ekvivalenssirelaation $\equiv'$ joukosta S joukkoon R, on siis

$$\begin{aligned}
&\prod_{i=1}^{s} \sum_{j=b_i}^{\infty} \frac{\lambda_j (j-1)! \left(1 + O(\rho(n)) + o\left(\frac{1}{n^{\frac{3}{4}}\sqrt{\rho(n)}}\right)\right)^{b_i - 1}}{m^{b_i-1}(j-b_i)!} \\
&= \left(\prod_{i=1}^{s} \sum_{j=b_i}^{\infty} \frac{\lambda_j (j-1)!}{(j-b_i)!}\right) \prod_{i=1}^{s} \frac{\left(1 + O(\rho(n)) + o\left(\frac{1}{n^{\frac{3}{4}}\sqrt{\rho(n)}}\right)\right)^{b_i - 1}}{m^{b_i-1}} \\
&= \left(\prod_{i=1}^{s} \sum_{j=b_i}^{\infty} \frac{\lambda_j (j-1)!}{(j-b_i)!}\right) \left(\frac{1 + O(\rho(n)) + o\left(\frac{1}{n^{\frac{3}{4}}\sqrt{\rho(n)}}\right)}{m}\right)^{\sum_{i=1}^{s} b_i - 1} \\
&= \left(\prod_{i=1}^{s} \sum_{j=b_i}^{\infty} \frac{\lambda_j (j-1)!}{(j-b_i)!}\right) \left(\frac{1 + O(\rho(n))}{m}\right)^{|S|-s} \\
&= \left(\prod_{i=1}^{s} \sum_{j=b_i}^{\infty} \frac{\lambda_j (j-1)!}{(j-b_i)!}\right) \left(\frac{1 + O(\rho(n))}{m}\right)^{|S|} \cdot \frac{m^s}{\left(1 + O(\rho(n)) + o\left(\frac{1}{n^{\frac{3}{4}}\sqrt{\rho(n)}}\right)\right)^{s}} \\
&= \left(\prod_{i=1}^{s} \sum_{j=b_i}^{\infty} \frac{\lambda_j (j-1)!}{(j-b_i)!}\right) \cdot \left(\frac{1 + O(\rho(n))}{m}\right)^{|S|} \cdot m^s \\
&= \left(\prod_{i=1}^{s} \sum_{j=b_i}^{\infty} \frac{\lambda_j (j-1)!}{(j-b_i)!}\right) \cdot \frac{1 + O\left(\sqrt{\rho(n)}\right)}{m^{|S|-s}}.
\end{aligned}$$

Edelleen todennäköisyys, että f säilyttää relaation M' joukosta S joukkoon R on

$$\frac{1}{\prod_{j=1}^{C}(m - 2j + 1)} = \frac{1}{m^C \left(1 - O(\frac{C}{m})\right)^C}.$$

Koska $C \leq \dfrac{|S|}{2}$, on tämä todennäköisyys rajoituksen 4.1 ollessa voimassa

$$\frac{1}{m^C \left(1 - O\left(\frac{1}{n\sqrt{\rho(n)}}\right)\right)^{O\left(\frac{1}{\sqrt{\rho(n)}}\right)}} = \frac{1 + O(\rho(n))}{m^C},$$

joten apulause on todistettu. $\square$

6.3 Puuprosessien ehdollisista todennäköisyyksistä

Tutkitaan nyt todennäköisyyttä, että $\mathcal{W} \xrightarrow{f} \mathcal{N}$, ehdolla $\mathcal{S} \overset{f}{\cong} \mathcal{R}$. Sitä varten paljastetaan ekvivalenssiluokat $[f(x_i)]$ edellä kuvaillussa järjestyksessä. Näytetään siis, aloittaen sukupolvesta 0, että kun $1 \leq i \leq s$, luokille $[f(x_i)]$ pätee, että kaikilla $j \geq b_i - c_i$, ehdolla $\mathcal{S} \overset{f}{\cong} \mathcal{R}$, todennäköisyys, että luokalla $[f(x_i)]$ on j jälkeläistä joukossa $\mathcal{V}_{f(x_i)}/\equiv$ on likimäärin λ_{j+c_i}. Toisaalta joukon $\mathcal{V}_{f(x_i)}/\equiv$ myöhemmässä sukupolvessa todennäköisyys, että luokalla $[f(y)]$ on j jälkeläistä, on likimäärin λ_{j+1}.

Tämän vuoksi puuprosessilla $\mathcal{P}$ on s yksilöä sukupolvessa 0, ja nämä yksilöt ovat puiden $\mathcal{P}_1, \ldots, \mathcal{P}_s$ juuret. Puun $\mathcal{P}_i$ juurella on haarautumistodennäköisyys $p_i(j) = \lambda_{j+c_i}$ ja myöhempien sukupolvien yksilöillä on haarautumistodennäköisyydet $p(j) = \lambda_{j+1}$. Tämä eroaa hieman tavallisen haarautumisprosessin tapauksesta, koska sukupolven 0 yksilöiden haarautumistodennäköisyydet eroavat nyt muista yksilöistä. Nyt sukupolven $t > 0$ jälkeläisten lukumäärän odotusarvo on

$$\mu = \sum_{i=1}^{s} \sum_{j=1}^{\infty} j p_i(j) \left(\sum_{j=1}^{\infty} j p(j) \right)^{t-1} . \tag{6.1}$$

Jos $\mathcal{W} \xrightarrow{f} \mathcal{N}$, niin kaikkien kokoonpanon $\mathcal{W}$ solmujen maksimisyvyys on r. Tällainen solmu kuuluu joukon $\mathcal{W}/\equiv'$ sellaiseen $\equiv'$-ekvivalenssiluokkaan, jonka syvyys on $\lfloor r/2 \rfloor$ tai $\lceil r/2 \rceil$. Voidaan olettaa, että r on parillinen, sillä kaikki joukon $\mathcal{W}/\equiv'$ ekvivalenssiluokat voidaan saavuttaa $r/2$ siirtymällä, kun siirtymiä ei käytetä ekvivalenssiluokan sisällä liikkumiseen. Tällöin kuvauksen f tarkastelussa voidaan keskittyä kyseisen puuprosessin ensimmäisen $r/2$ sukupolven arvioimiseen. Sovitaan, että merkintä $\mathcal{W} \to \mathcal{P} \restriction \frac{r}{2}$ tarkoittaa sellaista puuprosessiin liittyvää tapahtumaa, että $\mathcal{W}_{x_i}/\equiv'$ on isomorfinen puun $\mathcal{P}_i \restriction \frac{r}{2}$ kanssa, kun $1 \leq i \leq s$. Merkitään lisäksi todennäköisyyttä, että satunnaisesta n-kokoonpanosta puhuttaessa tapahtuma A toteutuu ehdolla B, $\mathbb{P}_c(A \mid B, n)$.

Apulause 6.8. Jos $|W| \leq \dfrac{1}{\sqrt{\rho(n)}}$, niin

$$\mathbb{P}_c\left(\mathcal{W} \xrightarrow{f} \mathcal{N} \mid \mathcal{S} \overset{f}{\cong} \mathcal{R}, n\right) = \mathbb{P}\left(\mathcal{W} \to \mathcal{P} \restriction \frac{r}{2}\right)\left(1 + O(\sqrt{\rho(n)})\right).$$

Todistus. Paljastetaan jälleen yksitellen ekvivalenssiluokat $[f(x_i)]$ käyttäen edellä kuvailtua järjestystä. Aloitetaan tarkastelu sukupolvesta 0 ja oletetaan, että $1 \leq i \leq s$ ja että $|[f(x_i)]| = a_i$. Selvästi $f([x_i] \cap S) \subseteq [f(x_i)]$. Tarkastellaan tapahtumaa, että $|[f(x_i)]| = a_i$ ja että kaikki solmut $y \in [f(x_i)]$, jotka eivät ole paritettu millekään solmulle $z \in R$ relaatiolla M, ovat paritettu jollekin solmulle $z \notin \bigcup_{j=1}^{i} [f(x_j)]$. Käyttäen samaa ideaa, kuin apulauseen 6.7 todistuksessa, havaitsemalla, että todennäköisyys, että $|[f(x_h)]| = a_i$, on lukujen

$$\frac{d_{a_i}(n) - i + 1}{n - i + 1} \quad \text{ja} \quad \frac{d_{a_i}(n)}{n - i + 1}$$

välillä, voidaan todeta, että kyseessä olevan tapahtuman todennäköisyys on $\lambda_{a_i}(1 + O(\rho(n)))$.

Koska luokalla $[x_i]$ on $a_i - c_i$ jälkeläistä joukossa $\mathcal{W}/\equiv'$, todennäköisyys, että f on laajennettu luokkaan $[x_i]$ siten, että se on konsistentti ehdon $\mathcal{W} \xrightarrow{f} \mathcal{N}$ määritelmän kanssa, on

$$p_i(a_i - c_i)(1 + O(\rho(n))).$$

Jatketaan tarkastelua joukon $\mathcal{W}/\equiv'$ myöhempiin sukupolviin ja oletetaan, että $i > s$. Tarkastellaan tapahtumaa, että $|[f(x_i)]| = a_i$ sekä että kaikki solmut $y \in [f(x_i)]$, jotka eivät ole paritettu millekään solmulle $z \in \bigcup_{j=1}^{i-1}[f(x_j)]$ relaatiolla M, ovat paritettu jollekin solmulle $z \notin \bigcup_{j=1}^{i}[f(x_j)]$. Samaan tapaan kuin edellä, voidaan todeta, että tämän tapahtuman todennäköisyys on

$$\lambda_{a_i}(1 + O(\rho(n)))$$

ja todennäköisyys, että f on laajennettu luokkaan $[x_i]$ siten, että se on konsistentti ehdon $\mathcal{W} \xrightarrow{f} \mathcal{N}$ määritelmän kanssa, on

$$p_i(a_i - 1)(1 + O(\rho(n))).$$

Kun koko kuvaus f on paljastettu, saadaan tapahtuman $\mathcal{W} \xrightarrow{f} \mathcal{N}$ ehdolla $\mathcal{S} \overset{f}{\cong} \mathcal{R}$ todennäköisyydeksi

$$\begin{aligned}
&\prod_{i=1}^{s} p_i(a_i - c_i)(1 + O(\rho(n))) \cdot \prod_{i=s+1}^{w} p(a_i - 1)(1 + O(\rho(n))) \\
= &\prod_{i=1}^{s} p_i(a_i - c_i) \cdot \prod_{i=s+1}^{w} p(a_i - 1) \cdot (1 + O(\rho(n)))^w \\
= &\mathbb{P}\left(\mathcal{W} \to \mathcal{P} \restriction \frac{r}{2}\right)\left(1 + O\left(\sqrt{\rho(n)}\right)\right).
\end{aligned}$$

$\square$

Määritelmä 6.9. Olkoon $(S, \equiv, M)$ osittainen kokoonpano. Joukon S permutaation φ sanotaan olevan kokoonpanon $(S, \equiv, M)$ *automorfismi*, mikäli kaikille $x, y \in S$ pätee

$$x \equiv y, \text{ jos ja vain jos } \varphi(x) \equiv \varphi(y),$$

sekä

$$x\ M\ y, \text{ jos ja vain jos } \varphi(x)\ M\ \varphi(y).$$

.

Määritelmä 6.10. Olkoot kuvaukset $f_i\colon S \to R$, $i = 1, 2$, bijektioita. Sanotaan, että kuvaukset f_1 ja f_2 ovat ekvivalentit, mikäli $f_1 \circ f_2^{-1}$ on automorfismi.

Seuraus 6.11. *On olemassa osittaisesta kokoonpanosta $\mathcal{S}$ riippuva vakio α, jolle pätee*

$$\mathbb{P}_c(\mathcal{N}(R, r) \cong \mathcal{W}, n) = \alpha\left(\prod_{i=1}^{s}\sum_{j=b_i}^{\infty}\frac{(j-1)!\lambda_j}{(j-b_j)!}\right) \cdot \frac{1 + O\left(\sqrt{\rho(n)}\right)}{m^{|S|+C-s}} \cdot \mathbb{P}\left(\mathcal{W} \to \mathcal{P} \restriction \frac{r}{2}\right).$$

Todistus. Olkoon β osittaisen kokoonpanon $\mathcal{S}$ automorfismien lukumäärä ja α bijektioiden $f\colon S \to R$, ekvivalenssiluokkien lukumäärä. Tällöin $\alpha = \frac{s!}{\beta}$ ja mille tahansa bijektiolle $f\colon S \to R$ pätee

$$\mathbb{P}_c\left(\mathcal{N}(R, r) \cong \mathcal{W}, n\right) = \alpha \mathbb{P}_c(\mathcal{S} \overset{f}{\cong} \mathcal{R}, n) \cdot \mathbb{P}_c(\mathcal{W} \xrightarrow{f} \mathcal{N} \mid \mathcal{S} \overset{f}{\cong} \mathcal{R}, n),$$

jolloin apulauseista 6.7 ja 6.8 seuraa haluttu yhtälö. $\square$

6.4 Satunnaiskokoonpanon asymptoottinen käyttäytyminen

Seuraavaksi näytetään, että lähes kaikki n-kokoonpanot ovat k-rikkaita sekä k-yksinkertaisia, kun n on tarpeeksi suuri. Tulosten todistamiseen tarvitaan vielä joitakin todennäköisyysteorian aputuloksia:

Apulause 6.12. (Markovin epäyhtälö) Olkoon $X : \Omega \to \mathbb{R}$ epänegatiivinen diskreetti satunnaismuuttuja, jolle $\mathbb{E}(X)$ on olemassa. Nyt kaikille $a > 0$ pätee

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

Todistus. Olkoon $A = \{\omega \in \Omega \mid X(\omega) \geq a\}$. Nyt satunnaismuuttujan X odotusarvolle pätee

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(X = X(\omega)) = \sum_{\omega \in A} X(\omega)\mathbb{P}(X = X(\omega)) + \sum_{\omega \notin A} X(\omega)\mathbb{P}(X = X(\omega)).$$

Satunnaismuuttujan X epänegatiivisuuden nojalla $\sum_{\omega \notin A} X(\omega)\mathbb{P}(X = X(\omega)) \geq 0$, mistä seuraa

$$\mathbb{E}(X) \geq \sum_{\omega \in A} X(\omega)\mathbb{P}(X = X(\omega)) \geq \sum_{\omega \in A} \mathbb{P}(X = X(\omega)) \cdot a = a\mathbb{P}(X \geq a),$$

ja väite on todistettu. □

Apulause 6.13. (Tšebyšovin epäyhtälö) Olkoon $a > 0$ ja X diskreetti satunnaismuuttuja, jolle $\mu = \mathbb{E}(X)$ sekä $\sigma^2 = \mathrm{Var}(X)$ ovat olemassa. Tällöin

$$\mathbb{P}(|X - \mu| \geq a) \leq \frac{\sigma^2}{a^2}.$$

Todistus. Markovin epäyhtälön nojalla

$$\mathbb{P}(|X - \mu| \geq a) = \mathbb{P}(|X - \mu|^2 \geq a^2) \leq \frac{\mathbb{E}(X - \mu)^2}{a^2} = \frac{\sigma^2}{a^2}.$$

□

Apulause 6.14. Olkoon $\mathcal{U}$ satunnainen n-kokoonpano. Tällöin $\lim_{n \to \infty} \mathbb{P}_c(\mathcal{U} \text{ on } k\text{-rikas}, n) = 1$.

Todistus. Oletetaan, että $r = 3^k$. Nyt apulauseen 5.32 nojalla riittää osoittaa, että todennäköisyydellä, joka lähestyy lukua 1, mille tahansa r-syvyiselle puulle $\mathcal{T}$ on olemassa vähintään k kappaletta erillisiä r-säteisiä ympäristöjä rakenteessa $\mathcal{U}$, jotka ovat relaatiossa $\sim_k$ puun $\mathcal{T}$ kanssa. Koska on olemassa vain äärellinen määrä korkeintaan r-syvyisten puiden $\sim_k$-ekvivalenssiluokkia, voidaan kiinnittää puu $\mathcal{T}$ ja näyttää, että asymptoottisesti lähes varmasti on olemassa vähintään k erillistä r-säteistä ympäristöä, jotka ovat isomorfisia puun $\mathcal{T}$ kanssa. Tällöin erityisesti jokaiselle korkeintaan $k - 1$-alkioiselle joukolle $S \subseteq U$ voidaan löytää määritelmän 5.38 mukainen solmu $x \in U$, joka on tarpeeksi etäällä joukosta S.

Olkoon $R \subseteq U$, $|R| \leq \dfrac{1}{\sqrt[4]{\rho(n)}}$ sekä $\mathcal{N} = \mathcal{N}(R, r)$. Olkoon lisäksi f joukon R identiteettikuvaus ja $\mathbf{A}$ niiden keskitettyjen kokoonpanojen $\mathcal{W} = (W, \equiv', M', R)$ joukko, joiden syvyys on korkeintaan r sekä $|W| \leq \dfrac{1}{\rho(n)}$, $\equiv' \restriction R = \emptyset$, $M' \restriction R = \emptyset$ ja joilla on kanoninen hajotelma

$\{(R,\emptyset,\emptyset),\mathcal{W}_x \mid x \in R\}$. Tällä määrittelyllä saadaan ”eristettyä” alkuperäisestä kokoonpanosta erillisiä puita, joita seuraavaksi tarkastellaan. Olkoon $\mathbf{B} \subset \mathbf{A}$ sellainen joukko, joka koostuu niistä keskitetyistä kokoonpanoista $\mathcal{W}$, joille pätee $\mathcal{W}_x \cong \mathcal{T}$ vähintään k solmulle $x \in R$.

Todennäköisyys, että on olemassa vähintään k erillistä puun $\mathcal{T}$ kanssa isomorfista ympäristöä, on vähintään

$$
\begin{aligned}
&\mathbb{P}_c\left((\exists x_1,\dots,x_k \in R)\left(\bigwedge_{i\neq j} N(x_i,r)\cap N(x_j,r)=\emptyset \wedge \bigwedge_{i\leq i\leq k}\mathcal{N}(x_i,r)\cong\mathcal{T}\right),n\right)\\
&\geq \mathbb{P}_c\left((\exists\mathcal{W}\in\mathbf{B})(\mathcal{W}\xrightarrow{f}\mathcal{N}),n\right)\\
&= \mathbb{P}_c\left((\exists\mathcal{W}\in\mathbf{A})(\mathcal{W}\xrightarrow{f}\mathcal{N}),n\right)-\mathbb{P}_c\left((\exists\mathcal{W}\in\mathbf{A}\setminus\mathbf{B})(\mathcal{W}\xrightarrow{f}\mathcal{N}),n\right).
\end{aligned}
$$

Nyt apulauseen 6.8 nojalla pätee

$$
\begin{aligned}
\mathbb{P}_c\left((\exists\mathcal{W}\in\mathbf{A})(\mathcal{W}\xrightarrow{f}\mathcal{N}),n\right) &= \mathbb{P}\left((\exists\mathcal{W}\in\mathbf{A})\left(\mathcal{W}\to\mathcal{P}\restriction\frac{r}{2}\right)\right)\left(1+O(\sqrt{\rho(n)})\right)\\
&\geq \mathbb{P}\left(\left|\mathcal{P}\restriction\frac{r}{2}\right|\leq\frac{1}{2\sqrt{\rho(n)}}\right)\left(1+O(\sqrt{\rho(n)})\right),
\end{aligned}
$$

sillä jokaista sellaista puuta $\mathcal{P}$ kohden, jolla on t yksilöä ensimmäisissä $r/2$ sukupolvessaan, on olemassa sellainen keskitetty kokoonpano $\mathcal{W}$, että $|W| \leq 2t$ ja $\mathcal{W}\to\mathcal{P}\restriction\frac{r}{2}$. Merkitään määritelmän 4.5 mukaisesti $\sum_{i=0}^{\infty} i\lambda_i = \Lambda$. Nyt yhtälön 6.1 nojalla

$$
\begin{aligned}
\mathbb{E}\left(\left|\mathcal{P}\restriction\frac{r}{2}\right|\right) &\leq \frac{r}{2}|R|\sum_{j=1}^{\infty} j\lambda_j\left(\sum_{j=1}^{\infty} j+\lambda_{j+1}\right)^{\frac{r}{2}-1}\\
&< \frac{r}{2}|R|\Lambda^{\frac{r}{2}}\\
&= O\left(\frac{1}{\sqrt[4]{\rho(n)}}\right).
\end{aligned}
$$

Nyt Markovin epäyhtälön nojalla

$$
\mathbb{P}\left(\left|\mathcal{P}\restriction\frac{r}{2}\right|>\frac{1}{2\sqrt{\rho(n)}}\right)=O\left(\sqrt[4]{\rho(n)}\right),
$$

joten

$$
\mathbb{P}_c\left((\exists\mathcal{W}\in\mathbf{A})(\mathcal{W}\xrightarrow{f}\mathcal{N}),n\right)=1-O\left(\sqrt[4]{\rho(n)}\right).
$$

Käyttäen uudelleen apulausetta 6.8, saadaan

$$
\begin{aligned}
&\mathbb{P}_c\left((\exists\mathcal{W}\in\mathbf{A}\setminus\mathbf{B})(\mathcal{W}\xrightarrow{f}\mathcal{N}),n\right)\\
&\leq \mathbb{P}\left(\left|\left\{x\in R \mid \mathcal{T}\to\mathcal{P}_x\restriction\frac{r}{2}\right\}\right|<k\right)\left(1+O(\sqrt{\rho(n)})\right).
\end{aligned}
$$

Nyt induktiolla luvun r suhteen ja astejonojen tasaisuutta soveltaen voidaan osoittaa, että $\mathbb{P}\left(\mathcal{T} \to \mathcal{P}_x \restriction \frac{r}{2}\right) > 0$. Koska tapahtumat $\{\mathcal{T} \to \mathcal{P}_x \restriction \frac{r}{2}\}$ ovat riippumattomia toisistaan kaikilla $x \in R$, saadaan

$$\mathbb{E}\left(\left|\left\{x \in R \mid \mathcal{T} \to \mathcal{P}_x \restriction \frac{r}{2}\right\}\right|\right) = \Theta\left(\frac{1}{\sqrt[4]{\rho(n)}}\right),$$

ja

$$\mathrm{Var}\left(\left|\left\{x \in R \mid \mathcal{T} \to \mathcal{P}_x \restriction \frac{r}{2}\right\}\right|\right) = \Theta\left(\frac{1}{\sqrt[4]{\rho(n)}}\right).$$

Tällöin Tšebyšovin epäyhtälön nojalla

$$\mathbb{P}\left(\left|\left\{x \in R \mid \mathcal{T} \to \mathcal{P}_x \restriction \frac{r}{2}\right\}\right| < k\right) = O\left(\sqrt[4]{\rho(n)}\right),$$

sekä

$$\mathbb{P}_c\left((\exists \mathcal{W} \in \mathbf{A} \setminus \mathbf{B})((\mathcal{W} \xrightarrow{f} \mathcal{N}), n\right) = O\left(\sqrt[4]{\rho(n)}\right),$$

mistä seuraa haluttu tulos. □

Apulause 6.15. Olkoon $\mathcal{U}$ mielivaltainen n-kokoonpano. Tällöin

$$\lim_{n\to\infty} \mathbb{P}_c(\mathcal{U} \text{ on } k\text{-yksinkertainen}) = 1.$$

Todistus. Riittää osoittaa, että kaikille kiinnitetyille luvuille s ja r, pätee

$$\lim_{n\to\infty} \mathbb{P}_c\left(\exists R(R \text{ on sykli, jonka pituus on vähintään } 2s \wedge \mathcal{N}(R,r) \text{ ei ole yksisyklinen}), n\right) = 0.$$

Olkoon $\mathbf{A}$ niiden keskitettyjen kokoonpanojen $\mathcal{W} = (W, \equiv', M', S)$ joukko, joiden syvyys on korkeintaan r ja joille pätee $|W| \leq \frac{1}{\sqrt{\rho(n)}}$, ja missä S on jokin sykli, jonka pituus on $2s$. Tarkastellaan bijektiota $f\colon S \to U$ ja määritellään $R = f(S)$, $\mathcal{S} = (S, \equiv', M')$, $\mathcal{R} = (R, \equiv, M)$ ja $\mathcal{N} = \mathcal{N}(R,r)$. Lisäksi olkoon Λ kuten määritelmässä 4.5. Nyt

$$\begin{aligned}
&\mathbb{P}_c\left(\mathcal{S} \overset{f}{\cong} \mathcal{R} \wedge \mathcal{N} \text{ ei ole yksisyklinen}, n\right) \\
&\leq \mathbb{P}_c\left(\mathcal{S} \overset{f}{\cong} \mathcal{R}, n\right) - \mathbb{P}_c\left(\mathcal{S} \overset{f}{\cong} \mathcal{R} \wedge (\exists \mathcal{W} \in \mathbf{A})(\mathcal{W} \xrightarrow{f} \mathcal{N}), n\right).
\end{aligned}$$

Nyt

$$\begin{aligned}
&\mathbb{P}_c\left(\mathcal{S} \overset{f}{\cong} \mathcal{R} \wedge (\exists \mathcal{W} \in \mathbf{A})(\mathcal{W} \xrightarrow{f} \mathcal{N}), n\right) \\
&= \mathbb{P}_c\left(\mathcal{S} \overset{f}{\cong} \mathcal{R}, n\right) \cdot \mathbb{P}_c\left((\exists \mathcal{W} \in \mathbf{A})(\mathcal{W} \xrightarrow{f} \mathcal{N}) \,\middle|\, \mathcal{S} \overset{f}{\cong} \mathcal{R}, n\right).
\end{aligned}$$

Apulauseen 6.7 nojalla

$$\begin{aligned}
\mathbb{P}_c\left(\mathcal{S} \overset{f}{\cong} \mathcal{R}, n\right) &= \left(\sum_{j=2}^{\infty} (j-1)\lambda_j\right)^s \cdot \frac{1 + O\left(\sqrt{\rho(n)}\right)}{m^{2s}} \\
&= \frac{(\Lambda - 1)^s + O\left(\sqrt{\rho(n)}\right)}{m^{2s}}.
\end{aligned}$$

Apulauseen 6.8 nojalla ja käyttäen samaa menetelmää, kuin apulauseen 6.14 todistuksessa, saadaan

$$\begin{aligned}
&\mathbb{P}_c\left((\exists \mathcal{W} \in \mathbf{A})(\mathcal{W} \xrightarrow{f} \mathcal{N}) \,\middle|\, \mathcal{S} \overset{f}{\cong} \mathcal{R}, n\right) \\
&= \mathbb{P}\left((\exists \mathcal{W} \in \mathbf{A})(\mathcal{W} \to \mathcal{P} \restriction \frac{r}{2}\right)\left(1 + O\left(\sqrt{\rho(n)}\right)\right) \\
&= 1 - O\left(\sqrt[4]{\rho(n)}\right).
\end{aligned}$$

Tällöin

$$\mathbb{P}_c\left(\mathcal{S} \overset{f}{\cong} \mathcal{R} \wedge \mathcal{N} \text{ ei ole yksisyklinen}, n\right) = \frac{O\left(\sqrt[4]{\rho(n)}\right)}{m^{2s}},$$

ja koska meillä on vain m^{2s} tapaa valita f, saadaan

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{P}_c(\exists R(R \text{ on sykli, jonka pituus on } 2s \wedge \mathcal{N}(R, r) \text{ ei ole yksisyklinen}),\ n) \\
&= O\left(\sqrt[4]{\rho(n)}\right).
\end{aligned}$$

□

Apulause 6.16. Olkoon $r \geq 0$ ja τ jokin sellaisten korkeintaan r-syvyisten osittaisten kokoonpanojen $\sim_k$-ekvivalenssiluokka, joilla on vain yksi sykli. Tällöin on olemassa sellainen vakio γ, että

$$\lim_{n\to\infty} \mathbb{E}(|\{R \subseteq U \mid \mathcal{N}(R, r) \in \tau\}|) = \gamma$$

Todistus. Olkoon $2s$ luokan τ jäsenten syklien pituus ja $\mathbf{A}$ niiden korkeintaan r-syvyisten kokoonpanojen $\mathcal{W} = (W, \equiv', M', S)$ joukko, joille pätee $|W| \leq \frac{1}{\rho(n)}$, ja missä S on jokin sykli, jonka pituus on $2s$. Olkoon lisäksi $\mathbf{B}$ niiden luokan τ kokoonpanojen joukko, joiden juurijoukko on S. Tällöin joukolle $R \subseteq U$, jolle pätee $|R| = 2s$, on voimassa

$$\begin{aligned}
&\mathbb{P}_c(\mathcal{N}(R, r) \in \tau, n) \\
&= \mathbb{P}_c((\exists \mathcal{W} \in \mathbf{A} \cap \mathbf{B})(\mathcal{N}(R, r) \cong \mathcal{W}), n) + \mathbb{P}_c((\exists \mathcal{W} \in \mathbf{B} \setminus \mathbf{A})(\mathcal{N}(R, r) \cong \mathcal{W}), n).
\end{aligned}$$

Nyt seurauksen 6.11 nojalla

$$\begin{aligned}
&\mathbb{P}_c((\exists \mathcal{W} \in \mathbf{A} \cap \mathbf{B})(\mathcal{N}(R, r) \cong \mathcal{W}), n) \\
&= (2s-1)!(\Lambda - 1)^s \cdot \frac{1 + O\left(\sqrt{\rho(n)}\right)}{m^{2s}} \cdot \mathbb{P}\left((\exists \mathcal{W} \in \mathbf{A} \cap \mathbf{B})\left(\mathcal{W} \to \mathcal{P} \restriction \frac{r}{2}\right)\right),
\end{aligned}$$

sekä

$$\begin{aligned}
&\mathbb{P}\left((\exists \mathcal{W} \in \mathbf{A} \cap \mathbf{B})\left(\mathcal{W} \to \mathcal{P} \restriction \frac{r}{2}\right)\right) \\
&= \mathbb{P}\left((\exists \mathcal{W} \in \mathbf{B})\left(\mathcal{W} \to \mathcal{P} \restriction \frac{r}{2}\right)\right) - \mathbb{P}\left((\exists \mathcal{W} \in \mathbf{B} \setminus \mathbf{A})\left(\mathcal{W} \to \mathcal{P} \restriction \frac{r}{2}\right)\right),
\end{aligned}$$

ja

$$\mathbb{P}\left((\exists \mathcal{W} \in \mathbf{B} \setminus \mathbf{A})\left(\mathcal{W} \to \mathcal{P} \restriction \frac{r}{2}\right)\right) = O\left(\sqrt[4]{\rho(n)}\right).$$

Käyttäen samaa menetelmää kuin apulauseen 6.14 todistuksessa, saadaan

$$\begin{aligned}&\mathbb{P}((\exists \mathcal{W} \in \mathbf{B} \setminus \mathbf{A})(\mathcal{N}(R,r) \cong \mathcal{W}), n)\\ &\leq \mathbb{P}_c(\mathcal{S} \cong \mathcal{R}, n) - \mathbb{P}_c((\exists \mathcal{W} \in \mathbf{A})(\mathcal{N}(R,r) \cong \mathcal{W}), n).\end{aligned}$$

Nyt soveltamalla seurausta 6.11, saadaan

$$\mathbb{P}_c(\mathcal{S} \cong \mathcal{R}, n) = (2s-1)!(\Lambda - 1)^s \cdot \frac{1 + O\left(\sqrt{\rho(n)}\right)}{m^{2s}},$$

sekä

$$\begin{aligned}&\mathbb{P}_c((\exists \mathcal{W} \in \mathbf{A})(\mathcal{N}(R,r) \cong \mathcal{W}), n)\\ &= (2s-1)!(\Lambda - 1)^s \cdot \frac{1 + O\left(\sqrt{\rho(n)}\right)}{m^{2s}} \cdot \mathbb{P}\left((\exists \mathcal{W} \in \mathbf{A})\left(\mathcal{W} \to \mathcal{P} \upharpoonright \frac{r}{2}\right)\right).\end{aligned}$$

Lisäksi

$$\mathbb{P}\left((\exists \mathcal{W} \in \mathbf{A})\left(\mathcal{W} \to \mathcal{P} \upharpoonright \frac{r}{2}\right)\right) = 1 - \frac{O\left(\sqrt[4]{\rho(n)}\right)}{m^{2s}},$$

joten saadaan

$$\mathbb{P}_c(\mathcal{N}(R,r) \in \tau, n) = (2s-1)!(\Lambda - 1)^s \cdot \frac{1 + O\left(\sqrt{\rho(n)}\right)}{m^{2s}} \cdot \mathbb{P}\left((\exists \mathcal{W} \in \mathbf{B})\left(\mathcal{W} \to \mathcal{P} \upharpoonright \frac{r}{2}\right)\right).$$

Koska meillä on vain $\binom{m}{2s}$ tapaa valita R, voidaan valita

$$\gamma = \frac{(\Lambda - 1)^s \cdot \mathbb{P}\left((\exists \mathcal{W} \in \mathbf{B})\left(\mathcal{W} \to \mathcal{P} \upharpoonright \frac{r}{2}\right)\right)}{2s},$$

jolloin apulause on todistettu. □

6.5 Suppenemislaki satunnaiskokoonpanolle

Tutkielman päätuloksen todistamiseksi tarvitaan vielä seuraavat apulauseet:

Apulause 6.17. (Boolen epäyhtälö) Olkoot A_i, missä $i = 1, 2, \ldots$, tapahtumia ja $\mathbb{P}(A_i)$ on tapahtuman A_i todennäköisyys. Tällöin

$$\mathbb{P}\Big(\bigcup_{i \geq 1} A_i\Big) \leq \sum_{i \geq 1} \mathbb{P}(A_i).$$

Todistus. Todistetaan väite induktiolla luvun $n \in \mathbb{N}$ suhteen.

Jos $n = 1$, niin $\mathbb{P}(A_1) \leq \mathbb{P}(A_1)$.

Tehdään induktio-oletus, että väite pätee kun $n = k$, eli

$$\mathbb{P}\Big(\bigcup_{i=1}^{k} A_i\Big) \leq \sum_{i=1}^{k} \mathbb{P}(A_i).$$

Nyt summa- ja erotusperiaatteen nojalla $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$, joten käyttäen hyväksi yhdisteen liitännäisyyttä saadaan

$$\mathbb{P}\Big(\bigcup_{i=1}^{k+1} A_i\Big) = \mathbb{P}\Big(\bigcup_{i=1}^{k} A_i\Big) + \mathbb{P}(A_{k+1}) - \mathbb{P}\Big(A_{k+1} \cap \bigcup_{i=1}^{k} A_i\Big).$$

Nyt koska

$$\mathbb{P}\Big(A_{k+1} \cap \bigcup_{i=1}^{k} A_i\Big) \geq 0,$$

saadaan todennäköisyyden epänegatiivisuuden nojalla

$$\mathbb{P}\Big(\bigcup_{i=1}^{k+1} A_i\Big) \leq \mathbb{P}\Big(\bigcup_{i=1}^{k} A_i\Big) + \mathbb{P}(A_{k+1}),$$

mistä seuraa

$$\mathbb{P}\Big(\bigcup_{i=1}^{k+1} A_i\Big) \leq \sum_{i=1}^{k} (\mathbb{P}(A_i) + \mathbb{P}(A_{k+1}) = \sum_{i=1}^{k+1} \mathbb{P}(A_i),$$

ja induktioperiaatteen nojalla väite on todistettu.

□

Boolen epäyhtälö yleistyy tunnetuiksi Bonferronin epäyhtälöiksi, joihin lukija voi perehtyä lisää artikkelissa [5].

Apulause 6.18. (Summa- ja erotusperiaate, [7]) Olkoon I äärellinen joukko ja olkoon A_i äärellinen joukko kaikilla $i \in I$. Tällöin joukko $\bigcup_{i \in I} A_i$ on äärellinen ja

$$\Big|\bigcup_{i \in I} A_i\Big| = \sum_{\emptyset \neq J \subset I} (-1)^{|J|+1} \Big|\bigcap_{j \in J} A_j\Big|$$

Todistus. Merkitään $A = \bigcup_{i \in I} A_i$ sekä $I_x = \{i \in I \mid x \in A_i\}$ kaikilla $x \in A$. Huomataan, että $I_x \neq \emptyset$. Nyt nähdään

$$|A| = \sum_{x \in A} 1 = \sum_{x \in A} \Big(1 - (1 + (-1))^{|I_x|}\Big) = \sum_{x \in A} \sum_{\emptyset \neq K \subset I_x} (-1)^{|K|+1}.$$

Ryhmitellään viimeisessä summassa olevat termit uudelleen kokoamalla kaikilla $J \subset I$ yhteen ne termit, jotka vastaavat joukkoa J. Huomataan, että jokaisella $x \in A$ on voimassa $J \subset I_x$ täsmälleen silloin, kun $x \in \bigcap_{i \in J} A_i =: A_J$. Nyt saadaan

$$\sum_{x \in A} \sum_{\emptyset \neq K \subset I_x} (-1)^{|K|+1} = \sum_{\emptyset \neq J \subset I} \sum_{x \in A_J} (-1)^{|J|+1},$$

mistä seuraa haluttu lauseke, sillä

$$\sum_{\emptyset \neq J \subset I} \sum_{x \in A_J} (-1)^{|J|+1} = \sum_{\emptyset \neq J \subset I} (-1)^{|J|+1} \sum_{x \in A_j} 1 = \sum_{\emptyset \neq J \subset I} (-1)^{|J|+1} |A_J|.$$

□

Käytetään seuraavaksi summa- ja erotusperiaatetta sekä Bonferronin epäyhtälöitä ja viimeistellään todistus väitteelle, että jokaiselle satunnaiskokoonpanolle ja ensimmäisen kertaluvun logiikan lauseelle φ on olemassa raja-arvo $\lim_{n\to\infty} \mathbb{P}_c(\varphi, n)$. Tämä menetelmä on kuvailtu tarkemmin artikkelissa [9].

Olkoon seuraavassa tarkastelussa τ_i, $i \in I$, indeksöinti kaikille sellaisille yksisyklisten osittaisten kokoonpanojen $\sim_k$-ekvivalenssiluokille, joissa luokan τ_i jäsenten syvyys on $r_i \leq 3^{k-1}$ ja syklin pituus on $s_i \leq 2\cdot 3^{k-1}$. Olkoon lisäksi γ_i apulauseen 6.16 mukainen vakio, joka liittyy luokkaan τ_i. Tällöin minkä tahansa k-rikkaan ja k-yksinkertaisen kokoonpanon $\sim_k$-luokkaa kuvastaa vektori $(u_i)_{i\in I}$, missä $0 \leq u_i \leq k$ on niiden ympäristöjen lukumäärä, jotka kuuluvat luokkaan τ_i. Tällöin määritelmän 5.41 mukaan luokat τ_1 ja τ_2 ovat samat, mikäli niitä kuvastavat vektorit ovat samat. Lisäksi jos $u_i = k$, niin määritelmän 5.41 mukaan on olemassa vähintään k ympäristöä, jotka kuuluvat luokkaan τ_i. Olkoon nyt $\mathcal{U}$ n-kokoonpano, jolla on solmujoukko $U = \{1, \ldots, m\}$. Merkitään kaikille $i \in I$, $C_i(\mathcal{U}) = \{R \subseteq U \mid \mathcal{N}(R, 3^{k-1}) \in \tau_i\}$ ja $\overline{C(\mathcal{U})} = (C_i(\mathcal{U}))_{i\in I}$.

Olkoon $\overline{S} = (S_i)_{i\in I}$, missä $S_i \subseteq \{R \subseteq U \mid |R| = s_i\}$. Määritellään

$$E^{\geq}(\overline{S}) = \{\mathcal{U} \mid C_i(\mathcal{U}) \supseteq S_i \text{ kun } i \in I\},$$

sekä

$$E^{=}(\overline{S}) = \{\mathcal{U} \mid \overline{C(\mathcal{U})} = \overline{S}\}.$$

Toisin sanoen $E^{\geq}(\overline{S})$ on niiden n-kokoonpanojen joukko, joilla jokainen $R \in S_i$ on sykli jossakin τ_i-luokan 3^{k-1}-säteisessä ympäristössä, ja $E^{=}(\overline{S})$ on niiden n-kokoonpanojen joukko, joilla S_i on täsmälleen edellä mainittujen syklien joukko. Olkoon vektori $\overline{u} = (u_i)_{i\in I}$ ja

$$L(\overline{u}, n) = \sum_{\overline{S} \,\mid\, |S_i| = u_i} \mathbb{P}_c(E^{\geq}(\overline{S}), n),$$

sekä olkoon kaikille $J \subseteq I$

$$M(J, \overline{u}, n) = \bigcup_{\substack{\forall i\in J(\overline{S}\,:\,|S_i|=u_i),\\ \forall i\in I\setminus J(\overline{S}\,:\,|S_i|\geq u_i)}} E^{=}(\overline{S}).$$

Toisin sanoen $M(J, \overline{u}, n)$ on niiden n-kokoonpanojen joukko, joilla on täsmälleen u_i ympäristöä luokassa τ_i kullakin $i \in J$ ja vähintään u_i ympäristöä luokassa τ_i kullakin $i \in I \setminus J$.

Käytetään yllä esiteltyjä määritelmiä ja merkintöjä seuraavan lauseen todistamiseen:

Lause 6.19. *Olkoon φ ensimmäisen kertaluvun logiikan lause. Tällöin on olemassa raja-arvo $\lim_{n\to\infty} \mathbb{P}_c(\varphi, n)$.*

Todistus. Kokoonpanon paikalliset ominaisuudet riittävät määrittämään, onko jokin kokoonpanoa koskeva ensimmäisen kertaluvun logiikan lause totta. Toisaalta kokoonpanon paikalliset ominaisuudet määrittävät kokoonpanon $\sim_k$-ekvivalenssiluokan. Tällöin riittää osoittaa, että todennäköisyys, että kokoonpano kuuluu tiettyyn $\sim_k$-ekvivalenssiluokkaan, suppenee. Toisin sanoen riittää näyttää, että $\mathbb{P}_c(M(J, \overline{u}, n), n)$ suppenee, kun n kasvaa rajatta.

Merkitään mille tahansa I-ulotteiselle vektorille $\overline{v}$, että $\overline{v} \geq \overline{u}$, mikäli $v_i \geq u_i$ kaikilla $i \in I$. Nyt summa- ja erotusperiaatteen nojalla saadaan

$$\lim_{n\to\infty} \mathbb{P}_c(M(J, \overline{u}, n), n) = \sum_{\overline{v}\geq\overline{u}} (-1)^{\sum(v_i - u_i)} \left(\prod_{i\in J} \binom{v_i}{u_i}\right) \left(\prod_{i\in I\setminus J} \binom{v_i - 1}{u_i - 1}\right) \cdot \lim_{n\to\infty} L(\overline{v}, n),$$

sekä

$$\lim_{n\to\infty} L(\overline{v},n) = \prod_{i\in I} \frac{\gamma_i^{v_i}}{v_i!}.$$

Tästä seuraa

$$\lim_{n\to\infty} \mathbb{P}_c(M(J,\overline{u},n),n) = \prod_{i\in J}\sum_{v\geq u_i}(-1)^{v-u_i}\frac{\gamma_i^v}{u_i!(v-u_i)!}\cdot\prod_{i\in I\setminus J}\sum_{v\geq u_i}(-1)^{v-u_i}\frac{\gamma_i^v}{v(u_i-1)!(v-u_i)!}.$$

Jos $i \in J$, niin

$$\sum_{v\geq u_i}(-1)^{v-u_i}\frac{\gamma_i^v}{u_i!(v-u_i)!} = \frac{\gamma_i^{u_i}e^{-\gamma_i}}{u_i!}.$$

Toisaalta jos $i \in I \setminus J$, niin

$$\begin{aligned}
\sum_{v\geq u_i}(-1)^{v-u_i}\frac{\gamma_i^v}{u_i!(v-u_i)!} &= \sum_{v\geq u_i}(-1)^{v-u_i}\left(\sum_{w<u_i}(-1)^{u_i-w-1}\binom{v}{w}\right)\frac{\gamma_i^v}{v!}\\
&= 1-\sum_{v<u_i}\left(\sum_{w\leq v}(-1)^{v-w}\binom{v}{w}\right)\frac{\gamma_i^v}{v!}-\sum_{v\geq u_i}(-1)^{v-u_i}\left(\sum_{w<u_i}(-1)^{u_i-w}\binom{v}{w}\right)\frac{\gamma_i^v}{v!}\\
&= 1-\sum_{w<u_i}\left(\sum_{v\geq w}(-1)^{v-w}\binom{v}{w}\frac{\gamma_i^v}{v!}\right)\\
&= 1-\sum_{w<u_i}\frac{\gamma_i^w e^{-\gamma_i}}{w!},
\end{aligned}$$

eli $\mathbb{P}_c(M(J,\overline{u},n),n)$ suppenee, kun n kasvaa rajatta. □

Nyt lauseista 5.12 ja 6.19 seuraa tutkielman päätulos:

Seuraus 6.20. *Olkoon G satunnainen n-verkko ja φ ensimmäisen kertaluvun logiikalla määriteltävä verkon ominaisuus. Tällöin on olemassa raja-arvo $\lim_{n\to\infty}\mathbb{P}(\varphi,n)$.*

Lähteet

[1] Clauset, Aaron; Shalizi, Cosma Rohilla; Newman, M. E. J.: Power-Law Distributions in Empirical Data. SIAM Review, vol. 51, no. 4, pp. 661-703, 2009.

[2] Diestel, Reinhard: Graph Theory. Springer-Verlag, Heidelberg, 2016.

[3] Ehrenfeucht, A.: An application of games to the completeness problem for formalized theories. Fund. Math. 49, pp. 129–141, 1961.

[4] Gaifman, H.: On local and non-local properties. In Proceedings of the Gerbrand Symposium Logic Colloquium, 1981

[5] Galambos Janos: Bonferroni inequalities. The Annals of Probability, vol. 5, no. 4, pp. 577-581, 1977.

[6] Harris T. E.: The theory of branching processes. Dover Publications, Inc., New York, 1989

[7] Junnila Heikki: Diskreettiä matematiikkaa-luentomoniste

[8] Lynch James F.: Convergence law for random graphs with specified degree sequence. ACM Transactions on Computational Logic (TOCL), pp. 727–748, 2005.

[9] Lynch James F.: Probabilities of sentences about very sparse random graphs. Random Struct. Alg. 3, pp. 33–53, 1992.

[10] McKay B. D.: Asymptotics for symmetric 0-1 matrices with prescribed row sums. Ars Combinatorica 19A, pp. 15-25, 1985

[11] Molloy, Michael; Reed, Bruce: A critical point for random graphs with a given degree sequence. Random Structures Algorithms 6, no. 2-3, pp. 161-179, 1995.

[12] Spencer Joel: The Strange Logic of Random Graphs. Springer Publishing Company, 2010.

[13] Wasserman, L: 36-705 Intermediate Statistics, Lecture notes.

[14] Weisstein, Eric W: Bonferroni Inequalities. MathWorld–A Wolfram Web Resource.