Lyhyt johdatus joukko-oppiin ja relaatioihin

Tommi Syrjänen

1 Johdanto

Tämän oppaan tarkoituksena on esittää lyhyt tiivistelmä joukko-opin ja relaatioiden perusteista. Esitys seuraa pääpiirteissään kirjan *Lewis, Papadimitriou: Elements of the Theory of Computation, 1998* ensimmäistä lukua. Teksti on tarkoitettu lähinnä muistin virkistämiseksi, ja kaikki todistukset on sivuutettu.

2 Joukot

1. Matemaattisesti **joukko** (*set*) on kokoelma **alkioita** (*element*). Esimerkiksi englanninkielisen aakkoston vokaalien joukkoa V merkitään seuraavasti:

$$V = \{a, e, i, o, u, y\} \ .$$

2. Mikäli alkio a on joukon S jäsen, merkitään sitä $a \in S$. Päinvastaisessa tapauksessa merkitään $a \notin S$. Esimerkiksi ylläolevalle joukolle V pätee:

$$\begin{array}{l} e \in V \\ k \notin V \ . \end{array}$$

3. Joukko A on joukon B **osajoukko** (*subset*), mikäli kaikki A:n alkiot ovat myös joukossa B. Tästä käytetään merkintää $A \subseteq B$. Esimerkiksi:

$$\begin{array}{l} \{a, b\} \subseteq \{a, b, c, d\} \\ \{a, b\} \not\subseteq \{a, c, d\} \ . \end{array}$$

Joukot A ja B ovat samat, mikäli $A \subseteq B$ ja $B \subseteq A$:

$$\begin{array}{l} \{a, b\} \subseteq \{b, a\} \text{ ja} \\ \{b, a\} \subseteq \{a, b\} \text{ , joten} \\ \{a, b\} = \{b, a\} \ . \end{array}$$

Kuten ylläolevasta esimerkistä huomataan, joukon alkioiden järjestyksellä ei ole väliä, kuten ei myöskään sillä, että jokin alkio esiintyy joukossa monta kertaa:

$$\{a, b, a, b, b\} = \{a, b\} \ .$$

4. Mikäli $A \subseteq B$, mutta $B \not\subseteq A$, on joukko A joukon B **aito osajoukko** (*proper subset*), ja siitä käytetään merkintää $A \subset B$.

5. Joukkoa, jossa ei ole yhtään alkiota kutsutaan **tyhjäksi joukoksi** (*empty set*), ja siitä käytetään merkintää $\emptyset$. Kaikille joukoille S pätee:

$$\emptyset \subseteq S \ .$$

6. Joukon S kaikkien osajoukkojen joukkoa kutsutaan **potenssijoukoksi** (*power set*), ja siitä käytetään merkintää 2^S. Kirjallisuudessa esiintyy myös merkintä $\mathcal{P}(S)$. Esimerkiksi:

$$2^{\{1,2\}} = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\} \ .$$

Joukon alkioina voi siis olla myös toisia joukkoja. Erityisesti kannattaa huomata, että

$$\{\emptyset\} \neq \emptyset \ .$$

7. Joukko S voidaan määritellä joko luettelemalla kaikki sen alkiot tai jonkin toisen joukon avulla. Esimerkiksi parittomien luonnolisten lukujen joukko O voidaan määritellä seuraavasti:

$$O = \{x \mid x \in \mathbb{N} \text{ ja } x \text{ ei ole jaollinen luvulla } 2\ \} \ .$$

Varsinkin amerikkalaisessa kirjallisuudessa käytetään usein pystyviivan tilalla kaksoispistettä.

2.1 Joukko-operaatiot

Tavallisimmat joukko-opilliset operaatiot ovat:

1. **unioni** (*union*):

$$\{1, 2, 3\} \cup \{1, 4\} = \{1, 2, 3, 4\}$$

2. **leikkaus** (*intersection*)

$$\{1,2,3\} \cap \{1,4\} = \{1\}$$

3. **erotus** (*difference*)

$$\{1,2,3\} - \{1,4\} = \{2,3\}$$

Joukkojen A ja B erotuksesta käytetään myös merkintää $A \setminus B$.

3 Relaatiot ja funktiot

1. Joukkojen A ja B **karteesinen tulo** (*Cartesian product*) on joukko:

$$A \times B = \{(x,y) \mid x \in A, y \in B\} \ ,$$

missä kukin (x,y) on **järjestetty pari** (*ordered pair*). Esimerkiksi

$$\{1,2,3\} \times \{1,4\} = \{(1,1),(1,4),(2,1),(2,4),(3,1),(3,4)\}$$

2. Kuten jo nimestä voi päätellä, järjestetyt parit (a,b) ja (b,a) eivät ole sama pari. Samoin $(a,(b,c)) \neq ((a,b),c)$.

3. Joukkojen A ja B välille muodostettu **relaatio** (*relation*) R on osajoukko karteesisesta tulosta $A \times B$:

$$R \subseteq A \times B \ .$$

4. Merkinnän $(a,b) \in R$ lisäksi käytetään usein merkintää aRb, etenkin tavallisten matemaattisten relaatioiden tapauksessa. Esimerkiksi $(a,b) \in \, <$ merkitään tuttuun tapaan $a < b$.

5. Relaatio voi olla muodostettu myös useamman kuin kahden joukon suhteen, esim.:

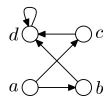
$$R \subseteq A \times B \times C \ .$$

Mikäli joukkoja on kaksi, on kyseessä **binäärirelaatio** (*binary relation*).

6. Binäärirelaatio joukolta itselleen voidaan esittää suunnatuna graafina. Esimerkiksi relaatiota $R \subseteq \{a,b,c,d\} \times \{a,b,c,d\}$:

$$R = \{(a,b),(a,c),(b,d),(c,d),(d,d)\}$$

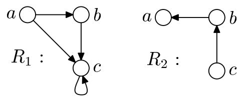
vastaa graafi:



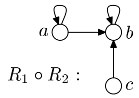
7. Kahden relaation $R_1 \subseteq A \times B$ ja $R_2 \subseteq B \times C$ **kompositio** (*composition*) $R_1 \circ R_2 \subseteq A \times C$ määritellään seuraavasti:

$$R_1 \circ R_2 = \{(a,c) \mid a \in A, c \in C, \exists b \in B : (a,b) \in R_1 \text{ ja } (b,c) \in R_2\}$$

Esimerkiksi relaatioiden:



kompositio $R_1 \circ R_2$ on:

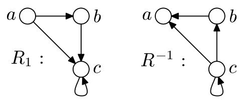


Komposition kaari (a,a) saadaan yhdistämällä kaaret (a,b) ja (b,a), kaari (a,b) kaarista (a,c) ja (c,b), kaari (c,b) kaarista (c,c) ja (c,b) sekä kaari (b,b) kaarista (b,c) ja (c,b).

8. Relaation $R \subseteq A \times B$ **käänteisrelaatio** $R^{-1} \subseteq B \times A$ on relaatio:

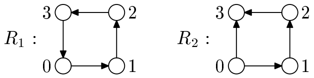
$$R^{-1} = \{(b,a) \mid (a,b) \in R\}$$

Esimerkiksi:



9. Funktio (*function*) f on relaatio $f \subseteq A \times B$, jossa jokaista **lähtöavaruuden** (*domain*) A alkiota a kohden löytyy täsmälleen yksi **kuva-avaruuden** (*image*) alkio b siten, että $(a,b) \in f$. Yleensä käytetään merkintöjä $f : A \mapsto B$ ja $f(a) = b$.

Tarkastellaan kahta relaatiota R_1 ja R_2:

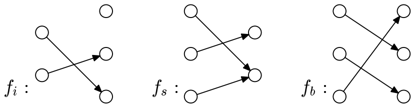


Relaatio R_1 on itseasiassa funktio $f : \mathbb{Z}_4 \mapsto \mathbb{Z}_4$,

$$f(x) = x + 1 \mod 4 \ .$$

Sitävastoin R_2 ei ole funktio, sillä alkiolla 0 on kaksi kuvaa ja alkiolla 3 ei ole yhtään.

10. Funktio $f : A \mapsto B$ on **injektiivinen** (*one-to-one*), mikäli kaikille erillisille alkiolle $a, a' \in A$ pätee $f(a) \neq f(a')$. Vastaavasti f on **surjektiivinen** (*onto*), mikäli kaikille $b \in B$ on olemassa $a \in A$ siten, että $f(a) = b$. Funktiota, joka on sekä injektiivinen että surjektiivinen, kutsutaan **bijektioksi** (*bijection*).



Näistä funktioista f_i on injektiivinen, f_s surjektiivinen ja f_b bijektio.

3.1 Relaatioiden luokittelua

11. Binäärirelaatiot voidaan luokitella niiden ominaisuuksien mukaan. Tärkeimmät ominaisuudet ovat: refleksiivisyys, symmetrisyys, transitiivisuus ja sarjallisuus.

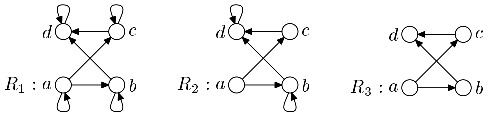
12. Binäärirelaatio $R \subseteq A \times A$ on **refleksiivinen** (*reflexive*), mikäli

$$\forall a \in A : (a, a) \in R \ .$$

Refleksiivinen relaation graafiesityksessä on kaari jokaisesta solmusta takaisin itseensä. Vastaavasti relaatio on antirefleksiivinen, mikäli

$$\forall a \in A : (a, a) \notin R \ .$$

Tarkastellaan seuraavia kolmea relaatiota:



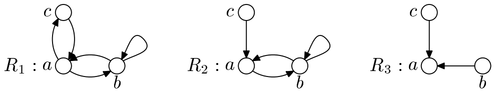
Näistä R_1 on refleksiivinen ja R_3 on antirefleksiivinen, sillä siinä ei ole yhtään refleksiivistä kaarta. Sitävastoin R_2 ei ole kumpaakaan.

13. Binäärirelaatio $R \subseteq A \times A$ on **symmetrinen** (*symmetric*), mikäli

$$\forall a, b \in A : (a, b) \in R \Rightarrow (b, a) \in R \ .$$

Symmetristä relaatiota kuvaavassa graafissa on jokaisella kaarella paluukaari.

Tarkastellaan taas kolmea relaatiota:



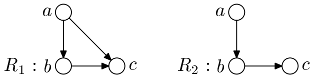
Tässä R_1 on symmetrinen, R_3 antisymmetrinen ja R_2 ei ole kumpaakaan.

14. Binäärirelaatio $R \subseteq A \times A$ on **transitiivinen** (*transitive*), mikäli

$$\forall a, b, c \in A : (a, b) \in R \wedge (b, c) \in R \Rightarrow (a, c) \in R$$

Mikäli transitiivisen relaation graafissa on jokin polku solmusta a solmuun b, täytyy siinä olla myös suora kaari (a, b).

Allaolevista relaatioista R_1 on transitiivinen, mutta R_2 ei ole, sillä kaari (a, c) puuttuu.



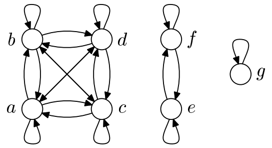
15. Binäärirelaatio $R \subseteq A \times A$ on **sarjallinen** (*serial*), mikäli

$$\forall a \in A, \exists b \in A : (a, b) \in R \ .$$

Sarjallisen relaation graafin kaikista solmuista lähtee vähintään yksi kaari.

16. Binäärirelaatio $R \subseteq A \times A$ on **ekvivalenssirelaatio** (*equivalence relation*), mikäli se on sekä refleksiivinen, transitiivinen että symmetrinen. Ekvivalenssirelaatio jakaa joukon A **ekvivalenssiluokkiin** (*equivalence class*).

Allaoleva ekvivalenssirelaatio jakaa joukon $\{a, b, c, d, e, f, g\}$ kolmeen ekvivalenssiluokkaan:



Luokat ovat $\{a,b,c,d\}$, $\{e,f\}$ ja $\{g\}$.

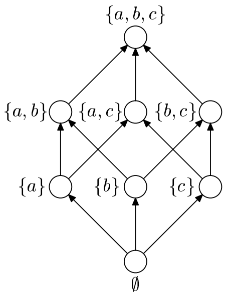
17. Binäärirelaatio $R \subseteq A \times A$ on **osittaisjärjestys** (*partial order*), mikäli se on refleksiivinen, transitiivinen ja kaikille erillisille alkioille $a, b \in A$ joko $(a,b) \notin R$ tai $(b,a) \notin R$. Alkio a on osittaisjärjestyksen R **minimi** (*minimum*), mikäli

$$\forall b \in A : (a = b) \vee (b,a) \notin R \ .$$

Vastaavasti **maksimi** (*maximum*) on alkio a, jolle pätee:

$$\forall b \in A : (a = b) \vee (a,b) \notin R \ .$$

Esimerkiksi mielivaltaisen joukon S potenssijoukko 2^S muodostaa osittaisjärjestyksen $\subseteq$-relaation yli. Alla on esitetty tämä joukolle $S = \{a,b,c\}$:[1]



Järjestyksen ainoa minimi on $\emptyset$ ja maksimi $\{a,b,c\}$.

18. Mikäli osittaisjärjestyksessä R kaikille pareille $a, b \in A$ joko $(a,b) \in R$ tai $(b,a) \in R$, on kyseessä **täysjärjestys** (*total order*).

Esimerkiksi tavalliseen tapaan määritelty relaatio $\leq \subset \mathbb{N} \times \mathbb{N}$ on täysjärjestys.

[1]Kuvasta on jätetty selvyyden vuoksi pois refleksiiviset ja transitiiviset kaaret. Esimerkiksi relaatioon kuuluu myös kaari $(\{a\}, \{a,b,c\})$, sillä $\{a\} \subseteq \{a,b\} \subseteq \{a,b,c\}$.

3.2 Esimerkkejä relaatiotyypeistä

19. Seuraavissa esimerkeissä käytetään tarkasteltavana joukkona kaikkien ihmisten joukkoa P.

20. Määritellään relaatio $R_1 \subset P \times P$:

$$R_1 = \{(a,b) \mid \ a\text{:lla ja } b\text{:llä on samat vanhemmat } \}$$

R_1 on refleksiivinen, symmetrinen (jos a:lla on samat vanhemmat kuin b:llä, niin myös b:llä on samat vanhemmat kuin a:lla) ja transitiivinen (jos a on b:n veli ja b on c:n sisko, niin a on c:n veli), joten se on myös ekvivalenssirelaatio. Yhden pariskunnan lapset muodostavat aina yhden ekvivalenssiluokan.

21. Määritellään $R_2 \subset P \times P$:

$$R_2 = \{(a,b) \mid b \text{ on } a\text{:n äiti}\}$$

R_2 on antirefleksiivinen (kukaan ei ole itsensä äiti), antisymmetrinen (kukaan ei ole äitinsä äiti) ja sarjallinen (kaikilla on äiti). Koska kaikilla henkilöillä on täsmälleen yksi äiti, relaatio voidaan tulkita myös funktiona

$$f(a) = \ a\text{:n äiti.}$$

Funktio ei kuitenkaan ole injektiivinen, sillä naisella voi olla useita lapsia, eikä surjektiivinen, sillä kaikki ihmiset eivät ole äitejä.

22. Määritellään $R_3 \subset P \times P$:

$$R_3 = \{(a,b) \mid \ b \text{ on } a\text{:n esi-isä } \}$$

R_3 on antirefleksiivinen, antisymmetrinen, sarjallinen ja transitiivinen (jos a on b:n esi-isä ja b on c:n esi-isä, niin a on myös c:n esi-isä).

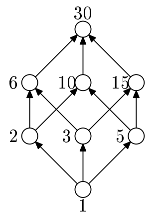
23. Siirrytään tarkastelemaan luonnollisten lukujen joukkoa $\mathbb{N} = \{0, 1, 2, \ldots\}$.

24. Relaatio $\preccurlyeq \subset \mathbb{N} - \{0\} \times \mathbb{N} - \{0\}$:

$$\preccurlyeq = \{(a,b) \mid \ b \text{ on jaollinen } a\text{:lla } \}$$

on refleksiivinen ja transitiivinen. Lisäksi tilanne $a \preccurlyeq b$ ja $b \preccurlyeq a$ on mahdollinen vain, jos $a = b$. Näin ollen $\preccurlyeq$ on osittaisjärjestys. Järjestyksellä on yksikäsitteinen minimialkio 1.

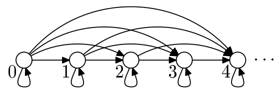
Esimerkiksi lukujoukon $\{1, 2, 3, 5, 6, 10, 15, 30\}$ jaollisuusrelaatio on[2]:



25. Viimeisenä esimerkkinä tarkastellaan relaatiota $\leq\, \subset \mathbb{N} \times \mathbb{N}$:

$$\leq\, = \{(a, b) \mid \text{ } a \text{ on pienempi tai yhtäsuuri kuin } b\}$$

Kuten $\preccurlyeq$, myös $\leq$ on osittaisjärjestys. Lisäksi kaikilla pareilla $a, b \in \mathbb{N}$ pätee $a \leq b$ tai $b \leq a$, joten kyseessä on täysjärjestys.



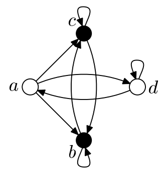
4 Sulkeumat

1. Olkoon $R \subseteq D \times D$ binäärirelaatio ja joukko $B \subseteq D$. Tällöin B on **suljettu** (*closed under*) relaation R suhteen, mikäli

$$\forall b \in B, \forall c \in D : (b, c) \in R \rightarrow c \in B \quad .$$

Toisin sanoen, graafiesityksessä yhdestäkään joukkoon B kuuluvasta solmusta ei saa lähteä kaarta solmuun, joka ei kuulu siihen.

Tarkastellaan esimerkiksi alla olevaa relaatiota R, missä $D = \{a, b, c, d\}$, $A = \{a, d\}$ ja $B = \{b, c\}$:



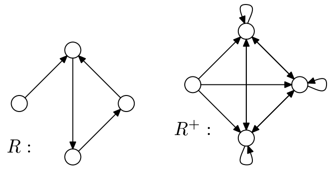
[2] Myös tästä kuvasta puuttuvat refleksiiviset ja transitiiviset kaaret.

Joukko B on suljettu R:n suhteen, sillä solmuista b ja c ei johda yhtään kaarta solmuihin a ja d. Sitävastoin A ei ole suljettu, sillä esimerkiksi $(a, b) \in R$, mutta $b \notin A$.

2. Joukko voi olla suljettu myös jonkin funktion suhteen. Esimerkiksi luonnollisten lukujen joukko $\mathbb{N}$ on suljettu yhteenlaskun suhteen, sillä $a + b \in \mathbb{N}$ aina kun $a \in \mathbb{N}$ ja $b \in \mathbb{N}$. Se ei kuitenkaan ole suljettu vähennyslaskun suhteen, sillä esimerkiksi $1 - 3 = -2 \notin \mathbb{N}$.

3. Relaation $R \subseteq A \times A$ **transitiivinen sulkeuma** (*transitive closure*) R^+ on pienin transitiivinen relaatio, johon sisältyvät kaikki R:n parit.

4. Esimerkiksi alla on relaatio R ja sen transitiivinen sulkeuma R^+:

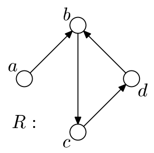


5. Transitiivinen sulkeuma voidaan muodostaa seuraavanlaisella algoritmilla:

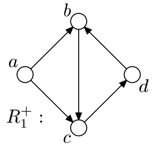
$R^+ := R$
while $\exists a_i, a_j, a_k \in A : (a_i, a_j), (a_j, a_k) \in R^+ \wedge (a_i, a_k) \notin R^+$ **do**
    $R^+ := R^+ \cup \{(a_i, a_k)\}$
end while

Algoritmissa lähdetään liikkeelle alkuperäisestä relaatiosta R ja etsitään siitä kahden askeleen mittaisia polkuja, jotka rikkovat transitiivisuusehdon.

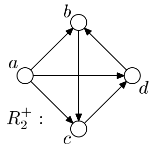
6. Esimerkki. Muodostetaan yllä esitetyn relaation R transitiivinen sulkeuma:



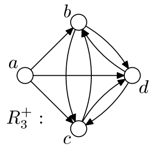
Relaatiossa on kaaret (a, b) ja (b, c), joten kaari (a, c) täytyy lisätä siihen:



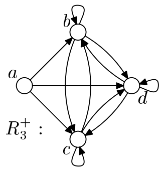
Tämän lisäyksen jälkeen R^+:ssä on kaaret (a, c) ja (c, d), joten myös (a, d) täytyy lisätä:



Koska solmut b, c ja d muodostavat silmukan, täytyy niiden välillä olla kaaret kumpaankin suuntaan.



Lopuksi täytyy vielä lisätä refleksiiviset kaaret silmukassa mukana oleville solmuille:



7. Transitiivisen sulkeuman lisäksi voidaan määritellä samalla tavoin myös **symmetrinen** ja **refleksiivinen sulkeuma**.

8. Käytännön kannalta tärkein on graafin R refleksiivinen ja transitiivinen sulkeuma, sillä se sisältää kaikki R:n polut.

$$R^* = \{(a, b) \mid \text{graafissa } R \text{ on polku solmusta } a \text{ solmuun } b\}$$

5 Äärelliset ja äärettömät joukot

1. Joukon S **mahtavuus** (*cardinality*) $|S|$ on sen sisältämien alkioiden määrä. Joukot A ja B ovat yhtä mahtavia (*equinumerous*), mikäli niiden välille voidaan muodostaa bijektio $f : A \mapsto B$.

2. Joukko S on **äärellinen** (*finite*), mikäli se on yhtä mahtava joukon $\{1, \ldots, n\}$ kanssa jollakin $n \in \mathbb{N}$. Joukko, joka ei ole äärellinen, on **ääretön** (*infinite*).

3. Joukko S on **numeroituva** (*countable*), mikäli se on äärellinen tai yhtä mahtava luonnollisten lukujen joukon $\mathbb{N}$ kanssa, muussa tapauksessa se on **ylinumeroituva** (*uncountable*).

4. Esimerkiksi parittomien luonnollisten lukujen joukko

$$O = \{1, 3, 5, \ldots\}$$

on numeroituva. Tässä tapauksessa bijektio $f : \mathbb{N} \mapsto O$ voidaan määritellä seuraavasti:

$$f(n) = 2n + 1 \ .$$

5. Numeroituvia joukkoja ovat myös kokonaislukujen $\mathbb{Z}$ ja rationaalilukujen $\mathbb{Q}$ joukot. Lisäksi kahden numeroituvan joukon karteesinen tulo on aina numeroituva.

6. Reaalilukujen $\mathbb{R}$ ja kompleksilukujen $\mathbb{C}$ joukot sekä luonnollisten lukujen joukon $\mathbb{N}$ potenssijoukko $2^{\mathbb{N}}$ ovat ylinumeroituvia.

6 Todistuksista

1. Halutaan todistaa, että jokin väite $P(n)$ pätee kaikille luonnollisille luvuille, esimerkiksi, että

$$\sum_{i=1}^{n} i = \frac{n^2 + n}{2}$$

kaikilla $n \geq 0$.

2. Matemaattisen **induktion** (*induction*) perusajatuksena on todistaa väite kahdessa osassa:

1. Todistetaan, että väite pätee tapauksessa $P(0)$. (**perustapaus**, *basic case*);
2. Todistetaan, että kaikilla n, $P(n) \to P(n+1)$. (**induktio-askel**, *induction step*).

Näiden osien yhdistämisestä seuraa se, että väite pätee kaikilla $n \in \mathbb{N}$: Perustapauksesta seuraa, että $P(0)$ on tosi. Siitä voidaan induktio-askeleen avulla näyttää, että $P(1)$ on tosi, josta puolestaan seuraa, että $P(2)$ on tosi, jne. Induktio-askel todistetaan yleensä **induktio-oletuksen** (*induction hypothesis*) avulla.

3. Todistetaan ylläoleva väite induktiolla.

Perustapaus. Kun $n = 0$, $\sum_{i=1}^{n} i = 0 = \frac{0^2+0}{2}$, joten väite pitää paikkansa.

Induktio-oletus. Oletetaan, että on olemassa jokin $k \in \mathbb{N}$ siten, että väittämä pitää paikkansa kaikilla $n \leq k$.

Induktio-askel. Tarkastellaan tapausta $n = k + 1$. Tällöin summa on

$$\sum_{i=1}^{k+1} i = 1 + 2 + \cdots + k + (k+1) \ .$$

Induktio-oletuksen perusteella $\sum_{i=1}^{k} i = \frac{k^2+k}{2}$, joten yhtälö sievenee muotoon:

$$\begin{aligned}
\sum_{i=1}^{k+1} i &= \frac{k^2+k}{2} + (k+1) \\
&= \frac{k^2+k+2k+2}{2} \\
&= \frac{(k^2+2k+1)+(k+1)}{2} \\
&= \frac{(k+1)^2+(k+1)}{2} \ .
\end{aligned}$$

4. Toinen tietojenkäsittelyteoriassa hyödyllinen todistusmenetelmä on niin kutsuttu "**kyyhkyslakkaperiaate**" (*pigeonhole principle*). Formaalisti periaate määritellään seuraavasti:

> *Mikäli A ja B ovat äärellisiä joukkoja ja $|A| > |B|$, ei ole olemassa injektiivistä funktiota $f : A \mapsto B$.*

Nimi "kyyhkyslakkaperiaate" tulee teoreemaa havainnollistavasta esimerkistä:

> Mikäli $n+1$ kyyhkystä varten on n pesää, täytyy ainakin yhteen pesään mennä vähintään 2 kyyhkystä.

Tässä teoreeman mainitsema injektiivinen funktio olisi tapa jakaa kaikki kyyhkyset pesiin siten, että kuhunkin pesään tulisi korkeintaan yksi kyyhkynen.