AUTOMAATTISET JONOT JA COBHAMIN LAUSE

Mikael Nuutila

Pro gradu -tutkielma
Elokuu 2021

Tarkastajat:
Prof. Jarkko Kari
Dos. Aleksi Saarela

MATEMATIIKAN JA TILASTOTIETEEN LAITOS

TURUN YLIOPISTO
Matematiikan ja tilastotieteen laitos

NUUTILA, MIKAEL: Automaattiset jonot ja Cobhamin lause
Pro gradu -tutkielma, 61 s.
Matematiikka
Elokuu 2021

Tutkielman tarkoituksena on esitellä Thijmen J. P. Krebsin uusi aikaisempaa lyhyempi todistus k-automaattisia jonoja koskevalle Cobhamin lauseelle. Oletetaan, että $k \geq 2$ ja $h \geq 2$ ovat multiplikatiivisesti riippumattomia luonnollisia lukuja. Tällöin Cobhamin lauseen mukaan äärellisen aakkoston jono on sekä $k-$ että h-automaattinen silloin ja vain silloin kun se on lopulta jaksollinen. Cobhamin lauseen lisäksi perehdytään joihinkin k-automaattisten jonojen perusominaisuuksiin. Tarkastellaan k-automaattisia jonoja erityisesti automaattis-teoreettiselta kannalta. Pääpaino on niissä tuloksissa joita tarvitaan Cobhamin lauseen todistuksessa.

Esitetään tarpeelliset pohjatiedot formaalisten kielten ja säännöllisten kielten teoriasta. Perehdytään myös lukujärjestelmiin, erityisesti tutkielman aiheen kannalta keskeisiin k-kantajärjestelmiin. Määritellään äärelliset funktioautomaatit, automaattiset funktiot, äärelliset transduktorit ja μ-transduktorit. Määritellään k-automaattiset jonot syöttämällä jonon indeksien k-kantaesityksiä funktioautomaatille. Osoitetaan, että lopulta jaksollinen jono on aina k-automaattinen. Muodostetaan μ-transduktori, joka laskee normalisaation kannassa k. Käytetään tätä transduktoria todistamaan, että jono on k-automaattinen silloin ja vain silloin kun se on (k, D)-automaattinen. Oletetaan, että $k \geq 2$ ja $h \geq 2$ ovat multiplikatiivisesti riippumattomia kokonaislukuja. Osoitetaan, että tällöin jono on lopulta jaksollinen, mikäli se on $(k, D_k)-$ ja (h, D_h)-automaattinen. Cobhamin lause seuraa.

Asiasanat: k-automaattiset jonot, funktioautomaatit, äärelliset transduktorit, Cobhamin lause, formaalit kielet, k-kantajärjestelmät

Sisältö

1 Johdanto

Tämän tutkielman päätavoitteena on esitellä *k-automaattisten jonojen* teorian perustuloksia ja lisäksi aikaisempaa yksinkertaisempi Thijmen J. P. Krebsin artikkelissa [1] esittämä todistus *Cobhamin lauseelle.*

Olkoon Δ jokin äärellinen *aakkosto.* Aakkoston Δ *jonoja* on ylinumeroituva määrä, joten kaikille jonoille ei ole olemassa äärellistä matemaattista kuvailua, eikä niitä voi edes listata. Tutkimuksen kannalta on hyödyllistä rajoittua tarkastelemaan niitä jonoja, joita voidaan kuvailla matemaattisesti. Voidaan rajoittua esimerkiksi jonoihin, jotka voidaan muodostaa käyttämällä *deterministisiä äärellisiä automaatteja.* Oletetaan, että on annettu deterministinen äärellinen automaatti $\mathcal{A}$. Määritellään jono $\mathbf{a} = (a_n)$ käyttämällä automaattia $\mathcal{A}$: syötetään indeksin $n \in \mathbb{N}$ jokin k-kantaesitys automaatille $\mathcal{A}$. Luettuaan kyseisen k-kantaesityksen $\mathcal{A}$ päätyy johonkin tilaan. Käyttämällä *tulostusfunktiota* voidaan tästä tilasta muodostaa jonon jäsen a_n. Sellaisia jonoja, jotka voidaan muodostaa edellä kuvatulla tavalla, kutsutaan *k-automaattisiksi jonoiksi.*

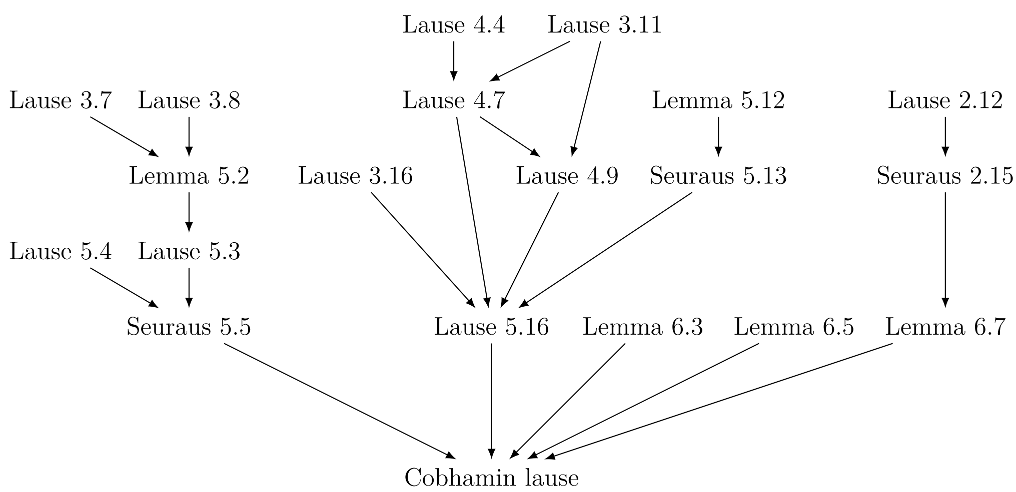
Luvussa 2 tarkastellaan tarpeellisia pohjatietoja. Alaluvussa 2.1 esitellään lyhyesti *formaalit kielet, jonot ja morfismit.* Alaluku 2.1 pohjautuu lähteisiin [2, 3, 4]. Alaluvussa 2.2 esitellään *äärellisten automaattien* ja *säännöllisten kielten* käsitteitä ja tuloksia tarpeellisilta osin. Alaluku 2.2 perustuu luentomonisteeseen [4]. Lukijan oletetaan tuntevan äärellisten automaattien ja säännöllisten kielten teoriaa alaluvussa 2.2 esitetyltä osalta. Alaluvussa 2.3 esitellään lyhyesti puolirenkaiden *lukujärjestelmät.* Erikoistapauksena tarkastellaan luonnollisten lukujen lukujärjestelmiä, erityisesti k-automaattisten jonojen kannalta tärkeitä *k-kantajärjestelmiä.*

Luvussa 3 esitellään k-automaattisten jonojen teorian kannalta tärkeät äärellisten automaattien laajennukset: *funktioautomaatit* ja *transduktorit.* Funktioautomaatit ovat automaatteja jotka tulostavat symbolin luettuaan syötesanan, kun taas transduktorit tulostavat saadun syötesanan perusteella muodostetun *tulostesanan.* Funktioautomaatteja käsittelevä alaluku pohjautuu Jean-Paul Allouchen ja Jeffrey Shallitin kirjaan [2], transduktoreita käsittelevä alaluku perustuu myös M. Lothaire -ryhmän kirjaan [3].

Tutkielman pääaihe, k-automaattiset jonot, esitellään luvussa 4. Osoitetaan, että k-automaattiset jonot voidaan määritellä käyttäen mitä tahansa indeksin *tavallista k-kantaesitystä.* Määritellään myös niin sanottu (k, D)-automaattisuus. Varsinaisiin k-automaattisia jonoja koskeviin tuloksiin perehdytään luvussa 5. Osoitetaan, että *lopulta jaksolliset* jonot ovat aina k-automaattisia. Todistetaan *Cobhamin pieni lause*, joka karakterisoi k-automaattiset jonot *k-uniformisten* morfismien *kiintopisteiden koodauksina.* Määritellään jonon k-ydin. Osoitetaan, että jono on k-automaattinen silloin ja vain silloin kun sen k-ydin on äärellinen. Alaluvussa 5.3 täydennetään Krebsin todistusta osoittamalla, että jono on k-automaattinen silloin ja vain sillion kun se on (k, D)-automaattinen.

Luvussa 6 tarkastellaan tarpeellisia aputuloksia, jonka jälkeen todistetaan *Cobhamin lause* käyttäen Krebsin menetelmää. Cobhamin lauseen nojalla mikäli k ja h ovat *multiplikatiivisesti riippumattomia* kantalukuja, jono (a_n) on sekä k-automaattinen että h-automaattinen silloin ja vain silloin kun se on lopulta jaksollinen. Ensimmäisenä Cobhamin lauseen todisti Alan Cobham vuonna 1969.

Kuvan 1 kaavio kuvaa Cobhamin lauseeseen johtavien tulosten välisiä seuraussuhteita.



Kuva 1: Keskeisimpien lauseiden, lemmojen ja seurauksien väliset seuraussuhteet.

2 Perusteet

2.1 Formaalit kielet

2.1.1 Määritelmiä ja merkintöjä

Symboliaakkosto tai lyhyemmin *aakkosto* on äärellinen ja epätyhjä joukko. Aakkostolta ei edellytetä mitään algebrallista rakennetta, se on vain joukko alkioita. Esimerkiksi seuraavat joukot ovat aakkostoja:

- $\{0, 1\}$
- $\{a, b, c\}$
- $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F\}$
- $\{\cdot, \times, \otimes\}$

Aakkostoista käytetään yleensä merkintöjä Σ, Δ ja Γ. Aakkoston alkioita kutsutaan *symboleiksi* tai *kirjaimiksi.*

Äärellisiä symbolijonoja kutsutaan *sanoiksi.* Esimerkiksi *abbbaabb* on aakkoston $\Sigma = \{a, b\}$ sana. Sanan w pituus $|w|$ on siinä esiintyvien symbolien määrä. Esimerkiksi $|\alpha\upsilon\tau o\mu\alpha\tau o\nu| = 9$.

Sanojen u ja v *katenaatiolla* $u \cdot v = uv$ tarkoitetaan sitä sanaa w, joka saadaan kun kirjoitetaan sanat u ja v peräkkäin. Esimerkiksi

$$11111 \cdot 100101 = 11111100101.$$

Määritellään *tyhjä sana* ε, jossa ei ole yhtäkään symbolia. Toisin sanoen $|\varepsilon| = 0$. Tällöin $\varepsilon w = w\varepsilon = w$ aina kun w on jokin sana.

Olkoot $w_1, w_2, \ldots, w_m$ jonkin aakkoston symboleita. Sanan $w = w_1 w_2 \cdots w_m$ *peilaus* w^R on sana $w_m w_{m-1} \cdots w_2 w_1$. Selvästi $(uv)^R = v^R u^R$.

Oletetaan, että $n \in \mathbb{N}$ on luonnollinen luku. Määritellään sanan w potenssi

$$w^n = \underbrace{ww \cdots ww}_{n \text{ kpl}}.$$

Erityisesti $w^0 = \varepsilon$ aina kun $w \in \Sigma^*$.

Määritellään n-pituisten aakkoston Σ sanojen joukko

$$\Sigma^n = \{a_1 a_2 \cdots a_n \mid a_i \in \Sigma \text{ aina kun } i = 1, 2, \ldots, n\}$$

aina kun $n > 0$, ja määritellään $\Sigma^0 = \{\varepsilon\}$ Määritellään myös kaikkien aakkoston Σ sanojen joukko Σ^*. Tyhjä sana on jokaisen aakkoston sana.

Sana $u \in \Sigma^*$ on sanan $w \in \Sigma^*$ *etuliite*, jos on olemassa sellainen sana $v \in \Sigma^*$, että $w = uv$. Vastaavasti u on sanan w *loppuliite*, mikäli on olemassa sellainen sana $v \in \Sigma$, että $w = vu$. Sana y on sanan $w \in \Sigma^*$ *osasana*, jos on olemassa sellaiset sanat $x \in \Sigma^*$ ja $z \in \Sigma^*$, että $w = xyz$.

Aakkoston Σ *(formaali) kieli* L on joukon Σ^* osajoukko. Jos L ja K ovat kieliä merkitään

$$LK = \{uv \mid u \in L,\ v \in K\}$$

ja

$$L^n = \{w_1 w_2 \cdots w_n \mid w_i \in L\}.$$

Määritellään kielten L ja K *osamäärä*

$$L/K = \{u \in \Sigma^* \mid \text{on olemassa sellainen sana } v \in K, \text{ että } uv \in L\}.$$

Kieli L/K sisältää siis ne sanat jotka saadaan kielen L sanoista poistamalla niiden lopusta kielen K sanojen esiintymät.

Määritellään unaarinen operaatio *Kleenen tähti*

$$L^* = \bigcup_{i=0}^{\infty} L^i.$$

Otetaan käyttöön lyhennysmerkintä $w^* = \{w\}^*$ aina kun w on aakkoston Σ sana.

2.1.2 Äärettömät sanat ja jonot

Oletetaan, että A ja B ovat joukkoja. Käytetään funktioiden $B \longrightarrow A$ joukosta merkintää A^B.

Tässä tutkielmassa luonnollisten lukujen joukolla $\mathbb{N}$ tarkoitetaan joukkoa $\{0, 1, 2, 3, 4, \ldots\}$, ja joukosta $\mathbb{N} \setminus \{0\}$ käytetään merkintää $\mathbb{N}_{>0}$.

Joukon $A^{\mathbb{N}}$ alkiot ovat joukon A *jonoja.* Jokainen funktio $\mathbf{a} \in A^{\mathbb{N}}$ ilmoittaa kuinka luonnolliset luvut kuvautuvat joukon A alkioiksi. Asettamalla nämä alkiot peräkkäin saadaan jono

$$\mathbf{a} = \mathbf{a}(0)\mathbf{a}(1)\mathbf{a}(2)\mathbf{a}(3) \cdots = a_0 a_1 a_2 a_3 \cdots .$$

Alkioita a_i kutsutaan jonon $\mathbf{a}$ *jäseniksi.* Selkeyden vuoksi jonon $\mathbf{a}$ jäsenet voidaan erotella toisistaan pilkuilla.

Voidaan merkitä myös $\mathbf{a} = (a_n)$. Tätä merkintätapaa käytetään erityisesti silloin kun halutaan ilmoittaa yleisen jäsenen lauseke eksplisiittisesti. Esimerkiksi neliölukujen jono on $(n^2) = 0, 1, 4, 9, 16, \ldots$. Selkeyden vuoksi alaindeksiin voidaan merkitä minkä muuttujan suhteen jonon jäsenet on indeksöity. Esimerkiksi luvulla $k \in \mathbb{N}_{>0}$ jaolliset luvut muodostavat lukujonon $\mathbf{a}_k = (ki)_{i \in \mathbb{N}}$.

Joskus on selkeyden kannalta parasta käyttää merkintätapaa

$$\mathbf{a}_n = a_n.$$

Olkoon Σ jokin aakkosto. Joukon $\Sigma^{\mathbb{N}}$ alkiot ovat aakkoston Σ *oikealle äärettömiä sanoja* tai *aakkoston* Σ *jonoja.*

Jonoa $\mathbf{a} = (a_n)$ kutsutaan *lopulta jaksolliseksi* mikäli on olemassa sellaiset kokonaisluvut $p \geq 1$ ja $m \in \mathbb{N}$ että $a_i = a_{i+p}$ aina kun $i \geq m$. Lukua p kutsutaan jonon $\mathbf{a}$ *jaksoksi* tai *jaksonpituudeksi.* Jos voidaan valita $m = 0$, jono $\mathbf{a}$ on *jaksollinen.*

Olkoon $w \in \Sigma^* \setminus \{\varepsilon\}$ jokin sana. Käytetään jaksollisesta äärettömästä sanasta $wwwwwwww\ldots$ merkintää w^∞.

2.1.3 Morfismit ja morfiset sanat

Olkoot Σ ja Δ aakkostoja. Funktio $f : \Sigma^* \longrightarrow \Delta^*$ on *morfismi*, jos

$$f(uv) = f(u)f(v)$$

aina kun u ja v ovat aakkoston Σ sanoja. Morfismi määräytyy yksikäsitteisesti siitä, miten se kuvaa aakkoston Σ symbolit. Funktio $f : \Sigma \longrightarrow \Delta^*$ voidaan nimittäin laajentaa morfismiksi $\Sigma^* \longrightarrow \Delta^*$: jos $w = a_1a_2\cdots a_n \in \Sigma^n$, niin $f(w) = f(a_1)f(a_2)\cdots f(a_n)$, eli sanan w kuva määräytyy symbolien kuvista. Toisaalta jos morfismi $g : \Sigma^* \longrightarrow \Delta^*$ on toinen funktion f laajennus, niin aina kun $w = a_1\cdots a_n \in \Sigma^n$ yhtälöketju

$$g(w) = g(a_1)g(a_2)\cdots g(a_n) = f(a_1)f(a_2)\cdots f(a_n) = f(w)$$

on voimassa. Toisin sanoen $f = g$, eli laajennus on yksikäsitteinen.

Tarkastellaan kuinka morfismi f kuvaa tyhjän sanan. Koska $f(\varepsilon) = f(\varepsilon\varepsilon) = f(\varepsilon)f(\varepsilon)$, yhtälö $|f(\varepsilon)| = |f(\varepsilon)f(\varepsilon)| = 2|f(\varepsilon)|$ on voimassa. Tällöin $|f(\varepsilon)| = 0$, joten $f(\varepsilon) = \varepsilon$.

Määritelmä 2.1. Olkoon $f : \Sigma^* \longrightarrow \Delta^*$ morfismi. Jos on olemassa sellainen luonnollinen luku $k \geq 1$, että $|f(a)| = k$ aina kun $a \in \Sigma$ on symboli, sanotaan että f on *k-uniforminen.* 1-uniformisia morfismeja kutsutaan *koodauksiksi.*

Määritelmä 2.2. Oletetaan, että $\varphi : \Sigma^* \longrightarrow \Sigma^*$ on k-uniformi morfismi. Olkoon $a \in \Sigma$. Oletetaan lisäksi, että on olemassa sellainen sana $x \in \Sigma^{k-1}$, että

$$\varphi(a) = ax.$$

Tällöin sanotaan, että morfismi φ on *symbolia a pidentävä.*

Morfismin $\varphi : \Sigma^* \longrightarrow \Gamma^*$ määrittelyjoukko voidaan laajentaa joukoksi $\Sigma^* \cup \Sigma^{\mathbb{N}}$, jos asetetaan $\varphi(\mathbf{a}) = \varphi(a_0)\varphi(a_1)\varphi(a_2)\varphi(a_3)\cdots$ aina kun $\mathbf{a} = (a_n) \in \Sigma^{\mathbb{N}}$. Jatkossa oletetaan, että tämä laajennus on tehty, eikä sitä mainita eksplisiittisesti. Oletetaan, että $\Gamma = \Sigma$. Sellaista sanaa $\mathbf{x} \in \Sigma^* \cup \Sigma^{\mathbb{N}}$, joka toteuttaa yhtälön $\varphi(\mathbf{x}) = \mathbf{x}$ kutsutaan morfismin φ *kiintopisteeksi.*

Lause 2.3. Olkoon $k \geq 2$ kokonaisluku, ja $\varphi : \Sigma^* \longrightarrow \Sigma^*$ symbolia $a \in \Sigma$ pidentävä k-uniformi morfismi. Olkoon $x \in \Sigma^{k-1}$ se sana, joka toteuttaa yhtälön $\varphi(a) = ax$. Tällöin ääretön sana

$$\varphi^\infty(a) = ax\varphi(x)\varphi^2(x)\varphi^3(a)\varphi^4(a)\cdots$$

on yksikäsitteinen morfismin φ symbolilla a alkava kiintopiste.

Todistus. Ensinnäkin

$$\begin{aligned}\varphi(\varphi^\infty(a)) &= \varphi(ax\varphi(x)\varphi^2(x)\varphi^3(a)\cdots) \\ &= \varphi(a)\varphi(x)\varphi^2(x)\varphi^3(x)\varphi^4(a)\cdots \\ &= ax\varphi(x)\varphi^2(x)\varphi^3(a)\varphi^4(a)\cdots \\ &= \varphi^\infty(a),\end{aligned}$$

joten $\varphi^\infty(a)$ todella on kiintopiste.

Oletetaan sitten, että $\mathbf{b} = b_0b_1b_2\cdots$ on morfismin φ jokin toinen symbolilla a alkava kiintopiste. Olkoon $i \geq 1$ pienin sellainen indeksi, jossa sanat $\varphi^\infty(a)$ ja $\mathbf{b}$ eroavat, ja olkoon $j \geq 0$ suurin sellainen luku, että $k^j < i$. Tällöin sana $\varphi^j(a) \in \Sigma^{k^j}$ on sanan $\mathbf{b}$ etuliite. Mutta nyt $\varphi^{j+1}(a)$ on sanojen $\varphi^\infty(a)$ ja $\varphi(\mathbf{b}) = \mathbf{b}$ yhteinen etuliite. Yhteisen etuliitteen $\varphi^{j+1}(a)$ pituus on $k^{j+1} > i$. Indeksin i määritelmän nojalla sanat $\varphi^\infty(a)$ ja $\mathbf{b}$ kuitenkin eroavat indeksissä i, mikä on ristiriita. Symbolilla a alkava kiintopiste on siis yksikäsitteinen. □

Lause 2.4. Oletetaan, että sana $\mathbf{a} = a_0a_1a_2a_3a_4\cdots$ on k-uniformin morfismin φ kiintopiste. Tällöin

$$\varphi(a_i) = a_{ki}a_{ki+1}\cdots a_{ki+k-1}.$$

Todistus. Koska $\mathbf{a}$ on k-uniformin morfismin φ kiintopiste

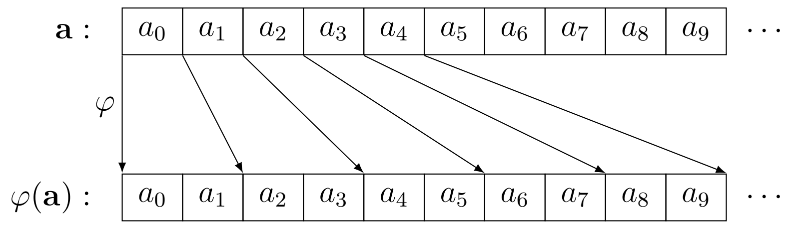
$$\varphi(a_0a_1\cdots a_i) = a_0a_1a_2\cdots a_{k(i+1)-1}.$$

Koska φ on k-uniforminen ja $\varphi(a_0a_1\cdots a_i) = \varphi(a_0a_1\cdots a_{i-1})\varphi(a_i)$, alkio a_i kuvautuu sanan $\varphi(\mathbf{a})$ siksi loppuliitteeksi, jonka pituus on k. Siksi

$$\varphi(a_i) = a_{ki}a_{ki+1}\cdots a_{k(i+1)-1}.$$

□

Kuvassa 2 havainnollistetaan kuinka 2-uniformi morfismi kuvaa kiintopisteensä.



Kuva 2: Kiintopisteen $\mathbf{a} = a_0a_1a_2a_3a_4a_5\cdots$ kuvautuminen 2-uniformisessa morfismissa φ.

2.2 Äärelliset automaatit ja säännölliset kielet

Äärelliset automaatit ovat eräs tapa mallintaa automaatteja tai algoritmeja matemaattisesti. Äärelliset automaatit koostuvat syötenauhasta ja ohjausyksiköstä. Syötenauha syöttää ohjausyksikölle symboleja jostakin symboliaakkostosta. Peräkkäin syötetyt symbolit muodostavat syötesanoja. Ohjausyksikkö sisältää äärellisen määrän muistia, jonka sisältöä mallinnetaan äärellisellä joukolla *tiloja.* Ohjausyksikkö sisältää ohjeet tilan muuttamiseen. Myös näitä ohjeita on äärellinen määrä. Ohjausyksikkö *lukee* syötenauhan symboleja yksi kerrallaan. Ohjausyksikkö vaihtaa sisäisen tilansa luetun symbolin ja kulloisenkin tilan perusteella noudattaen ohjausyksikön ohjeita. Osa tiloista on *hyväksyviä tiloja.* Äärellinen automaatti *hyväksyy* lukemansa sanan, mikäli ohjausyksikön tila on jokin hyväksyvistä tiloista. Muutoin se *hylkää* kyseisen sanan. Äärellisen automaatin $\mathcal{A}$ hyväksymien sanojen joukkoa kutsutaan *automaatin $\mathcal{A}$ hyväksymäksi* tai *tunnistamaksi kieleksi.* Käytetään automaatin $\mathcal{A}$ tunnistamasta kielestä merkintää $L(\mathcal{A})$. *Säännölliset* kielet ovat kieliä jotka ovat jonkin äärellisen automaatin hyväksymiä.

Determinististen äärellisten automaattien (*deterministic finite automata*), epädeterminististen äärellisten automaattien (*non-deterministic finite automata*) ja epädeterminististen äärellisten epsilon-automaattien (*nondeterministic finite automata with ε-transitions*), sekä säännöllisten kielten ja -lausekkeiden perusominaisuudet oletetaan tunnetuiksi. Käytetään näistä käsitteistä lyhennysmerkintöjä DFA, NFA ja ε-NFA. Äärellisten automaattien teoriaa ja säännöllisiä kieliä on käsitelty luentomonisteessa [4].

Lause 2.5. Oletetaan että L, L_1 ja L_2 ovat säännöllisiä kieliä. Olkoon lisäksi K kieli. Tällöin myös seuraavat kielet ovat säännöllisiä:

(a) $L_1 \cup L_2$
(b) $L_1 \cap L_2$
(c) $\bar{L}$
(d) L^*

(e) L^R
(f) L/K

Todistus. Kohta (e) todistetaan artikkelissa [5]. Muut todistetaan luentomonisteessa [4]. □

Lause 2.6. Olkoon $f : \Sigma^* \longrightarrow \Delta^*$ morfismi, ja olkoon $L \subseteq \Sigma^*$ säännöllinen kieli. Tällöin myös $f(L) \subseteq \Delta^*$ on säännöllinen.

Todistus. Luentomonisteessa [4] on esitetty todistus käyttämällä säännöllisiä lausekkeita. Kirjassa [2] on luonnos todistuksesta, jossa käytetään myöhemmin määriteltäviä *transduktoreita.* □

Lause 2.7 (*Kleenen lause*)**.** Kieli L on säännöllinen silloin ja vain silloin kun sille on olemassa säännöllinen lauseke.

Todistus. Todistetaan luentomonisteessa [4]. □

Lause 2.8. Jokainen äärellinen kieli on säännöllinen.

Todistus. Oletetaan, että $L = \{w_1, w_2, \dots, w_n\} \subseteq \Sigma^*$ on äärellinen kieli. Tällöin kielellä L on säännöllinen lauseke $w_1 + w_2 + \cdots + w_n$, joten se on Kleenen lauseen nojalla säännöllinen. □

2.3 Lukujärjestelmät

Tämä alaluku pohjautuu kirjaan [2] ja artikkeliin [6].

2.3.1 Yleiset lukujärjestelmät

Tavallisesti lukuja on tapana merkitä käyttämällä kymmenjärjestelmää. Kymmenjärjestelmässä lukuja merkitään symboliaakkoston $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ äärellisillä tai äärettömillä sanoilla. Kyseisen symboliaakkoston alkioita kutsutaan *numeroiksi.*

Yleisemmin lukujärjestelmällä tarkoitetaan tapaa ilmaista jonkin puolirenkaan S alkioita aakkoston $D \subseteq S$ sanoina.

Määritelmä 2.9. Olkoon S puolirengas ja $\mathbf{u} = (u_i) \in S^{\mathbb{N}}$ jokin jono. Olkoon D puolirenkaan S äärellinen osajoukko. Puolirenkaan S *lukujärjestelmä* $\mathcal{N}$ on kolmikko $\mathcal{N} = (\mathbf{u}, D, R)$. Jonoa $\mathbf{u}$ kutsutaan *kantajonoksi,* ja joukkoa D *numeroiksi.* Kieli $R \subseteq D^*$ on *hyvinmääriteltyjen esitysten joukko* tai *sallittujen esitysten joukko.*

Määritelmä 2.10. Oletetaan, että S on puolirengas. Olkoon $\mathcal{N} = (\mathbf{u}, D, R)$ puolirenkaan S lukujärjestelmä. Määritellään kuvaus

$$[\ ]_{\mathbf{u}} : D^* \longrightarrow S$$

asettamalla

$$[w]_{\mathbf{u}} = \sum_{i=0}^{t} a_i u_i$$

aina kun $w = a_t a_{t-1} \cdots a_1 a_0 \in D^{t+1}$. Kuvausta $[\ ]_{\mathbf{u}}$ kutsutaan lukujärjestelmän $\mathcal{N}$ *arvotukseksi.* Sana w on luvun $[w]_{\mathbf{u}}$ *esitys järjestelmässä* $\mathcal{N}$. Nolla-alkiolla on aina esitys ε.

Määritelmä 2.11. Olkoon $\mathcal{N}$ puolirenkaan S lukujärjestelmä. Jos jokaisella alkiolla $s \in S$ on esitys järjestelmässä $\mathcal{N}$, sanotaan että $\mathcal{N}$ on *täysi.* Jos jokaisella alkiolla $s \in S$ on enintään yksi esitys järjestelmässä $\mathcal{N}$, sanotaan että $\mathcal{N}$ on *yksikäsitteinen.* Jos $\mathcal{N}$ on sekä täysi että yksikäsitteinen, sanotaan että se on *täydellinen.*

2.3.2 Luonnollisten lukujen lukujärjestelmiä

Tässä tutkielmassa tarkastellaan vain puolirengasta $\mathbb{N}$.

Oletetaan, että on kiinnitetty aidosti kasvava kantajono $\mathbf{u} = (u_i)$. Oletetaan lisäksi, että $u_0 = 1$, ja että jono $(v_i) = (u_{i+1}/u_i)_i$ on ylhäältä rajoitettu. Tavoitteena on muodostaa hyvinmääriteltyjen esitysten joukko R ja numerojoukko D niin, että $(\mathbf{u}, D, R)$ on täydellinen. Määritellään ensin niin sanottu *ahne algoritmi* $A_{\mathbf{u}}$, joka saatuaan syötteeksi luvun $n \in \mathbb{N}_{>0}$ toimii seuraavasti:

$t := 0$
while $u_{t+1} \leq n$ **do**
  $t := t + 1$
end while
$n_t := n$
for $i = t$ to 0 **do**
  $a_i := \lfloor n_i / u_i \rfloor$
  $n_{i-1} := n_i - a_i u_i$
end for
return $a_t a_{t-1} \cdots a_1 a_0$

Toisin sanoen algoritmi $A_{\mathbf{u}}$ etsii suurimman sellaisen indeksin t, että $u_t \leq n$. Tämän jälkeen se määrittelee luvut $a_0, a_1, \ldots, a_t$ ja $n_{-1}, n_0, \ldots, n_t$ käyttämällä rekursioyhtälöä

$$\begin{cases} n_t = n \\ a_i = \lfloor n_i/u_i \rfloor \\ n_{i-1} = n_i - a_i u_i. \end{cases}$$

Osoitetaan induktiolla, että $n_j = \sum_{i=0}^{j} a_i u_i$ aina kun $j \geq 0$. Ensinnäkin $a_0 = \lfloor n_0/u_0 \rfloor = n_0/u_0$, sillä $u_0 = 1$. Siten $n_0 = a_0 u_0$. Oletetaan sitten, että $n_j = \sum_{i=0}^{j} a_i u_i$. Tällöin koska $n_j = n_{j+1} - a_{j+1} u_{j+1}$,

$$n_{j+1} = a_{j+1} u_{j+1} + n_j = a_{j+1} u_{j+1} + \sum_{i=0}^{j} a_i u_i = \sum_{i=0}^{j+1} a_i u_i. \tag{1}$$

Erityisesti $n = n_t = \sum_{i=0}^{t} a_i u_i = [a_t a_{t-1} \cdots a_1 a_0]_{\mathbf{u}}$, joten ahneen algoritmin tuottama sana $A_{\mathbf{u}}(n) = a_t a_{t-1} \cdots a_1 a_0$ on luvun n esitys.

Asetetaan nyt hyvinmääriteltyjen esitysten joukoksi

$$R = \{A_{\mathbf{u}}(n) \mid n \in \mathbb{N}_{>0}\} \cup \{\varepsilon\},$$

joka sisältää ahneen algoritmin tuottamat esitykset ja lisäksi luvun 0 esityksen ε. Jokaisella luonnollisella luvulla on siis esitys.

Koska jono $(v_i) = (u_{i+1}/u_i)_i$ on ylhäältä rajoitettu, supremum $c = \sup_{i \geq 0} v_i$ on olemassa. Olkoon $i < t$. Tällöin

$$\begin{aligned} n_i &= n_{i+1} - a_{i+1} u_{i+1} \\ &= n_{i+1} - \left\lfloor \frac{n_{i+1}}{u_{i+1}} \right\rfloor u_{i+1} \\ &= \left(\frac{n_{i+1}}{u_{i+1}} - \left\lfloor \frac{n_{i+1}}{u_{i+1}} \right\rfloor \right) u_{i+1} \\ &< u_{i+1}, \end{aligned}$$

sillä lattiafunktion määritelmän nojalla $\frac{n_{i+1}}{u_{i+1}} - \left\lfloor \frac{n_{i+1}}{u_{i+1}} \right\rfloor < 1$. Näin ollen $a_i = \lfloor n_i/u_i \rfloor \leq n_i/u_i < u_{i+1}/u_i$. Tällöin $a_i < c \leq \lceil c \rceil$. Numerojoukoksi voidaan siis valita $D = \Sigma_{\lceil c \rceil} = \{0, 1, \ldots, \lceil c \rceil - 1\}$. Muodostettu lukujärjestelmä $(\mathbf{u}, D, R)$ on täysi.

Lause 2.12. Onkoon $1 = u_0 < u_1 < u_2 < \cdots$ aidosti kasvava jono. Tällöin jokaisella luonnollisella luvulla $n \in \mathbb{N}_{>0}$ on yksikäsitteinen esitys muodossa

$$n = \sum_{i=0}^{s} a_i u_i,$$

jossa $a_s \neq 0$, ja joka toteuttaa epäyhtälön

$$a_0u_0 + a_1u_1 + \cdots + a_iu_i < u_{i+1} \tag{2}$$

aina kun $i \geq 0$.

Todistus. Muodostetaan esitys $n = \sum_{i=0}^{s} a_iu_i$ käyttämällä ahnetta algoritmia. Yhtälön (1) nojalla $n_j = \sum_{i=0}^{j} a_iu_i$, ja toisaalta tiedetään että $n_j < u_{j+1}$. Riittää osoittaa, että esitys on yksikäsitteinen. Oletetaan sitä vastoin, että esitys ei ole yksikäsitteinen. Toisin sanoen oletetaan, että sillä on myös toinen esitys

$$n = \sum_{i=0}^{t} b_iu_i.$$

Olkoon $r = \max\{s, t\}$. Jos tarpeen vaatiessa lisätään lyhyemmän esityksen alkuun nollia, saadaan kaksi esitystä

$$n = \sum_{i=0}^{r} a_iu_i = \sum_{i=0}^{r} b_iu_i$$

Olkoon j suurin niistä indekseistä, joissa $a_j \neq b_j$. Tällöin summat $\sum_{i=j+1}^{r} a_iu_i$ ja $\sum_{i=j+1}^{r} b_iu_i$ ovat yhtäsuuret, joten $\sum_{i=0}^{j} a_iu_i = \sum_{i=0}^{j} b_iu_i$. Vähentämällä edelliset summalausekkeet toisistaan saadaan yhtälö $\sum_{i=0}^{j}(a_i - b_i)u_i = 0$, joka voidaan muokata muotoon

$$(a_j - b_j)u_j = \sum_{i=0}^{j-1}(b_i - a_i)u_i.$$

Koska $a_j \neq b_j$, voidaan yleisyyttä loukkaamatta olettaa, että $a_j - b_j > 0$. Tällöin

$$u_j \leq (a_j - b_j)u_j = \sum_{i=0}^{j-1}(b_i - a_i)u_i \leq \sum_{i=0}^{j-1} b_iu_i,$$

mikä on ristiriidassa epäyhtälön (2) kanssa. Esitys on siis yksikäsitteinen.

□

2.3.3 k-kantajärjestelmät

Määritelmä 2.13. Oletetaan, että $k \geq 2$ on kokonaisluku. Käytetään merkintää

$$\Sigma_k = \{0, 1, \ldots, k-1\}.$$

Määritelmä 2.14. Olkoon $k \geq 2$ kokonaisluku. Käytetään merkintää

$$C_k = (\Sigma_k \setminus \{0\})\Sigma_k^* \cup \{\varepsilon\}.$$

Seuraus 2.15. Olkoon $k \geq 2$ kokonaisluku. Valitaan kantajonoksi $\mathbf{u}_k = (k^i)$. Tällöin lukujärjestelmä $(\mathbf{u}_k, \Sigma_k, C_k)$ on täydellinen.

Todistus. Koska $\sup_{i \in \mathbb{N}} k^{i+1}/k^i = k$, ahneen algoritmin tuottama esitys on aina aakkoston Σ_k sana.

Oletetaan, että $a_t a_{t-1} \cdots a_1 a_0 \in (\Sigma_k \setminus \{0\})\Sigma_k^*$. Tällöin

$$\sum_{i=0}^{t} a_i k^i \leq (k-1) \sum_{i=0}^{t} k^i = k^{t+1} - 1 < k^{t+1},$$

joten epäyhtälö (2) on aina voimassa. Sana $a_t a_{t-1} \cdots a_1 a_0$ on luvun $n = \sum_{i=0}^{t} a_i k^i$ esitys, joten lauseen 2.12 nojalla se yhtyy ahneen algoritmin antamaan esitykseen. Siten

$$\{A_{\mathbf{u}_k}(n) \mid n \in \mathbb{N}_{>0}\} \cup \{\varepsilon\} = (\Sigma_k \setminus \{0\})\, \Sigma_k^* \cup \{\varepsilon\} = C_k.$$

Esityksen yksikäsitteisyyden nojalla luonnollisten lukujen lukujärjestelmä $(\mathbf{u}_k, \Sigma_k, C_k)$ on täydellinen.

□

Määritelmä 2.16. Seurauksen 2.15 nojalla ahne algoritmi tuottaa jokaiselle kokonaisluvulle $n > 0$ yksikäsitteisen esityksen muodossa $n = \sum_{i=0}^{t} a_i k^i$, missä $a_t \neq 0$. Tätä esitystä kutsutaan luvun n *kanoniseksi k-kantaesitykseksi*, ja siitä käytetään merkintää $(n)_k$. Määritellään $(0)_k = \varepsilon$.

Jos höllennetään vaatimusta lukujärjestelmän yksikäsitteisyydestä, voidaan hyvinmääriteltyjen esitysten joukoksi valita myös Σ_k^*. Lukujärjestelmä $(\mathbf{u}_k, \Sigma_k, \Sigma_k^*)$ on täysi, muttei täydellinen. Luvun n kaikki esitykset ovat muotoa $0^i(n)_k$, missä $i \in \mathbb{N}$. Lukujärjestelmää $(\mathbf{u}_k, \Sigma_k, \Sigma_k^*)$ kutsutaan *tavalliseksi k-kantajärjestelmäksi.*

Oletetaan että $D \subseteq \mathbb{Z}$ on joukko, joka sisältää aakkoston Σ_k. Lukujärjestelmää $(\mathbf{u}_k, D, D^*)$ kutsutaan *k-kantajärjestelmäksi.*

Määritellään funktio $[\ ]_k : \mathbb{N}^* \longrightarrow \mathbb{N}$ asettamalla

$$[a_t a_{t-1} \cdots a_1 a_0]_k = \sum_{i=0}^{t} a_i k^i.$$

Jos u ja v ovat aakkoston D sanoja,

$$[uv]_k = \sum_{i=0}^{|u|-1} u_i k^{i+|v|} + \sum_{i=0}^{|v|-1} v_i k^i = [u]_k \cdot k^{|v|} + [v]_k.$$

Luvuilla voi k-kantajärjestelmässä olla useita eri esityksiä, vaikka ne alkaisivatkin nollasta eriävällä luvulla. Esimerkiksi

$$15 = [(15)_2]_2 = [1111]_2 = [311]_2 = [303]_2 = [63]_2 = [71]_2.$$

Oletetaan, että $w \in D^*$. Jos $[w]_k = n$, sanotaan että w on luvun *k-kantaesitys,* tai että w *on luvun* n *esitys kannassa* k. Jos $D = \Sigma_k$, voidaan sanoa, että w on luvun n *tavallinen k-kantaesitys.*

Tähän mennessä olemme tarkastelleet lukujen *eniten merkitsevästä numerosta alkavia* k-kantaesityksiä. Luvun k-kantaesityksen ensimmäinen numero on nimittäin arvotuksessa suurimman kantaluvun potenssin kerroin. Voidaan määritellä myös *vähiten merkitsevästä numerosta alkavat* k-kantajärjestelmät: luvun $n \in \mathbb{N}$ vähiten merkitsevästä numerosta alkava k-kantaesitys on sellainen sana $w = a_0a_1 \cdots a_n$, että $n = \sum_{i=0}^{n} a_i k^i$. Toisin sanoen vähiten merkitsevästä numerosta alkava k-kantaesitys on eniten merkitsevästä numerosta alkavan k-kantaesityksen peilaus. Jokaisella luonnollisella luvulla on siis vähiten merkitsevästä numerosta alkava k-kantaesitys.

Määritelmä 2.17. Olkoon $\nu : D^* \longrightarrow \Sigma_k^*$ funktio. Jos $[w]_k = [\nu(w)]_k$ aina kun $w \in D^*$, sanotaan että ν on *normalisaatio.* Normalisaatio siis kuvaa luvun n esityksen sen tavalliseksi k-kantaesitykseksi.

3 Äärellisten automaattien laajennuksia

Äärelliset automaatit saavat syötteekseen syöteaakkoston sanan, jonka ne joko hyväksyvät tai hylkäävät. Tässä luvussa tarkastelemme kahta eri äärellisten automaattien muunnelmaa, *deterministisiä äärellisiä funktioautomaatteja* ja *äärellisiä transduktoreita*. Nekin saavat syötteeksi sanan, mutta äärellisistä automaateista poiketen ne tulostavat joko symbolin tai tulostesanan. Ne ovat yksi tapa mallintaa sellaisia funktioita, joiden laskemiseen vaaditaan vain äärellinen määrä muistia ja toimintaohjeita, ja joiden määrittelyjoukko on syötesanojen joukko.

3.1 Äärelliset funktioautomaatit

Tarkastellaan ensin niin sanottuja *deterministisiä äärellisiä funktioautomaatteja.* Ne toimivat sanaa lukiessaan vastaavalla tavalla kuin deterministiset äärelliset automaatit, mutta luettuaan koko syötteen ne tulostavat symbolin. Äärelliset funktioautomaatit määrittelevät funktion syötesanojen joukolta symboliaakkostolle. Tällaiset funktiot muodostavat *automaattisten funktioiden* luokan.

3.1.1 Äärellisen funktioautomaatin määritelmä

Määritelmä 3.1. *Deterministinen äärellinen funktioautomaatti* tai lyhyemmin *äärellinen funktioautomaatti* tai *funktioautomaatti* on kuusikko

$$\mathcal{M} = (Q, \Sigma, \delta, q_0, \Delta, \tau).$$

Äärellinen joukko Q on funktioautomaatin $\mathcal{M}$ *tilajoukko*, Σ on *syöteaakkosto*, $\delta : Q \times \Sigma \to Q$ on *transitiofunktio* ja $q_0 \in Q$ on *alkutila.* Näillä on sama merkitys kuin determinististen äärellisten automaattien vastaavilla komponenteilla. Transitiofunktio δ määrää sen, kuinka tila muuttuu funktioautomaatin $\mathcal{M}$ lukiessa yksittäisiä symboleja. Jos $\mathcal{M}$ on tilassa $q \in Q$ ja se lukee symbolin $a \in \Sigma$, se siirtyy tilaan $\delta(q, a) = r \in Q$. Tällaista siirtymää tilasta toiseen kutsutaan *transitioksi* tai tarkemmin *a-transitioksi.* Jos tilasta q on a-transitio tilaan r voidaan merkitä $q \xrightarrow{a} r$.

Oletetaan, että $a_1, a_2, \ldots, a_n \in \Sigma$ ovat symboleja ja $w = a_1 a_2 \cdots a_n$ on sana. Kun $\mathcal{M}$ lukee sanan w, se lukee kunkin sanan w symbolin yksi kerrallaan järjestyksessä ja muuttaa tilaansa transitiofunktion mukaan

$$q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} q_2 \xrightarrow{a_3} \cdots \xrightarrow{a_{n-1}} q_{n-1} \xrightarrow{a_n} q_n.$$

Transitiofunktio δ voidaan laajentaa funktioksi $\delta : Q \times \Sigma^* \to Q$ käyttämällä indukiota samalla tavoin kuin deterministisillä automaateilla:

1. Aina kun $q \in Q$, määritellään ε-transitio

$$\delta(q, \varepsilon) = q$$

2. Aina kun $w = va \in \Sigma^*$, $a \in \Sigma$ on sana

$$\delta(q, w) = \delta(\delta(q, v), a).$$

Jos $\delta(q, w) = r$, voidaan merkitä $q \overset{w}{\rightsquigarrow} r$.

Äärellisistä automaateista poiketen funktioautomaateilla ei ole lopputilojen joukkoa, vaan edellisten lisäksi *tulostejoukko* Δ ja *tulostusfunktio* $\tau : Q \to \Delta$. Luettuaan sanan w funktioautomaatti päätyy tilaan $\delta(q_0, w)$, ja se tulostaa tulosteen $\tau(\delta(q_0, w))$.

Transitiofunktion laajennuksen induktiivinen määritelmä on käytännössä varsin kömpelö, joten osoitetaan seuraava lause:

Lause 3.2. Olkoon $\mathcal{M} = (Q, \Sigma, \delta, q_0, \Delta, \tau)$ funktioautomaatti ja $q \in Q$ jokin tila. Oletetaan lisäksi, että u ja v ovat aakkoston Σ sanoja. Tällöin

$$\delta(q, uv) = \delta(\delta(q, u), v).$$

Todistus. Olkoon $u \in \Sigma^*$ mielivaltainen sana. Todistetaan induktiolla sanan $v \in \Sigma^*$ suhteen, että yhtälö

$$\delta(q, uv) = \delta(\delta(q, u), v) \tag{3}$$

on voimassa.

1° Olkoon $v = \varepsilon$. Tällöin

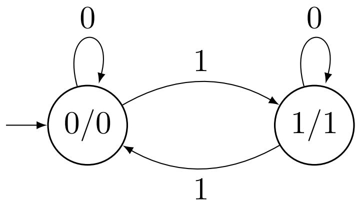
$$\delta(q, uv) = \delta(q, u) = \delta(\delta(q, u), \varepsilon) = \delta(\delta(q, u), v).$$

2° Oletetaan, että yhtälö (3) on voimassa sanalle $v \in \Sigma^*$. Olkoon $a \in \Sigma$ symboli. Tällöin

$$\begin{aligned}\delta(q, uva) &= \delta(\delta(q, uv), a) \\ &= \delta(\delta(\delta(q, u), v), a) \\ &= \delta(\delta(q, u), va).\end{aligned}$$

Yhtälö (3) on siis voimassa aina kun $v \in \Sigma^*$, ja koska u oli mielivaltainen, se on voimassa aina kun u ja v ovat aakkoston Σ sanoja.

□



Kuva 3

Määritelmä 3.3. Oletetaan, että $\mathcal{M} = (Q, \Sigma, \delta, q_0, \Delta, \tau)$ on äärellinen funktioautomaatti. Funktioautomaatti $\mathcal{M}$ määrää funktion $f_{\mathcal{M}} : \Sigma^* \to \Delta$, joka kuvaa sanan $w \in \Sigma^*$ alkioksi $\tau(\delta(q_0, w))$. Funktiota $f_{\mathcal{M}}$ kutsutaan *funktioautomaatin $\mathcal{M}$ määräämäksi funktioksi.* Sanotaan myös, että *funktioautomaatti $\mathcal{M}$ laskee funktion $f_{\mathcal{M}}$.*

Huomautus 3.4. Äärelliset automaatit voidaan tulkita funktioautomaateiksi, jotka tulostavat luvun 1 silloin kun syötesana on kyseisen automaatin tunnistamassa kielessä, ja luvun 0 muulloin. Tällöin funktioautomaatin tuloste tulkitaan väitteen ”annettu syöte kuuluu automaatin määräämään säännölliseen kieleen” totuusarvoksi. Kyseinen funktioautomaatti saadaan äärellisestä automaatista asettamalla tulostusfunktioksi automaatin lopputilajoukon F karakteristisen funktion $\tau = \chi_F$.

Äärellisiä funktioautomaatteja voi kuvata suunnatuilla graafeilla lähes samaan tapaan kuin äärellisiä automaatteja. Pistejoukon Q sijasta pistejoukkona käytetään joukkoa $\{q/\tau(q) \,|\, q \in Q\}$. Suunnattu viiva $(q/\tau(q), a, r/\tau(r))$ kuuluu graafiin mikäli $\delta(q, a) = r$.

Esimerkki 3.5. Muodostetaan funktioautomaatti, joka saa syötteenään binäärisanoja, eli aakkoston $\{0, 1\}$ sanoja, ja joka laskee syötesanan bittien summan modulo 2. Asetetaan tilajoukoksi joukko $\{0, 1\}$. Kulloinenkin tila ilmoittaa jo luetun sanan etuliitteen bittisumman pariteetin. Alkutilaksi asetetaan 0, sillä bittisumma on 0 kun yhtäkään bittiä ei vielä ole luettu. Bitin 0 lukeminen ei muuta tilaa, kun taas bitin 1 lukeminen muuttaa. On muodostettu automaatti $(\{0, 1\}, \{0, 1\}, \delta, 0, \{0, 1\}, \tau)$, missä $\delta(0, 0) = \delta(1, 1) = 0$, $\delta(0, 1) = \delta(1, 0) = 1$, ja $\tau : \{0, 1\} \to \{0, 1\}$ on identiteettifunktio. Kyseistä funktioautomaattia kuvaa kuvan 3 graafi.

3.1.2 Funktioautomaattien määräämät kielet

Äärelliset automaatit jakavat syötesanat kahteen kieleen, kyseisen automaatin hyväksymään kieleen ja sen komplementtiin. Säännöllisten kielten määritelmän nojalla äärellisen automaatin tunnistama kieli on säännöllinen, ja lauseen 2.5 nojalla äärellisen automaatin tunnistaman kielen komplementti on säännöllinen. Olkoon $\mathcal{A}$ äärellinen automaatti ja $\mathcal{M}$ huomautuksen 3.4 tulkinnan mukainen funktioautomaatti. Tällöin kielet $f_{\mathcal{M}}^{-1}\big(\{1\}\big) = L(\mathcal{A})$ ja $f_{\mathcal{M}}^{-1}\big(\{0\}\big) = \overline{L(\mathcal{A})}$ ovat säännöllisiä.

Tämä havainto yleistyy. Määritellään automaatin hyväksymän kielen yleistys funktioautomaateille.

Määritelmä 3.6. Olkoon $\mathcal{M} = (Q, \Sigma, \delta, q_0, \Delta, \tau)$ funktioautomaatti. Oletetaan, että $d \in \Delta$. Määritellään kieli

$$L_d(\mathcal{M}) = f_{\mathcal{M}}^{-1}\big(\{d\}\big) = \{w \in \Sigma^* \mid \tau(\delta(q_0, w)) = d\}.$$

Lause 3.7. Olkoon $\mathcal{M} = (Q, \Sigma, \delta, q_0, \Delta, \tau)$ äärellinen funktioautomaatti. Tällöin kieli $L_d(\mathcal{M})$ on säännöllinen aina kun $d \in \Delta$.

Todistus. Muodostetaan jokaiselle tilalle $q \in Q$ DFA $\mathcal{A}_q = (Q, \Sigma, \delta, q_0, \{q\})$, joka hyväksyy kaikki ne sanat jotka luettuaan $\mathcal{M}$ päätyy tilaan q. Tällöin

$$L_d(\mathcal{M}) = \bigcup_{q \in \tau^{-1}(\{d\})} L(\mathcal{A}_q),$$

sillä yhtälön oikea puoli koostuu kaikista niistä sanoista, jotka luettuaan $\mathcal{M}$ on missä tahansa sellaisessa tilassa jonka τ kuvaa alkioksi d. Säännöllisten kielten $L(\mathcal{A}_q)$ unioni on säännöllinen, joten $L_d(\mathcal{M})$ on säännöllinen.

□

Äärelliset funktioautomaatit siis osittavat syöteavaruuden Σ^* säännöllisiin osiin. Toisaalta, mikäli Σ^* voidaan osittaa säännöllisiin kieliin, voidaan muodostaa äärellinen funktioautomaatti joka määrää kyseisen osituksen.

Lause 3.8. Oletetaan, että Σ ja $\Delta = \{d_1, \ldots, d_n\}$ ovat äärellisiä aakkostoja. Oletetaan lisäksi, että kielet $L_1, \ldots, L_n$ ovat säännöllisiä ja pareittain alkiovieraita, ja että

$$\Sigma^* = \bigcup_{1 \leq i \leq n} L_i.$$

Tällöin on olemassa sellainen funktioautomaatti $\mathcal{M} = (Q, \Sigma, \delta, q_0, \Delta, \tau)$, että $L_{d_i}(\mathcal{M}) = L_i$ aina kun $1 \leq i \leq n$.

Todistus. Muodostetaan jokaista indeksiä $i \in \{1, \ldots, n\}$ kohti säännöllisen kielen L_i tunnistava deterministinen äärellinen automaatti $\mathcal{A}_i = (Q_i, \Sigma, \delta_i, q_i, F_i)$. Asetetaan funktioautomaatin $\mathcal{M}$ tilajoukoksi $Q = Q_1 \times \cdots \times Q_n$. Valitaan alkutilaksi $q_0 = (q_1, \ldots, q_n)$. Tarkoituksena on, että funktioautomaatti $\mathcal{M}$ laskee kaikkien automaattien $\mathcal{A}_1, \ldots, \mathcal{A}_n$ tilan samanaikaisesti. Siksi määritellään

$$\delta((q^1, \ldots, q^n), a) = (\delta_1(q^1, a), \delta_2(q^2, a), \ldots, \delta_n(q^n, a))$$

aina kun $(q^1, \ldots, q^n) \in Q$ ja $a \in \Sigma$.

Määritellään funktio $\tau : Q \to \Sigma$ asettamalla jokaista $(q^1, \ldots, q^n) \in Q$ kohti

$$\tau(q^1, \ldots, q^n) = \begin{cases} d_j, & \text{jos on tarkalleen yksi sellainen indeksi } j \text{, että } q^j \in F_j, \\ d_1 & \text{muulloin,} \end{cases}$$

ja asetetaan τ funktioautomaatin $\mathcal{M}$ tulostusfunktioksi.

Koska $L_1, \ldots, L_n$ muodostavat kielen Σ^* osituksen, jokaista syötesanaa $w \in \Sigma^*$ kohti tarkalleen yksi automaateista $\mathcal{A}_1, \ldots \mathcal{A}_n$ päätyy lopputilaan. Olkoon $j \in \{1, \ldots, n\}$ mielivaltainen. Tällöin seuraava päättelyketju on voimassa:

$$\begin{aligned} w \in L_j &\iff \delta_j(q_j, w) \in F_j \\ &\iff \delta((q_1, \ldots, q_n), w) = (r_1, \ldots, r_n) \text{ ja } r_i \in F_i \text{ vain kun } i = j \\ &\iff \tau(\delta((q_1, \ldots, q_n), w)) = d_j \\ &\iff w \in L_{d_j}(\mathcal{M}). \end{aligned}$$

Toisin sanoen $L_j = L_{d_j}(\mathcal{M})$.

$\square$

3.1.3 Automaattiset funktiot

Määritellään seuraavaksi *automaattiset funktiot*, eli funktiot jotka voidaan laskea jollakin funktioautomaatilla:

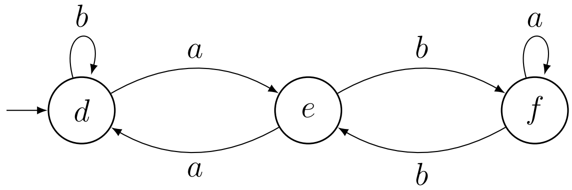
Määritelmä 3.9. Olkoon $f : \Sigma^* \to \Delta$ funktio. Jos on olemassa sellainen funktioautomaatti $\mathcal{M} = (Q, \Sigma, \delta, q_0, \Delta, \tau)$, että $f \equiv f_{\mathcal{M}}$, sanotaan että f on *automaattinen.*

Määritelmä 3.10. Olkoon $f : \Sigma^* \to \Delta$ funktio. Määritellään funktion f *peilausfunktio* $f^R : \Sigma^* \longrightarrow \Delta$ asettamalla

$$f^R(w) = f(w^R)$$

aina kun $w \in \Sigma^*$.

Tarkastellaan suunnatun graafin

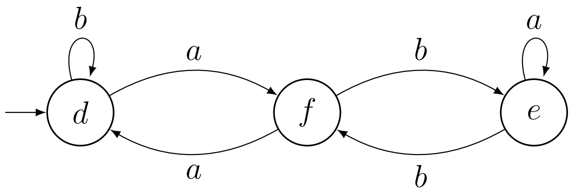


kuvaamaa funktioautomaattia $\mathcal{M}$. Selkeyden vuoksi solmut on nimetty vain tulostusfunktion arvoilla. Kun funktioautomaatti lukee sanan *abab*, se tulostaa symbolin e. Kuljetaan solmusta e alkaen graafin viivoja väärään suuntaan sanan $abab^R = baba$ ilmoittamassa järjestyksessä. Tällöin päädytään solmuun d, joka esittää funktioautomaatin alkutilaa.

Toimitaan kuten edellä, mutta lisäksi nimetään jokainen solmu jonka kautta kuljetaan nimellä e. Tällöin alkutilan uudeksi nimeksi tulee e. Voidaan ajatella, että tieto tulostusfunktion arvosta e voidaan kuljettaa alkutilaa esittävään solmuun.

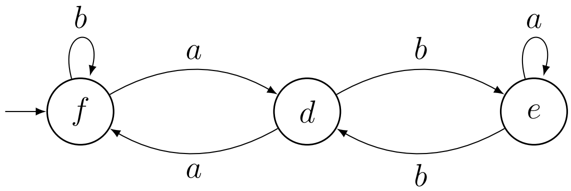
Kuvitellaan, että emme tiedä mihin tilaan funktioautomaatti $\mathcal{M}$ päätyy kun se lukee sanan *abab*, mutta tunnemme sen tulostusfunktion. Siirretään jokaisen solmun nimi niihin solmuihin, joihin päästään siirtymällä sanan *baba* symbolien ilmoittamia viivoja väärän suuntaan. Tällöin alkutilaa kuvaavan solmun uusi nimi ilmoittaa mihin tilaan funktioautomaatti päätyy lukiessaan sanan *abab*.

Aloitetaan siirtämällä jokaisen solmun nimi siihen solmuun, johon pääsee kulkemalla symbolilla b merkittyä viivaa väärään suuntaan. Saadaan seuraava graafi:



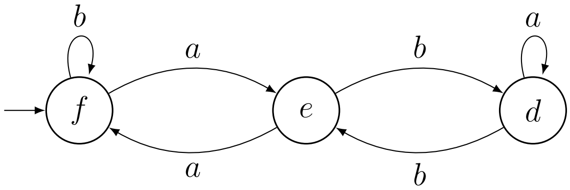
.

Siirretään tämän graafin solmujen nimet nyt siihen solmuun, johon pääsee kulkemalla symbolilla a merkittyä viivaa väärään suuntaa. Saadaan graafi



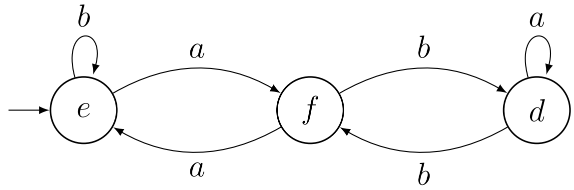
.

Siirretään solmujen nimet uudelleen symbolilla b merkittyjä viivoja pitkin jolloin saadaan graafi



,

josta saadaan edelleen graafi



siirtämällä nimet symbolilla a merkittyjä viivoja pitkin. Nyt alkutilan uusi nimi on e eli se symboli, jonka alkuperäinen funktioautomaatti tulostaa luettuaan sanan *abab*.

Hyödynnetään tätä ajatusta seuraavan lauseen todistuksessa:

Lause 3.11. Olkoon $f : \Sigma^* \longrightarrow \Delta$ automaattinen funktio. Tällöin myös funktion f peilausfunktio f^R on automaattinen.

Todistus. Olkoon $\mathcal{M} = (Q, \Sigma, \delta, q_0, \Delta, \tau)$ funktioautomaatti, joka laskee funktion f. Muodostetaan funktioautomaatti $\mathcal{M}'$, joka laskee funktion f^R. Asetetaan funktioautomaatin $\mathcal{M}'$ tilajoukoksi Δ^Q eli funktioiden $Q \to \Delta$ joukko. Jokainen funktio $g \in \Delta^Q$ vastaa tapaa nimetä funktioautomaattia $\mathcal{M}$ kuvaavan graafin solmut käyttämällä tulosteaakkoston symboleja, voidaan nimittäin tulkita, että solmu $q \in Q$ on nimetty symbolilla $g(q)$.

Asetetaan alkutilaksi funktio τ. Määritellään transitiofunktio $\delta' : \Delta^Q \times \Sigma \longrightarrow \Delta^Q$ seuraavasti: oletetaan, että $g \in \Delta^Q$. Olkoon $h \in \Delta^Q$ funktio, joka saadaan asettamalla $h(q) = g(\delta(q, a))$ aina kun $q \in Q$. Olkoon $\delta'(g, a) = h$. Funktio h saadaan funktiosta g siirtämällä jokaisen solmun nimi symbolilla a merkittyjä viivoja vastakkaiseen suuntaan.

Määritellään tulostusfunktio $\tau' : \Delta^Q \longrightarrow \Delta$ asettamalla

$$\tau'(g) = g(q_0)$$

aina kun $g \in \Delta^Q$. Tulostusfunktio tulostaa siis aina alkutilan nimen.

Osoitetaan induktiolla, että aina kun w on aakkoston Σ sana $\delta'(\tau, w) = h$, missä h on funktio, joka on määritelty asettamalla $h(q) = \tau(\delta(q, w^R))$.

1° Olkoon $w = \varepsilon$. Tällöin

$$\delta'(\tau, w) = \delta'(\tau, \varepsilon) = \tau.$$

Aina kun $q \in Q$, $\tau(q) = \tau(\delta(q, \varepsilon^R))$, joten väite on voimassa.

2° Oletetaan, että väite pätee sanalle $w \in \Sigma^*$, eli että $\delta'(\tau, w) = g$, missä g on funktio jolle pätee $g(q) = \tau(\delta(q, w^R))$ aina kun $q \in Q$. Olkoon $a \in \Sigma$ symboli. Tällöin

$$\delta'(\tau, wa) = \delta'(\delta'(\tau, w), a) = \delta'(g, a) = h,$$

missä h on funktio, joka kuvaa alkion $q \in Q$ alkioksi $g(\delta(q, a))$. Tällöin

$$\begin{aligned} h(q) &= g(\delta(q, a)) \\ &= \tau(\delta(\delta(q, a), w^R)) \\ &= \tau(\delta(q, aw^R)) \\ &= \tau(\delta(q, (wa)^R)). \end{aligned}$$

Väite pätee siis myös sanalle wa, eli induktioaksiooman nojalla kaikille kielen Σ^* sanoille.

Oletetaan, että $w \in \Sigma^*$ on sana. Tällöin

$$\tau'(\delta'(\tau, w)) = \tau'(g) = g(q_0) = \tau(\delta(q_0, w^R)) = f(w^R),$$

joten funktioautomaatti $\mathcal{M}' = (\Delta^Q, \Sigma, \delta', \tau, \Delta, \tau')$ laskee funktion f^R.

$\square$

3.2 Äärelliset transduktorit

3.2.1 Äärelliset transduktorit ja μ-transduktorit

Toinen äärellisten automaattien laajennus on *äärelliset transduktorit.* Kuten äärelliset automaatit ja äärelliset funktioautomaatitkin, äärelliset transduktorit koostuvat ohjausyksiköstä ja syötenauhasta. Ohjausyksikkö sisältää äärellisen määrän tiloja. Lisäksi äärellinen transduktori sisältää tulostenauhan. Äärellinen transduktori lukee syötenauhalta syötesanan, ja se tulostaa tulostenauhalle tulostesanan. Se, miten transduktori muodostaa syötesanalle tulostesanan, voidaan määritellä eri tavoin. Allouche ja Shallit määrittelevät transduktorit kirjassa [2] niin, että ne lukevat syötesanan symboli kerrallaan ja tulostavat tulostenauhalle jokaista symbolia vastaavan sanan, ja syötesanan tulostesana on näiden sanojen katenaatio. Lothaire määrittelee kirjassa [3] transduktorit yleisemmin. Emme kuitenkaan tarvitse tätä yleisempää käsitettä vaan esittelemme ainoastaan Allouchen ja Shallitin määritelmän, joka vastaa lähinnä Lothairen *sekventaalista transduktoria (sequential transducer).*

Määritelmä 3.12 (Äärellinen transduktori)**.** *Äärellinen transduktori* tai lyhyemmin *transduktori* on kuusikko

$$\mathcal{T} = (Q, \Sigma, \delta, q_0, \Delta, \lambda).$$

Äärellinen joukko Q on transduktorin $\mathcal{T}$ tilajoukko, Σ on syöteaakkosto, $\delta : Q \times \Sigma \longrightarrow Q$ on transitiofunktio, $q_0 \in Q$ on alkutila, Δ on tulosteaakkosto ja $\lambda : Q \times \Sigma \longrightarrow \Delta^*$ on tulostusfunktio.

Kun transduktori $\mathcal{T}$ on tilassa $q \in Q$ ja se lukee symbolin $a \in \Sigma$, se siirtyy transitiofunktion mukaisesti tilaan $\delta(q, a)$. Samalla se tulostaa tulostenauhalle sanan $\lambda(q, a)$. Jos $\delta(q, a) = r$ ja $\lambda(q, a) = b \in \Delta^*$, voidaan merkitä

$$q \xrightarrow{a/b} r.$$

Transitiofunktio δ voidaan laajentaa funktioksi $\delta : Q \times \Sigma^* \longrightarrow Q$ samoin kuin äärelliset automaatit ja funktioautomaatit. Tulostusfunktio voidaan laajentaa funktioksi $Q \times \Sigma^* \longrightarrow \Delta^*$ asettamalla

1. $\lambda(q, \varepsilon) = \varepsilon$ aina kun $q \in Q$.

2. Oletetaan, että $\lambda(q, w)$ on määritelty sanalle $w \in \Sigma^*$. Olkoon $a \in \Sigma$ symboli. Määritellään

$$\lambda(q, wa) = \lambda(q, w)\lambda(\delta(q, w), a)$$

aina kun $q \in Q$.

Lause 3.13. Oletetaan että $\mathcal{T} = (Q, \Sigma, \delta, q_0, \Delta, \lambda)$ on äärellinen transduktori. Tällöin aina kun u ja v ovat kielen Σ^* sanoja ja $q \in Q$ jokin tila

(a) $\delta(q, uv) = \delta(\delta(q, u), v)$, ja

(b) $\lambda(q, uv) = \lambda(q, u)\lambda(\delta(q, u), v)$

Todistus. Kohta (a) todistetaan samoin kuin lause 3.2. Olkoon u mielivaltainen aakkoston Σ sana. Todistetaan kohta (b) induktiolla sanan v suhteen.

1° Oletetaan että $v = \varepsilon$. Tällöin

$$\begin{aligned} \lambda(q, uv) &= \lambda(q, u) \\ &= \lambda(q, u)\underbrace{\lambda(\delta(q, u), \varepsilon)}_{\varepsilon} \\ &= \lambda(q, u)\lambda(\delta(q, u), v), \end{aligned}$$

joten (b) on voimassa.

2° Oletetaan sitten että $\lambda(q, uv) = \lambda(q, u) \cdot \lambda(\delta(q, u), v)$. Olkoon $a \in \Sigma$ symboli. Tällöin $\lambda(q, uva) = \lambda(q, uv) \cdot \lambda(\delta(q, uv), a)$. Induktio-oletuksen nojalla

$$\lambda(q, uv) \cdot \lambda(\delta(q, uv), a) = \lambda(q, u) \cdot \lambda(\delta(q, u), v) \cdot \lambda(\delta(q, uv), a).$$

Koska $\delta(q, uv) = \delta(\delta(q, u), v)$,

$$\lambda(\delta(q, u), v) \cdot \lambda(\delta(q, uv), a) = \lambda(\delta(q, u), v) \cdot \lambda(\delta(\delta(q, u), v), a).$$

Toisaalta tulostusfunktion λ laajennuksen määritelmän nojalla

$$\lambda(\delta(q, u), v) \cdot \lambda(\delta(\delta(q, u), v), a) = \lambda(\delta(q, u), va),$$

joten yhtälö $\lambda(q, uva) = \lambda(q, u)\lambda(\delta(q, u), va)$ on voimassa.

Näin ollen $\lambda(q, uv) = \lambda(q, u)\lambda(\delta(q, u), v)$ aina kun u ja v ovat aakkoston Σ sanoja.

□

Äärellinen transduktori lukee syötesanan symboli kerrallaan. Aina kun se lukee yhden kirjaimen, se tulostaa tulostenauhalle sanan. Jos nämä sanat asetetaan peräkkäin saadaan syötesanaa vastaava tulostesana. Voidaan ajatella, että jokainen transduktori määrittelee funktion syöteaakkoston sanoilta tulosteaakkoston sanoihin.

Määritelmä 3.14. Olkoon $\mathcal{T} = (Q, \Sigma, \delta, q_0, \Delta, \lambda)$ äärellinen transduktori. Määritellään funktio

$$\mathcal{T} : \Sigma^* \longrightarrow \Delta^*$$

asettamalla $\mathcal{T}(w) = \lambda(q_0, w)$ aina kun w on aakkoston Σ sana.

Olkoon $w = w_1 w_2 \cdots w_n \in \Sigma^n \subseteq \Sigma^*$ mielivaltainen sana, ja oletetaan että

$$q_0 \xrightarrow{w_1/v_1} r_1 \xrightarrow{w_2/v_2} r_2 \xrightarrow{w_3/v_3} \cdots \xrightarrow{w_{n-1}/v_{n-1}} r_{n-1} \xrightarrow{w_n/v_n} r_n.$$

Tällöin $\mathcal{T}(w) = v_1 v_2 \cdots v_n \in \Delta^*$.

Myös transduktoireita voi kuvata suunnatuilla graafeilla. Solmujoukkona käytetään tilajoukkoa kuten äärellisten automaattien tapauksessa. Jos $\delta(q, a) = r$ ja $\lambda(q, a) = b$, niin solmusta q on suunnattu viiva nimeltä a/b solmuun r.

Määritelmä 3.15 (μ-transduktori)**.** Oletetaan, että $\mathcal{T} = (Q, \Sigma, \delta, q_0, \Delta, \lambda)$ on transduktori, ja että $\mu : Q \longrightarrow \Delta^*$ funktio. Määritellään *μ-transduktori*

$$\mathcal{T}_\mu = (\mathcal{T}, \mu).$$

Kun μ-transduktori $\mathcal{T}_\mu$ lukee sanan $w \in \Sigma^*$, se toimii samoin kuin transduktori $\mathcal{T}$. Luettuaan sanan w se päätyy tilaan $\delta(q_0, w) = q$, minkä jälkeen se liittää tulostesanan loppuun sanan $\mu(q)$. Toisin sanoen

$$\mathcal{T}_\mu(w) = \mathcal{T}(w)\mu(\delta(q_0, w)).$$

3.2.2 Normalisaation $D^* \longrightarrow \Sigma_k^*$ laskeva μ-transduktori

Oletetaan että D on jokin aakkostom, joka sisältää aakkoston Σ_k. Seuraavaksi osoitetaan, että on olemassa normalisaation $D^* \longrightarrow \Sigma_k^*$ laskeva μ-transduktori, kun sekä syötesanat että tulostesanat tulkitaan luonnollisten lukujen k-kantaesityksiksi alkaen vähiten merkitsevästä numerosta. Tätä tulosta tarvitaan myöhemmin lauseen 5.16 todistuksessa. Lausetta 5.16 taas tarvitaan edelleen *Cobhamin lauseen* todistuksessa.

Normalisoidaan ensin sana $2354 = 4532^R \in \{0, 1, 2, 3, 4, 5\}^*$, joka on eräs luvun 100 esitys yleisessä 3-kantajärjestelmässä. Tarkastellaan summaa

$$[4532^R]_3 = 4 \cdot 1 + 5 \cdot 3 + 3 \cdot 9 + 2 \cdot 27.$$

Jaetaan ykkösten kerroin 4 kantaluvulla 3 käyttäen jakoalgoritmia, jolloin saadaan yhtälö

$$4 \cdot 1 + 5 \cdot 3 + 3 \cdot 9 + 2 \cdot 27 = (1 \cdot 3 + 1) \cdot 1 + 5 \cdot 3 + 3 \cdot 9 + 2 \cdot 27.$$

Kutsutaan jakoalgoritmin tuottamaa osamäärää *muistiluvuksi.* Jakojäännös 1 on aakkoston Σ_k symboli. Luku 4 saadaan siis hajotettua summaksi, jonka summattavina on muistinumeron ja kantaluvun tulo sekä aakkoston Σ_k jäsen. Muistiluku voidaan siirtää kantaluvun ensimmäisen potenssin kertoimeen käyttämällä distributiivilakeja:

$$(1 \cdot 3 + 1) \cdot 1 + 5 \cdot 3 + 3 \cdot 9 + 2 \cdot 27 = 1 \cdot 1 + (5 + 1) \cdot 3 + 3 \cdot 9 + 2 \cdot 27.$$

Kantaluvun ensimmäisen potenssin kerroin 6 voidaan hajottaa vastaavalla tavalla käyttämällä jakoalgoritmia. Uudeksi muistinumeroksi saadaan 2 ja jakojäännökseksi 0, joten

$$1 \cdot 1 + (5 + 1) \cdot 3 + 3 \cdot 9 + 2 \cdot 27 = 1 \cdot 1 + (2 \cdot 3 + 0) \cdot 3 + 3 \cdot 9 + 2 \cdot 27.$$

Siirtämällä muistinumero 2 kantaluvun toisen potenssin kertoimeen ja käyttämällä jälleen jakoalgoritmia saadaan yhtälö

$$1 \cdot 1 + 0 \cdot 3 + (3 + 2) \cdot 9 + 2 \cdot 27 = 1 \cdot 1 + 0 \cdot 3 + (1 \cdot 3 + 2) \cdot 9 + 2 \cdot 27$$

ja edelleen

$$1 \cdot 1 + 0 \cdot 3 + 2 \cdot 9 + (2 + 1) \cdot 27 = 1 \cdot 1 + 0 \cdot 3 + 2 \cdot 9 + (1 \cdot 3 + 0) \cdot 27.$$

Lopulta saadaan yhtälö

$$4 \cdot 1 + 5 \cdot 3 + 3 \cdot 9 + 2 \cdot 27 = 1 \cdot 1 + 0 \cdot 3 + 2 \cdot 9 + 0 \cdot 27 + 1 \cdot 81,$$

toisin sanoen

$$[4532^R]_3 = 1 \cdot 3^4 + [1020^R]_3 = [10201^R]_3. \tag{4}$$

Luku 10201 on siis luvun 2354 normalisaatio. Huomionarvoista on, että $1 \cdot 3^4$ on viimeinen muistinumero kerrottuna kantaluvun potenssilla, jonka eksponentti on sanan 4532 pituus.

Todistetaan seurava lause:

Lause 3.16. Oletetaan, että D aakkoston Σ_k sisältävä aakkosto. Tällöin on olemassa μ-transduktori, joka laskee normalisaation $D^* \longrightarrow \Sigma^*$.

Todistus. Olkoon $m = \max\{|d-a| \mid a \in \Sigma_k,\ d \in D\}$, ja merkitään $\gamma = \frac{m}{k-1}$. Määritellään mahdollisten muistinumeroiden joukko

$$Q = \{s \in \mathbb{Z} \mid |s| < \gamma\}.$$

Olkoon $s \in Q$ ja $d \in D$. Jakoalgoritmin nojalla on olemassa sellaiset yksikäsitteiset luvut $a \in \Sigma_k$ ja s', että

$$s + d = s'k + a. \tag{5}$$

Asetetaan jokaista paria $(s, d) \in Q \times D$ kohti

$$\delta(s, d) = s'$$

ja

$$\lambda(s, d) = a.$$

Tällöin

$$s + d = \delta(s, d) \cdot k + \lambda(s, d). \tag{6}$$

Koska $|s| < \gamma$ ja $|d - a| \leq m$

$$|s'| = \frac{|s + d - a|}{k} \leq \frac{|s| + |d - a|}{k} \leq \frac{|s| + m}{k} < \frac{\gamma + m}{k}.$$

Toisaalta

$$\begin{aligned}
\frac{\gamma + m}{k} &= \frac{\gamma}{k} + \frac{m}{k} \\
&= \frac{m}{k(k-1)} + \frac{m(k-1)}{k(k-1)} \\
&= \frac{mk}{k(k-1)} \\
&= \frac{m}{k-1} \\
&= \gamma,
\end{aligned}$$

joten myös $|s'| < \gamma$, eli toisin sanoen $s' \in Q$. Lukujen $s' \in Q$ ja $a \in \Sigma_k$ yksikäsitteisyyden nojalla saadaan hyvinmääritellyt funktiot $\delta : Q \times D \longrightarrow Q$ ja $\lambda : Q \times D \longrightarrow \Sigma_k$.

Olkoon nyt $\mathcal{T} = (Q, D, \delta, 0, \Sigma_k, \lambda)$ transduktori. Osoitetaan induktiolla, että

$$[w^R]_k = \delta(0, w) \cdot k^{|w|} + [\lambda(0, w)^R]_k \tag{7}$$

aina kun $w \in D^*$. Vertaa yhtälöä (7) yhtälöön (4). Ensinnäkin

$$[\varepsilon]_k = 0 = \delta(0, \varepsilon) \cdot k + [\lambda(0, \varepsilon)]_k,$$

sillä $\delta(0, \varepsilon) = 0$ ja $\lambda(0, \varepsilon) = \varepsilon$.

Oletetaan sitten, että $[w^R] = \delta(0, w) \cdot k^{|w|} + [\lambda(0, w)^R]_k$. Olkoon $d \in D$ symboli. Käytetään sanasta $\lambda(0, w)$ lyhennysmerkintää a ja merkitään tilaa $\delta(0, w)$ symbolilla s. Induktio-oletus on siis $[w^R]_k = s \cdot k^{|w|} + [a^R]_k$. Tällöin

$$\begin{aligned}
[(wd)^R]_k = [dw^R]_k &= d \cdot k^{|w|} + [w^R]_k \\
&= d \cdot k^{|w|} + s \cdot k^{|w|} + [a^R]_k \\
&= (s + d) \cdot k^{|w|} + [a^R]_k.
\end{aligned}$$

Koska $s = \delta(0, w)$ yhtälön (6) nojalla

$$\begin{aligned}
s + d &= \delta(s, d) \cdot k + \lambda(s, d) \\
&= \delta(0, wd) \cdot k + \lambda(s, d),
\end{aligned}$$

joten

$$\begin{aligned}
(s + d) \cdot k^{|w|} + [a^R]_k &= \Big(\delta(0, wd) \cdot k + \lambda(s, d)\Big) \cdot k^{|w|} + [a^R]_k \\
&= \delta(0, wd) \cdot k^{|wd|} + \lambda(s, d) \cdot k^{|w|} + [a^R]_k. \\
&= \delta(0, wd) \cdot k^{|wd|} + [\lambda(s, d)a^R]_k \\
&= \delta(0, wd) \cdot k^{|wd|} + [(a\lambda(s, d))^R]_k.
\end{aligned}$$

Koska $a\lambda(s, d) = \lambda(0, w) \cdot \lambda(\delta(0, w), d) = \lambda(0, wd)$ saadaan lopulta yhtälö

$$[(wd)^R]_k = \delta(0, wd) \cdot k^{|wd|} + [\lambda(0, wd)^R]_k,$$

joten yhtälö (7) on voimassa aina kun $w \in D^*$.

Määritellään nyt funktio $\mu : Q \longrightarrow \Sigma_k^*$ asettamalla

$$\mu(q) = (q)_k^R$$

aina kun $q \in Q$. Nyt jos μ-transduktori $\mathcal{T}_\mu$ saa syötteekseen sanan $w \in D^*$, se tulostaa sanan $\lambda(0,w)(\delta(0,w))_k^R$, joka esittää lukua

$$\left[\left(\lambda(0,w) \cdot (\delta(0,w))_k^R\right)^R\right]_k = [(\delta(0,w))_k \cdot \lambda(0,w)^R]_k.$$

Koska λ tulostaa vain yksittäisiä symboleja $|\lambda(0,w)| = |w|$, joten

$$[(\delta(0,w))_k \cdot \lambda(0,w)^R]_k = \delta(0,w)_k \cdot k^{|w|} + [\lambda(0,w)^R]_k.$$

Yhtälön (7) nojalla

$$[\mathcal{T}_\mu(w)^R]_k = [w^R]_k,$$

joten $\mathcal{T}_\mu$ laskee normalisaation $D^* \longrightarrow \Sigma_k^*$, kun sekä syötesanat että tulostesanat tulkitaan luonnollisten lukujen k-kantaesityksiksi alkaen pienimmästä merkitsevästä numerosta.

$\square$

4 Automaattiset jonot

Tässä luvussa tarkastellaan niin sanottuja *automaattisia jonoja,* jotka ovat keskeisessä asemassa tässä tutkielmassa. Säännölliset kielet ovat mahdollisesti äärettömiä kieliä, jotka jokin äärellinen automaatti tunnistaa, kun taas automaattiset funktiot ovat äärellisen funktioautomaatin laskemia funktioita, joiden lähtöjoukko on ääretön. Vastaavasti *k-automaattiset jonot* ovat äärettömiä jonoja, joiden mikä tahansa jäsen voidaan muodostaa antamalla jonon jäsenen indeksin k-kantaesitys jonkin funktioautomaatin syötteeksi. Tällä tavoin voidaan määritellä äärettömiä jonoja käyttämällä äärellistä laskentamallia.

4.1 Automaattisen jonon määritelmä

Jatkossa käsitellään etupäässä funktioautomaatteja, joiden syöteaakkostona on $\Sigma_k = \{0, 1, \ldots, k-1\}$, missä $k \geq 2$ on luonnollinen luku. Tavallisesti kielen Σ_k^* sanat tulkitaan luonnollisten lukujen k-kantaesityksiksi.

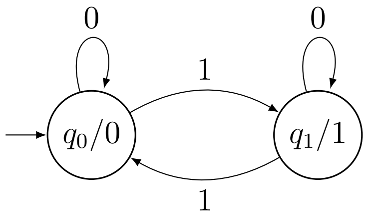
Määritelmä 4.1. Olkoon $k \geq 2$ luonnollinen luku. Sanotaan, että funktioautomaatti $\mathcal{M} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ on *k-funktioautomaatti.*

Määritelmä 4.2 (*k-automaattiset jonot*)**.** Olkoon Δ äärellinen aakkosto. Oletetaan, että on olemassa sellainen äärellinen k-funktioautomaatti $\mathcal{M} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$, että $a_n = \tau(\delta(q_0, w))$ aina kun $n \in \mathbb{N}$ ja $w \in \Sigma_k^*$ on sana jolle pätee $[w]_k = n$. Tällöin sanotaan, että jono (a_n) on *k-automaattinen,* ja että k-funktioautomaatti $\mathcal{M}$ *generoi* jonon (a_n).

Esimerkki 4.3 (*Thuen-Morsen jono*)**.** Muodostetaan aakkoston $\{0, 1\}$ jono $\mathbf{t} = (t_n)$ asettamalla

$$t_n = \begin{cases} 0 \text{ jos sanassa } (n)_2 \text{ on parillinen määrä symboleita } 1 \\ 1 \text{ jos sanassa } (n)_2 \text{ on pariton määrä symboleita } 1. \end{cases}$$

Tällöin $\mathbf{t} = 0110100110010110\ldots$. Tätä jonoa kutsutaan *Thuen-Morsen jonoksi.* Olkoon $\mathcal{M}$ kuvan 4 funktioautomaatti. Funktioautomaatti $\mathcal{M}$ on tilassa q_0 silloin kun siihen asti luettujen symbolien 1 määrä on parillinen. Muutoin se on tilassa q_1. Jos $\mathcal{M}$ lukee symbolin 0, luettujen symbolien 1 määrä ei muutu, joten tilakaan ei muutu. Jos taas $\mathcal{M}$ lukee symbolin 1, symbolien 1 lukumäärän pariteetti ja siten tila muuttuu. Alkutila on q_0, sillä yhtäkään symbolia ei ole vielä luettu.



Kuva 4: 2-funktioautomaatti joka generoi Thuen-Morsen jonon.

4.2 Automaattisen jonon määritelmän ekvivalentteja muunnelmia ja laajennuksia

4.2.1 Indeksit kanonisessa k-kantaesityksessä ja alkaen vähiten merkitsevästä numerosta

Tässä alaluvussa tarkastellaan k-automaattisten jonojen määritelmän muunnelmia. Osoitamme, että kyseiset muunnelmat määrittelevät tarkalleen samat jonot.

Tarkastellaan ensin muunnelmaa, jossa jonon jäsenten indeksit on syötettävä kanonisina k-kantaesityksinä.

Lause 4.4. Jono (a_n) on k-automaattinen jos ja vain jos on olemassa sellainen k-funktioautomaatti $\mathcal{M} = (Q, \Sigma_k, q_0, \Delta, \tau)$, että $a_n = \tau(\delta(q_0, (n)_k))$ aina kun $n \in \mathbb{N}$. Lisäksi voidaan valita sellainen $\mathcal{M}$, että $\delta(q_0, 0) = q_0$.

Todistus. " $\Longrightarrow$ ": Olkoon $\mathcal{M} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ k-funktioautomaatti, joka generoi jonon (a_n) määritelmän 4.2 mielessä. Olkoon $n \in \mathbb{N}$. Koska $(n)_k$ on eräs luvun n esitys kannassa k määritelmän nojalla $a_n = \tau(\delta(q_0, (n)_k))$.

"$\Longleftarrow$": Olkoon $\mathcal{M} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ sellainen k-funktioautomaatti, että $a_n = \tau(\delta(q_0, (n)_k))$ aina kun $n \in \mathbb{N}$. Muodostetaan uusi tilajoukko $Q' = Q \cup \{q_0'\}$ lisäämällä joukkoon Q uusi alkutila q_0'. Määritellään transitiofunktio $\delta' : Q' \times \Sigma_k \to Q'$ ja tulostusfunktio $\tau' : Q' \to \Delta$ asettamalla

$$\begin{aligned}
\delta'(q, a) &= \delta(q, a), \text{ aina kun } q \in Q \text{ ja } a \in \Sigma_k, \\
\delta'(q_0', a) &= \begin{cases} \delta(q_0, a), \text{ kun } a \neq 0 \\ q_0', \text{ kun } a = 0 \end{cases}, \\
\tau'(q_0') &= \tau(q_0) \text{ ja} \\
\tau'(q) &= \tau(q) \text{ kun } q \in Q.
\end{aligned}$$

Toisin sanoen k-automaattiin $\mathcal{M}'$ lisätään uusi alkutila, joka kuvautuu uudessa tulostusfunktiossa samoin kuin alkuperäinen alkutila

k-funktioautomaatissa $\mathcal{M}$. Alkuperäinen ja uusi alkutila ovat uuden tulostusfunktion kannalta samat. Uuteen alkutilaan lisätään 0-transitio $q_0' \xrightarrow{0} q_0'$, jonka tarkoitus on poistaa syötteen alussa mahdollisesti olevien nollien vaikutus. Symbolit $a \neq 0$ muuttavat tilan q_0' tilaksi $\delta(q_0, a)$ eli samaksi tilaksi kuin $\mathcal{M}$ muuttaa oman alkutilansa.

Olkoon $n > 0$ luonnollinen luku. Olkoon sana $w \in \Sigma_k^*$ luvun n jokin k-kantaesitys. Tällöin $w = 0^i(n)_k$ jollain $i \geq 0$. Nyt

$$\begin{aligned}\tau'(\delta'(q_0', w)) &= \tau'(\delta'(q_0', 0^i(n)_k)) \\ &= \tau'(\delta'(q_0', (n)_k)),\end{aligned}$$

ja edelleen koska sanan $(n)_k$ ensimmäinen symboli on nollasta eroava

$$\begin{aligned}\tau'(\delta'(q_0', (n)_k)) &= \tau'(\delta(q_0, (n)_k)) \\ &= \tau(\delta(q_0, (n)_k)) \\ &= a_n,\end{aligned}$$

eli k-funktioautomaatti $\mathcal{M}'$ generoi k-automaattisen jonon (a_n).

Huomataan, että edellä määritelty k-funktioautomaatti $\mathcal{M}'$ toteuttaa ehdon $\tau'(\delta'(q_0', (n)_k)) = a_n$, aina kun $n \in \mathbb{N}$ ja $\delta'(q_0', 0) = q_0'$. Lauseen 4.4 k-funtioautomaatiksi $\mathcal{M}$ voidaan valita $\mathcal{M}'$.

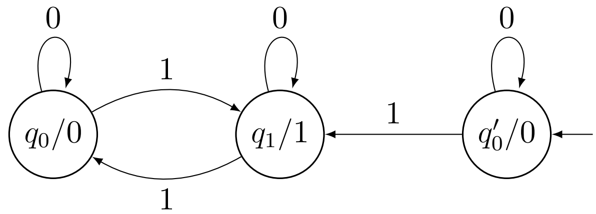
□

Huomautus 4.5. Edellisen lauseen nojalla mikäli k-funktioautomaatti $\mathcal{M} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ generoi jonon (a_n) voidaan olettaa, että $\delta(q_0, 0) = q_0$. Lisäksi voidaan olettaa, että aina kun $\mathcal{M}$ saa syötteeksi jonkin luvun $n \in \mathbb{N}$ tavallisen k-kantaesityksen se päätyy tilaan $\delta(q_0, (n)_k)$. Jokainen luvun n tavallinen k-kantaesitys on nimittäin muotoa $w = 0^i(n)_k$, missä $i \geq 0$. Tällöin $\delta(q_0, w) = \delta(q_0, 0^i(n)_k) = \delta(q_0, (n)_k)$.

Esimerkki 4.6. Muokataan esimerkin 4.3 funktioautomaattia edellisessä lauseessa kuvatulla menetelmällä niin, että muokattu funktioautomaatti sisältää 0-transition alkutilasta takaisin alkutilaan. Saadaan kuvan 5 funktioautomaatti.

Aikaisemmat määritelmät käyttivät jonon (a_n) jäsenten indeksöintiin luvun n eniten merkitsevästä numerosta alkavia k-kantaesityksiä. Automaattiset jonot voidaan määritellä myös käyttäen vähiten merkitsevästä numerosta alkavia kantaesityksiä.

Lause 4.7. Seuraavat väitteet ovat ekvivalentteja:



Kuva 5: 2-funktioautomaatti joka generoi Thuen-Morsen jonon, ja joka sisältää 0-transition $q_0' \xrightarrow{0} q_0'$.

(a) Jono (a_n) on k-automaattinen

(b) On olemassa sellainen k-funktioautomaatti $\mathcal{M} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$, että jos $n \in \mathbb{N}$ ja $[w]_k = n$, niin $a_n = \tau(\delta(q_0, w^R))$

(c) On olemassa sellainen k-funktioautomaatti $\mathcal{M}' = (Q', \Sigma_k, \delta', q_0', \Delta, \tau')$, että $a_n = \tau'(\delta'(q_0', (n)_k^R))$ aina kun $n \in \mathbb{N}$.

Todistus. "(a) $\implies$ (b)": Olkoon (a_n) k-automaattinen jono. Lauseen 3.11 nojalla funktio $f^R : \Sigma_k^* \to \Delta$, joka kuvaa sanan $w \in \Sigma^*$ peilauksen symboliksi $a_{[w]_k}$, on automaattinen. Siten on olemassa funktioautomaatti $\mathcal{M} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ joka laskee funktion f^R. Olkoon $n \in \mathbb{N}$ luku ja $w \in \Sigma^*$ sen k-kantaesitys. Tällöin $\tau(\delta(q_0, w^R)) = a_{[w]_k} = a_n$, joten (b) on voimassa.

"(b) $\implies$ (c)": Luvun n kanoninen k-kantaesitys $(n)_k$ on eräs luvun n tavallinen k-kantaesitys, joten jos k-funktioautomaatti toteuttaa kohdan (b) se toteuttaa myös kohdan (c).

"(c) $\implies$ (a)": Oletetaan, että k-funktioautomaatti $\mathcal{M}'$ toteuttaa ehdon (c). Tällöin funktio $f_{\mathcal{M}'}$ on automaattinen. Funktio $f_{\mathcal{M}'}$ kuvaa luvun n kanonisen k-kantaesityksen peilauksen symboliksi a_n. Lauseen 3.11 nojalla $f_{\mathcal{M}'}^R$ on automaattinen, joten on olemassa k-funktioautomaatti $\mathcal{M} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ joka laskee sen. Funktio $f_{\mathcal{M}'}^R$ kuvaa luvun n kanonisen k-kantaesityksen symboliksi a_n, joten $\tau(\delta(q_0, (n)_k)) = a_n$. Lauseen 4.4 nojalla jono (a_n) on k-automaattinen eli siis (a) on voimassa.

$\square$

4.2.2 Jonon (k, D)-automaattisuus

Määritellään seuraavaksi jonon (k, D)-automaattisuus. Aakkosto D on jokin aakkoston Σ_k sisältävä aakkosto. Sen sijaan, että funktioautomaatille syötettäisiin jonon indeksien tavallisia k-kantaesityksiä, (k, D)-automaattisen jonon generoiva funktioautomaatti saa syötteeksi k-kantaesityksiä aakkostosta D.

Määritelmä 4.8 ((k,D)*-automaattiset jonot*)**.** Olkoon D aakkosto, joka sisältää aakkoston Σ_k. Sanotaan, että jono (a_n) on (k,D)*-automaattinen*, jos on olemassa sellainen funktioautomaatti $\mathcal{M} = (Q, D, \delta, q_0, \Delta, \tau)$, että

$$a_n = \tau(\delta(q_0, w))$$

aina kun $w \in D^*$ on luvun n k-kantaesitys.

Myös (k,D)-automaattisuus voidaan määritellä käyttämällä vähiten merkitsevästä numerosta alkavia indeksien esityksiä.

Lause 4.9. Jono (a_n) on (k,D)-automaattinen jos ja vain jos on olemassa sellainen funktioautomaatti $\mathcal{M} = (Q, D, \delta, q_o, \Delta, \tau)$, että

$$a_n = \tau(\delta(q_0, w^R))$$

aina kun $w \in D^*$ on luvun $n \in \mathbb{N}$ k-kantaesitys.

Todistus. "$\Longrightarrow$": Todistetaan samoin kuin edellisen lauseen implikaatio (a) $\Longrightarrow$ (b).

"$\Longleftarrow$": Olkoon $\mathcal{M} = (Q, D, \delta, q_o, \Delta, \tau)$ sellainen funktioautomaatti, että

$$a_n = \tau(\delta(q_0, w^R))$$

aina kun $w \in D^*$ on luvun $n \in \mathbb{N}$ k-kantaesitys. Funktio $f_{\mathcal{M}}$ on automaattinen, joten lauseen 3.11 nojalla on olemassa funktioautomaatti $\mathcal{M}'$ joka laskee funktion $f_{\mathcal{M}}^R$. Siten jono (a_n) on (k,D)-automaattinen.

$\square$

Myöhemmin lauseessa 5.16 osoitamme, että jono on (k,D)-automaattinen tarkalleen silloin kun se on k-automaattinen.

5 Automaattisten jonojen ominaisuuksia

5.1 Jaksolliset ja lopulta jaksolliset jonot

Tarkastellaan tässä alaluvussa jaksollisten ja lopulta jaksollisten jonojen k-automaattisuutta.

5.1.1 Jonon k-säikeet

Määritellään ensin hyödyllinen käsite, jonon *k-säikeet.*

Määritelmä 5.1. Olkoon $\mathbf{a} = (a_n) \in \Delta^*$ ääretön sana, $k \geq 2$ kokonaisluku ja $d \in \Delta$. Määritellään *k-säie* $I_k(\mathbf{a}, d) = \{(n)_k \,|\, a_n = d\}$.

Jonon $\mathbf{a}$ k-säie $I_k(\mathbf{a}, d)$ sisältää siis niiden indeksien n kanoniset k-kantaesitykset, joissa jonon jäsen on d.

Seuraavaksi osoitetaan, että jono $\mathbf{a}$ on k-automaattinen jos ja vain jos sen kaikki k-säikeet ovat säännöllisiä.

Lemma 5.2. Jono $\mathbf{a} = (a_n)$ on k-automaattinen jos ja vain jos säikeet $I_k(\mathbf{a}, d)$ ovat säännöllisiä aina kun $d \in \Delta$.

Todistus. Oletetaan ensin, että jono $\mathbf{a} = (a_n)$ on k-funktioautomaatin $\mathcal{M} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ generoima. Olkoon $d \in \Delta$ symboli. Lauseen 3.7 nojalla kieli $L_d(\mathcal{M})$ on säännöllinen. Säie $I_k(\mathbf{a}, d)$ koostuu niistä kielen $L_d(\mathcal{M})$ sanoista, jotka alkavat nollasta eroavalla symbolilla, joten

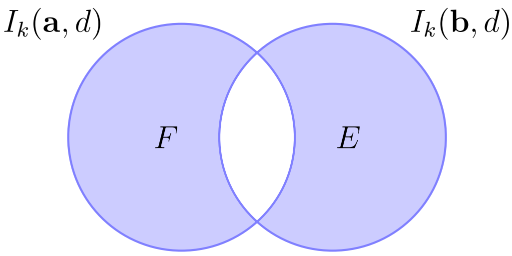
$$I_k(\mathbf{a}, d) = L_d(\mathcal{M}) \cap ((\Sigma_k \setminus \{0\})\Sigma_k^*).$$

Kahden säännöllisen kielen leikkaus on säännöllinen, joten $I_k(\mathbf{a}, d)$ on säännöllinen.

Olkoon $\Delta = \{d_1, \ldots, d_r\}$. Oletetaan, että säie $I_k(\mathbf{a}, d)$ on säännöllinen aina kun $d \in \Delta$. Muodostetaan jokaista $d \in \Delta$ kohti säännöllinen kieli $I'_k(\mathbf{a}, d) = 0^* I_k(\mathbf{a}, d)$. Kielet $I'_k(\mathbf{a}, d)$ osittavat kielen Σ_k^*, sillä jokainen $w \in \Sigma_k^*$ on muotoa $w = 0^i u$, missä $i \geq 0$ on alun nollien määrä ja u kielen $(\Sigma_k \setminus \{0\})\Sigma_k^*$ sana. Nyt $u \in I_k(\mathbf{a}, a_{[u]_k})$, joten $w \in I'_k(\mathbf{a}, a_{[u]_k})$. Lisäksi säikeet ovat pareittain alkiovieraita, joten niin ovat myös kielet $I'_k(\mathbf{a}, d)$. Lauseen 3.8 nojalla on olemassa sellainen k-funktioautomaatti $\mathcal{M} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$, että $L_d(\mathcal{M}) = I'_k(\mathbf{a}, d)$ aina kun $d \in \Delta$.

Oletetaan, että $v \in \Sigma_k^*$ ja $[v]_k = n \in \mathbb{N}$. Tällöin $v \in I'_k(\mathbf{a}, a_n) = L_{a_n}(\mathcal{M})$, joten $\tau(\delta(q_0, v)) = a_n$. Siten jono (a_n) on k-automaattinen.

$\square$



Kuva 6: $I_k(\mathbf{b}, d) = I_k(\mathbf{a}, d) \setminus F \cup E$

5.1.2 Melkein yhtäsuuret jonot

Jos kaksi jonoa eroavat vain äärellisen monessa indeksissä, sanotaan että kyseiset jonot ovat *melkein yhtäsuuria.* Osoitetaan, että jos kaksi jonoa ovat melkein yhtäsuuret ja toinen niistä on k-automaattinen, niin silloin molemmat jonot ovat k-automaattisia.

Lause 5.3. Oletetaan, että $\mathbf{a} = (a_n)$ on k-automaattinen, ja $\mathbf{b} = (b_n)$ jono. Jos $\mathbf{a}$ ja $\mathbf{b}$ ovat melkein yhtäsuuria, niin myös $\mathbf{b}$ on k-automaattinen.

Todistus. Lemman 5.2 nojalla aina kun $d \in \Delta$ säie $I_k(\mathbf{a}, d)$ on säännöllinen. Osoitetaan, että kaikki säikeet $I_k(\mathbf{b}, d)$ ovat säännöllisiä. Olkoon $d \in \Delta$ mielivaltainen ja olkoon

$$E = I_k(\mathbf{b}, d) \setminus I_k(\mathbf{a}, d)$$

niiden sanojen $(n)_k$ joukko, jotka kuuluvat säikeeseen $I_k(\mathbf{b}, d)$, mutta eivät säikeeseen $I_k(\mathbf{a}, d)$. Joukko E esittää niitä indeksejä $n \in \mathbb{N}$, joissa $b_n = d \neq a_n$. Koska jono $\mathbf{b}$ eroaa jonosta $\mathbf{a}$ äärellisen monessa kohdassa, joukko E on äärellinen ja siten lauseen 2.8 nojalla säännöllinen.

Vastaavasti määritellään äärellinen ja siten säännöllinen kieli

$$F = I_k(\mathbf{a}, d) \setminus I_k(\mathbf{b}, d).$$

Kahden säännöllisen kielen erotus on säännöllinen, joten $I_k(\mathbf{a}, d) \setminus F$ on säännöllinen. Tällöin

$$I_k(\mathbf{b}, d) = I_k(\mathbf{a}, d) \setminus F \cup E$$

on kahden säännöllisen kielen unionina säännöllinen. Kuva 6 havainnollistaa tätä yhtälöä.

Koska $d \in \Delta$ oli mielivaltainen, kaikki jonon $\mathbf{b}$ säikeet ovat säännöllisiä. Lemman 5.2 nojalla $\mathbf{b}$ on k-automaattinen.

$\square$

5.1.3 Jaksollisten ja lopulta jaksollisten jonojen k-automaattisuus

Osoitetaan, että jaksolliset ja lopulta jaksolliset jonot ovat aina k-automaattisia. Lopulta jaksollisten jonojen k-automaattisuutta käytetään myöhemmin *Cobhamin lauseen* todistuksessa.

Lause 5.4. Oletetaan, että jono $\mathbf{a} = (a_n)$ on jaksollinen. Tällöin se on k-automaattinen.

Todistus. Olkoon jonon $\mathbf{a}$ jaksonpituus $p \in N_{>0}$. Tarkoitus on muodostaa k-funktioautomaatti $\mathcal{M}$, jolla on p tilaa jotka ilmaisevat kulloinkin luetun syötteen esittämän luvun jäännöksen modulo p. Olkoon $\mathcal{M} = (\Sigma_p, \Sigma_k, \delta, 0, \Delta, \tau)$ k-funktioautomaatti, missä transitiofunktio $\delta : \Sigma_p \times \Sigma_k \longrightarrow \Sigma_p$ on määritelty asettamalla

$$\delta(q, b) = (kq + b) \pmod p$$

aina kun $q \in \Sigma_p$ ja $b \in \Sigma_k$. Osoitetaan induktiolla pituuden r suhteen, että luettuaan sanan $w = w_0 w_1 \cdots w_r$ funktioautomaatti $\mathcal{M}$ on tilassa $[w]_k \pmod p$.

1° Olkoon $w = \varepsilon$. Luettuaan sanan w funktioautomaatti $\mathcal{M}$ päätyy tilaan $0 = [\varepsilon]_k$.

2° Oletetaan sitten, että luettuaan sanan w funktioautomaatti $\mathcal{M}$ on tilassa $q = [w]_k \pmod p$. Olkoon $b \in \Sigma_k$ symboli. Koska $[wb]_k = [w]_k \cdot k + b$,

$$\begin{aligned}\delta(0, wb) &= \delta(q, b) \\ &= q \cdot k + b \pmod p \\ &= [w]_k \cdot k + b \pmod p \\ &= [wb]_k \pmod p.\end{aligned}$$

Siten jos k-funktioautomaatti $\mathcal{M}$ lukee luvun $n \in \mathbb{N}$ esityksen w, se päätyy tilaan $n \pmod p$. Määritellään nyt tulostusfunktio $\tau : \Sigma_p \longrightarrow \Delta$ asettamalla

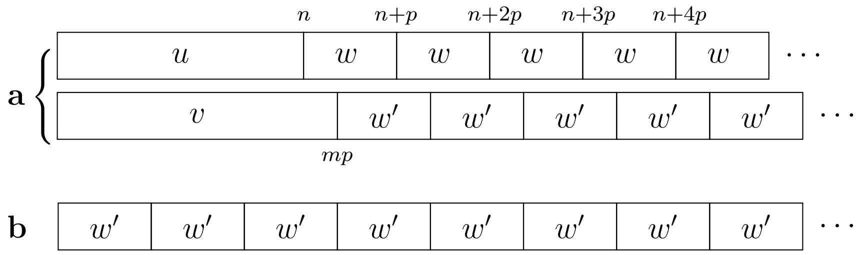
$$\tau(q) = a_q \text{ aina kun } q \in \Sigma_p$$

aina kun $q \in \Sigma_p$. Tällöin $\tau(\delta(0, w)) = a_h$, missä $h \equiv n \pmod p$. Koska $\mathbf{a}$ on jaksollinen ja sen jaksonpituus on p,

$$\tau(\delta(0, w)) = a_h = a_n$$

eli jono $\mathbf{a}$ on k-automaattinen.

□



Kuva 7: Jonot **a** ja **b**.

Seuraus 5.5. Jos jono $\mathbf{a} = (a_n)$ on lopulta jaksollinen, niin se on k-automaattinen.

Todistus. Jono **a** on lopulta jaksollinen, joten on olemassa jaksonpituus $p \in \mathbb{N}_{>0}$ ja sellainen luonnollinen luku n, että $a_{i+p} = a_i$ aina kun $i \geq n$. Merkitään $m = \left\lceil \frac{n}{p} \right\rceil$, jolloin $mp \geq n$. Olkoon $u \in \Delta^n$ sanan **a** etuliite. Koska **a** on lopulta jaksollinen, on olemassa sellainen sana $w \in \Delta^p$, että $\mathbf{a} = uw^\infty$. Toisaalta jos $v \in \Delta^{mp}$ on sanan **a** etuliite, on olemassa sellainen sana $w' \in \Delta^p$, että $\mathbf{a} = vw'^\infty$. Muodostetaan jaksollinen sana $\mathbf{b} = w'^\infty$, missä etuliite v on korvattu samanpituisella sanalla w'^m. Jono **b** on jaksollinen, joten lauseen 5.4 nojalla se on k-automaattinen. Sanat **a** ja **b** eroavat toisistaan vain äärellisen pituisen etuliitteen osalta, joten lauseen 5.3 nojalla **a** on k-automaattinen. Kuva 7 esittää jonoja **a** ja **b** ja niiden hajotelmia sanojen u, v, w ja w' katenaatioiksi.

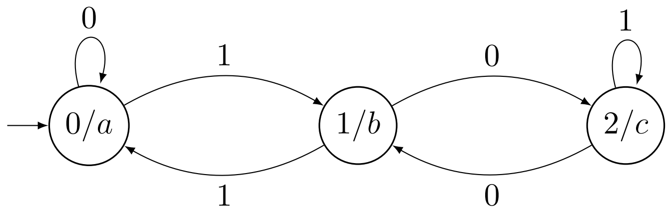
□

Esimerkki 5.6. Jono $\mathbf{x}_{abc} = abcabcabcabcabcabc\ldots$ on jaksollinen, joten se on 2-automaattinen. Muodostetaan 2-funktioautomaatti joka generoi sen. Tilajoukkona on $\Sigma_3 = \{0, 1, 2\}$. Koska

$$
\begin{aligned}
0 \cdot 2 + 0 &\equiv 0 \pmod 3 \\
0 \cdot 2 + 1 &\equiv 1 \pmod 3 \\
1 \cdot 2 + 0 &\equiv 2 \pmod 3 \\
1 \cdot 2 + 1 &\equiv 0 \pmod 3 \\
2 \cdot 2 + 0 &\equiv 1 \pmod 3 \\
2 \cdot 2 + 1 &\equiv 2 \pmod 3
\end{aligned}
$$

lauseen 5.4 todistuksen nojalla kuvan 8 funktioautomaatti generoi jonon $abcabcabc\ldots$.

Esimerkki 5.7. Seurauksen 5.3 nojalla myös jono **b**, joka eroaa jostakin k-automaattisesta jonosta **a** äärellisen monessa kohdassa, on



Kuva 8: 2-funktioautomaatti joka generoi jonon $\mathbf{x}_{abc}$

k-automaattinen. Voidaan ajatella, että jonon **b** generoi k-funktioautomaatti, joka "muistaa" sen äärellisen etuliitteen, jossa **a** ja **b** eroavat. Koska kyseinen etuliite on äärellinen, sen "muistamiseen" tarvitaan vain äärellinen määrä tiloja ja transitioita.

Muodostetaan 2-funktioautomaatti, joka generoi jonon

$$\tilde{\mathbf{x}}_{abc} = cbbcbaabcabcabcabcabc\ldots.$$

Jono $\tilde{\mathbf{x}}_{abc}$ eroaa jonosta $\mathbf{x}_{abc}$ äärellisen etuliitteen *cbbcba* osalta. Kuvan 9 suunnattu graafi esittää 2-funktioautomaattia, jonka transitiofunktio on osittainen. Kun se saa syötteenään luvun $n \in \{0,1,2,3,4,5\}$ binääriesityksen, se päätyy tilaan $\bar{n}$.

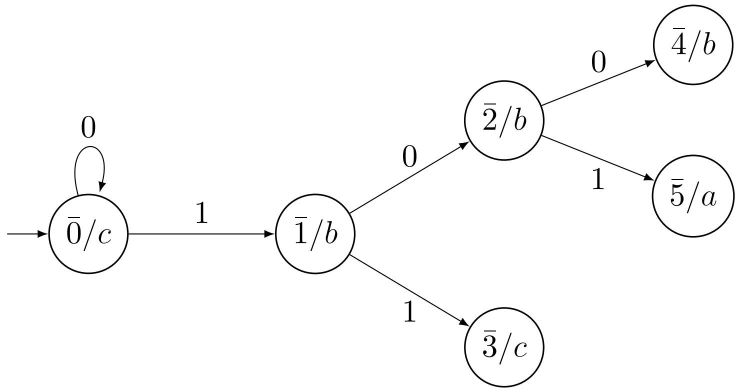
Koska $|cbbcba| = 6 \equiv 0 \pmod 3$, voidaan käyttää kuvan 8 funktioautomaattia generoimaan jonon $\tilde{\mathbf{x}}_{abc}$ alkiot indeksistä 6 alkaen. Lisätään tiloihin $\bar{3}$, $\bar{4}$ ja $\bar{5}$ transitiot samalla logiikalla kuin lauseessa 5.4, toisin sanoen jos $q \in \{0,1,2,3,4,5\}$ ja $p \in \{0,1,2\}$, lisätään transitio

$$\bar{q} \xrightarrow{a} p$$

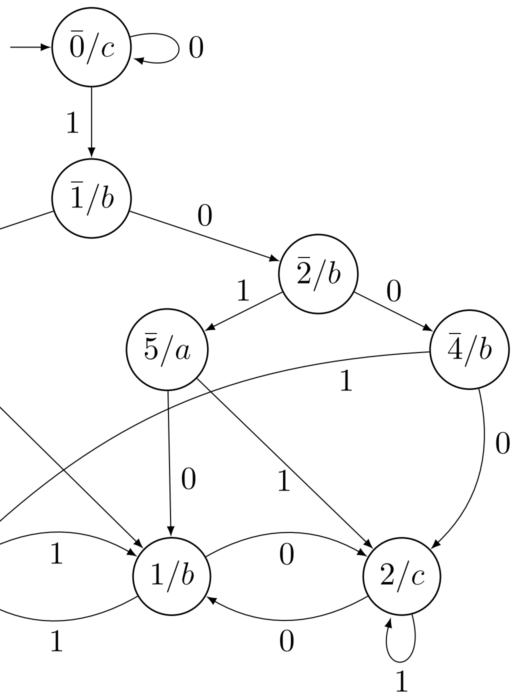
silloin ja vain silloin kun $2 \cdot q + a \equiv p \pmod 3$. Koska

$$\begin{aligned}
3 \cdot 2 + 0 &\equiv 6 \equiv 0 \pmod 3\\
3 \cdot 2 + 1 &\equiv 7 \equiv 1 \pmod 3\\
4 \cdot 2 + 0 &\equiv 8 \equiv 2 \pmod 3\\
4 \cdot 2 + 1 &\equiv 9 \equiv 0 \pmod 3\\
5 \cdot 2 + 0 &\equiv 10 \equiv 1 \pmod 3\\
5 \cdot 2 + 1 &\equiv 11 \equiv 2 \pmod 3,
\end{aligned}$$

saadaan kuvan 10 funktioautomaatti, joka generoi jonon $\tilde{\mathbf{x}}_{abc}$.



Kuva 9: Suunnattu graafi, joka "muistaa" sanan *cbbcba*.



Kuva 10: 2-funktioautomaatti, joka generoi jonon *cbbcbaabcabcabc*...

5.2 Jonon koodaus ja jonojen karteesinen tulo

Seuraavaksi osoitetaan, että k-automaattisen jonon koodaus on myös k-automaattinen.

Lause 5.8. Olkoon $\mathbf{a} = (a_n)$ k-automaattinen jono joukossa Δ, ja olkoon $\rho : \Delta \longrightarrow \Gamma$ koodaus. Tällöin myös jono $\rho(\mathbf{a}) = (\rho(a_n))$ on k-automaattinen.

Todistus. Koska (a_n) on k-automaattinen, lauseen 4.4 nojalla on olemassa sellainen k-automaatti $\mathcal{M} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$, että $a_n = \tau(\delta(q_0, (n)_k))$ aina kun $n \in \mathbb{N}$. Muodostetaan automaatti $\mathcal{M}' = (Q, \Sigma_k, \delta, q_0, \Gamma, \rho \circ \tau)$. Tällöin yhtälöketju

$$(\rho \circ \tau)(\delta(q_0, (n)_k)) = \rho(\tau(\delta(q_0, (n)_k))) = \rho(a_n) = \rho(\mathbf{a})_n,$$

on voimassa aina kun $n \in \mathbb{N}$. Jono $\rho(\mathbf{a})$ on siis k-funktioautomaatin $\mathcal{M}'$ generoima.

□

Määritellään kahden jonon $\mathbf{a} = (a_n)$ ja $\mathbf{b} = (b_n)$ *karteesinen tulo*

$$\mathbf{a} \times \mathbf{b} = ((a_n, b_n)_n).$$

Osoitetaan että kahden k-automaattisen jonon karteesinen tulo on myös k-automaattinen.

Lause 5.9. Oletetaan, että $\mathbf{a} = (a_n) \in \Delta_a^\infty$ ja $\mathbf{b} = (b_n) \in \Delta_b^\infty$ ovat kaksi k-automaattista jonoa. Tällöin myös jono $\mathbf{a} \times \mathbf{b}$ on k-automaattinen.

Todistus. Oletetaan, että k-funktioautomaatti $\mathcal{M}_a = (Q_a, \Sigma_k, \delta_a, q_a, \Delta_a, \tau_a)$ generoi jonon $\mathbf{a}$, ja että $\mathcal{M}_b = (Q_b, \Sigma_k, \delta_b, q_b, \Delta_b, \tau_b)$ generoi jonon $\mathbf{b}$.

Muodostetaan k-funktioautomaatti

$$\mathcal{M}_{a\times b} = (Q_a \times Q_b, \Sigma_k, \delta, (q_a, q_b), \Delta_a \times \Delta_b, \tau),$$

joka laskee automaattien $\mathcal{M}_a$ ja $\mathcal{M}_b$ tilan samanaikaisesti. Määritellään transitiofunktio $\delta : (Q_a \times Q_b) \times \Sigma_k \longrightarrow Q_a \times Q_b$ asettamalla

$$\delta((q, r), a) = (\delta_a(q, a), \delta_b(r, a))$$

aina kun $(q, r) \in Q_a \times Q_b$ ja $a \in \Sigma_k$.

Määritellään vastaavasti tulostusfunktio $\tau : Q_a \times Q_b \longrightarrow \Delta_a \times \Delta_b$ niin, että tilan ensimmäinen komponentti kuvataan tulostusfunktiolla τ_a, ja toinen komponentti tulostusfunktiolla τ_b. Toisin sanoen määritellään

$$\tau((q, r)) = (\tau_a(q), \tau_b(r))$$

aina kun $(q, r) \in Q_a \times Q_b$.

Olkoon $n \in \mathbb{N}$ luonnollinen luku, ja olkoon $w \in \Sigma^*$ luvun n k-kantaesitys. Tällöin

$$\delta((q_a, q_b), w) = (\delta_a(q_a, w), \delta_b(q_b, w)).$$

Siten

$$\begin{aligned}
\tau(\delta((q_a, q_b), w)) &= \tau((\delta_a(q_a, w), \delta_b(q_b, w))) \\
&= (\tau_a(\delta_a(q_a, w)), \tau_b(\delta_b(q_b, w))) \\
&= (a_n, b_n),
\end{aligned}$$

joten k-automaatti $\mathcal{M}_{a\times b}$ generoi jonon $\mathbf{a} \times \mathbf{b} = ((a_n, b_n))$.

□

Lauseista 5.8 ja 5.9 seuraa, että mikäli jonon $\mathbf{c}$ jäsenet on muodostettu k-automaattisten jonojen $\mathbf{a} = (a_n)$ ja $\mathbf{b} = (b_n)$ jäsenistä käyttämällä jotakin funktiota, myös $\mathbf{c}$ on k-automaattinen.

Seuraus 5.10. Oletetaan, että $\mathbf{a} = (a_n) \in \Delta_a^\infty$ ja $\mathbf{b} = (b_n) \in \Delta_b^\infty$ ovat kaksi k-automaattista jonoa. Oletetaan lisäksi, että $f : \Delta_a \times \Delta_b \longrightarrow \Delta$ on funktio. Tällöin jono $f(\mathbf{a} \times \mathbf{b}) = (f(a_n, b_n))$ on k-automaattinen.

Todistus. Lauseen 5.9 nojalla jono $\mathbf{a} \times \mathbf{b}$ on k-automaattinen. Funktio f on koodaus aakkostosta $\Delta_a \times \Delta_b$ tulosteaakkostoon Δ, joten lauseen 5.8 nojalla jono $f(\mathbf{a} \times \mathbf{b})$ on k-automaattinen.

□

Esimerkki 5.11. Oletetaan, että jonot $\mathbf{a} = (a_n)$ ja $\mathbf{b} = (b_n)$ ovat k-automaattisia. Tällöin seurauksen 5.10 nojalla myös niiden summajono $\mathbf{a} + \mathbf{b} = (a_n + b_n)$ ja tulojono $\mathbf{a} \cdot \mathbf{b} = (a_n b_n)$ ovat k-automaattisia.

5.3 Funktioautomaatin ja transduktorin yhdiste ja jonon (k, D)-automaattisuus

Tässä alaluvussa tarkastellaan sellaisia jonoja, joiden jäsenet voidaan laskea syöttämällä jäsenen indeksin k-kantaesitys transduktorille, ja edelleen transduktorin tulostama sana jollekin k-funktioautomaatille. Havaitaan, että näin määritellyt jonot ovat k-automaattisia. Sama pätee jos transduktorin sijasta käytetään μ-transduktoria. Todistetaan tätä μ-transduktoria käyttämällä, että jono on k-automaattinen silloin ja vain silloin kun se on (k, D)-automaattinen, minkä Krebs jättää todistuksessaan lukijan tehtäväksi.

5.3.1 Syötesanan transduktio

Aloitetaan todistamalla, että automaattisen funktion ja transduktorin yhdiste on automaattinen, mikä voidaan tehdä muokkaamalla kirjan [2] lauseen 6.8.6 todistuksen funktioautomaattia.

Lemma 5.12. Olkoon $\mathcal{M} = (Q, \Sigma, \delta, q_0, \Delta, \tau)$ funktioautomaatti ja $\mathcal{T} = (Q_T, D, \delta_T, q_T, \Sigma, \lambda)$ transduktori. Tällöin on olemassa funktioautomaatti $\mathcal{M}'$, joka luettuaan syötesanan $w \in D^*$ tulostaa sen symbolin, jonka $\mathcal{M}$ tulostaa luettuaan sanan $\mathcal{T}(w)$. Toisin sanoen

$$f_{\mathcal{M}'}(w) = (f_{\mathcal{M}} \circ \mathcal{T})(w) = f_{\mathcal{M}}(\mathcal{T}(w)) = \tau(\delta(q_0, \mathcal{T}(w))).$$

Todistus. Tavoitteena on muodostaa funktioautomaatti $\mathcal{M}'$, joka laskee funktioautomaatin $\mathcal{M}$ ja transduktorin $\mathcal{T}$ tilat samanaikaisesti. Kun luetaan syötesana w, halutaan, että $\mathcal{M}$ saa syötteeksi sanan $\mathcal{T}(w)$, jolloin saadaan tulosteeksi $\tau(\delta(q_0, \mathcal{T}(w)))$. Tämä voidaan toteuttaa seuraavasti: oletetaan, että $\mathcal{M}$ on tilassa p ja $\mathcal{T}$ tilassa q, ja että luetaan symboli $a \in D$. Päivitetään tila p siksi tilaksi, johon $\mathcal{M}$ päätyy tilasta p lukemalla sanan $\lambda(q, a)$. Samalla on päivitettävä transduktorin $\mathcal{T}$ tilaa a-transition mukaisesti. Syötesanan w symbolit siis muunnetaan yksi kerrallaan transduktorin $\mathcal{T}$ tulostesanoiksi, jotka sitten annetaan automaatille $\mathcal{M}$ syötteeksi. Samalla päivitetään transduktorin $\mathcal{T}$ tilaa.

Sekä funktioautomaatin $\mathcal{M}$ että transduktorin $\mathcal{T}$ tilat on pidettävä muistissa samanaikaisesti, joten valitaan tilajoukoksi $Q \times Q_T$. Alkutilaksi valitaan (q_0, q_T).

Määritellään transitiofunktio $\delta' : (Q \times Q_T) \times D \longrightarrow Q \times Q_T$ ja tulostusfunktio $\tau' : Q \times Q_T \longrightarrow \Delta$ asettamalla

$$\begin{aligned} \delta'\Big((p, q)\,,\; a\Big) &= \Big(\delta(p, \lambda(q, a))\,,\; \delta_T(q, a)\Big) \text{ ja} \\ \tau'((p, q)) &= \tau(p) \end{aligned}$$

aina kun $(p, q) \in Q \times Q_T$ ja $a \in D$.

Muodostetaan sitten funktioautomaatti

$$\mathcal{M}' = (Q \times Q_T, D, \delta', (q_0, q_T), \Delta, \tau').$$

Osoitetaan induktiolla, että aina kun $w \in D^*$ on sana

$$\delta'\Big((q_0, q_T)\,,\; w)\Big) = \Big(\delta(q_0, \mathcal{T}(w))\,,\; \delta_T(q_T, w)\Big). \tag{8}$$

1° Oletetaan että $w = \varepsilon$. Tällöin

$$\begin{aligned}
\delta'\Big((q_0, q_T)\ ,\ w\Big) &= \delta'\Big((q_0, q_T)\ ,\ \varepsilon\Big) \\
&= \Big(\delta(q_0, \lambda(q_T, \varepsilon))\ ,\ \delta_T(q_T, \varepsilon)\Big) \\
&= \Big(\delta(q_0, \mathcal{T}(\varepsilon))\ ,\ \delta_T(q_T, \varepsilon)\Big) \\
&= \Big(\delta(q_0, \mathcal{T}(w))\ ,\ \delta_T(q_T, w)\Big).
\end{aligned}$$

2° Oletetaan sitten, että sana $w \in D^*$ toteuttaa yhtälön (8). Olkoon $a \in D$ symboli. Tällöin

$$\begin{aligned}
\delta'\Big((q_0, q_T), wa\Big) &= \delta'\Big(\delta'\Big((q_0, q_T), w\Big)\ ,\ a\ \Big) \\
&= \delta'\Big(\Big(\delta_M(q_0, \mathcal{T}(w))\ ,\ \delta_T(q_T, w)\Big)\ ,\ a\ \Big) \\
&= \Big(\delta\Big(\delta(q_0, \mathcal{T}(w))\ ,\ \lambda(\delta_T(q_T, w), a)\Big)\ ,\ \delta_T(\delta_T(q_T, w), a)\ \Big) \\
&= \Big(\delta\Big(q_0\ ,\ \mathcal{T}(w)\lambda(\delta_T(q_T, w), a)\ \Big)\ ,\ \delta_T(q_T, wa)\ \Big) \\
&= \Big(\delta(q_0\ ,\ \mathcal{T}(wa))\ ,\ \delta_T(q_T, wa)\ \Big),
\end{aligned}$$

joten sana wa toteuttaa yhtälön (8).

Olkoon w aakkoston D sana. Tällöin

$$\begin{aligned}
\tau'\Big(\delta'\Big((q_0, q_T)\ ,\ w\Big)\Big) &= \tau'\Big(\delta(q_0, \mathcal{T}(w))\ ,\ \delta_T(q_T, w)\Big) \\
&= \tau\Big(\delta(q_0, \mathcal{T}(w))\Big).
\end{aligned}$$

Siten $f_{\mathcal{M}'}(w) = \tau(\delta(q_0, \mathcal{T}(w)))$ aina kun $w \in D^*$, toisin sanoen funktio $f_{\mathcal{M}} \circ \mathcal{T}$ on automaattinen.

□

Vastaava on voimassa myös, jos transduktorin sijasta käytettään μ-transduktoria:

Seuraus 5.13. Olkoon $\mathcal{M} = (Q, \Sigma, \delta, q_0, \Delta, \tau)$ funktioautomaatti ja $\mathcal{T} = (Q_T, D, \delta_T, q_T, \Sigma, \lambda)$ transduktori. Olkoon $\mu : Q_T \longrightarrow \Sigma^*$ funktio. Tällöin on olemassa funktioautomaatti $\mathcal{M}'$, joka luettuaan syötesanan $w \in D^*$ tulostaa sen symbolin, jonka $\mathcal{M}$ tulostaa luettuaan sanan $\mathcal{T}_\mu(w)$. Toisin sanoen $f_{\mathcal{M}'}(w) = f_{\mathcal{M}}(\mathcal{T}_\mu(w)) = \tau(\delta(q_0, \mathcal{T}_\mu(w)))$.

Todistus. Muodostetaan $\mathcal{M}' = (Q \times Q_T, D, \delta', (q_0, q_T), \Delta, \tau')$ muutoin samoin kuin lemmassa 5.12, paitsi että määritellään tulostusfunktio uudelleen asettamalla

$$\tau'((p, q)) = \tau(\delta(p, \mu(q))).$$

Toisin sanoen automaatin $\mathcal{M}$ tulostusfunktiota ei käytetä suoraan lopulliseen tilaan, vaan annetaan automaatin $\mathcal{M}$ ensin lukea vielä sana $\mu(q) \in \Sigma^*$. Jos nyt w on aakkoston D sana, niin

$$\begin{aligned}\tau'(\delta'((q_M, q_T), w)) &= \tau'(\delta(q_0, \mathcal{T}(w)), \delta_T(q_T, w)) \\ &= \tau\Big(\delta\Big(\ \delta(q_0, \mathcal{T}(w))\ ,\ \mu(\delta_T(q_T, w))\ \Big)\Big) \\ &= \tau\Big(\delta\Big(\ q_0,\ \mathcal{T}(w)\mu(\delta_T(q_T, w))\ \Big)\Big) \\ &= \tau(\delta_M(q_0, \mathcal{T}_\mu(w))).\end{aligned}$$

□

Käyttämällä lemmaa 5.12 ja seurausta 5.13 voidaan todistaa seuraavat kaksi lausetta:

Lause 5.14. Oletetaan että k-funktioautomaatti $\mathcal{M} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ generoi jonon (a_n). Olkoon $\mathcal{T} = (Q_T, \Sigma_k, \delta_T, q_T, \Sigma_k, \lambda)$ äärellinen transduktori. Muodostetaan jono (b_n) asettamalla

$$b_n = a_{[\mathcal{T}((n)_k)]_k}.$$

Tällöin jono (b_n) on k-automaattinen.

Todistus. Sovelletaan edellistä lemmaa k-funktioautomaattiin $\mathcal{M}$ ja transduktoriin $\mathcal{T}$. Ei voida taata sitä, että transduktori $\mathcal{T}$ kuvaisi luvun n jokaisen k-kantaesityksen sellaiseksi tilaksi, jonka tulostusfunktio τ kuvaisi alkioksi $\tau(\delta(q_0, \mathcal{T}((n)_k)))$. Jotta saataisiin yksikäsitteinen jono, on siis käytettävä syötteenä vain kanonisia esityksiä. Lemman 5.12 nojalla on olemassa sellainen k-funktioautomaatti $\mathcal{M}' = (Q', \Sigma_k, \delta', q_0', \Delta, \tau')$, että

$$\tau'(\delta'(q_0', (n)_k)) = \tau(\delta(q_0, \mathcal{T}((n)_k)))$$

aina kun $n \in \mathbb{N}$. Toisaalta $\tau(\delta(q_0, \mathcal{T}((n)_k))) = a_{[\mathcal{T}((n)_k)]_k} = b_n$, joten jono (b_n) on lauseen 4.4 nojalla k-automaattinen.

□

Lause 5.15. Olkoon (a_n) k-funktioautomaatin $\mathcal{M} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ generoima k-automaattinen jono. Oletetaan, että $\mathcal{T} = (Q_T, \Sigma_k, \delta_T, q_T, \Sigma_k, \lambda)$ on äärellinen transduktori, ja että $\mu : Q_T \longrightarrow \Sigma_k^*$ on funktio. Muodostetaan jono (b_n) asettamalla

$$b_n = a_{[\mathcal{T}_\mu((n)_k)]_k}.$$

Tällöin jono (b_n) on k-automaattinen.

Todistus. Todistetaan samoin kuin edellinen lause, mutta sovelletaan seurausta 5.13 funktioautomaattiin $\mathcal{M}$ ja μ-transduktoriin $\mathcal{T}_\mu$. □

5.3.2 k-automaattisuuden ja (k, D)-automaattisuuden ekvivalenttius

Todistetaan seuraavaksi, että jono **a** on k-automaattinen silloin ja vain silloin kun se on (k, D)-automaattinen. Tämä tarkoittaa sitä, että on yhdentekevää määritelläänkö automaattiset jonot syöttämällä funktioautomaateille tavallisia k-kantaesityksiä vai mielivaltaisia k-kantaesityksiä. Tätä tulosta käytetään myöhemmin *Cobhamin lauseen* todistuksessa.

Lause 5.16. Olkoon $\mathbf{a} = (a_n)$ jono. Olkoon $D \subseteq \mathbb{N}$ sellainen joukko, että $\Sigma_k \subseteq D$. Jono **a** on k-automaattinen jos ja vain jos se on (k, D)-automaattinen.

Todistus. Jos jono **a** on (k, D)-automaattinen, se on myös k-automaattinen, sillä $\Sigma_k \subseteq D$. Riittää siis osoittaa, että jonon **a** k-automaattisuudesta seuraa sen (k, D)-automaattisuus. Oletetaan, että **a** on k-automaattinen. Lauseen 4.7 nojalla on olemassa sellainen k-funktioautomaatti $\mathcal{M} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$, että

$$a_n = \tau(\delta(q_0, w^R)) \tag{9}$$

aina kun $w \in \Sigma_k^*$ on luvun $n \in \mathbb{N}$ tavallinen k-kantaesitys.

Olkoon $\mathcal{T}_\mu$ lauseen 3.16 kielen D^* sanat normalisoiva μ-transduktori. Tällöin yhtälö

$$[\mathcal{T}_\mu(w)^R]_k = [w^R]_k \tag{10}$$

on voimassa aina kun $w \in D^*$ ja $[w^R]_k = n \in \mathbb{N}$.

Lemman 5.13 nojalla on olemassa sellainen funktioautomaatti $\mathcal{M}' = (Q \times Q_T, D, \delta', q_0', \Delta, \tau')$, että

$$\tau'(\delta'(q_0', w)) = \tau(\delta(q_0, \mathcal{T}_\mu(w)))$$

aina kun $w \in D^*$.

Valitaan $n \in \mathbb{N}$ ja jokin sen k-kantaesitys $w \in D^*$. Tällöin

$$\tau'(\delta'(q_0', w^R)) = \tau(\delta(q_0, \mathcal{T}_\mu(w^R))),$$

ja yhtälön (9) nojalla

$$\tau(\delta(q_0, \mathcal{T}_\mu(w^R))) = a_{[\mathcal{T}_\mu(w^R)^R]_k}.$$

Yhtälön (10) nojalla

$$[\mathcal{T}_\mu(w^R)^R]_k = [w]_k = n,$$

joten $\tau'(\delta'(q_0', w^R)) = a_n$. Lauseen 4.9 nojalla **a** on (k, D)-automaattinen.

□

Kuten aiemmin lauseissa 4.4 ja 4.7 havaittiin, jos funktioautomaateille syötetään lukujen tavallisia k-kantaesityksiä alkaen suurimmasta tai pienimmästä numerosta, päädytään k-automaattisiin jonoihin. Lauseiden 4.9 ja 5.16 nojalla sama pätee, kun funktioautomaateille syötetään mitä tahansa k-kantaesityksiä. Tässä mielessä k-automaattisten jonojen määritelmä on hyvin perustavanlaatuinen: kun muodostetaan jonoja syöttämällä funktioautomaateille lukujen k-kantaesityksiä, päädytään väistämättä k-automaattisiin jonoihin.

5.4 Morfismien kiintopisteet ja Cobhamin pieni lause

5.4.1 Cobhamin pieni lause

Seuraavaksi tarkastellaan automaattisten jonojen tutkimuksen erästä perustulosta, *Cobhamin pientä lausetta.* Sen mukaan k-automaattiset jonot ovat tarkalleen k-uniformisten morfismien kiintopisteiden koodauksia.

Lause 5.17 (*Cobhamin pieni lause*)**.** Jono $\mathbf{a} \in \Delta^{\mathbb{N}}$ on k-automaattinen silloin ja vain silloin kun se on jonkin k-uniformin morfismin kiintopisteen koodaus.

Todistus. Olkoon $\mathbf{w} = w_0 w_1 w_2 \cdots \in \Gamma^{\mathbb{N}}$ jonkin k-uniformin morfismin $\varphi : \Gamma^* \longrightarrow \Gamma^*$ kiintopiste. Oletetaan, että $\rho : \Gamma \longrightarrow \Delta$ on koodaus. Muodostetaan jono $\mathbf{a} = \rho(\mathbf{w})$. Osoitetaan, että $\mathbf{a}$ on k-automaattinen.

Muodostetaan k-funktioautomaatti

$$\mathcal{M} = (\Gamma, \Sigma_k, \delta, w_0, \Delta, \rho),$$

missä transitiofunktio on määritelty asettamalla

$$\delta(q, b) = \varphi(q)_b.$$

Osoitetaan induktiolla luvun $n \in \mathbb{N}$ suhteen, että aina kun $n \in \mathbb{N}$

$$w_n = \delta(w_0, (n)_k). \tag{11}$$

1° Olkoon $n = 0$. Tällöin $\delta(w_0, (n)_k) = \delta(w_0, \varepsilon) = w_0$, kuten väitettiinkin.

2° Oletetaan sitten, että $w_i = \delta(w_0, (i)_k)$ aina kun $i < n$. Oletetaan, että $(n)_k = n_1 n_2 \cdots n_t$. Merkitään $n' = [n_1 n_2 \cdots n_{t-1}]_k$. Tällöin $n = kn' + n_t$. Kun $\mathcal{M}$ lukee sanan $(n)_k$, se päätyy tilaan

$$\begin{aligned} \delta(w_0, (n)_k) &= \delta(w_0, n_1 n_2 \cdots n_t) \\ &= \delta(\delta(w_0, n_1 n_2 \cdots n_{t-1}), n_t) \\ &= \delta(\delta(w_0, (n')_k), n_t). \end{aligned}$$

Induktio-oletuksen nojalla $\delta(w_0,(n')_k) = w_{n'}$, joten saadaan yhtäsuuruus

$$\delta(w_0,(n)_k) = \delta(w_{n'}, n_t),$$

joten transitiofunktion määritelmän nojalla

$$\delta(w_0,(n)_k) = \varphi(w_{n'})_{n_t}.$$

Lauseen 2.4 nojalla

$$\begin{aligned}\varphi(w_{n'})_{n_t} &= (w_{kn'}w_{kn'+1}\cdots w_{k(n'+1)-1})_{n_t}\\ &= w_{kn'+n_t}\\ &= w_n.\end{aligned}$$

Yhtälö (11) on siis voimassa aina kun $n \in \mathbb{N}$.

Koska tulostusfunktio on koodaus ρ, luettuaan sanan $(n)_k$ funktioautomaatti $\mathcal{M}$ tulostaa

$$\rho(\delta(w_0,(n)_k)) = \rho(w_n) = a_n,$$

eli jono $\mathbf{a}$ on lauseen 4.4 nojalla k-automaattinen.

Oletetaan sitten, että jono $\mathbf{a}$ on k-automaattinen, ja että sen generoi k-funktioautomaatti $\mathcal{M} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$. Lauseen 4.4 nojalla voidaan olettaa, että $\delta(q_0, 0) = q_0$. Määritellään morfismi $\varphi : Q^* \longrightarrow Q^*$ asettamalla

$$\varphi(q) = \delta(q,0)\delta(q,1)\cdots\delta(q,k-1).$$

Nyt $\varphi(q_0) = \delta(q_0,0)\delta(q_0,1)\cdots\delta(q_0,k-1)$. Koska $\delta(q_0,0) = q_0$, morfismi φ on symbolia q_0 pidentävä. Lemman 2.4 nojalla $\mathbf{w} = w_0w_1w_2\cdots = \varphi^\infty(q_0)$ on morfismin φ kiintopiste.

Osoitetaan induktiolla sanan $y \in \Sigma_k^*$ pituuden suhteen, että aina kun $y \in \Sigma_k^*$

$$\delta(q_0, y) = \mathbf{w}_{[y]_k}. \tag{12}$$

1° Oletetaan, että $|y| = 0$. Tällöin $\delta(q_0,y) = \delta(q_0,\varepsilon) = q_0 = w_0 = w_{[y]_k}$. Toisin sanoen yhtälö (12) on voimassa.

2° Oletetaan, että yhtälö (12) on voimassa aina kun $|y| < i$. Oletetaan, että $x \in \Sigma_k^{i-1}$ ja $a \in \Sigma_k$. Tällöin $\delta(q_0, xa) = \delta(\delta(q_0,x),a)$, joten koska $x \in \Sigma_k^{i-1}$ induktio-oletuksen nojalla

$$\delta(q_0, xa) = \delta(w_{[x]_k}, a).$$

Morfismin φ määritelmän nojalla $\delta(w_{[x]_k}, a) = \varphi(w_{[x]_k})_a$. Lemman 2.4 nojalla

$$\varphi(w_{[x]_k})_a = w_{k[x]_k+a} = w_{[xa]_k},$$

joten $\delta(q_0, xa) = w_{[xa]_k}$.

Yhtälö (12) on siis voimassa aina kun y on aakkoston Σ_k sana.

Jono $\mathbf{a}$ on k-automaattinen, joten

$$a_n = \tau(\delta(q_0, (n)_k)) = \tau(w_{[(n)_k]_k}) = \tau(w_n).$$

Näin ollen $\mathbf{a} = \tau(\mathbf{w})$, eli jono $\mathbf{a}$ on morfismin φ kiintopisteen $\mathbf{w}$ koodaus.

□

5.4.2 Jonon k^m-automaattisuus

Lause 5.18. Oletetaan, että $m \in \mathbb{N}_{>0}$. Tällöin jono $\mathbf{a} = (a_n) \in \Delta^{\mathbb{N}}$ on k-automaattinen jos ja vain jos se on k^m-automaattinen.

Todistus. "$\Longrightarrow$": Oletetaan, että jono $\mathbf{a}$ on k-automaattinen. Cobhamin pienen lauseen nojalla $\mathbf{a}$ voidaan ilmaista jonkin morfismin $\varphi : \Sigma^* \longrightarrow \Sigma^*$ kiintopisteen koodauksena. Jos kyseinen kiintopiste alkaa symbolilla $a \in \Sigma$, se on $\varphi^\infty(a)$. Olkoon $\tau : \Sigma \longrightarrow \Delta$ sellainen koodaus, että $\mathbf{a} = \tau(\varphi^\infty(a))$.

Käytetään morfismista φ^m merkintää γ. Sana $\gamma^i(a)$ on sekä sanan $\varphi^\infty(a)$ että sanan $\gamma^\infty(a)$ etuliite aina kun $i \in \mathbb{N}$. Sanoilla $\varphi^\infty(a)$ ja $\gamma^\infty(a)$ on siis mielivaltaisen pitkiä yhteisiä etuliitteitä, joten $\gamma^\infty(a) = \varphi^\infty(a)$. Siten $\mathbf{a} = \tau(\gamma^\infty(a))$. Cobhamin pienen lauseen nojalla $\mathbf{a}$ on k^m-automaattinen, sillä γ on k^m-uniformi.

"$\Longleftarrow$": Oletetaan, että $\mathbf{a}$ on k^m-automaattinen. Lauseen 5.2 nojalla säikeet $I_d = I_{k^m}(\mathbf{a}, d) = \{(n)_{k^m} \mid a_n = d\}$ ovat säännöllisiä. Osoitetaan, että säikeet $I_k(\mathbf{a}, d)$ ovat säännöllisiä. Määritellään morfismi $\beta : \Sigma_{k^m}^* \longrightarrow \Sigma_k^*$, joka kuvaa luvun $j \in \Sigma_{k^m}$ sen tavalliseksi m-pituiseksi k-kantaesitykseksi. Lauseen 2.6 nojalla kielet $\beta(I_d)$ ovat säännöllisiä.

Kielen $\beta(I_d)$ sanojen alussa voi olla nollia. Muodostetaan kieli

$$(\beta(I_d)^R/0^*)^R, \tag{13}$$

joka koostuu sanoista jotka on saatu poistamalla jokin määrä nollia kielen $\beta(I_d)$ sanojen alusta. Säännöllisten kielten joukko on suljettu peilauksen ja osamäärän suhteen, joten kieli (13) on säännöllinen. Nollasta eriävillä symboleilla alkavien sanojen kieli $\{\varepsilon\} \cup (\Sigma_k \setminus \{0\})\Sigma_k^*$ on säännöllinen, joten

$$(\beta(I_d)^R/0^*)^R \cap (\varepsilon \cup (\Sigma_k \setminus \{0\})\Sigma_k^*) = \{(n)_k \mid a_n = d\} = I_k(\mathbf{a}, d)$$

on säännöllinen. Lauseen 5.2 nojalla jono $\mathbf{a}$ on k-automaattinen.

□

Esimerkki 5.19. Thuen-Morsen sana on 2-automaattinen, joten se on myös 2^m-automaattinen aina kun $m \geq 1$.

5.5 Jonon k-ydin ja k-automaattisuus

Olemme määritelleet k-automaattiset jonot käyttäen funktioautomaatteja ja k-uniformisten morfismien kiintopisteiden koodauksia. k-automaattiset jonot voidaan karakterisoida vielä kolmannella tavalla: k-automaattiset jonot ovat jonoja joiden *k-ydin* on äärellinen.

Määritelmä 5.20. Määritellään jonon $\mathbf{a} = (a_n) \in \Delta^{\mathbb{N}}$ k-ydin

$$\mathrm{Ker}_k(\mathbf{a}) = \{(a_{k^i \cdot n+j})_n \mid i \in \mathbb{N} \text{ ja } 0 \leq j < k^i\}.$$

Jonon $\mathbf{a} = (a_n)$ 2-ydin $\mathrm{Ker}_2(\mathbf{a})$ koostuu siis osajonoista

$$\begin{array}{rl}
i = 0,\ j = 0: & \boldsymbol{a}_0\boldsymbol{a}_1\boldsymbol{a}_2\boldsymbol{a}_3\boldsymbol{a}_4\boldsymbol{a}_5\boldsymbol{a}_6\boldsymbol{a}_7\boldsymbol{a}_8\boldsymbol{a}_9\boldsymbol{a}_{10}\boldsymbol{a}_{11}\cdots \\
i = 1,\ j = 0: & \boldsymbol{a}_0 a_1\boldsymbol{a}_2 a_3\boldsymbol{a}_4 a_5\boldsymbol{a}_6 a_7\boldsymbol{a}_8 a_9\boldsymbol{a}_{10} a_{11}\cdots \\
i = 1,\ j = 1: & a_0\boldsymbol{a}_1 a_2\boldsymbol{a}_3 a_4\boldsymbol{a}_5 a_6\boldsymbol{a}_7 a_8\boldsymbol{a}_9 a_{10}\boldsymbol{a}_{11}\cdots \\
i = 2,\ j = 0: & \boldsymbol{a}_0 a_1 a_2 a_3\boldsymbol{a}_4 a_5 a_6 a_7\boldsymbol{a}_8 a_9 a_{10} a_{11}\cdots \\
i = 2,\ j = 1: & a_0\boldsymbol{a}_1 a_2 a_3 a_4\boldsymbol{a}_5 a_6 a_7 a_8\boldsymbol{a}_9 a_{10} a_{11}\cdots \\
i = 2,\ j = 2: & a_0 a_1\boldsymbol{a}_2 a_3 a_4 a_5\boldsymbol{a}_6 a_7 a_8 a_9\boldsymbol{a}_{10} a_{11}\cdots \\
i = 2,\ j = 3: & a_0 a_1 a_2\boldsymbol{a}_3 a_4 a_5 a_6\boldsymbol{a}_7 a_8 a_9 a_{10}\boldsymbol{a}_{11}\cdots \\
 & \vdots
\end{array}$$

Lause 5.21. Olkoon $k \geq 2$. Jono $\mathbf{u} = (u_n) \in \Delta^{\mathbb{N}}$ on k-automaattinen jos ja vain jos $\mathrm{Ker}_k(\mathbf{u})$ on äärellinen.

Todistus. " $\Longrightarrow$ ": Oletetaan, että $\mathbf{u}$ on k-automaattinen. Lauseen 4.7 nojalla on olemassa sellainen k-funktioautomaatti $\mathcal{M} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$, että aina kun $n \in \mathbb{N}$ ja $t \in \mathbb{N}$

$$u_n = \tau(\delta(q_0, (n)_k^R 0^t)),$$

sillä $(n)_k^R 0^t$ on eräs luvun n vähiten merkitsevästä numerosta alkava k-kantaesitys.

On tarkoitus osoittaa, että k-ytimeen $\mathrm{Ker}_k(\mathbf{u})$ kuuluva osajono $\mathbf{u}^{ij} = (u_{k^i n+j})_n$ määräytyy siitä, mihin tilaan $\mathcal{M}$ päätyy luettuaan syötteen i ensimmäistä symbolia. Oletetaan, että $w \in \Sigma_k^i$ on i-pituinen luvun j esitys kannassa k. Sana w on luvun j kanoninen k-kantaesitys, jonka alkuun on lisätty riittävä määrä nollia. Toisin sanoen on olemassa sellainen $t \in \mathbb{N}$, että $w = 0^t(j)_k$. Oletetaan, että luettuaan sanan w^R funktioautomaatti $\mathcal{M}$ päätyy tilaan q.

Tarkastellaan kuinka funktioautomaatti $\mathcal{M}$ kuvaa luvun $k^i \cdot n + j$ esityksen. Olkoon ensin $n > 0$. Tällöin

$$[(n)_k w]_k = [(n)_k]_k \cdot k^{|w|} + [w]_k = k^i \cdot n + j,$$

joten $(k^i \cdot n + j)_k = (n)_k w$. Tällöin

$$\begin{aligned}\delta(q_0, (k^i \cdot n + j)_k^R) &= \delta(q_0, w^R (n)_k^R) \\ &= \delta(\delta(q_0, w^r), (n)_k^R).\end{aligned}$$

Koska $\mathcal{M}$ päätyy tilaan q luettuaan sanan w^R

$$\delta(\delta(q_0, w^R), (n)_k^R) = \delta(q, (n)_k^R),$$

joten $u_{k^i \cdot n+j} = \tau(\delta(q_0, (k^i \cdot n + j)_k^R)) = \tau(\delta(q, (n)_k^R))$.

Olkoon sitten $n = 0$. Tällöin $(k^i \cdot n + j)_k = (j)_k$, joten

$$\begin{aligned}\tau(\delta(q_0, (k^i \cdot n + j)_k^R)) &= \tau(\delta(q_0, (j)_k^R)) \\ &= \tau(\delta(q_0, w^R)) \\ &= \tau(\delta(q, \varepsilon)) \\ &= \tau(\delta(q, (n)_k^R)).\end{aligned}$$

Siten aina kun n on luonnollinen luku

$$\mathbf{u}_n^{ij} = u_{k^i \cdot n+j} = \tau(\delta(q_0, (k^i \cdot n + j)_k^R)) = \tau(\delta(q, (n)_k^R)),$$

joten lauseen 4.7 nojalla k-automaatti $(Q, \Sigma_k, \delta, q, \Delta, \tau)$ generoi jonon $\mathbf{u}^{ij}$. Tiloja q on äärellinen määrä, joten niin on myös jonoja $\mathbf{u}^{ij}$. Siten $\mathrm{Ker}_k(\mathbf{u})$ on äärellinen.

”$\Longleftarrow$”: Oletetaan sitten, että $\mathrm{Ker}_k(\mathbf{u})$ on äärellinen. Määritellään aakkoston Σ_k sanojen binäärirelaatio $\equiv$ seuraavasti: $x \equiv y$ jos ja vain jos $u_{k^{|x|} \cdot n+[x]_k} = u_{k^{|y|} \cdot n+[y]_k}$ aina kun $n \in \mathbb{N}$. Toisin sanoen $x \equiv y$ silloin ja vain silloin kun $\mathbf{u}^{|x|[x]_k} = \mathbf{u}^{|y|[y]_k}$. Jonojen yhtäsuuruus on ekvivalenssirelaatio, joten myös $\equiv$ on ekvivalenssirelaatio.

Olkoon $Q = \{[x] \mid x \in \Sigma_k^*\}$ relaation $\equiv$ ekvivalenssiluokkien joukko. Ydin $\mathrm{Ker}_k(\mathbf{u})$ on äärellinen, joten myös Q on äärellinen. Määritellään siirtymäfunktio $\delta : Q \times \Sigma_k \longrightarrow Q$ ja tulostusfunktio $\tau : Q \longrightarrow \Delta$ asettamalla

$$\begin{aligned}\delta([x], a) &= [ax] \\ \tau([x]) &= u_{[x]_k}\end{aligned}$$

aina kun $x \in \Sigma_k^*$ ja $a \in \Sigma_k$. Tulostusfunktio siis kuvaa ekvivalenssiluokan $[x]$ sanan x määräämän osajonon ensimmäiseksi symboliksi.

On tarkistettava, että δ ja τ ovat hyvinmääriteltyjä. Oletetaan, että $[x] = [y]$. Olkoon $n \in \mathbb{N}$. Tällöin

$$u_{k^{|ax|} \cdot n+[ax]_k} = u_{k^{|x|+1} \cdot n+a \cdot k^{|x|}+[x]_k} = u_{k^{|x|} \cdot (kn+a)+[x]_k}.$$

ja vastaavasti

$$u_{k^{|ay|}\cdot n+[ay]_k} = u_{k^{|y|}\cdot(kn+a)+[y]_k}.$$

Koska $[x]=[y]$

$$u_{k^{|x|}\cdot(kn+a)+[x]_k} = u_{k^{|y|}\cdot(kn+a)+[y]_k},$$

joten

$$u_{k^{|ax|}\cdot n+[ax]_k} = u_{k^{|ay|}\cdot n+[ay]_k}.$$

Koska n on mielivaltainen luonnollinen luku, $[ax]=[ay]$. Transitiofunktio δ on siis hyvinmääritelty.

Myös tulostusfunktio on hyvinmääritelty, sillä

$$u_{[x]_k} = u_{k^{|x|}\cdot 0+[x]_k} = u_{k^{|y|}\cdot 0+[y]_k} = u_{[y]_k}.$$

Osoitetaan induktiolla että

$$\delta([\varepsilon], w^R) = [w] \tag{14}$$

aina kun $w \in \Sigma_k^*$. Ensinnäkin $\delta([\varepsilon],\varepsilon)=[\varepsilon]$. Oletetaan että $\delta([\varepsilon],w^R)=[w]$. Tällöin

$$\begin{aligned}
\delta([\varepsilon],(aw)^R) &= \delta([\varepsilon], w^R a)\\
&= \delta(\delta([\varepsilon], w^R), a)\\
&= \delta([w], a)\\
&= [aw],
\end{aligned}$$

eli siis yhtälö (14) on voimassa aina kun $w \in \Sigma_k^*$.

Muodostetaan k-funktioautomaatti $\mathcal{M} = (Q, \Sigma_k, \delta, [\varepsilon], \Delta, \tau)$. Funktioautomaatti $\mathcal{M}$ generoi jonon $\mathbf{u}$, sillä $\tau(\delta([\varepsilon], w^R)) = \tau([w]) = u_{[w]_k}$. Jono $\mathbf{u}$ on k-automaattinen.

□

6 Cobhamin lause

Tässä luvussa tarkastellaan tärkeää *Cobhamin lausetta.* Esittelemme sille Thijmen J. P. Krebsin aiempaa helpomman ja lyhyemmän todistuksen, joka on artikkelista [1].

Määritellään *multiplikatiivinen riippuvuus* ja *multiplikatiivinen riippumattomuus:*

Määritelmä 6.1. Olkoot k ja h kokonaislukuja. Sanotaan, että k ja h ovat *multiplikatiivisesti riippuvia*, jos on olemassa sellaiset aidosti positiiviset kokonaislukueksponentit n ja m, että $k^n = h^m$. Jos k ja h eivät ole multiplikatiivisesti riippuvia, sanotaan että ne ovat *multiplikatiivisesti riippumattomia.*

Tiedetään, että lopulta jaksolliset jonot ovat aina k-automaattisia (lause 5.5). Cobhamin lause antaa ehdon sille, milloin automaattinen jono on lopulta jaksollinen. Oletetaan, että $k \geq 2$ ja $h \geq 2$ ovat kaksi multiplikatiivisesti riippumatonta kantalukua. Cobhamin lauseen mukaan jono $\mathbf{a}$ on sekä $k-$ että h-automaattinen silloin ja vain silloin kun se on lopulta jaksollinen. Cobhamin lausetta voidaan käyttää siihen, että jono osoitetaan lopulta jaksolliseksi, tai mikäli tiedetään että jono ei ole lopulta jaksollinen sitä käyttämällä voidaan osoittaa ettei jono ole h-automaattinen.

Lause 6.2 (Cobham)**.** Oletetaan, että luvut $k \in \mathbb{N}$ ja $h \in \mathbb{N}$ ovat multiplikatiivisesti riippumattomia. Tällöin jono $(a_n)_{n \in \mathbb{N}}$ on $k-$ ja h-automaattinen jos ja vain jos se on lopulta jaksollinen.

Cobhamin lauseen todisti ensimmäisenä Alan Cobham vuonna 1969. Vaikka lauseen väite onkin yksinkertainen, sen todistus on vaikea. Useat matemaatikot ovat yrittäneet löytää sille helpomman todistuksen.

6.1 Aputuloksia

Lemma 6.3. Olkoon $\varepsilon > 0$ aidosti positiivinen reaaliluku. Oletetaan, että $k \geq 2$ ja $h \geq 2$ ovat luonnollisia lukuja. Tällöin on olemassa sellaiset aidosti positiiviset kokonaisluvut m ja n, että $|k^m - h^n| \leq \varepsilon h^n$.

Todistus. Voidaan olettaa että $k \geq h$, sillä muutoin voidaan käyttää luvun k sijasta lukua $\kappa = k^l$, missä $l \in \mathbb{N}_{>0}$ on niin suuri, että $k^l \geq h$. Mikäli on olemassa sellainen positiivinen kokonaisluku μ, että $|\kappa^\mu - h^n| \leq \varepsilon h^n$, niin voidaan valita $m = l\mu$, jolloin $|k^m - h^n| = |k^{l\mu} - h^n| = |\kappa^\mu - h^n| \leq \varepsilon h^n$.

Muodostetaan lukujono $(a_i)_{i \geq 0}$ asettamalla $a_i = \max\{j \in \mathbb{N} \mid k^i \geq h^j\}$. Jonon (a_i) määritelmän nojalla

$$k^i \geq h^{a_i} \tag{15}$$

aina kun $i \in \mathbb{N}$, joten
$$k^i h^{-a_i} \geq 1. \tag{16}$$

Toisaalta $k^i h^{-a_i} < h$. Tehdään vastaoletus, eli oletetaan että $k^i h^{-a_i} \geq h$. Tällöin $k^i \geq h^{a_i+1}$, mikä on ristiriidassa luvun a_i maksimaalisuuden kanssa. Toisin sanoen aina kun $i \in \mathbb{N}$
$$k^i h^{-a_i} \in [1, h). \tag{17}$$

Osoitetaan, että jono (a_i) on aidosti kasvava. Koska $k \geq h$, kertomalla epäyhtälön (15) vasen puoli luvulla k, ja oikea puoli luvulla h, saadaan epäyhtälö $k^{i+1} \geq h^{a_i+1}$, joten
$$a_i + 1 \in \{j \in \mathbb{N} \mid k^{i+1} \geq h^j\}.$$
Siten
$$a_i + 1 \leq \max\{j \in \mathbb{N} \mid k^{i+1} \geq h^j\} = a_{i+1},$$
eli jono $(a_i)_{i \geq 0}$ on aidosti kasvava.

Etsitään nyt sellaiset luonnolliset luvut $x < y$, että $|k^y h^{-a_j} - k^x h^{-a_x}| \leq \varepsilon$. Oletetaan, että on olemassa sellaiset luvut x ja y, että $k^x h^{-a_x} = k^y h^{-a_y}$. Tällöin $|k^y h^{-a_j} - k^x h^{-a_x}| = 0 < \varepsilon$.

Jos taas $k^x h^{-a_x} \neq k^y h^{-a_y}$ aina kun $x \neq y$, jono $(k^i h^{-a_i})_i$ saa äärettömän monta eri arvoa, jotka ovat sisältymisrelaation (17) nojalla välillä $[1, h)$. Väli $[1, h)$ voidaan osittaa $\left\lceil \frac{h-1}{\varepsilon} \right\rceil$ yhtä suureen peräkkäiseen osaväliin. Tällöin kunkin osavälin pituus on pienempi kuin ε. Lokeroperiaatteen nojalla jotkin kaksi lukujonon $(k^i h^{-a_i})_i$ jäsenistä kuuluvat samaan osaväliin. Olkoot nämä kaksi jäsentä $k^x h^{-a_x}$ ja $k^y h^{-a_y}$. Voidaan olettaa, että $x < y$. Tällöin $|k^y h^{-a_y} - k^x h^{-a_x}| \leq \varepsilon$, sillä osavälin pituus on lyhyempi kuin ε.

Kertomalla epäyhtälö
$$|k^y h^{-a_y} - k^x h^{-a_x}| \leq \varepsilon$$
puolittain luvulla $k^{-x} h^{a_y}$ saadaan epäyhtälö
$$|k^{-x} h^{a_y} (k^y h^{-a_y} - k^x h^{-a_x})| = |k^{y-x} - h^{a_y - a_x}| \leq \varepsilon k^{-x} h^{a_y}.$$
Jonon (a_x) määritelmän nojalla $k^{-x} \leq h^{-a_x}$, joten
$$|k^{y-x} - h^{a_y - a_x}| \leq \varepsilon k^{-x} h^{a_y} \leq \varepsilon h^{a_y - a_x}.$$
Valitaan nyt $m = y - x$ ja $n = a_y - a_x$, jolloin epäyhtälö
$$|k^m - h^n| \leq \varepsilon h^n$$
on voimassa. Koska $x < y$, luku m on aidosti positiivinen. Jono (a_i) on aidosti kasvava, joten myös n on aidosti positiivinen.

$\square$

Määritelmä 6.4. Olkoon (a_n) jono aakkoston A alkioita, ja olkoon $I \subseteq \mathbb{N}$ väli. Sanotaan, että jonolla (a_n) on välillä I paikallinen jakso $p \geq 1$, mikäli $a_n = a_{n+p}$ aina kun $(n, n+p) \in I \times I$.

Lemma 6.5. Oletetaan, että jonolla (a_n) on paikallinen jakso p välillä $I \subseteq \mathbb{N}$, ja paikallinen jakso q välillä $J \subseteq \mathbb{N}$. Jos lisäksi $|I \cap J| \geq p + q$, niin jonolla (a_n) on paikallinen jakso p välillä $I \cup J$.

Todistus. Välien I ja J leikkaus on epätyhjä, joten $I \cup J$ on väli. Valitaan mikä tahansa pari $(n, n+p) \in (I \cup J) \times (I \cup J)$. Oletetaan ensin, että $(n, n+p) \in I \times I$. Tällöin $a_n = a_{n+p}$, sillä jonolla (a_n) on paikallinen jakso p joukossa I.

Oletetaan sitten, että $(n, n+p) \notin I \times I$ eli siis joko $n \in J \setminus I$ tai $n+p \in J \setminus I$. Koska $|I \cap J| \geq p + q$, välin $I \setminus J$ etäisyys välistä $J \setminus I$ on vähintään $p + q$. Koska $|n - (n+p)| = p < p + q$, sekä n että $n + p$ kuuluvat väliin J.

Jokin q:sta ensimmäisestä välin $I \cap J$ alkiosta on kongruentti luvun n kanssa modulo q. Olkoon m kyseinen luku. Jonolla (a_n) on paikallinen jakso q välillä J, joten koska n ja m eroavat toisistaan jakson q kokonaislukumonikerran verran,

$$a_n = a_m. \tag{18}$$

Toisaalta koska $|I \cap J| \geq p+q$ myös $m+p \in I \cap J$, joten $(m, m+p) \in I \times I$. Jonolla (a_n) on paikallinen jakso p välillä I, joten

$$a_m = a_{m+p}. \tag{19}$$

Lisäksi $m + p \equiv n + p \pmod{q}$ ja $(m+p, n+p) \in J \times J$, joten

$$a_{m+p} = a_{n+p}. \tag{20}$$

Yhdistämällä yhtälöt (18), (19) ja (20) saadaan yhtälöketju

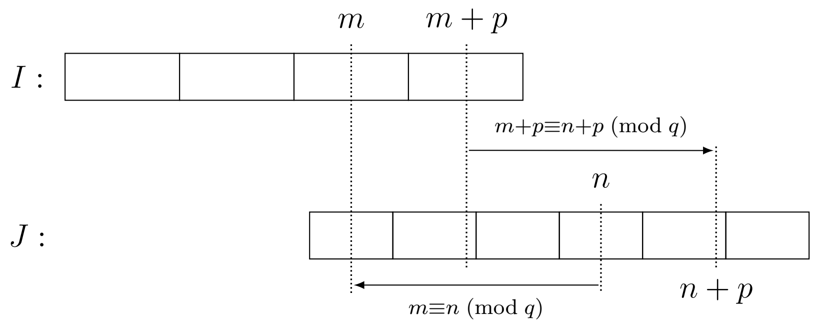
$$a_n = a_m = a_{m+p} = a_{n+p},$$

joten $a_n = a_{n+p}$ myös siinä tapauksessa että $(n, n+p) \notin I \times I$. Kuva 11 havainnollistaa tätä tilannetta.

On siis osoitettu, että $a_n = a_{n+p}$ aina kun $n \in I \cup J$. Toisin sanoen jonolla (a_n) on paikallinen jakso p välillä $I \cup J$.

□

Määritelmä 6.6. Olkoon $x \geq 0$ reaaliluku ja $r \in [0, x]$. Määritellään r-säteinen x-keskinen pallo $B(x, r) = \{y \in \mathbb{N} \,|\, |x - y| \leq r\}$ joukossa $\mathbb{N}$. Olkoon $k \in \mathbb{N}$ luonnollinen luku. Käytetään merkitään $D_k = B(k, k) = \{0, 1, \ldots, 2k\}$.



Kuva 11: Tapaus $(n, n+p) \notin I \times I$. Laatikoiden leveydet esittävät jaksojen p ja q pituuksia.

Lemma 6.7. Olkoon $k \geq 2$ kokonaisluku. Tällöin $D_{k^n} = B(k^n, k^n) \subseteq [D_k^n]_k$ aina kun $n \in \mathbb{N}_{>0}$. Toisin sanoen jokainen välin D_{k^n} luku voidaan ilmaista n-pituisena k-kantaesityksenä käyttämällä symboliaakkostona aakkostoa D_k.

Todistus. Tavallisessa k-kantajärjestelmässä sanoilla, joiden pituus on n, voi esittää kaikki lukua k^n pienemmät luonnolliset luvut. Siten

$$\{0, 1, 2, \cdots, k^n - 1\} = \Big[\{d_1 d_2 \cdots d_n \,|\, d_i \in \Sigma_k\}\Big]_k = [\Sigma_k^n]_k \subseteq [D_k^n]_k. \tag{21}$$

Tarkastellaan kielen $(k + \Sigma_k)\Sigma_k^{n-1} \subseteq D_k^n$ sanojen arvotuksia kannassa k :

$$\begin{aligned}
[(k + \Sigma_k)\Sigma_k^{n-1}]_k &= \Big[\{(k + d_1)d_2 \cdots d_n \,|\, d_i \in \Sigma_k\}\Big]_k \\
&= \left\{(k + d_1)k^{n-1} + \sum_{i=0}^{n-2} d_i k^i \,\middle|\, d_i \in \Sigma_k\right\} \\
&= \left\{k^n + \sum_{i=0}^{n-1} d_i k^i \,\middle|\, d_i \in \Sigma_k\right\} \\
&= k^n + [\Sigma_k^n]_k \\
&= k^n + \{0, 1, 2, \cdots, k^n - 2, k^n - 1\} \\
&= \{k^n, k^n + 1, \ldots, 2k^n - 1\}.
\end{aligned}$$

Näin ollen

$$\{k^n, k^n + 1, \ldots, 2k^n - 1\} = [(k + \Sigma_k)\Sigma_k^{n-1}]_k \subseteq [D_k^n]_k. \tag{22}$$

Lisäksi sanan $(2k - 1)(k - 1) \cdots (k - 1)k \in D_k^n$ arvotus on

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

joten luvulla $2k^n$ on kieleen D_k^n kuuluva k-kantaesitys. Toisin sanoen

$$2k^n \in [D_k^n]_k. \tag{23}$$

Sisältymisrelaatioiden (21), (22) ja (23) nojalla

$$\begin{aligned} D_{k^n} &= \{0, 1, \ldots, 2k^n\} \\ &= \{0, 1, \ldots k^n - 1\} \cup \{k^n, k^n + 1, \ldots, 2k^n - 1\} \cup \{2k^n\} \\ &\subseteq [D_k^n]_k. \end{aligned}$$

□

6.2 Cobhamin lause

Lause 6.8 (*Cobhamin lause*)**.** Oletetaan, että luvut $k, h \in \mathbb{N}$ ovat multiplikatiivisesti riippumattomia. Tällöin jono $\mathbf{a} = (a_n)$ on $k-$ ja h-automaattinen jos ja vain jos se on lopulta jaksollinen.

Todistus. Ensinnäkin seurauksen 5.5 nojalla lopulta jaksollinen jono on myös $k-$ ja h-automaattinen. Riittää siis olettaa, että jono $\mathbf{a}$ on sekä k-automaattinen että h-automaattinen. Tällöin lauseen 5.16 nojalla aina kun $l \in \{k, h\}$ jono $\mathbf{a}$ on (l, D_l)-automaattinen. Siten kun $l \in \{k, h\}$, on olemassa funktioautomaatti $\mathcal{M}_l = (Q_l, D_l, \delta_l, q_l, \Delta, \tau_l)$, joka generoi jonon $\mathbf{a}$.

Määritellään jokaista $q \in Q_l$ kohti kieli $L_{lq} = \{w \in D_l^* \mid \delta_l(q_l, w) = q\}$, eli niiden sanojen joukko, jotka luettuaan automaatti $\mathcal{M}_l$ päätyy tilaan q. Oletetaan, että $w \in L_{lq}$ ja $v \in D_l^*$. Tällöin

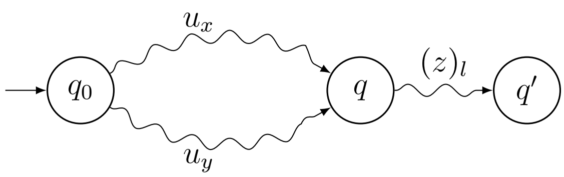
$$a_{[wv]_l} = \tau_l(\delta_l(q_l, wv)) = \tau_l(\delta_l(q, v)). \tag{24}$$

Oletetaan, että $x, y \in [L_{lq}]_l$ ja $z \in [D_l^n]_l$. Tällöin luvun x jokin l-kantaesitys u_x kuuluu kieleen L_{lq}, kuten myös luvun y jokin l-kantaesitys u_y. Sana $u_x(z)_l$ esittää lukua $xl^n + z$ ja sana $u_y(z)_l$ lukua $yl^n + z$. Käyttämällä yhtälöä (24) kahdesti saadaan yhtälöketju

$$\begin{aligned} a_{xl^n+z} &= \tau_l(\delta_l(q_l, u_x(z)_l)) \\ &= \tau_l(\delta_l(q, (z)_l)) \\ &= \tau_l(\delta_l(q_l, u_y(z)_l)) \\ &= a_{yl^n+z}, \end{aligned}$$

joten aina kun $x, y \in [L_{lq}]_l$ ja $z \in [D_l^n]_l$

$$a_{xl^n+z} = a_{yl^n+z}. \tag{25}$$



Kuva 12: Funktioautomaatti $\mathcal{M}_l$ lukee sanat $u_x(z)_l$ ja $u_y(z)_l$. Koska $x, y \in [L_{lq}]_l$ funktioautomaatti päätyy tilaan q luettuaan sanat u_x ja u_y.

Kuvan 12 graafi havainnollistaa sitä, miten funktioautomaatin $\mathcal{M}_l$ tila muuttuu, kun se lukee sanat $u_x(z)_l$ ja $u_y(z)_l$.

Määritellään funktioautomaatin $\mathcal{M}_h$ tilojen osajoukko

$$Q_\infty = \{q \in Q_h \mid [L_{hq}]_h \text{ } \textit{on ääretön } \}.$$

Tila $q \in Q_h$ kuuluu joukkoon Q_∞, mikäli on olemassa äärettömän monta lukua joiden esitykset luettuaan funktioautomaatti $\mathcal{M}_h$ päätyy tilaan q.

Osoitetaan, että kokoelma $\{[L_{kt}]_k \mid t \in Q_k\}$ on kokonaislukujen joukon äärellinen peite. Olkoon $n \in \mathbb{N}$ mielivaltainen luonnollinen luku. Tällöin $\delta_k(q_k, (n)_k)$ on jokin tila $q \in Q_k$, joten $n \in [L_{kq}]_k \subseteq \bigcup_{t \in Q_k} [L_{kt}]_k$. Näin ollen $\mathbb{N} \subseteq \bigcup_{t \in Q_k} [L_{kt}]_k$, eli $\{[L_{kt}]_k \mid t \in Q_k\}$ on äärellinen peite.

Valitaan jokaista joukon Q_∞ alkiota s kohti sellainen $t \in Q_k$, että leikkaus $[L_{hs}]_h \cap [L_{kt}]_k$ on ääretön: aina kun $s \in Q_\infty$, joukko $[L_{hs}]_h \subseteq \mathbb{N}$ on ääretön. Peitteessä $\{[L_{kt}]_k \mid t \in Q_k\}$ on äärellisen monta jäsentä, joten johonkin peitteen jäseneen kuuluu äärettömän monta joukon $[L_{hs}]_h$ alkiota. Olkoon $t \in Q_k$ jokin sellainen tila, että $[L_{kt}]_k$ on jokin tällainen peitteen jäsen. Valitaan kaksi erisuurta lukua x_s ja y_s leikkauksesta $[L_{hs}]_h \cap [L_{kt}]_k$.

Määritellään kokonaisluku

$$\xi = \max\{x_s, y_s \mid s \in Q_\infty\} + 1.$$

Lemman 6.3 nojalla on olemassa sellaiset kokonaisluvut $n \in \mathbb{N}_{>0}$ ja $m \in \mathbb{N}_{>0}$, että

$$|k^m - h^n| < \frac{1}{6\xi} h^n,$$

eli toisin sanoen

$$\xi |k^m - h^n| < \frac{1}{6} h^n. \tag{26}$$

Erityisesti $h^n - k^m \leq |k^m - h^n| \leq \frac{1}{6\xi} h^n \leq \frac{1}{6} h^n$, sillä $\xi \geq 1$. Lisäämällä k^m ja vähentämällä $\frac{1}{6} h^n$ puolittain saadaan epäyhtälö

$$\frac{5}{6} h^n \leq k^m. \tag{27}$$

Oletetaan, että luvun $k^m - h^n$ merkki on sama kuin luvun $x_s - y_s$ merkki. Ellei näin ole, voidaan luvun x_s ja luvun y_s nimet vaihtaa päittäin. Määritellään

$$p_s = (x_s - y_s)(k^m - h^n) \geq 0.$$

Koska k ja h ovat multiplikatiivisesti riippumattomia $k^m - h^n \neq 0$, joten $p_s > 0$. Siten

$$\begin{aligned} p_s &= (x_s - y_s)(k^m - h^n) \\ &= |(x_s - y_s)(k^m - h^n)| \\ &= |(x_s - y_s)||(k^m - h^n)|. \end{aligned}$$

Koska x_s ja y_s ovat positiivisia, luvun ξ määritelmän nojalla $|x_s - y_s| \leq \xi$. Siten

$$|(x_s - y_s)||(k^m - h^n)| \leq \xi |k^m - h^n|.$$

Epäyhtälön (26) nojalla

$$p_s \leq \frac{h^n}{6}. \tag{28}$$

Osoitetaan, että aina kun $s \in Q_\infty$ ja $x \in [L_{hs}]_h$, jonolla $\mathbf{a}$ on paikallinen jakso p_s välillä $I_x = B((x+1)h^n, \frac{2}{3}h^n)$. Havaitaan, että

$$I_x = B\left((x+1)h^n, \frac{2}{3}h^n\right) = xh^n + B\left(h^n, \frac{2}{3}h^n\right).$$

Siten aina kun luvut y ja $y + p_s$ ovat välillä I_x, voidaan valita sellainen lukupari $(z, z + p_s) \in B\left(h^n, \frac{2}{3}h^n\right) \times B\left(h^n, \frac{2}{3}h^n\right)$, että $(y, y + p_s) = (xh^n + z, xh^n + z + p_s)$. Näin ollen paikallisen jaksollisuuden osoittamiseksi riittää osoittaa, että $a_{xh^n+z} = a_{xh^n+z+p_s}$ aina kun $(z, z + p_s) \in B\left(h^n, \frac{2}{3}h^n\right)^2$. Koska $B\left(h^n, \frac{2}{3}h^n\right) \subseteq B(h^n, h^n)$ lemman 6.7 nojalla z ja $z + p_s$ kuuluvat joukkoon $[D_h^n]_h$. Koska lisäksi $x \in [L_{hs}]_h$ ja $y_s \in [L_{hs}]_h$, yhtälön (25) nojalla

$$a_{xh^n+z} = a_{y_s h^n+z}.$$

Lisäämällä ja vähentämällä indeksistä $y_s k^m$ saadaan yhtälöketju

$$\begin{aligned} a_{y_s h^n+z} &= a_{y_s(k^m - k^m + h^n)+z} \\ &= a_{y_s k^m + z - y_s(k^m - h^n)}. \end{aligned}$$

Osoitetaan, että $z - y_s(k^m - h^n) \in [D_k^m]_k$, jolloin voidaan soveltaa yhtälöä (25) symboliin $a_{y_s k^m + z - y_s(k^m - h^n)}$. Luvun $z - y_s(k^m - h^n)$ etäisyys luvusta k^m on

$$\begin{aligned}|z - y_s(k^m - h^n) - k^m| &= |z - h^n + h^n - y_s(k^m - h^n) - k^m| \\ &= |z - h^n - (y_s + 1)(k^m - h^n)|.\end{aligned}$$

Kolmioepäyhtälön nojalla

$$|z - h^n - (y_s + 1)(k^m - h^n)| \leq |z - h^n| + (y_s + 1)|k^m - h^n|.$$

Koska $z \in B\left(h^n, \frac{2}{3}h^n\right)$

$$|z - h^n| + (y_s + 1)|k^m - h^n| \leq \frac{2}{3}h^n + (y_s + 1)|k^m - h^n|.$$

Luvun ξ määritelmän nojalla $(y_s + 1) \leq \xi$ joten

$$\frac{2}{3}h^n + (y_s + 1)|k^m - h^n| \leq \frac{2}{3}h^n + \xi|k^m - h^n|.$$

Epäyhtälön (26) nojalla

$$\frac{2}{3}h^n + \xi|k^m - h^n| < \frac{2}{3}h^n + \frac{1}{6}h^n = \frac{5}{6}h^n.$$

Siten epäyhtälön (27) nojalla $|z - y_s(k^m - h^n) - k^m| \leq \frac{5}{6}h^n \leq k^m$, eli toisin sanoen $z - y_s(k^m - h^n) \in B(k^m, k^m)$. Lemman 6.7 nojalla $B(k^m, k^m) \subseteq [D_k^m]_k$, joten $z - y_s(k^m - h^n) \in [D_k^m]_k$.

Käyttämällä nyt yhtälöä (25) saadaan yhtälö

$$a_{y_s k^m + z - y_s(k^m - h^n)} = a_{x_s k^m + z - y_s(k^m - h^n)},$$

josta saadaan muokkaamalla yhtälöketju

$$\begin{aligned}a_{x_s k^m + z - y_s(k^m - h^n)} &= a_{x_s(h^n + k^m - h^n) + z - y_s(k^m - h^n)} \\ &= a_{x_s h^n + x_s(k^m - h^n) + z - y_s(k^m - h^n)} \\ &= a_{x_s h^n + z + (x_s - y_s)(k^m - h^n)}.\end{aligned}$$

Määritelmän nojalla $(x_s - y_s)(k^m - h^n) = p_s$, joten

$$a_{x_s h^n + z + (x_s - y_s)(k^m - h^n)} = a_{x_s h^n + z + p_s}.$$

Koska $z + p_s \in [D_h^n]_h$ ja $x_s \in [L_{hs}]_h$, voidaan käyttää yhtälöä (25) alkioon $a_{x_s h^n + z + p_s}$ jolloin saadaan haluttu yhtäsuuruus

$$a_{xh^n + z} = a_{xh^n + z + p_s}.$$

Näin ollen jokaista tilaa $s \in Q_\infty$ ja lukua $x \in [L_{hs}]_h$ kohti jonolla $\mathbf{a}$ on paikallinen jakso p_s välillä I_x.

Olemme nyt muodostaneet jokaista tilaa $s \in Q_\infty$ kohti kokoelman $\{I_x \mid x \in [L_{hs}]_h\}$ välejä, joissa jonolla $\mathbf{a}$ on paikallinen jakso p_s. Seuraavaksi muodostamme jonon välejä, jotka peittävät kaikki luonnolliset luvut jostakin luvusta alkaen. Käytetään näitä välejä laajentamaan jonkin paikallisen jakson vaikutusalue kattamaan mielivaltaisen suuria lukuja.

Kokoelma $\{[L_{ht}]_h \mid t \in Q_h\}$ on luonnollisten lukujen äärellinen peite, mikä voidaan todistaa samalla tavalla kuin kantaluvulle k. Alikokoelma $\{[L_{ht}]_h \mid t \in Q_h \setminus Q_\infty\}$ peittää vain äärellisen määrän lukuja, joten on olemassa suurin luku, jonka sen jäsenet peittävät. Siksi on olemassa sellainen luku $m \in \mathbb{N}$, että kokoelma $\{[L_{ht}]_h \mid t \in Q_\infty\}$ peittää joukon $m+\mathbb{N}$. Olemme osoittaneet, että jokaista lukua $i \geq m$ kohti on olemassa paikallinen jakso p_i välillä I_i. Yhtälön (28) nojalla jokainen paikallinen jakso p_i on lyhyempi kuin $\frac{1}{6}h^n$.

Osoitetaan induktiolla, että aina kun $j \geq m$ jonolla $\mathbf{a}$ on paikallinen jakso p_m välillä $\bigcup_{i=m}^{j} I_i$.

1° Kun $j = m$ joukko

$$\bigcup_{i=m}^{j} I_i = \bigcup_{i=m}^{m} I_i = I_m,$$

on väli missä jonolla $\mathbf{a}$ on paikallinen jakso p_m.

2° Oletetaan, että $\bigcup_{i=m}^{j} I_i$ on väli, ja että jonolla $\mathbf{a}$ on paikallinen jakso p_m välillä $\bigcup_{i=m}^{j} I_i$. Jonolla $\mathbf{a}$ on paikallinen jakso p_{j+1} välillä $I_{j+1} = B((j+1)h^n + h^n, \frac{2}{3}h^n)$. Koska

$$\max I_j = (j+1)h^n + \left\lfloor \tfrac{2}{3}h^n \right\rfloor = \left\lfloor \left(j+1+\tfrac{2}{3}\right) h^n \right\rfloor = \left\lfloor \left(j+\tfrac{5}{3}\right) h^n \right\rfloor$$

on suurempi kuin

$$\min I_{j+1} = (j+1+1)h^n - \left\lfloor \tfrac{2}{3}h^n \right\rfloor = \left\lceil \left(j+\tfrac{4}{3}\right) h^n \right\rceil,$$

$\bigcup_{i=m}^{j+1} I_i$ todella on väli.

Lisäksi

$$\begin{aligned}
\left|\left(\bigcup_{i=m}^{j} I_i\right) \cap I_{j+1}\right| &\geq |I_j \cap I_{j+1}| \\
&= \max I_j - \min I_{j+1} + 1 \\
&= \left\lfloor \left(j + \tfrac{5}{3}\right) h^n \right\rfloor - \left(\left\lceil \left(j + \tfrac{4}{3}\right) h^n \right\rceil - 1\right) \\
&\geq \left\lfloor \left(j + \tfrac{5}{3}\right) h^n \right\rfloor - \left\lfloor \left(j + \tfrac{4}{3}\right) h^n \right\rfloor \\
&= \left\lfloor \left(j + \tfrac{5}{3}\right) h^n - \left(j + \tfrac{4}{3}\right) h^n \right\rfloor \\
&= \left\lfloor \tfrac{1}{3} h^n \right\rfloor \\
&\geq 2 \left\lfloor \tfrac{1}{6} h^n \right\rfloor \\
&\geq p_m + p_{j+1}.
\end{aligned}$$

Edellä viimeisin epäyhtälö johtuu epäyhtälöstä (28) ja siitä, että p_m ja p_{j+1} ovat kokonaislukuja. Lemman 6.5 nojalla jonolla $\mathbf{a}$ on paikallinen jakso p_m välillä $\bigcup_{i=m}^{j+1} I_i$.

Koska

$$\lim_{j \to \infty} \bigcup_{i=m}^{j} I_i = \min I_m + \mathbb{N},$$

jonolla $\mathbf{a}$ on paikallinen jakso p_m välillä $\min I_m + \mathbb{N}$, toisin sanoen $\mathbf{a}$ on lopulta jaksollinen. Olemme osoittaneet Cobhamin lauseen.

□

Esimerkki 6.9. Sanotaan, että sanassa $w \in \Sigma^*$ on *päällekkäisyys*, mikäli on olemassa sellainen symboli $a \in \Sigma$ ja sana $u \in \Sigma^*$, että $auaua$ on sanan w osasana. Kirjassa [7] osoitetaan, että Thuen-Morsen sanassa $\mathbf{t}$ ei ole päällekkäisyyksiä. Osoitetaan, että $\mathbf{t}$ ei ole lopulta jaksollinen. Jos $\mathbf{t}$ olisi lopulta jaksollinen, se sisältäisi osasanan www, missä $w \neq \varepsilon$ on aakkoston $\{0, 1\}$ sana. Oletetaan, että a on sanan w ensimmäinen symboli. Tällöin on olemassa sellainen sana $u \in \{0, 1\}^*$, että $w = au$. Sana $www = auauau$ on Thuen-Morsen sanan osasana, joten $\mathbf{t}$ sisältää päällekkäisyyden, mikä on ristiriita. Siten $\mathbf{t}$ ei ole lopulta jaksollinen.

Esimerkin 5.19 nojalla $\mathbf{t}$ on 2^m-automaattinen aina kun $m \geq 1$ on kokonaisluku. Cobhamin lauseen nojalla $\mathbf{t}$ on k-automaattinen vain silloin kun k on luvun 2 potenssi. Muut kantaluvut ovat nimittäin multiplikatiivisesti riippumattomia luvun 2 kanssa, ja $\mathbf{t}$ ei ole lopulta jaksollinen.

7 Lopuksi

Määrittelimme k-automaattiset jonot käyttämällä k-funktioautomaatteja. Funktioautomaatille syötettiin jonon indeksien tavallisia eniten merkitsevästä numerosta alkavia k-kantaesityksiä, ja saatu tuloste asetettiin jonon jäseneksi. Lauseiden 4.4, 4.7, 4.9 ja 5.16 nojalla k-automaattiset jonot voidaan määritellä syöttämällä funktioautomaateille jonon jäsenten indeksien mitä tahansa k-kantaesityksiä.

Seurauslauseen 5.5 nojalla jokainen lopulta jaksollinen jono on k-automaattinen riippumatta kantaluvusta k. Cobhamin lauseen nojalla jos jono $\mathbf{a}$ on $k-$ ja h-automaattinen ja k ja h ovat multiplikatiivisesti riippumattomia, $\mathbf{a}$ on lopulta jaksollinen. Tällöin Cobhamin lauseen ja lauseen 5.18 nojalla k-automaattinen jono $\mathbf{a}$ toteuttaa vain toisen seuraavista väitteistä:

1. Jono $\mathbf{a}$ on k-automaattinen olipa kantaluku $k \geq 2$ mikä tahansa kokonaisluku. Lauseen 5.5 ja Cobhamin lauseen nojalla tällaiset jonot ovat tarkalleen lopulta jaksollisia jonoja.

2. Jono $\mathbf{a}$ on k-automaattinen vain silloin kun k on jonkin kiinnitetyn luvun potenssi. Esimerkin 6.9 nojalla Thuen-Morsen jono on tällainen 2-automaattinen jono.

Kaikkien automaattisten jonojen joukko koostuu siis kahdesta osajoukosta, lopulta jaksollisista jonoista, ja niistä jonoista jotka eivät ole lopulta jaksollisia. Näistä edellinen joukko koostuu yksinkertaisemmista jonoista, nimittäin lopulta jaksollisista. Jälkimmäinen sisältää monimutkaisempia jonoja, mutta kuitenkin siinä määrin säännöllisiä, että niiden generoimiseksi riittää käyttää äärellisiä funktioautomaatteja.

Cobhamin pienen lauseen ja lauseen 5.21 nojalla jonon $\mathbf{a} = (a_n) \in \Delta^{\mathbb{N}}$ k-automaattisuudella on kolme eri ekvivalenttia tulkintaa:

1. On olemassa sellainen k-funktioautomaatti $\mathcal{M} = (Q, \Sigma, \delta, q_0, \Delta, \tau)$, että $a_{[w]_k} = \tau(\delta(q_0, w))$ aina kun w on aakkoston Σ sana.

2. On olemassa sellainen k-uniformi morfismi $\varphi : \Gamma^* \longrightarrow \Gamma^*$, sellainen koodaus $\rho : \Gamma \longrightarrow \Delta$ ja sellainen morfismin φ kiintopiste $\mathbf{w}$, että $\mathbf{a} = \rho(\mathbf{w})$.

3. Jonon $\mathbf{a}$ ydin $\mathrm{Ker}_k(\mathbf{a})$ on äärellinen.

Tässä tutkielmassa tarkasteltiin vain k-automaattisia jonoja, joiden jäsenten indeksöintiin käytettiin luonnollisia lukuja. Indeksijoukkona voitaisiin käytää myös joukkoa $\mathbb{N}^m$, tai jotain muuta algebrallista rakennetta, kuten Gaussin kokonaislukuja. Näitä ja muita k-automaattisten jonojen yleistyksiä on esitelty kirjassa [2].

Viitteet

[1] Thijmen J. P. Krebs. A More Reasonable Proof of Cobham's Theorem. *International Journal of Foundations of Computer Science*, 32(02):203–207, 2021.

[2] Jean-Paul Allouche and Jeffrey Shallit. *Automatic Sequences. Theory, Applications, Generalizations.* Cambridge University Press, 2003.

[3] M. Lothaire. *Algebraic combinatorics on words.* Encyclopedia of mathematics and its applications ; 90. Cambridge University Press, Cambridge, 2002.

[4] Jarkko Kari. *Automata and formal languages.* 2020.

[5] Juraj Šebej. Reversal of regular languages and state complexity. In *ITAT*, 2010.

[6] Aviezri S. Fraenkel. The use and usefulness of numeration systems. *Information and Computation*, 81(1):46–61, 1989.

[7] Jean Berstel, Aaron Lauve, Christophe Reutenauer, and Franco Saliola. *Combinatorics on Words: Christoffel Words and Repetition in Words.* 2008.