TURUN KAUPPAKORKEAKOULU
Tilastotieteen ja talous-
matematiikan laitos
Ilkka Virtanen

DISKREETIN ÄÄRELLISSOLMUISEN MARKOV-PROSESSIN TARKASTELU YLEISEN SYSTEEMITEORIAN ESIMERKKINÄ

1. Johdanto

Tässä kirjoituksessa tarkastellaan diskreetin äärellissolmuisen Markov-prosessin perusominaisuuksia ja käyttäytymistä yleisen systeemiteorian lähestymistapaa noudattaen. Samalla esitys on selvitys yleisen systeemiteorian alkeista helposti hallittavan matemaattisen mallin puitteissa esitettynä. Esityksen sisällöllisenä perustana on R. A. HOWARDIN erinomaisessa dynaamisia todennäköisyyssysteemejä käsittelevässä teoksessaan[1] esittämä diskreettien Markov-prosessien teoria. Howardin esittämät tulokset "käännetään" nyt yleisen systeemiteorian kielelle. Tässä käännöstyössä sovelletaan S. SALOVAARAN[2] ja P. MALASKAN[3] systeemiteorian perusteita käsittelevissä puheenvuoroissaan esittämiä formalismeja. Tarkastelut suoritetaan verraten korkealla abstraktiotasolla, joskin käsitteiden ja tulosten tulkintaan sekä teorian soveltamismahdollisuuksiin ainakin jossain määrin viitataan.

---

1. R. A. Howard: Dynamic Probabilistic Systems, Volume I: Markov Models, John Wiley & Sons, Inc., New York 1971
2. S. Salovaara: Systeemiteorian joukko-opillisista perusteista, esitelmä Helsingin Yliopiston tutkijakollokviossa (G. H. von Wright, O. E. Niitamo) 14.3.1968
3. P. Malaska: Kansantalouden malli systeemiteorian esimerkkinä, alustus Turun Kauppakorkeakoulun tutkijaseminaarissa 1973

2. Diskreetti äärellissolmuinen Markov-prosessi

2.1. Peruskäsitteet

Tarkastellaan tiettyä, toistaiseksi tarkemmin määrittelemätöntä ilmiötä ("reaalisysteemiä"), josta tehdään tai ainakin voidaan ajatella tehtävän havaintoja diskreetteinä ajanhetkinä $t_0, t_1, \ldots$. Yksikäsitteistä kuvausta ilmiöstä hetkellä t_n $(n=0,1,\ldots)$ kutsutaan tavallisesti ilmiön tilaksi tuolla hetkellä. Tila voidaan ilmaista esimerkiksi verbaalisena kuvauksena, tarkastelujen kannalta relevanttien muuttujien arvoina jne. Seuraavassa oletetaan, että ilmiöllä on vain äärellinen määrä mahdollisia tiloja. Numeroidaan nämä tilat $1,2,\ldots,N$.

Siirryttäessä jostakin ajanhetkestä t_n seuraavaan hetkeen t_{n+1} ilmiötä luonnehtivien tilasuureiden arvot yleensä muuttuvat, ilmiön tila muuttuu. Tällaista muutosta kutsutaan (tila-)siirtymäksi. Tarkasteltavan ilmiön oletetaan vielä olevan luonteeltaan stokastisen, jolloin siirtymät tapahtuvat tiettyjen todennäköisyyslakien, siirtymätodennäköisyyksien mukaisesti. Yleisessä tapauksessa siirtymätodennäköisyydet riippuvat siirtymän ajankohdasta t_n ja ilmiön tiloista ajanhetkinä $t_0, t_1, \ldots, t_n$. Kun merkitään symbolilla X suuretta, joka ilmoittaa ilmiön tilan numeron kullakin hetkellä t_n, ts.

(1) $X(t_n)$: ilmiön tila hetkellä t_n, $n=0,1,\ldots$, on $X(t_n)$,
missä $X(t_n)=1,2,\ldots,N$,

niin voidaan todeta, että $X(t_n)$ on satunnaissuure, jonka jakautuma määräytyy siirtymätodennäköisyyksien ja tiettyjen alkuehtojen perusteella. Siirtymätodennäköisyydet ovat yleisessä tapauksessa ehdollisia todennäköisyyksiä

$$(2) \quad P\left\{X(t_{n+1})=j \mid X(t_n)=i_n, X(t_{n-1})=i_{n-1}, \ldots, X(t_0)=i_0\right\},$$

missä $n=0,1,\ldots,\ 1 \leq i_0, i_1, \ldots, i_n, j \leq N$. Äärettömän monidimensionaalista satunnaissuuretta

$$(3) \quad X = (X(t_0), X(t_1), \ldots, X(t_n), \ldots),$$

johon liittyvät (2):n mukaiset siirtymien valintasäännöt, kutsutaan tarkasteltavaa ilmiötä kuvaavaksi stokastiseksi prosessiksi. Jokainen havaittu perättäisiä ajanhetkiä vastaavien tilojen jono

$$(4) \quad (x(t_0), x(t_1), \ldots, x(t_n), \ldots)$$

on tämän prosessin yksityinen realisaatio.

Lausekkeesta (2) huomataan, että prosessin hallitsemiseksi tarvittava informaatiomäärä on yleisessä tapauksessa tavattoman suuri. Tämän vuoksi tarkastelut usein rajataan vain Markov-prosesseja käsittäviksi: oletetaan, että prosessin tulevaan käyttäytymiseen vaikuttaa vain sen tarkasteluhetken tila. Tällä oletta-muksella on

$$(5) \quad P\left\{X(t_{n+1})=j \mid X(t_n)=i_n, X(t_{n-1})=i_{n-1}, \ldots, X(t_0)=i_0\right\} = P\left\{X(t_{n+1})=j \mid X(t_n)=i_n\right\}.$$

Markov-prosessin siirtymätodennäköisyyksiä voidaan siten merkitä

$$(6) \quad P\left\{X(t_{n+1})=j \mid X(t_n)=i\right\} = p_{ij}(t_n),\ 1 \leq i,j \leq N,\ n=0,1,\ldots$$

Siirtymätodennäköisyydet voidaan esittää myös matriisina

$$(7) \quad P(t_n) = \left\{p_{ij}(t_n)\right\} = \begin{bmatrix} p_{11}(t_n) & p_{12}(t_n) & \ldots & p_{1N}(t_n) \\ p_{21}(t_n) & p_{22}(t_n) & \ldots & p_{2N}(t_n) \\ \vdots & \vdots & & \vdots \\ p_{N1}(t_n) & p_{N2}(t_n) & \ldots & p_{NN}(t_n) \end{bmatrix},\ n=0,1,\ldots$$

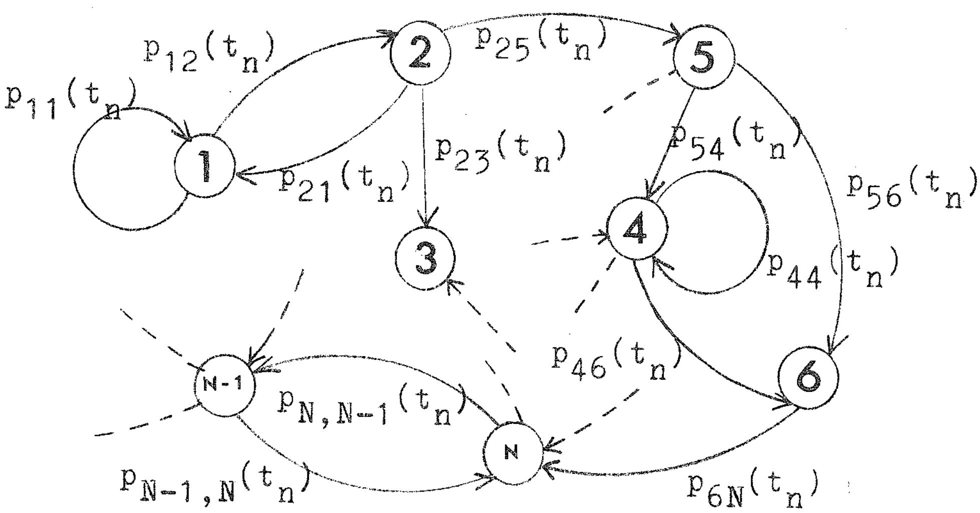
Jokainen matriisi $P(t_n)$, $n=0,1,\ldots$, on luonnollisesti stokastinen matriisi: niiden alkiot toteuttavat ehdon

(8) $$0 \leq p_{ij}(t_n) \leq 1,\quad 1 \leq i,j \leq N,\quad n=0,1,\ldots$$

ja rivisummat ovat = 1,

(9) $$\sum_{j=1}^{N} p_{ij}(t_n) = 1,\quad i=1,2,\ldots,N,\quad n=0,1,\ldots$$

Markov-prosessia havainnollistetaan tavallisesti verkoilla, kutakin siirtymätodennäköisyysmatriisia vastaa prosessin siirtymädiagrammi (kuva 1). Verkon solmut vastaavat ilmiön tiloja ja solmujen väliset haarat siirtymätodennäköisyyksiä.



Kuva 1. Ajanhetkeä t_n vastaava Markov-prosessin siirtymädiagrammi

Myöhemmin, kun siirrytään Markov-prosessin systeemiteoreettisiin tarkasteluihin, systeemin tila-käsitteellä on aivan erityinen merkityksensä. Sekaannuksen välttämiseksi luovutaan nyt edellä käyttöön otetusta ilmiön tila-nimityksestä ja kutsutaan tätä käsitettä seuraavassa solmuksi. Tälle nimitykselle on löydettävissä perusteluja esimerkiksi siirtymädiagrammista, jossa verkon solmu vastaa nyt ilmiön tai sitä kuvaavan prosessin solmua.

2.2. Markov-prosessin käyttäytymisestä[1]

Keskeisellä sijalla Markov-prosessin käyttäytymisen tarkastelussa ovat moniaskeliset siirtymätodennäköisyydet

(10) $\phi_{ij}(t_m,t_n) = P\{X(t_n)=j \mid X(t_m)=i\}$, $1 \leq i,j \leq N$, $0 \leq m < n$.

Moniaskelisten siirtymätodennäköisyyksien matriisi $\phi(t_m,t_n) = \{\phi_{ij}(t_m,t_n)\}$ voidaan lausua siirtymätodennäköisyysmatriisien tulona

(11) $\phi(t_m,t_n) = P(t_m)P(t_{m+1})\cdots P(t_{n-1})$, $0 \leq m < n$.

Matriisi voidaan myös hajoittaa osiin

(12) $\phi(t_m,t_n) = \phi(t_m,t_r)\phi(t_r,t_n)$, $0 \leq m < r < n$.

Siinä erikoistapauksessa, että siirtymätodennäköisyydet eivät riipu lainkaan siirtymän ajankohdasta, ts. kun on voimassa

(13) $p_{ij}(t_0) = p_{ij}(t_1) = \ldots = p_{ij}(t_n) = \ldots = p_{ij}$, $1 \leq i,j \leq N$,

saa yhtälö (11) muodon

(14) $\phi(t_m,t_n) = P^{n-m}$, $0 \leq m < n$,

missä on merkitty $P = \{p_{ij}\}$. Erityisesti on tarkastelun alkuhetkeä lähtökohtana pitäen (m=0):

(15) $\phi(t_0,t_n) = P^n$, $n=1,2,\ldots$

Moniaskeliset siirtymätodennäköisyydet kuvaavat prosessin siirtymistä solmusta toiseen tietyn aikavälin kuluessa. Kun halutaan informaatiota pelkästään siitä, missä solmussa prosessi jollakin tietyllä hetkellä on riippumatta sen "miehittämistä" aikaisem-

1. Tässä jaksossa olevat Markov-prosessin käyttäytymistä koskevat tulokset esitetään yleensä johtamatta tai todistamatta, niiden yksityiskohtaisesta johtamisesta ks. R.A. Howard mt. ss. 6-19 ja 511-516.

mista solmuista, käytetään <u>solmutodennäköisyyksiä</u>

(16) $\pi_i(t_n) = P\{X(t_n) = i\}$, $i=1,2,\ldots,N$, $n=0,1,\ldots$

Merkitään näiden muodostamaa vektoria

(17) $\big(\pi(t_n) = \pi_1(t_n), \pi_2(t_n), \ldots, \pi_N(t_n)\big)$, $n=0,1,\ldots$

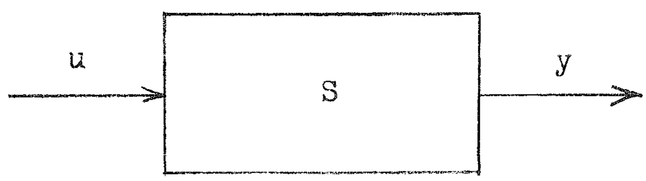
Ajanhetkeen t_n liittyvät solmutodennäköisyydet voidaan määrittää, kun tunnetaan solmutodennäköisyydet hetkellä t_0 sekä siirtymätodennäköisyydet hetkinä $t_0, t_1, \ldots, t_{n-1}$ tai moniaskeliset siirtymätodennäköisyydet $\phi_{ij}(t_0, t_n)$, $1 \leq i,j \leq N$:

(18) $\pi(t_n) = \pi(t_0)\phi(t_0,t_n) = \pi(t_0)P(t_0) \cdots P(t_{n-1})$, $n=1,2,\ldots$

3. <u>Markov-prosessin systeemiesitys</u>

3.1. <u>Systeemiesityksen joukko-opilliset perusteet</u>

Kuvassa 2 on esitetty systeemikaavion yleinen muoto:

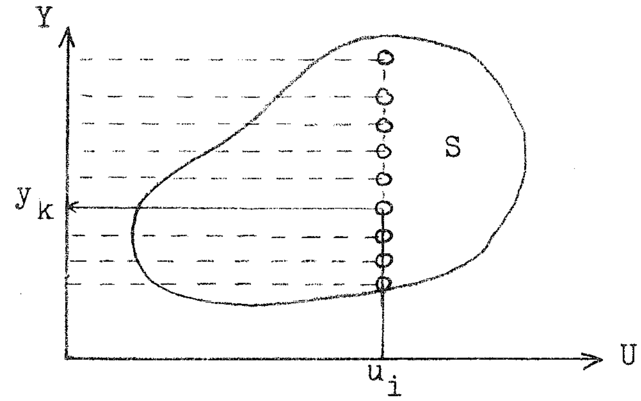


Kuva 2. Systeemin kaavioesitys

Tässä u on <u>heräte</u> eli <u>syöttösuure</u>, y <u>ulostulo</u> eli <u>tulossuure</u> ja S itse <u>systeemi</u>. Systeemi on täsmällisesti määritelty, kun tunnetaan mahdollisten herätteiden joukko $U = \{u\}$, mahdollisten ulostulojen joukko $Y = \{y\}$ sekä relaatio S, joka ilmoittaa, mitkä herätteet ja ulostulot voivat tämän systeemin puitteissa olla toisiaan vastaavia. Joukko-opillisesti systeemi on siis relaatio

(19) $\quad S = \{(u,y) \mid u \in U,\ y \in Y,\ u \rightarrow y\}$,

missä merkintä $u \rightarrow y$ tarkoittaa, että y voi esiintyä ulostulona, kun u esiintyy herätteenä. Kuvassa 3 y_k on eräs niistä ulostulomahdollisuuksista y, jotka voivat esiintyä ulostulona, kun herätteenä on u_i; muut u_i:tä vastaavat ulostulot on merkitty katkoviivoin.



Kuva 3. Systeemi joukko-opillisena relaationa

Konstruoidaan nyt Markov-prosessille edellä määritellyt perussuureet. Oletetaan, että ilmiötä tarkastellaan diskreetteinä ajanhetkinä $t_0, t_1, \ldots, t_T$, <u>havaintovälinä</u> on siten $\mathcal{T} = [t_0, t_T]$. Valitaan herätteeksi prosessin siirtymätodennäköisyysmatriisien jono

(20) $\quad u = (P(t_0), P(t_1), \ldots, P(t_{T-1}))$,

missä kukin $P(t_i)$, $i=0,1,\ldots,T-1$, on siis ehtojen (8) ja (9) mukainen NxN stokastinen matriisi. Olkoon seuraavassa $\mathcal{P}$ NxN stokastisten matriisien joukko. Herätteiden joukko on tällöin

(21) $\quad U(t_0, t_T) = \{u\} \subset \mathcal{P} \times \mathcal{P} \times \cdots \times \mathcal{P} = \mathcal{P}^T$.

Tulossuureeksi valitaan solmutodennäköisyysvektorien jono

(22) $\quad y = (\pi(t_1), \pi(t_2), \ldots, \pi(t_T))$.

Tulossuureiden joukko on siten

(23) $\quad Y(t_0, t_T) = \{y\} \subset \sum \times \sum \times \cdots \times \sum = \sum^T$,

missä Σ:lla on merkitty muotoa

$$(24) \qquad (x_1, x_2, \ldots, x_N),\ 0 \le x_i \le 1,\ \sum_{i=1}^{N} x_i = 1$$

olevien vektoreiden muodostamaa joukkoa. Selvästikin on $\Sigma \subset R^N$. Tulossuureen y ja herätteen u välillä on yhtälön (18) mukainen yhteys. Markov-prosessia vastaava systeemi on siten

$$(25) \qquad S = \{(u,y) \mid u \in U(t_0,t_T), y \in Y(t_0,t_T),\ (18) \text{ on voimassa, } t_n \in \mathcal{T}\}.$$

Relaatiossa S tiettyä herätettä voi siis vastata useita eri ulostuloja. Markov-prosessin tapauksessa erityisesti tiettyä siirtymätodennäköisyysmatriisijonon muodostamaa herätettä vastaava ulostulo on tarkalleen määrätty vasta, kun alkuhetken solmutodennäköisyysvektori $\pi(t_0)$ on tunnettu. Syytä siihen, että systeemi antaa samalla herätteellä eri kerroilla eri ulostulon, kutsutaan systeemin alkutilaksi. Sen jälkeen kun alkutila on kiinnitetty, heräte määrää ulostulon yksikäsitteisesti. Markov-prosessi esimerkissä alkutilana on selvästi hetkeen t_0 liittyvä solmutodennäköisyysvektori $\pi(t_0)$.

Kiinteällä alkutilalla $\pi^0 = \pi(t_0) \in \Sigma$ yhtälö (18) määrittelee siten relaation herätteiden joukolta ulostulojen joukolle

$$(26) \qquad \Gamma(\pi^0): U(t_0,t_T) \to Y(t_0,t_T), \text{ ts.}$$

$$(27) \qquad \Gamma(\pi^0) = \{(u,y) \mid u \in U(t_0,t_T), y \in Y(t_0,t_T),\ u \text{ ja } y \text{ tot. } (18)\text{:n}\},$$

joka on kuvaus, ts. jokaista $u \in U(t_0,t_T)$ vastaa yksi ja vain yksi (yhtälön (18) perusteella määräytyvä) $y \in Y(t_0,t_T)$. Tällä kuvauksella on selvästi ominaisuudet

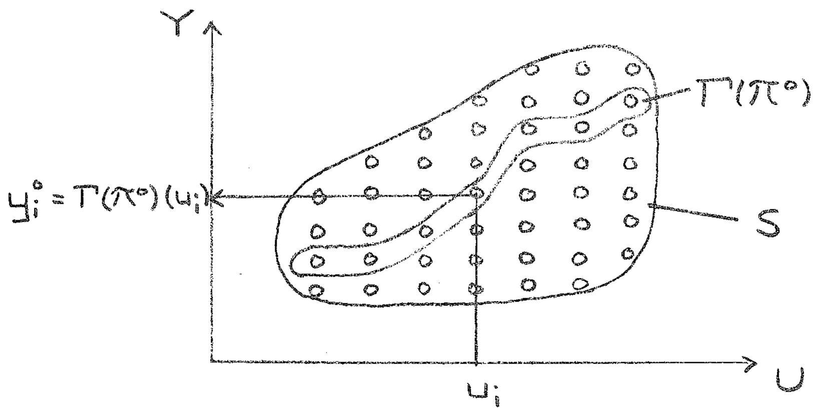
$$(28) \qquad \Gamma(\pi^0) \subset S,\ \forall\, \pi^0 \in \Sigma,$$

$$(29) \qquad \bigcup_{\pi^0 \in \Sigma} \Gamma(\pi^0) = S.$$

Joukkoa $\sum$, joka ilmoittaa alkutilan mahdolliset arvot, kutsutaan systeemin alkutila-avaruudeksi. Alkutilojen mukaan S on siten tullut jaetuksi osajoukkoihin $\Gamma(\pi^0)$, joista kukin on funktio ja jotka yhdessä peittävät S:n; S:lle on konstruoitu kuvauspeite. Kussakin kuvauspeitteen alkiossa $\Gamma(\pi^0)$ heräte määrää yksikäsitteisesti ulostulon, mikä voidaan vielä merkitä

(30) $y = \Gamma(\pi^0)(u)$.

Funktio $\Gamma(\pi^0)$ on kaavamaisesti esitetty kuvassa 4.



Kuva 4. Systeemi alkutila π^0 kiinnitettynä

Kuvausperhe

(31) $\mathcal{F} = \{\Gamma(\pi^0) \mid \pi^0 \in \sum\}$

on edellä esitetyn mukaisesti eräs S:n kuvauspeite. Kutakin alkutilaa $\pi^0 \in \sum$ vastaa tarkalleen yksi $\Gamma(\pi^0) \in \mathcal{F}$. Näin Γ voidaan tulkita myös kuvaukseksi alkutila-avaruudesta $\sum$ kuvausperheelle $\mathcal{F}$:

(32) $\Gamma\colon \sum \to \mathcal{F}$.

3.2. Systeemin tila mielivaltaisella ajanhetkellä

Seuraavassa tavoitteena on määritellä ja konstruoida systeemille S tila hetkellä $t_k \in \mathcal{T}$. Tällä tilakäsitteellä halutaan olevan vastaavat ominaisuudet kuin alkutilalla $\pi(t_0)$: kun tila hetkellä t_k on kiinnitetty, niin jokaista herätteen u väliä $[t_k, t_T]$ vastaavaa osaherätettä kohti on olemassa yksi tarkalleen määrätty ulostulon y osaulostulo välillä $[t_k, t_T]$. Tilaan liittyvät kuvaus ja kuvauspeite halutaan vielä muodostaa niin, että alkuperäisen systeemin vastaavia suureita on käytetty hyväksi.

Merkitään yleisesti

(33) $u_{t_m, t_n} = (P(t_m), \ldots, P(t_{n-1}))$, $0 \leq m < n \leq T$

(34) $y_{t_m, t_n} = (\pi(t_{m+1}), \ldots, \pi(t_n))$, $0 \leq m < n \leq T$

ts. u_{t_m, t_n} ja y_{t_m, t_n} ovat herätteestä $u \in U(t_0, t_T)$ ja ulostulosta $y \in Y(t_0, t_T)$ muodostetut, aikaväliä $[t_m, t_n]$ vastaavat osaheräte ja -ulostulo. Muodostetaan nyt S:n avulla kaksi uutta relaatiota erottamalla S:n heräte-ulostulopareista (u,y) toisaalta alkupäät t_0:sta t_k:hon ja toisaalta loppupäät t_k:sta t_T:hen:

(35) $S_{t_0, t_k} = \{(u_{t_0, t_k}, y_{t_0, t_k}) \mid (u, y) \in S\}$

(36) $S_{t_k, t_T} = \{(u_{t_k, t_T}, y_{t_k, t_T}) \mid (u, y) \in S\}$.

Ajatellaan nyt, että on siirrytty t_0:sta hetkeen $t_k \in \mathcal{T}$. S:n sijasta tarkastellaan S_{t_k, t_T}:tä. S_{t_k, t_T}:n kohdalla halutaan saada aikaan sama tilanne kuin S:n kohdalla edellä: S_{t_k, t_T} halutaan jakaa kuvauspeitteen muodostaviin osajoukkoihin (funktioihin) niin, että tämä jako samalla määrittelisi systeemin tilan hetkellä S. Tämä merkitsee, että taulukon 1 oikeanpuoleisessa sarakkeessa olevat suureet on määriteltävä siten, että niillä ajanhetkeen t_k liittyen on vastaavat ominaisuudet kuin taulukon vasemmanpuoleisen sarakkeen suureilla on t_0 lähtökohtana. Seuraavassa konstruoidaan tälle tehtävälle eräs luonnollinen ratkaisu.

$t = t_0$	$t = t_k$
1^o Relaatio: S	1^o Relaatio: S_{t_k,t_T}
2^o Alkutila: π^0	2^o Tila hetkellä t_k: π^k
3^o Alkutila-avaruus: Σ	3^o Tila-avaruus hetkellä t_k: Σ_k
4^o Kuvaus: $\Gamma(\pi^0) \subset S$	4^o Kuvaus: $\Gamma_k(\pi^k) \subset S_{t_k,t_T}$
5^o Kuvauspeite: $\mathcal{F} = \{\Gamma(\pi^0) \mid \pi^0 \in \Sigma\}$ $\bigcup_{\pi^0 \in \Sigma} \Gamma(\pi^0) = S$	5^o Kuvauspeite: $\mathcal{F}_k = \{\Gamma_k(\pi^k) \mid \pi^k \in \Sigma_k\}$ $\bigcup_{\pi^k \in \Sigma_k} \Gamma_k(\pi^k) = S_{t_k,t_T}$

Taulukko 1. Systeemin tila ja siihen liittyvät suureet hetkillä $t = t_0$ ja $t = t_k$.

1^o. Relaatio S_{t_k,t_T} tuli jo määritellyksi yhtälössä (36). Selvästikin relaatio on konstruoitu alkuperäiseen S:ään turvautuen, kuten tavoitteena oli.

2^o. Määritellään systeemin tila hetkellä t_k alkutilan π^0 ja välin $[t_0,t_k]$ osaherätteen u_{t_0,t_k} muodostamana parina:

$$(37) \quad \pi^k = (\pi^0, u_{t_0,t_k}), \ \pi^0 \in \Sigma, \ u \in U(t_0,t_T).$$

Ottamalla huomioon u_{t_0,t_k}:n määritelmä (33) saadaan Markov-prosessin tilaksi

$$(38) \quad \pi^k = (\pi^0, (P(t_0), P(t_1), \ldots, P(t_{k-1}))).$$

Tila hetkellä t_k muodostuu siis alkuhetken t_0 solmutodennäköisyysvektorista ja hetkien $t_0, \ldots, t_{k-1}$ siirtymätodennäköisyysmatriiseista.

3^o. Tilan π^k määrittelyn yhteydessä on hetkeen t_k liittyvä tila-avaruuskin tullut itse asiassa määriteltyä. Sitä voidaan merkitä

$$(39) \qquad \Sigma_k = \{(\pi^o, u_{t_0,t_k}) \mid \pi^o \in \Sigma,\ u \in U(t_0,t_T)\}$$
$$= \{(\pi^o, x) \mid \pi^o \in \Sigma,\ x \in U(t_0,t_k)\}.$$

4^o. Seuraavaksi konstruoidaan osaherätteiden joukolta $U(t_k,t_T)$ osaulostulojen joukolle $Y(t_k,t_T)$ relaatio, joka on S_{t_k,t_T}:n osajoukko ja kiinteällä tilalla π^k kuvaus. Olkoon tämän kohdan tarkastelujen ajan tila $\pi^k \in \Sigma_k$ kiinnitetty ja olkoon r mielivaltainen välin t_k,t_T osaheräte, ts.

$$(40) \qquad r = (P(t_k),\ldots,P(t_{T-1})).$$

Määritellään Γ_k nyt Γ:n avulla seuraavasti:

$$(41) \qquad \Gamma_k(\pi^k)(r) = \Gamma_k(\pi^o,x)(r)$$
$$= \Gamma_k(\pi^o,(P(t_0),\ldots,P(t_{k-1})))\ (P(t_k),\ldots,P(t_{T-1}))$$
$$= \left[\Gamma(\pi^o)\ (P(t_0),\ldots,P(t_{k-1}),P(t_k),\ldots,P(t_{T-1}))\right]_{t_k,t_T}$$
$$= \left[\Gamma(\pi^o)\ (x,r)\right]_{t_k,t_T},$$

missä merkintä $[\quad]_{t_k,t_T}$ tarkoittaa, että hakasuluissa olevasta ulostulosta otetaan vain väliä $[t_k,t_T]$ vastaava osa, ja (x,r):llä on merkitty vektoria, jonka komponentit on saatu asettamalla x:n ja r:n komponentit peräkkäin (k ja T-k komponenttiset vektorit on yhdistetty yhdeksi T komponenttiseksi vektoriksi). Koska Γ oli kuvaus, niin määritelmästä seuraa, että myös Γ_k on kuvaus. Selvästi on myös $\Gamma_k(\pi^k) \subset [\Gamma(\pi^o)]_{t_k,t_T}$ ja $[\Gamma(\pi^o)]_{t_k,t_T} \subset S_{t_k,t_T}$, joten

$$(42) \qquad \Gamma_k(\pi^k) \subset S_{t_k,t_T}.$$

5°. Lopuksi osoitetaan, että kuvausperhe $\mathcal{F}_k = \{\Gamma_k(\pi^k) | \pi^k \in \Sigma_k\}$ on S_{t_k,t_T}:n kuvauspeite. Edellä kohdassa 4° osoitettiin jo, että $\Gamma_k(\pi^k)$:t ovat kuvauksia ja S_{t_k,t_T}:n osajoukkoja. Vielä on siis osoitettava, että ne peittävät S_{t_k,t_T}:n, ts. että on voimassa

(43) $$\bigcup_{\pi^k \in \Sigma_k} \Gamma_k(\pi^k) = S_{t_k,t_T}.$$

Koska ehdon (42) nojalla on selvää, että $\bigcup_{\pi^k \in \Sigma_k} \Gamma_k(\pi^k) \subset S_{t_k,t_T}$, riittää (43):n oikeaksi todistamista varten osoittaa, että on myös $S_{t_k,t_T} \subset \bigcup_{\pi^k \in \Sigma_k} \Gamma_k(\pi^k)$. Tarkastellaan mielivaltaista osaheräte-ulostuloparia $(r,p) \in S_{t_k,t_T}$. Määritelmän (36) mukaan on olemassa $(u,y) \in S$ siten, että

(44) $$(r,p) = (u_{t_k,t_T}, y_{t_k,t_T}), \text{ ts.}$$

(45) $$r = u_{t_k,t_T} \text{ ja } p = y_{t_k,t_T}.$$

Koska $\bigcup_{\pi^0 \in \Sigma} \Gamma(\pi^0) = S$, on olemassa $\pi^0 \in \Sigma$ siten, että $(u,y) \in \Gamma(\pi^0)$ ts. että $y = \Gamma(\pi^0)(u)$. Valitaan nyt $x = u_{t_0,t_k}$ ja $\pi^k = (\pi^0, x)$. Tällöin on $u = (x,r)$ ja

(46) $$p = y_{t_k,t_T} = \left[\Gamma(\pi^0)(u)\right]_{t_k,t_T} = \left[\Gamma(\pi^0)(x,r)\right]_{t_k,t_T}$$
$$= \Gamma_k(\pi^0,x)(r) = \Gamma_k(\pi^k)(r),$$

jolloin $(r,p) \in \Gamma_k(\pi^k)$. Tästä seuraa, että $S_{t_k,t_T} \subset \bigcup_{\pi^k \in \Sigma_k} \Gamma_k(\pi^k)$. Kohdan 5° kaikki tulokset yhdistämällä nähdään, että kuvausperhe $\mathcal{F}_k = \{\Gamma_k(\pi^k) | \pi^k \in \Sigma_k\}$ muodostaa S_{t_k,t_T}:n kuvauspeitteen.

Kuvauspeitteen $\mathcal{F}_k$ konstruoinnissa oli parametrina π^k eli pari (π^0, u_{t_0,t_k}). Tätä paria sanotaan systeemin S tilaksi hetkellä t_k (alkutila π^0 ja heräte u kiinnitettynä). Systeemin tila hetkellä t_k koostuu siis alkutilasta π^0 ja herätteen historiasta väliltä $[t_0, t_k]$. Tämä merkitsee Markov-prosessin alkuhetken solmutodennäköisyysvektoria $\pi^0 = \pi(t_0)$ ja siirtymätodennäköisyysmatriisien jonoa $u_{t_0,t_k} = (P(t_0),\ldots,P(t_{k-1}))$. Markov-systeemin tilasta hetkellä t_k voidaan vielä todeta, että tilan tunteminen merkitsee samalla solmutodennäköisyyksien $\pi_i(t_k)$, $i=1,2,\ldots,N$, ja moniaskelisten siirtymätodennäköisyyksien $\phi_{ij}(t_0,t_k)$, $i,j=1,2,\ldots,N$ tuntemista, sillä onhan yhtälöiden (11) ja (18) nojalla

(47) $\pi(t_k) = \pi(t_0)P(t_0)\cdots P(t_{k-1})$ ja

(48) $\phi(t_0,t_k) = P(t_0)\cdots P(t_{k-1})$.

Sen sijaan käänteinen tulos ei yleisessä tapauksessa ole voimassa; solmutodennäköisyysvektorin $\pi(t_k)$ ja moniaskelisten siirtymätodennäköisyyksien matriisien $\phi(t_0,t_k)$ tunteminen ei riitä määrittämään systeemin tilaa hetkellä t_k.

Tarkastellaan tilan ominaisuuksia seuraavassa vielä lähemmin. Oletetaan alkutila $\pi^0 \in \Sigma$ ja heräte $u \in U(t_0,t_T)$ jälleen kiinnitetyiksi. Edellä esitettyjen tarkastelujen nojalla vastaa jokaista havaintovälin $\mathcal{T}$ ajanhetkeä t_k nyt yksi, tarkalleen määrätty tila $\pi^k = (\pi^0, u_{t_0,t_k}) \in \Sigma_k$. Tila voidaan siten tulkita kuvaukseksi joukolta $\mathcal{T}$ kaikkien mahdollisten tilojen joukolle $\bigcup_{k=1}^{T} \Sigma_k$:

(49) $\gamma : \mathcal{T} \rightarrow \bigcup_{k=1}^{T} \Sigma_k$,

missä

(50) $\gamma(t_k) = \pi^k = (\pi^0, u_{t_0,t_k}) \in \Sigma_k, \ \forall\, t_k \in \mathcal{T}$.

Tila on siten ulostuloon verrattavissa oleva suure, joka kehittyy täysin määrätyllä tavalla, kun systeemin alkutila on kiinnitetty ja systeemiin vaikuttaa tietty heräte. Tila voidaan tulkita myös systeemin muistiksi, siihen tallentuvat tiedot systeemin alkutilasta (π^0) ja systeemiin aikaisemmin vaikuttaneista eksogeenisista suureista (osaheräte väliltä $[t_0,t_k]$). Tästä syystä systeemin tuleva käyttäytyminen (ulostulo välillä $[t_k,t_T]$) voidaan johtaa tilasta tarkasteluhetkellä (π^0, u_{t_0,t_k}) ja systeemiin tulevaisuudessa vaikuttavista eksogeenisista tekijöistä (osaherätteestä välillä $[t_k, t_T]$). Hetken t_k tilalla π^k on siten täysin vastaavat ominaisuudet havaintoväliä $\mathcal{T}_k = [t_k, t_T]$ ajatellen kuin alkutilalla π^0 havaintovälillä $\mathcal{T} = [t_0, t_T]$. Havaintoväliä $\mathcal{T}_k$ koskevissa tarkasteluissa on vain siirryttävä taulukon 1 mukaisiin k-indeksoituihin suureisiin.

Tarkastellaan lopuksi tilan erästä varsin keskeistä ominaisuutta, ns. puoliryhmäominaisuutta. Ajatellaan $t_k \in \mathcal{T}$ kiinnitetyksi ja olkoon $t_{k'} \in \mathcal{T}$, $t_{k'} > t_k$ mielivaltainen. Hetkeen $t_{k'}$ liittyen voidaan nyt ajatella muodostettavaksi kaksi eri tilakäsitettä:

(a) käyttämällä alkuperäistä havaintoväliä $\mathcal{T} = [t_0, t_T]$ ja siihen liittyviä suureita, jolloin tila hetkellä $t_{k'}$ on $(\pi^0, u_{t_0, t_{k'}})$ tai

(b) siirtymällä uuteen havaintoväliin $\mathcal{T}_k = [t_k, t_T]$, jolloin tila hetkellä $t_{k'}$ on $(\pi^k, u_{t_k, t_{k'}}) = ((\pi^0, u_{t_0, t_k}), u_{t_k, t_{k'}})$.

On mahdollista osoittaa[1], että näin määritellyt tilakäsitteet ovat keskenään ekvivalentteja siinä mielessä, että ne määräävät saman kuvauspeitteen $S_{t_{k'}, t_T}$:lle (tilakäsitehän alunperin vastasi kuvauspeitteen muodostavan kuvausperheen määritystä). Markov-systeemissä (a)- ja (b)-kohtien mukaisten tilojen ekvivalenttisuus tuntuu varsin luonnolliselta, merkitseväthän ne kumpikin alkuhetken solmutodennäköisyysvektorin $\pi(t_0)$ ja siirtymätodennäköisyysmatriisien $P(t_0), \ldots, P(t_{k'})$ tallentamista muistiin tilaksi $\pi^{k'}$. Tilan puoliryhmäominaisuus merkitsee laajemmin vielä sitä, että systeemin tila mielivaltaisella hetkellä $t_k \in \mathcal{T}$ voidaan määrittää, kun tunnetaan

(a) alkutila π^0 ja osaheräte u_{t_0, t_k} tai

(b) tunnetaan tila jonakin hetkenä $t_s < t_k$ ja heräte välillä $[t_s, t_k]$.

1. Todistus sivuutetaan tässä yhteydessä esitystä liikaa paisuttavana, vrt. Salovaara mt. s. 11.

3.3. Systeemiesitysmuodon tulkinnasta

Edellisissä jaksoissa laadittiin diskreetille äärellissolmuiselle Markov-prosessille systeemiesitys yleisen systeemiteorian periaattein. Tässä esityksessä systeemiin liittyvät perussuureet saivat seuraavan (ei välttämättä kuitenkaan ainoan mahdollisen) sisällön. Systeemin herätteeksi valittiin siirtymätodennäköisyysmatriisien jono, ulostuloksi muodostui prosessin solmutodennäköisyysvektorien jono. Systeemin alkutilaksi määriteltiin alkuhetken solmutodennäköisyysvektori, tila yleisenä ajanhetkenä t_k koostui alkutilasta ja ajanhetkeen t_k mennessä systeemiin syötetystä herätteestä.

Mainituista suureista useimmilla on varsin luonnollinen tarkastelujen perustana olevaan reaalimaailman ilmiöön liittyvä tulkinta. Alkutila kuvaa ilmiön luonnetta tarkastelun alkuhetkellä t_0, antaahan solmutodennäköisyysvektori $\pi(t_0)$ jakautuman prosessin miehittämän solmun numeron ilmaisevalle satunnaissuureelle. Systeemin ulostulo kuvaa tämän jakautuman muuttumista ajan mukana, ts. ilmoittaa, mitkä ovat solmutodennäköisyydet kullakin ajanhetkellä.

Hankalimmin tulkittavissa on systeemin heräte, joka muodostuu jonosta siirtymätodennäköisyysmatriiseja. Heräte käsitetään yleensä puhtaasti eksogeeniseksi suureeksi, joko ohjattavissa tai ei-ohjattavissa olevaksi. Siirtymätodennäköisyysmatriisi taas usein mielletään kuvattavan ilmiön sisäiseksi, sen rakennetta ilmentäväksi suureeksi. Ei ole kuitenkaan vaikeata löytää esimerkkejä ilmiöistä, joiden käyttäytymisen kuvaamisessa siirtymätodennäköisyysmatriiseilla on eksogeeninen luonne. Olkoon tarkasteltavana ilmiönä esimerkiksi kuluttajan käyttäytyminen, kun hän suorittaa toistuvasti valintoja tietystä tuoteryhmästä. Markov-prosessin solmut ovat tällöin eri tuotteita, siirtymät näihin tuotteisiin kohdistuvia valintoja (joko samaan tuotteeseen peräkkäin kohdistuvia tai tuotteesta toiseen siirtymistä merkitseviä). Siirtymätodennäköisyysmatriisit ovat nyt selvästi eksogeenisia luonteeltaan, vaikuttavathan niiden muodostumiseen useat kuluttajan ympäristöstä peräisin olevat tekijät kuten tuotteiden hinta ja laatu, muoti, vuodenaika, kuluttajavalistus, mainonta jne. Nämä tekijät edustavat sekä ohjaus- että ei-ohjaussuureita.

Markov-systeemin rakenteeksi jää näin, sen jälkeen kun siirtymätodennäköisyysmatriisit on kiinnitetty herätteeksi, yhtälön (18) muotoon puettu laki. Tämä laki määrää kiinteällä alkutilalla ja herätteellä täysin yksikäsitteisesti ulostulon. Yhtälön (18) mukainen rakenne johtuu suoraan tarkasteltavan prosessin Markov-ominaisuudesta. Yhtälö (18) edellyttää ja on toisaalta seuraus olettamuksesta (5); prosessin seuraavaksi miehittämän solmun valintaan vaikuttavat vain valinnan ajankohta ja prosessin valintahetkellä miehittämä solmu.