PRIRODNO MATEMATIČKI FAKULTET, BEOGRAD, 1978.

Janković D. Slobodan

NEERGODIČKI IZVORI U TEORIJI INFORMACIJA

disertacija

PREDGOVOR

Tema ovoga rada je ispitivanje osobina stacionarnih izvora informacija, bez predpostavke o ergodičnosti. Naime, u matematičkoj Teoriji informacija, nastaloj 1948. godine radom C. E. Shannona, jedan od osnovnih pojmova jeste izvor informacija. Od samog početka, pažnja je uglavnom bila upravljena na stacionarne i, posebno, na ergodičke izvore informacija. Razlog tome su dosta pogodne osobine stacionarnih ergodičkih izvora, naročito poznata osobina $\mathcal{E}$, koja je potrebna kod dokazivanja osnovnih Shannonovih teorema o kodiranju i prenosu informacija kroz kanal sa šumovima.

Međutim, dok je stacionarnost dosta prirodna osobina izvora informacija, ergodičnost to ne mora biti, barem ne u istom stepenu.

Na taj način, ovaj rad predstavlja uopštenje dosadašnjih rezultata. Specijalno, pokazalo se da i neergodički stacionarni izvori imaju dosta pogodne osobine pa čak i, delimično modifikovanu, osobinu $\mathcal{E}$.

Osnovni rezultat u čitavom radu je teorem 1 iz tačke 5, koji nosi naziv teorem o entropijama. Čitav je rad sastavljen tako da odgovara tom cilju. Najpre da se taj teorem dokaže, a zatim da se izvuku njegove posledice.

Sam se rad sastoji od sedam tačaka koje se mogu podeliti u dva dela. Prve tri tačke su uvodne i sadrže poznate definicije i osobine izvora informacija /bez dokaza/. Preostale četiri tačke čine glavni deo rada i sadrže originalna razmatranja o stacionarnim izvorima, bez uslova ergodičnosti. Prema tome, razmatrani su isključivo stacionarni izvori informacija, a pridev "neergodički" se odnosi na nove osobine, koje ergodički izvori nemaju.

Pri tome su svi originalni rezultati - teoreme i leme - numerisani rednim brojevima u svakoj tački posebno. Suprotno tome, rezultati koji se navode nisu numerisani.

Sažeto, sadržaj pojedinih tačaka je sledeći:

Tačka 1 sadrži definiciju izvora informacija i to kao diskretnog izvora sa konačnim alfabetom. Uveden je pojam stacionarnosti a zatim su definisani Bernoullijev i izvor Markova.

Tačka 2 obuhvata ergodičnost. Sem osnovnih osobina, data je i geometrijska interpretacija ergodičnosti radi boljeg sagledavanja problema ergodičnosti odnosno neergodičnosti stacionarnih izvora.

U tački 3 je uveden pojam entropije, kako za prostu slučajnu promenljivu, tako i za stacionarni izvor informacija. Definisana je osobina $\mathcal{E}$ i naveden je poznati McMillanov teorem.

Te tri tačke ne sadrže originalne rezultate. Shodno tome, dokazi su davani samo tamo gde je to bilo neophodno. Nasuprot tome, sledeće četiri tačke sadrže dosada nerazmatranu materiju i svi dati dokazi su originalni.

Tačka 4, prva od tih, posvećena je dekompoziciji stacionarnih neergodičkih izvora. Pokazano je se svaka stacionarna mera verovatnoće izvora može prikazati kao linearna kombinacija ergodičkih merâ / u smislu geometrijske interpretacije /. Konstruisani su i primeri izvora čije dekompozicije imaju željeni broj članova - od jednog do kontinum mnogo.

Tačka 5 se bavi entropijama članova dekompozicije stacionarnih neergodičkih izvora, što je ključni deo u radu. Posle definisanja neophodnih pojmova dokazane su dve neophodne leme, a na osnovu njih i glavni rezultat rada: Teorem o entropijama, koji određuje entropije članova dekompozicije.

Posledice koje proizilaze iz teorema o entropijama nalaze se u tački 6. Tu je dokazano da se uslovna entropija stacionarnog izvora ne smanjuje ako je uslov neka σ-algebra invarijantnih skupova osnosno neka invarijantna slučajna promenljiva. Posle toga su dobiveni potrebni i dovoljni uslovi da bi stacionarni izvor informacija imao osobinu $\mathcal{E}$ i pokazano je da konstruisaniprimeri iz tačke 4 mogu dati neergodičke stacionarne izvore koji imaju osobinu $\mathcal{E}$. Na kraju ove tačke, pokazano je da svaki stacionarni izvor ima jednu osobinu koja predstavlja izvesno uopštenje osobine $\mathcal{E}$. Iz toga, specijalno, sledi da se važenje poznatih Shannonovih teorema može proširiti sa ergodičkih na proizvoljne stacionarne izvore.

Konačno, tačka 7 sadrži primenu predhodnih rezultata na izvore Markova. Posebno, pokazano je da je tada broj članova dekompozicije konačan i to ograničen brojem slova u alfabetu.

Korišćeni matematički aparat obuhvata sem uobičajnih metoda iz teorije verovatnoće i teorije informacija i aparaturu monotonih klasa, uslovnih verovatnoća i uslovnih očekivanja. Pri tome je, s obzirom na potpunost, uglavnom navođena knjiga Neveua, koja je navedena u literaturi.

Spisak literature, sem neposredno korišćenih dela, sadrži i manji broj naslova koji su od interesa za ovu oblast. U radu je literatura navođena, isključivo, imenom autora.

Na kraju, osećam obavezu da se zahvalim profesoru Zoranu Ivkoviću na pomoći i podsticanju koje mi je pružio prilikom moga rada.

Janković Slobodan,

novembra 1978.

SADRŽAJ

1. IZVORI INFORMACIJA

a. Definicija

U Teoriji informacija, jedan od najznačajnijih pojmova jeste izvor informacija. Najčešće se razmatraju diskretni izvori sa konačnim alfabetom, što će i ovde biti slučaj. Matematički, on se može opisati diskretnim slučajnim procesom

/1/ $\ldots X_{-2}, X_{-1}, X_0, X_1, X_2, \ldots$

koji uzima vrednosti u konačnom alfabetu α. Tu je α skup dopustivih signala, to jest slova, koji je po pretpostavci konačan, a X_t je slovo koje je izvor emitovao u trenutku $t \in I$. Pri tome se

$$I = \{\ldots -2, -1, 0, 1, 2, \ldots\}$$

radi jednostavnosti, uzima za skup svih trenutaka u kojima izvor emituje signale.

Slučajni proces /1/ se najpogodnije zadaje na odgovarajućem faznom prostoru verovatnoće $(\Omega, \mathcal{a}, \mu)$. Tu je $\Omega = \alpha^I$ skup svih trajektorije, to jest dvostruko beskonačnih nizova slova:

$$x = (\ldots x_{-1}, x_0, x_1, \ldots)$$

gde je $x_t \in \alpha$ za svako $t \in I$. $\mathcal{a}$ je σ-algebra na Ω generirana algebrom cilindara $\mathcal{C}$. Pri tome je svaki cilindar, to jest član iz $\mathcal{C}$, oblika

$$C_E = \{x;\ (x_{t_1}, \dots x_{t_n}) \in E\}$$

gde je $E \subset \alpha^n$ baza toga cilindra. Kako je svaki takav cilindar konačna disjunktna unija takozvanih tankih cilindara

$$u_{t,n} = \{x;\ x_t = a_0, \dots x_{t+n-1} = a_{n-1}\}$$

gde su $a_0, \dots a_{n-1}$ slova iz α, to se $\mathcal{C}$, pa i $\mathcal{A}$, može generirati i tim tankim cilindrima. Kako svaki tanak cilindar određuje konačan niz uzastopno emitovanih slova, prirodno je koristiti termin reč, što će nadalje biti slučaj.

Konačno, ako je $\mu(C_E)$ definisano kao verovatnoća događaja

$$(X_{t_1}, \dots X_{t_n}) \in E$$

tada je μ konačno-aditivna mera verovatnoće na $(\Omega, \mathcal{C})$. Može se pokazati da je μ i σ-aditivno /Neveu, str. 82/, što znači da se jednoznačno proširuje na meru verovatnoće na $(\Omega, \mathcal{A})$.

Pri tome, prema definiciji mere verovatnoće μ, koordinatni proces

/2/ $\quad \dots x_{-2}, x_{-1}, x_0, x_1, \dots$

ima iste funkcije raspodele kao i zadani proces /1/, te se izvor informacija može opisati i koordinantnim procesom /2/ odnosno i navedenim prostorom verovatnoće $(\Omega, \mathcal{A}, \mu)$. Obratno, ako je μ neka mera verovatnoće na $(\Omega, \mathcal{A}$, tada koordinatni proces /2/ opisuje jedan izvor informacija.

Na osnovu toga, može se dati i formalna definicija.

Definicija. Izvor informacija je par $X = [\alpha, \mu]$, gde je α konačan alfabet, a μ mera verovatnoće na

$(\Omega, \mathcal{A})$ gde je $\Omega = \alpha^I$ a $\mathcal{A}$ je σ-algebra generirana cilindrima.

Naravno, dovoljno je meru verovatnoće μ zadati na rečima $u_{t,n}$. Ako je

$$\mu_{t,n}[a_0, \dots a_{n-1}] = \mu\{x;\ x_t = a_0, \dots x_{t+n-1} = a_{n-1}\}$$

tada, očigledno, važi

$$\mu_{t,n}[a_0, \dots a_{n-1}] \geq 0$$

$$/3/ \qquad \sum_{a_0 \in \alpha} \mu_{t,1}[a_0] = 1$$

$$\sum_{a_n \in \alpha} \mu_{t,n+1}[a_0, \dots a_{n-1}, a_n] = \mu_{t,n}[a_0, \dots a_{n-1}]$$

Obratno, ako funkcije $\mu_{t,n}$ zadovoljavaju te uslove, one čine saglasne funkcije raspodele koordinatnog procesa /2/, te je sa

$$\mu(u_{t,n}) = \mu_{t,n}[a_0, \dots a_{n-1}]$$

definisana mera verovatnoće μ na $(\Omega, \mathcal{A})$.

b. Stacionarnost

Neka je T obostrano jednoznačna transformacija na Ω definisana sa

$$T(\dots x_{-1}, x_0, x_1, \dots) = (\dots x_0, x_1, x_2, \dots)$$

to jest $(Tx)_t = x_{t+1}$. Naziva se pomeraj i može se interpretirati kao translacija za jedinicu vremena unapred. Za ranije definisani cilindar C_E je

$$TC_E = \{x;\ (x_{t_1-1}, \dots x_{t_n-1}) \in E\}$$

Na osnovu toga, stacionarnost procesa /1/ jeste ekvivalentna sa $\mu(TC) = \mu(C)$ za svako $C \in \mathcal{C}$, to jest sa zahtevom da T čuva meru cilindara. No, sa time je ekvivalentno i da T uopšte čuva meru. Naime,

$$\mathcal{M} = \{A \in \mathcal{A};\ TA \in \mathcal{A},\ \mu(TA) = \mu(A)\}$$

čini monotonu klasu koja sadrži sve cilindre, te se može pokazati da je $\mathcal{M} = \mathcal{Q}$ /Neveu, str. 14/.

Na osnovu toga, stacionarnost izvora se može ovako definisati:

<u>Definicija</u>. Izvor informacija $X = [\alpha, \mu]$ je stacionaran ako je $\mu(TA) = \mu(A)$ za svako $A \in \mathcal{Q}$. Kaže se, takođe, da je tada i mera verovatnoće μ stacionarna.

U odnosu na funkcije $\mu_{t,n}$ stacionarnost znači da je $\mu_{t+1,n} = \mu_{t,n}$ što daje potreban i dovoljan uslov za stacionarnost izvora $X = [\alpha, \mu]$:

/4/ $$\sum_{a_{-1} \in \alpha} \mu_{t,n+1} [a_{-1}, a_0, \dots a_{n-1}] = \mu_{t,n} [a_0, \dots a_{n-1}]$$

S obzirom na primene, u Teoriji informacija se najčešće razmatraju stacionarni izvori, što će i ovde biti slučaj. Stoga je, u stacionarnom slučaju, indeks t nepotreban te će biti ispušten. Na primer, u_n i μ_n označavaju $u_{t,n}$ i $\mu_{t,n}$ ako je stacionarnost utvrđena.

c. Primeri izvora informacijâ

Neka je zadan konačan alfabet α i neka je $p = (p_a)_{a \in \alpha}$ vektor pozitivnih verovatnoća. Znači, $p_a > 0$, $\sum_{a \in \alpha} p_a = 1$. Stavljajući da je

$$\mu_{t,n} [a_0, \dots a_{n-1}] = p_{a_0} \cdots p_{a_{n-1}}$$

uslovi /3/ i /4/ biće zadovoljeni, te se dobija stacionarni izvor $X = [\alpha, \mu]$. Takav se izvor naziva Bernoullijev izvor, jer su tada koordinatne promenjljive u /2/ nezavisne.

U drugom slučaju, neka je takođe α konačan alfabet i $p = (p_a)_{a\in\alpha}$ vektor pozitivnih verovatnoća. Dalje, neka je $\Pi = [p_{ab}]_{a,b\in\alpha}$ matrica verovatnoća. Znači, $p_{ab} \geq 0$, $\sum_{b\in\alpha} p_{ab} = 1$. Stavljajući da je

$$\mu_{t,n}[a_0, \dots a_{n-1}] = p_{a_0} p_{a_0 a_1} \cdots p_{a_{n-2} a_{n-1}},$$

uslovi /3/ jesu zadovoljeni, dok je /4/ ekvivalentno sa $p\Pi = p$, to jest da je p stacionarni vektor za matricu Π. Dakle, dobija se izvor koji može biti stacionaran. Naziva se izvor Markova, jer su sada koordinatne promenjljive u /2/ Markovski zavisne.

2. ERGODIČNOST

a. Definicija

Neka je zadan stacionaran izvor $X=[\alpha,\mu]$. Skup $S\in\mathcal{a}$ je invarijantan ako je $\mu(S\,\Delta\, TS)=0$. Kolekcija $\mathcal{Y}$ svih invarijantnih skupova čini σ-podalgebru od $\mathcal{a}$. Prema Birkoffovom ergodičkom teoremu /Neveu, str.21o/, za svaku integrabilnu funkciju f na Ω,

/1/ $$\frac{1}{n}\sum_{\nu=0}^{n-1} f(T^{\nu}x) \xrightarrow[n\to\infty]{} E(f\,|\,\mathcal{Y}) \quad \text{skoro izvesno}$$

gde je $E(f|\mathcal{Y})$ uslovno očekivanje od f u odnosu na σ-algebru $\mathcal{Y}$. Od posebnog je značaja da granična vrednost bude konstanta što zahteva izvesna ograničenja za σ-algebru $\mathcal{Y}$.

<u>Definicija</u>. Stacionaran izvor informacija $X=[\alpha,\mu]$ je ergodičan ako je svaki invarijantan skup mere 0 ili 1, to jest ako je σ-algebra $\mathcal{Y}$ trivijalna. Takođe se kaže da je tada i mera verovatnoće μ ergodična.

Dakle, u ergodičkom slučaju, granična vrednost u /1/ je $E(f)$. Specijalno, ako je I_B indikator skupa $B\in\mathcal{a}$, biće

$$\frac{1}{n}\sum_{\nu=0}^{n-1} I_B(T^{\nu}x) \xrightarrow[n\to\infty]{} \mu(B) \quad \text{skoro izvesno}$$

Množeći sa $I_A(x)$ i integraleći po meri μ, dobija se

/2/ $$\frac{1}{n}\sum_{\nu=0}^{n-1} \mu(A\cap T^{-\nu}B) \xrightarrow[n\to\infty]{} \mu(A)\mu(B)$$

za $A, B\in\mathcal{a}$. Obratno, stavljajući za A i B invarijantan skup S, dobiće se $\mu(S) \longrightarrow \mu(S)\mu(S)$, to jest $\mu(S) = 0$ ili 1. Dakle je /2/ potreban i dovoljan uslov da stacionarni izvor bude ergodičan.

Za ergodičnost je dovoljno i da /2/ važi samo za $A, B \in \mathcal{C}$ pa, stoga, i samo za tanke cilindre - reči. Zaista, iz toga sledi da je za $C \in \mathcal{C}$

$$\mathcal{M}_C = \{A \in \mathcal{Q};\ \frac{1}{n}\sum_{0}^{n-1}\mu(A \cap T^{-\nu}C) \longrightarrow \mu(A)\mu(C)\} \supset \mathcal{C}$$

monotona klasa, te je $\mathcal{M}_C = \mathcal{Q}$. Slično tome, za $A \in \mathcal{Q}$,

$$\mathcal{M}_A = \{B \in \mathcal{Q};\ \frac{1}{n}\sum_{0}^{n-1}\mu(A \cap T^{-\nu}B) \longrightarrow \mu(A)\mu(B)\} \supset \mathcal{C}$$

je takođe monotona klasa, te je $\mathcal{M}_A = \mathcal{Q}$ što znači da /2/ važi i na $\mathcal{Q}$.

Za Bernoullijev izvor je

$$\mu(A \cap T^{-\nu}B) = \mu(A)\mu(T^{-\nu}B) = \mu(A)\mu(B)$$

za tanke cilindre A i B i dovoljno veliko ν, te je Bernoullijev izvor ergodičan.

Slično razmatranje pokazuje da je za stacionarni izvor Markova /2/ ekvivalentno sa

$$\frac{1}{n}\sum_{\nu=0}^{n-1} p_{a\ell}^{(\nu)} \xrightarrow[n\to\infty]{} p_\ell$$

gde je $p_{a\ell}^{(\nu)}$ element matrice Π^ν, to jest verovatnoća prelaza u ν koraka. Detaljnije o ergodičnosti izvora Markova biće reči u tački 7.

B. Geometrijska interpretacija

Ovde će biti govora o geometrijskoj interpretaciji ergodičnosti koja omogućava preglednije predstavljanje pojma ergodičnosti.

Za fiksirani konačni alfabet α, skup svih stacionarnih mera verovatnoća μ na $(\Omega, \mathcal{Q})$ jeste konveksan u smislu da sa μ_1 i μ_2 sadrži i $\lambda\mu_1+(1-\lambda)\mu_2$ za $0\leq \lambda\leq 1$. Ekstremne tačke toga konveksnog skupa su one mere verovatnoće μ za koje nije moguće

/3/ $\mu=\lambda\mu_1+(1-\lambda)\mu_2$ za $0<\lambda<1$ i različite stacionarne μ_1 i μ_2

Očigledno, ako stacionarna mera verovatnoće μ nije ergodična, tada je

$$\mu(\cdot)=\mu(S)\mu(\cdot|S)+\mu(S^c)\mu(\cdot|S^c)$$

za neki netrivijalni invarijantni skup S, pa /3/ važi. Obratno, ako važi /3/, tada se može pokazati da μ nije ergodično /Billingsley, str. 39/.

Dakle, ergodičke mere verovatnoće su vrhovi konveksnog skupa stacionarnih merâ. Stoga se može očekivati da se stacionarne neergodičke mere mogu prikazati kao linearne kombinacije ergodičkih merâ. O tome će biti govora u tački 4, gde će biti i preciznija formulacija.

3. ENTROPIJA

a. Shannonova entropija

Neka su $\xi, \eta, \zeta, \ldots$ slučajne promenljive na $(\Omega, \mathcal{A}, \mu)$ koje uzimaju konačno mnogo vrednosti. Radi sažetosti pisanja, neka $\mu(\xi)$ označava verovatnoće pojedinih vrednosti za ξ. Preciznije, ako su $\xi^{(i)}$ vrednosti za ξ, tada je

$$\mu(\xi) = \mu\{x;\ \xi = \xi^{(i)}\} \qquad \text{za} \quad \xi = \xi^{(i)}$$

i analogno za $\eta, \zeta, \ldots$ Slično tome,

$$\mu(\xi|\eta) = \frac{\mu(\xi,\eta)}{\mu(\eta)}$$

označava verovatnoće pojedinih vrednosti za ξ pod uslovom da je η uzelo neku svoju vrednost.

<u>Definicija</u>. Entropija od ξ je broj

$$H(\xi) = -\sum_{\xi} \mu(\xi) \log \mu(\xi) = E[-\log \mu(\xi)]$$

Uslovna entropija od ξ za datu vrednost od η je slučajna promenljiva

$$H(\xi \| \eta) = -\sum_{\xi} \mu(\xi|\eta) \log \mu(\xi|\eta) = E_{\xi}[-\log \mu(\xi|\eta)]$$

Uslovna entropija od ξ za dato η je broj

$$H(\xi|\eta) = \sum_{\eta} \mu(\eta) H(\xi \| \eta) = E[H(\xi \| \eta)]$$

Pri tome je osnova logaritama 2, a 0log0 se definiše kao 0.

Slobodno govoreći, $H(\xi)$ meri neodređenost odnosno informaciju koju sadrži slučajna promenljiva ξ, $H(\xi \| \eta)$ informaciju koju sadrži ξ ako je poznata realizovana vrednost za η, a $H(\xi | \eta)$ informaciju koju sadrži ξ ako se η već realizovalo.

Osnovne osobine ovih entropija se elementarno dokazuju:

/1/ Sve entropije su nenegativne

/2/ $H(\xi) \leq \log N_\xi$ gde je N_ξ broj vrednosti za ξ

/3/ $H(\xi | \eta) \leq H(\xi)$

/4/ $H(\xi, \eta) = H(\xi) + H(\eta | \xi)$

/5/ $H(\xi | \eta, \zeta) \leq H(\xi | \eta)$

/6/ $H(\xi, \eta | \zeta) = H(\xi | \zeta) + H(\eta | \xi, \zeta)$

b. Entropija izvora informacija

Za stacionarni izvor $X = [\alpha, \mu]$, informacija koju nose reči od n slova je

$$H_n = H(x_t, \dots x_{t+n-1})$$
$$= -\sum \mu_n[a_0, \dots a_{n-1}] \log \mu_n[a_0, \dots a_{n-1}]$$

gde se sumira po svim $a_0, \dots a_{n-1} \in \alpha$. Po osobinama /4/ i /6/ je

$$H_n = H(x_t) + H(x_{t+1} | x_t) + \dots H(x_{t+n-1} | x_t, \dots x_{t+n-2}),$$

Zbog stacionarnosti,

$$H'_\nu = H(x_{t+\nu-1} | x_t, \dots x_{t+\nu-2})$$

je, po osobini /5/, monotono nerastući niz. Kako je on odozdo ograničen nulom, to postoji njegova granica H. No tada i niz aritmetičkih sredina

$$\frac{H_n}{n} = \frac{1}{n}(H'_1 + H'_2 + \dots H'_n)$$

takođe monotono nerastuće konvergira ka istoj granici Zato sledeća definicija ima smisla.

Definicija. Entropija stacionarnog izvora $X=[\alpha,\mu]$ je broj

$$H(X) = \lim_{n\to\infty} \frac{1}{n} H(x_t, \ldots x_{t+n-1})$$

Slobodno govoreći, $H(X)$ je informacija koju stacionarni izvor daje, u proseku, svakog trenutka svoga emitovanja. Jasno, važi

$$0 \le H(X) \le H(x_t) \le \log N$$

gde je N broj slova u alfabetu α.

Za stacionarni izvor Markova je

$$H_n = -\sum_a p_a \log p_a - (n-1) \sum_{a,b} p_a p_{ab} \log p_{ab}$$

te je njegova entropija

$$H(X) = \lim \frac{H_n}{n} = -\sum_{a,b} p_a p_{ab} \log p_{ab}$$

Specijalno, ako je $p_{ab} = p_b$, radi se o Bernoullijevom izvoru, te je njegova entropija

$$H(X) = -\sum_a p_a \log p_a$$

c. Osobina $\mathcal{E}$

Neka je, za stacionarni izvor $X=[\alpha,\mu]$ i fiksirano t,

$$u_n(x) = \{x';\ x'_t = x_t, \ldots x'_{t+n-1} = x_{t+n-1}\}$$

Dakle, to je slučajna reč. Dalje, neka u_n označava vrednosti za $u_n(x)$, to jest konkretne reči. Njih ima $N^n = 2^{n \log N}$ gde je N broj slova u alfabetu α.

Definicija. Stacionarni izvor $X=[\alpha,\mu]$ sa entropijom $H(X)=H$ ima osobinu $\mathcal{E}$ ako za proizvoljne ε, $\delta>0$ postoji n tako da se sve reči u_n mogu podeliti u dve grupe tako da

-za reči iz prve, takozvane visoko verovatne grupe važi

/7/ $$2^{-n(H+\varepsilon)} < \mu(u_n) < 2^{-n(H-\varepsilon)}$$

-ukupna verovatnoća reči iz druge, takozvane nisko verovatne grupe jeste manja od δ.

Osobina $\mathcal{E}$ je od posebne važnosti u Teoriji informacija. Grubo govoreći, ona garantuje izvesnu "ujednačenost", naime da je $\mu(u_n) \approx 2^{-nH}$ za reči iz grupe ukupne verovatnoće bliske 1. Sa druge strane, broj reči u toj, visoko verovatnoj grupi je, prema /7/, manji od $2^{n(H+\varepsilon)}$ te je odnos toga broja prema broju svih reči

$$< 2^{n(H+\varepsilon-\log N)} \xrightarrow[n\to\infty]{} 0$$

za $H < \log N$ i dovoljno malo ε. Dakle, u tom slučaju, visoko verovatna grupa obuhvata relativno mali broj reči.

Definicija osobine $\mathcal{E}$ može se ekvivalentno napisati kao

$$\mu\{x;\ |-\tfrac{1}{n}\log\mu(u_n(x)) - H| < \varepsilon\} \geq 1-\delta$$

za proizvoljne $\varepsilon, \delta > 0$ i dovoljno veliko n, što označava konvergenciju u verovatnoći niza slučajnih promenljivih

$$h_n(x) = -\frac{1}{n}\log\mu(u_n(x))$$

ka entropiji H. Potrebne uslove za osobinu $\mathcal{E}$ daje

McMillanov teorem. Za stacionarni izvor $X=[\alpha,\mu]$ sa entropijom H

$$h_n(x) \xrightarrow[n\to\infty]{} h(x) \qquad \text{skoro izvesno}$$

gde je h invarijantna, to jest $\mathcal{S}$-merljiva funkcija sa očekivanjem $Eh = H$. U ergodičkom slučaju, $h(x) = H$, te X ima osobinu $\mathcal{E}$ /Guiaşu, str. 118/.

Treba primetiti dve stvari: Najpre, konvergencija o kojoj je reč je ne samo u verovatnoći, već i skoro izvesna. Drugo, klasa izvora sa osobinom $\mathcal{E}$ je šira od klase ergodičkih izvora. Primer neergodičkog izvora sa osobinom $\mathcal{E}$ biće dat u tački 6, gde će biti izvedeni potrebni i dovoljni uslovi za osobinu $\mathcal{E}$.

4. DEKOMPOZICIJA NEERGODIČKIH IZVORA

a. Uvod

Neka je $X=[\alpha,\mu]$ stacionarni neergodički izvor. Tada, postoji invarijantni netrivijalni skup S. Dalje, neka je

$$\mu_1(\cdot) = \mu(\cdot|S), \quad p_1 = \mu(S)$$
$$\mu_2(\cdot) = \mu(\cdot|S^c), \quad p_2 = \mu(S^c)$$

Odatle sledi da se stacionarna mera verovatnoće μ može prikazati kao linearna kombinacija tih dveju novih stacionarnih merâ:

$$\mu = p_1\mu_1 + p_2\mu_2$$

gde je $p_1>0$, $p_2>0$, $p_1+p_2=1$.

Ukoliko neka, ili obe, od merâ μ_1, μ_2 nije ergodička, postupak se može nastaviti. Tako će se μ prikazati kao linearna kombinacija od $\leq 2^2$ novih merâ. U sledećem koraku ih ima $\leq 2^3$. Nastavljajući taj postupak, dolazi se do jedne od sledećih situacija:

Prva, najjednostavnija mogućnost je da se posle konačno mnogo koraka došlo do linearne kombinacije od samo ergodičkih merâ:

$$\mu = p_1\mu_1 + \cdots p_k\mu_k$$

gde su $p_1, \cdots p_k$ pozitivni brojevi zbira 1.

Druga mogućnost je da se u svakom, sem konačno mnogo, gore opisanih koraka dobija jedna ergodička i jedna neergodička mera. U tom slučaju, mera μ je linearna kombinacija od beskonačno ali prebrojivo mnogo ergodičkih merâ:

$$\mu = p_1\mu_1 + p_2\mu_2 + \cdots$$

gde su $p_1, p_2, \ldots$ pozitivni brojevi zbira 1.

Treća mogućnost je da se u beskonačno mnogo koraka dobijaju po dve neergodičke mere. Razumno je očekivati da se i tako dolazi do ergodičkih merâ, ali da ih ima više od prebrojivo mnogo. Cilj je takvo tvrđenje dokazati i, šta više, dati primer izvora gde se ovakva mogućnost pojavljuje.

U svim tim slučajevima, predstavljanje stacionarne mere μ kao linearne kombinacije ergodičkih merâ zvaće se, radi jednostavnosti, dekompozicija te mere, a odgovarajuće ergodičke mere će se nazivati članovima te dekompozicije.

b. Tvrđenje

Najpre, neka je $X = [\alpha, \mu]$ zadani stacionarni izvor i

$$\mu_x(\cdot) = \mu(\cdot \mid \mathcal{S}), \quad x \in \Omega$$

uslovna verovatnoća u odnosu na σ-algebru invarijantnih skupova $\mathcal{S}$. Drugim rečima, μ_x je, kao funkcija od $x \in \Omega$, invarijantna - dakle $\mathcal{S}$-merljiva - funkcija koja zadovoljava

/1/ $$\mu(A \cap S) = \int_S \mu_x(A)\,\mu(dx) \qquad \text{za } A \in \mathcal{Q},\ S \in \mathcal{S}$$

Pokazaće se da su baš mere μ_x članovi dekompozicije mere μ. Dokaz će biti izveden u više etapa, formulisanih kao leme.

Najpre, s obzirom da μ_x nije jedinstveno određeno relacijom /1/ - jer su svake dve verzije jednake samo skoro izvesno - to je potrebno izabrati onu verziju koja jeste mera verovatnoće za svako $x \in \Omega$.

<u>Lema 1</u>. μ_x se može tako izabrati da bude mera verovatnoće na $(\Omega, \mathcal{A})$ za svako $x \in \Omega$.

Dokaz. Neka je za $i < j$,

$$C_{a_i \cdots a_j} = \{ x ;\ x_i = a_i, \ldots x_j = a_j \}$$

i neka je $\mathcal{C}_{i \ldots j}$ algebra generirana tim tankim cilindrima. Pri tome je

$$\bigcup_{i<j} \mathcal{C}_{i \ldots j} = \mathcal{C}$$

jer je svaki cilindar iz $\mathcal{C}$ konačna disjunktna unija nekih tankih cilindara. Za proizvoljno $x \in \Omega$ potrebno je odrediti μ_x tako da budu mere verovatnoće na $(\Omega, \mathcal{A})$.

U prvom koraku, neka μ_x bude verovatnoća na $\mathcal{C}_0$. Dakle,

$$\mu_x(C_{a_0}) \geq 0, \quad \sum_{a_0 \in \alpha} \mu_x(C_{a_0}) = 1$$

za svako $x \in \Omega$.

To je moguće. Na primer, jedan od načina da se do te verzije dođe je sledeći. Najpre, za neko $a_0' \in \alpha$ je

$$0 \leq \mu_x(C_{a_0'}) \leq 1$$

skoro izvesno, to jest za x iz nekog skupa μ-mere 1. Za preostale x, $\mu_x(C_{a_0'})$ se može tako izmeniti da gornja nejednakost važi. Pri tome se samo redukuje broj verzija za μ_x. Dalje se za neko drugo $a_0'' \in \alpha$ verzije za μ_x dalje redukuju tako da

$$0 \leq \mu_x(C_{a_0''}) \leq 1 - \mu_x(C_{a_0'})$$

važi za svako $x \in \Omega$. Nastavljajući taj postupak, dolazi se do traženih verzija.

U drugom koraku, od dobivenih verzija za μ_x, biraju se one koje jesu verovatnoće na $\mathcal{C}_{0,1}$. Dakle,

$$\mu_x(C_{a_0 a_1}) \geq 0, \quad \sum_{a_1 \in \alpha} \mu_x(C_{a_0 a_1}) = \mu_x(C_{a_0})$$

za svako $x \in \Omega$.

U trećem koraku, μ_x se određuje tako da bude verovatnoća na $\mathcal{C}_{-1,0,1}$, u četvrtom na $\mathcal{C}_{-1,0,1,2}$. Nastavljajući taj postupak, μ_x će biti mera verovatnoće na $\mathcal{C}$ za svako $x \in \Omega$ te se može jednoznačno proširiti na $\mathcal{a}$.

<u>Lema 2</u>. Mera μ je linearna kombinacija merâ μ_x u smislu da je

/2/ $\quad \mu(\cdot) = \int_\Omega \mu_x(\cdot)\, \mu(dx)$

to jest $\mu = E\mu_x$.

Dokaz. Dovoljno je staviti $S = \Omega$ u relaciji /1/.

<u>Lema 3</u>. μ_x je stacionarno skoro izvesno, to jest

$$\mu\{x;\ \mu_x \text{ je stacionarno}\} = 1$$

Dokaz. Neka je $C \in \mathcal{C}$ i $S \in \mathcal{S}$. Tada je, jer je μ stacionarno,

$$\mu(TC \cap S) = \mu(T(C \cap S)) = \mu(C \cap S)$$

pa je, prema /1/,

$$\int_S \mu_x(TC)\, d\mu = \int_S \mu_x(C)\, d\mu$$

što znači da je

$$\mu_x(TC) = \mu_x(C) \qquad \text{skoro izvesno}$$

Kako je algebra $\mathcal{C}$ prebrojiva, to je

$$\mu\{x;\ \mu_x(TC) \neq \mu_x(C) \quad \text{za neko } C \in \mathcal{C}\}$$
$$= \mu\Big(\bigcup_{C \in \mathcal{C}} \{x;\ \mu_x(TC) \neq \mu_x(C)\}\Big)$$

$$\leq \sum_{C\in\mathcal{C}} \mu\{x;\ \mu_x(TC) \neq \mu_x(C)\}$$

$$= 0$$

Dakle, postoji skup $\Omega_s \in \mathcal{A}$ mere 1, tako da za $x \in \Omega_s$ μ_x čuva meru cilindara. No, kao što je pokazano u tački 1.b, to znači da su te mere μ_x stacionarne.

Lema 4. μ_x je ergodičko skoro izvesno, to jest

$$\mu\{x;\ \mu_x \text{ je ergodičko}\} = 1$$

Dokaz. Neka je $x \in \Omega_s$ i $A, B \in \mathcal{C}$. Tada, prema ergodičkom teoremu,

$$\frac{1}{n}\sum_{\nu=0}^{n-1} I_B(T^{\nu}x') \xrightarrow[n\to\infty]{} E(I_B|\mathcal{G})$$

skoro izvesno. Množeći sa $I_A(x')$ i integraleći po meri μ_x, dobija se

$$\frac{1}{n}\sum_{\nu=0}^{n-1} \mu_x(A\cap T^{-\nu}B) \xrightarrow[n\to\infty]{} E\left[I_A \cdot E(I_B|\mathcal{G})\ |\ \mathcal{G}\right]$$

No, $E(I_B|\mathcal{G})$ je $\mathcal{G}$-merljivo, pa je desna strana jednaka

$$E(I_B|\mathcal{G})\cdot E(I_A|\mathcal{G}) = \mu_x(B)\cdot\mu_x(A)$$

skoro izvesno. Dakle, postoji $\Omega_e \in \mathcal{A}$ tako da je $\Omega_e \subset \Omega_s$ i $\mu(\Omega_e) = 1$ i da za $x \in \Omega_e$ μ_x zadovoljava kriterijum ergodičnosti - tačka 2.a, relacija /2/ - za cilindre A, B. Slično kao i u dokazu predhodne leme, dokazuje se da to μ_x zadovoljava kriterijum ergodičnosti za svaki par cilindara. Prema tome, μ_x je ergodično za $x \in \Omega_e$.

Lema 5. Postoji verzija za μ_x koja je stacionarna i ergodička mera verovatnoće za svako $x \in \Omega$.

Dokaz. Po lemama 1,3 i 4 μ_x je, za $x \in \Omega_e$ stacionarna i ergodička mera verovatnoće. Za $x \in \Omega_e^c$, neka je μ_x uvek jedna ista stacionarna i ergodička mera verovatnoće. Na primer, jedna od μ_x, $x \in \Omega_e$. Na taj način, dobivena je tražena verzija.

Skupljene zajedno, leme 2 i 5 daju potrebno tvrđenje:

<u>Teorem 1</u>. /teorem o dekompoziciji/ Za zadani stacionarni izvor $[\alpha, \mu]$ postoji kolekcija izvora $[\alpha, \mu_x]$, $x \in \Omega$ koji su svi stacionarni i ergodički i pri tome je $\mu = E\mu_x$.

Drugim rečima, neergodička stacionarna mera μ se može prikazati kao linearna kombinacija ergodičkih merâ μ_x. Ukoliko je μ ergodičko, ta dekompozicija je trivijalna jer je tada $\mu_x = \mu$ za svako $x \in \Omega$.

c. Primeri dekompozicije

Da bi dalja razmatranja o dekompoziciji neergodičkih izvora imala dovoljno smisla, potrebno je dati primere izvora sa potrebnim brojem članova dekompozicije. Sledeća konstrukcija će obuhvatiti takve primere.

Najpre, neka je $(U, \mathcal{U}, \mathcal{P})$ neki prostor verovatnoće. Konkretno, neka je U poluotvoreni interval $[0, 1)$ sa realne prave, $\mathcal{U}$ Borelova σ-algebra na njemu, a $\mathcal{P}$ Lebesgueova mera. Dalje, neka je konačan alfabet α fiksiran i neka, za svako $u \in U$, $[\alpha, \mu^u]$ bude stacionaran ergodički izvor i neka je pri tome $\mu^u(A)$, za svako $A \in \mathcal{a}$, $\mathcal{U}$-merljiva funkcija po u koja uzima različite vrednosti za različite vrednosti argumenta u.

To se sve može postići, na primer, ovako. Neka je $p^u = (p_a^u)_{a \in \alpha}$, $u \in U$ kolekcija vektora pozitivnih verovatnoća, tako da je svaka p_a^u $\mathcal{U}$-merljiva funkcija po u. Dalje, mere μ^u neka su definisane sa

$$\mu^u\{x;\, x_t = a_0, \dots x_{t+n-1} = a_{n-1}\} = p_{a_0}^u \cdots p_{a_{n-1}}^u$$

Sada su svi izvori $[\alpha, \mu^u]$, $u \in U$ Bernoullijevog tipa, te su svi stacionarni i ergodički. Pri tome je $\mu^u(C)$ $\mathcal{U}$-merljivo za svako $C \in \mathcal{C}$. Dalje,

$$\mathcal{M}_{\mathcal{U}} = \{A \in \mathcal{A};\ \mu^u(A) \text{ je } \mathcal{U}\text{-merljivo}\}$$

jeste monotona klasa koja sadrži sve cilindre, te je $\mathcal{M}_{\mathcal{U}} = \mathcal{A}$, što znači da je $\mu^u(A)$ $\mathcal{U}$-merljivo za svako $A \in \mathcal{A}$.

Verovatnoće p_a^u se mogu izabrati tako da budu obostrano jednoznačne funkcije po u. Na primer, mogu se uzeti monotone funkcije. Tada su, jasno, mere μ^u različite za različite vrednosti u.

Dakle, svi potrebni uslovi za izvore $[\alpha, \mu^u]$ su postignuti. Sada se može definisati izvor $[\alpha, \mu]$ sa

$$\mu(A) = \int_U \mu^u(A)\, \mathcal{P}(du) \quad \text{za } A \in \mathcal{A}$$

Očigledno je μ stacionarna mera verovatnoće, jer su takve i μ^u, $u \in U$. Definisana je baš kao linearna kombinacija ergodičkih merâ μ^u i pokazaće se da su upravo μ^u članovi njene dekompozicije.

<u>Teorem 2</u>. Za svako $u \in U_0$, gde je U_0 skup Lebesgueove mere 1, postoji $x \in \Omega$ tako da je $\mu_x = \mu^u$, te merâ μ_x ima barem koliko i merâ μ^u, to jest, u ovom slučaju, kontinuum mnogo.

Dokaz. Najpre, po ergodičkom teoremu, za svako $C \in \mathcal{C}$

$$\frac{1}{n}\sum_{\nu=0}^{n-1} I_C(T^\nu x) \xrightarrow[n\to\infty]{} \mu_x(C)$$

μ-skoro izvesno, to jest, za $x \in \Omega_0$ gde je $\mu(\Omega_0) = 1$. No, kako su mere μ^u ergodičke, to za svako $C \in \mathcal{C}$

$$\frac{1}{n}\sum_{\nu=0}^{n-1} I_C(T^\nu x) \xrightarrow[n\to\infty]{} \mu^u(C)$$

μ^u - skoro izvesno, to jest za $x \in \Omega^u$ gde je $\mu^u(\Omega^u)=1$ za svako $u \in U$. Sada, neka je

$$\Omega_0^u = \Omega_0 \cap \Omega^u$$

Kako je

$$1 = \mu(\Omega_0) = \int_U \mu^u(\Omega_0)\, \mathcal{P}(du)$$

to je $\mu^u(\Omega_0)=1$ $\mathcal{P}$-skoro izvesno, to jest za $u \in U_0$ gde je $\mathcal{P}(U_0) = 1$. Sada je očigledno da je i $\mu^u(\Omega_0^u)=1$ za $u \in U_0$.

Konačno, ako $x \in \Omega_0^u$ za $u \in U_0$, tada, za proizvoljno $C \in \mathcal{C}$,

$$\frac{1}{n}\sum_{\nu=0}^{n-1} I_C(T^\nu x) \xrightarrow[n\to\infty]{} \mu_x(C) \qquad \mu^u(C)$$

te je $\mu_x = \mu^u$, što dokazuje tvrđenje.

Ako se pridruživanje $u \longleftrightarrow x$, definisano sa $x \in \Omega_0^u$ za $u \in U_0$, obeleži sa $u = \phi x$, tada je, eksplicitno, $\mu_x = \mu^{\phi x}$. Skup svih tako dobivenih merâ μ_x se ne mora poklopiti sa skupom svih merâ μ_x, ali ipak obuhvata njihov bitan deo. Naime, skup odgovarajućih x-ova obuhvata

$$\Omega_{U_0} = \operatorname*{esssup}_{u \in U_0} \Omega_0^u$$

Kako je $\mu^u(\Omega_0^u)=1$ $\mathcal{P}$-skoro uzvesno, to je i $\mu^u(\Omega_{U_0}) = 1$ $\mathcal{P}$-skoro izvesno što daje $\mu(\Omega_{U_0}) = 1$, na osnovu definicije mere μ. Dakle, sa $\mu_x = \mu^{\phi x}$, $x \in \phi^{-1} U_0$ dobijaju se ako ne sve, onda skoro izvestan deo od svih merâ μ_x.

Sada je lako modifikovati ovaj primer da bi se dobila dekompozicija sa manje članova. Modifikacija se odnosi prvenstveno na σ-algebru $\mathcal{U}$ i to tado da se $\mathcal{U}$ generira deobom intervala $U = [0,1)$ na konačan ili beskonačan broj podintervala U_i. Pri tome mora i svako p_a^u, kao funkcija od u, biti konstanta na svakom od tih podintervala. Tada je i μ^u funkcija indeksa podintervala U_i za koji $u \in U_i$. Tada je

$$\mu(A) = \sum_i \mu^i(A)\mu(U_i)$$

gde je $\mu^i = \mu^u$ za $u \in U_i$.

Prema ranijem, sada za svako i postoji $x \in \Omega_0^u$ za $u \in U_i$, tako da je $\mu_x = \mu^i$. Kao i ranije, drugih merâ μ_x skoro izvesno nema, a mogu se i zameniti jednom od dobivenih. Znači, merâ μ^x ima koliko i merâ μ^i, te sada dekompozicija ima konačno ili beskonačno prebrojivo mnogo članova.

5. ENTROPIJE ČLANOVA DEKOMPOZICIJE

a. Definicije

Sem ranije uvedenih pojmova i oznaka, potrebno je uvesti i izvestan broj novih. Neka, kao i ranije, $H = H(X)$ označava entropiju stacionarnog izvora $X=[\alpha,\mu]$. Dalje, neka H_x označava entropije članova dekompozicije, to jest izvora $[\alpha,\mu_x]$, $x\in\Omega$. Analogno kao i u tački 3.a, H_x se može označiti i sa $H(X\|\mathcal{S})$, pa se može uvesti i

$$H(X|\mathcal{S}) = E\,H(X\|\mathcal{S}) = E\,H_x$$

Opštije, ako je $\mathcal{S}'\subset\mathcal{S}$, neka je

$$\mu'_x(\cdot) = \mu(\cdot|\mathcal{S}'),\quad x\in\Omega$$

Kao i u tački 4.b, može se pokazati da su μ'_x stacionarne mere verovatnoće za svako $x\in\Omega$, odnosno da postoji takva verzija. Naravno, one ne moraju biti ergodične, čak ni skoro izvesno.

Dalje, neka je $H'_x = H(X\|\mathcal{S}')$ entropija izvora $[\alpha,\mu'_x]$ za $x\in\Omega$ i

$$H(X|\mathcal{S}') = E\,H(X\|\mathcal{S}') = E\,H'_x$$

uslovna entropija izvora X za dato $\mathcal{S}'\subset\mathcal{S}$. Specijalno je

$$H(X|\mathcal{S}_0) = H(X)$$

gde je $\mathcal{S}_0 = \{\emptyset,\Omega\}$. Takođe, ako je $\mathcal{S}_k$ konačna algebra sa k invarijantnih netrivijalnih atoma $S_1, \ldots S_k$, biće

$$H(X|\mathcal{S}_k) = \sum_{k=1}^{k} \mu(S_k) H(X|S_k)$$

gde je $H(X|S_k) = H(X\|\mathcal{S}_k)$ za $x \in S_k$ entopija izvora $[\alpha, \mu(\cdot|S_k)]$.

$H(X|\mathcal{S}')$ se može interpretirati kao srednja količina informacije koju izvor X daje u jedinici vremena pod uslovom da se realizovao neki događaj iz $\mathcal{S}'$.

Sada, kako su sve mere μ_x ergodičke, skup

$$\Omega_x = \{ y ; \ \frac{1}{n} \sum_{\nu=0}^{n-1} I_c(T^\nu y) \xrightarrow[n\to\infty]{} \mu_x(C) \text{ za svako } C \in \mathcal{C} \}$$

zadovoljava $\mu_x(\Omega_x) = 1$ te se može nazvati nosiocem mere μ_x. Pri tome su svaka dva nosioca ili jednaka ili disjunktna zavisno od toga da li su odgovarajuće mere jednake ili različite. Naime, ako je $\Omega_x \cap \Omega_{x'}$ neprazno, tada je $\mu_x = \mu_{x'}$ te je i $\Omega_x = \Omega_{x'}$, dok u slučaju da su Ω_x i $\Omega_{x'}$ disjunktni, odgovarajuće mere moraju biti različite. Šta više, prema ergodičkom teoremu, za svako $C \in \mathcal{C}$

$$\frac{1}{n} \sum_{\nu=0}^{n-1} I_c(T^\nu y) \xrightarrow[n\to\infty]{} E(I_c|\mathcal{S}) = \mu_y(C)$$

μ-skoro izvesno, što znači da $x \in \Omega_x$ μ-skoro izvesno.

Konačno, neka je kao i ranije

$$u_n(y) = \{ x ; \ x_t = y_t, \ldots x_{t+n-1} = y_{t+n-1} \}$$

$$h(y) = -\lim_{n\to\infty} \frac{1}{n} \log \mu(u_n(y))$$

Prema McMillanovom teoremu, h je invarijantna funkcija i $Eh = H$. Konkretne vrednosti za slučajnu reč $u_n(y)$ biće označene sa u_n.

Na kraju, nekoliko napomena. Slova Ω sa nekim dodatnim oznakama će biti rezervisana za skupove čija je neka mera 1, a slova S, takođe sa eventualnim dodatnim oznakama, će označavati isključivo invarijantne skupove. Izraz "skoro izvesno" ili skraćeno "s. i." će se osnositi isključivo na meru μ, dok će se za druge mere koristiti odgovarajući prefiks, na primer: μ_x--skoro izvesno.

b. Leme

Osnovni cilj je odrediti entropije H_x članova dekompozicije stacionarnog izvora $X=[\alpha,\mu]$. U tu svrhu, najpre je potrebno dokazati dve leme. Prva od njih je jedno uopštenje osobine /5/ iz tačke 3.a za Shannonovu entropiju.

Lema 1. Ako je $\mathcal{S}'\subset\mathcal{S}''$, tada je $H(X|\mathcal{S}')\geq H(X|\mathcal{S}'')$.

Napomena: Tvrđenje leme će se koristiti samo za $\mathcal{S}', \mathcal{S}''\subset\mathcal{S}$, zbog čega su korišćene takve oznake, mada lema važi i za proizvoljne σ-podalgebre od $\mathcal{Q}$.

Dokaz. Uz očiglednu oznaku $\mu_x''(\cdot)=\mu(\cdot|\mathcal{S}'')$ koristiće se činjenica da je $e^{-1}\log e$ integrabilna majoranta za $-\mu_x''\log\mu_x''$, zatim Jensenova nejednakost u obliku

$$E(-\mu_x''\log\mu_x''|\mathcal{S}')\leq -E(\mu_x''|\mathcal{S}')\log E(\mu_x''|\mathcal{S}')$$

/Neveu, str. 55 i 122/, kao i sledeće dve osobine uslovnih očekivanja:

$$E(E(\cdot|\mathcal{S}'))=E(\cdot)$$
$$E(E(\cdot|\mathcal{S}'')|\mathcal{S}')=E(\cdot|\mathcal{S}') \quad \text{s. i.}$$

Polazeći od definicije uslovne entropije izvora i koristeći navedene činjenice, dobija se

$$
\begin{aligned}
H(X\,|\,\varphi'') &= E\Big[-\lim_n \frac{1}{n}\sum_{u_n}\mu''_x(u_n)\log\mu''_x(u_n)\Big] \\
&= E\Big[\,E\Big(-\lim_n \frac{1}{n}\sum_{u_n}\mu''_x(u_n) \\
&\qquad\qquad \log\mu''_x(u_n)\,\Big|\,\varphi'\Big)\Big] \\
&= E\Big[\lim_n \frac{1}{n}\sum_{u_n} E\Big(-\mu''_x(u_n) \\
&\qquad\qquad \log\mu''_x(u_n)\,\Big|\,\varphi'\Big)\Big] \\
&\le E\Big[-\lim_n \frac{1}{n}\sum_{u_n} E\big(\mu''_x(u_n)\,\|\,\varphi'\big) \\
&\qquad\qquad \log E\big(\mu''_x(u_n)\,\|\,\varphi'\big)\Big] \\
&= E\Big[-\lim_n \frac{1}{n}\sum_{u_n}\mu'_x(u_n)\log\mu'_x(u_n)\Big] \\
&= H(X\,|\,\varphi')
\end{aligned}
$$

čime je lema dokazana.

Specijalno, s obzirom da je $\varphi \supset \varphi_0$, dobija se sledeća nejednakost:

/1/ $\qquad E H_x \le H$

Sledeća će lema biti od osnovne važnosti jer sadrži suštinski odnos koji je potreban za određivanje entropijâ H_x.

<u>Lema 2.</u> μ-skoro izvesno po x i μ_x-skoro izvesno po y važi

$$H_x \ge h(y)$$

Primedba: Samo tvrđenje leme nije jednostavno, kao što neće biti ni njen dokaz. Treba obratiti pažnju da u parovima (x, y) za koje važi gornja nejednakost y na određen način zavisi od x. Formalno, tvrđenje leme se može napisati kao

$$\mu\{x;\ \mu_x\{y;\ H_x \ge h(y)\} = 1\} = 1$$

Dokaz. Neka je

$$S_x = \{y;\ H_x < h(y)\}$$
$$S = \{x;\ \mu_x(S_x) > 0\}$$

Kako je funkcija $h(y)$ merljiva i invarijantna, to je i skup S_x merljiv i invarijantan za svako $x \in \Omega$. Prema tome, definicija skupa S ima smisla. Jasno je da se $\mu_x(S_x) > 0$ može zameniti sa $\mu_x(S_x) = 1$, jer su mere μ_x ergodične. Sada je potrebno dokazati da je i skup S merljiv, a tada je očigledno i da je on invarijantan.

Zaista, skup

$$\{(x,y);\ H_x - \ell(y)\}$$

je merljiv u proizvod prostoru $(\Omega \times \Omega,\ \mathcal{A} \otimes \mathcal{A})$ jer je funkcija $H_x - \ell(y)$ $\mathcal{A}\otimes\mathcal{A}$-merljiva. No, očigledno je sekcija /sečenje/ toga skupa po x-koordinati upravo S_x. Na osnovu toga, sledi da je

$$\mu_x(S_x) = \int_\Omega I_{S_x}(y)\,\mu_x(dy)$$

$\mathcal{A}$-merljiva funkcija po x /Neveu, str. 74/, te je i skup S $\mathcal{A}$-merljiv, što je i bilo potrebno dokazati. Znači, pstoji $\mu(S)$.

Lako je proveriti da je S komplement skupa x-ova navedenih u primedbi posle iskaza leme. Stoga je potrebno dokazati da je $\mu(S) = 0$. Neka je, suprotno tome, $\mu(S) > 0$.

Sada je cilj dokazati da predpostavka $\mu(S) > 0$ dovodi do kontradikcije.

Najpre, kako je svako μ_x ergodično, dok μ to ne mora biti, iz McMillanovog teorema sledi da

$$/2/ \qquad -\frac{1}{n}\log\frac{\mu(u_n(y))}{\mu_x(u_n(y))} \xrightarrow[n\to\infty]{} \ell(y) - H_x$$

μ i μ_x-skoro izvesno po y. Kako $\mu(\Omega^\circ) = 1$ neophodno daje $\mu_x(\Omega^\circ) = 1$ μ-skoro izvesno po x, to relacija /2/ važi μ-skoro izvesno po x i μ_x-skoro izvesno po y, to jest za

$$x\in\Omega^\circ, \quad y\in\Omega^\circ_x$$

pri čemu je

$$\mu(\Omega^\circ)=1, \quad \mu_x(\Omega^\circ_x)=1 \quad /\text{za } x\in\Omega^\circ/$$

Dalje, neka je

$$S^\circ = S\cap\Omega^\circ$$

$$S^\circ_x = S_x\cap\Omega^\circ_x$$

Tada je

$$\mu(S^\circ)=\mu(S)>0$$

$$\mu_x(S^\circ_x)=\mu_x(S_x)=1 \qquad /\text{za } x\in S\cap\Omega^\circ=S^\circ/$$

Sada kako je $S^\circ\subset S, \Omega^\circ$ i $S^\circ_x\subset S_x, \Omega^\circ_x$, to za $x\in S^\circ$ i $y\in S^\circ_x$ važi $H_x<h(y)$ i relacija /2/, što zajedno daje

$$\frac{\mu(u_n(y))}{\mu_x(u_n(y))}\xrightarrow[n\to\infty]{} 0 \quad \text{za } x\in S^\circ,\ y\in S^\circ_x$$

Prema Fatouovoj lemi /Neveu, str. 42/, odatle se dobija

$$\begin{aligned}\int_\Omega \liminf_n \frac{\mu_z(u_n(y))}{\mu_x(u_n(y))}\,\mu(dz) &\le \liminf_n \int_\Omega \frac{\mu_z(u_n(y))}{\mu_x(u_n(y))}\,\mu(dz)\\ &= \liminf_n \frac{\mu(u_n(y))}{\mu_x(u_n(y))}\\ &=0\end{aligned}$$

za $x\in S^\circ$, $y\in S^\circ_x$, te mora biti i

$$/3/ \qquad \liminf_{n\to\infty} \frac{\mu_z(u_n(y))}{\mu_x(u_n(y))}=0$$

μ -skoro izvesno po z. Dakle, za

$$x\in S^\circ, \quad y\in S^\circ_x, \quad z\in\Omega^1$$

gde je $\mu(\Omega^1)=1$, važi relacija /3/.

Nadalje, neka je

$$S^1=S^\circ\cap\Omega^1, \quad \Omega^1_x=\Omega^\circ_x\cap S^1$$

Tada je

$$\mu(S^1)=\mu(S^\circ)>0$$

i, zbog,

$$\mu(S^1) = \int_{S^1} \mu_x(S^1)\mu(dx)$$

važi

$$\mu_x(S^1) = 1 \qquad \text{za} \quad x \in S^2 \subset S^1$$

gde je

$$\mu(S^2) = \mu(S^1) > 0$$

Zato je i

$$\mu_x(\Omega_x^1) = 1 \qquad \text{za} \quad x \in S^2$$

Iz $z \in \Omega_x^1$ sledi $z \in \Omega_x$ što daje, na osnovu definicije nosioca Ω_x, $\mu_z = \mu_x$ za $z \in \Omega_x^2 \subset \Omega_x^1$ gde je $\mu_x(\Omega_x^2) = 1$ za $x \in S^2$. Dakle,

/4/ $$x \in S^2, \quad z \in \Omega_x^2 \implies \mu_z = \mu_x$$

Odnos između pojedinih skupova pregledno je prikazan na sledećoj shemi:

$$\begin{array}{ccccccc}
 & & & & \Omega^1 & & \Omega_x^0 \subset \Omega_x \\
 & & & & \cup & & \cup \\
S & \supset & S^0 & \supset & S^1 & \supset & \Omega_x^1 \\
 & & & & \cup & & \cup \\
 & & & & S^2 & & \Omega_x^2
\end{array}$$

Sada, neka

/5/ $$x \in S^2 \quad \text{i} \quad y \in S_x^0, \quad z \in \Omega_x^2 \quad /\text{za } x \in S^2/$$

Takva trojka sigurno postoji jer je

$$\mu(S^2) > 0 \ \text{i}$$

$$\mu_x(S_x^0) = 1, \quad \mu_x(\Omega_x^2) = 1 \quad /\text{za } x \in S^2/$$

te su ti skupovi neprazni. Prema gornjoj shemi, pretpostavka /5/ implicira

$$x \in S^0, \quad y \in S_x^0, \quad z \in \Omega^1$$

a to, prema ranijem razmatranju, implicira relaciju /3/.

Sa druge strane, predpostavka /5/ implicira i

$$x \in S^2, \quad z \in \Omega_x^2$$

što, prema /4/, daje $\mu_z = \mu_x$, a to je u suprotnosti sa relacijom /3/.

Dakle, iz moguće pretpostavke /5/ slede dve uzajamno kontradiktorne posledice, što obara polaznu hipotezu da je $\mu(S) > 0$. Zato je $\mu(S) = 0$, čime je lema u potpunpsti dokazana.

Treba napomenuti da će se iskazi ovih dveju lema kasnije poboljšati, kao i relacija /1/. To će biti urađeno u tački 6. U ovom obliku, kao što su navedene, one se koriste radi dokazivanja osnovnog teorema.

c. Osnovni teorem

Sada se može pristupiti dokazivanju osnovnog teorema, koji je jedan od glavnih ciljeva ovog rada. Taj teorem određuje entropije H_x izvora $[\alpha, \mu_x]$, to jest članova dekompozicije stacionarnog izvora $X = [\alpha, \mu]$.

Teorem 1. /teorem o entropijama/ Za svaki stacionaran izvor je

$$H_x = h(x) \qquad \text{skoro izvesno.}$$

Dokaz. Neka je

$$\Omega_x^* = \{y;\ H_x \geq h(y)\}$$
$$\Omega^* = \{x;\ \mu_x(\Omega_x^*) = 1\}$$

Prema lemi 2 je

$$\mu(\Omega^*) = 1 \quad \text{i} \quad \mu_x(\Omega_x^*) = 1 \quad za \quad x \in \Omega^*$$

Dalje, neka je

$$h^*(x) = \operatorname{ess\,sup}\{h(y);\ y \in \Omega_x^*\}$$

Tada je $h^*(x) \leq H_x$ za $x \in \Omega^*$. Najpre je potrebno dokazati da tu važi jednakost skoro izvesno, to jest, da je skup

$$S^* = \{x;\ l^*(x) < H_x\}$$

mere nula.

Pre svega je

$$\begin{aligned} E(l\,|\,\varphi) &= \int_\Omega l(y)\mu_x(dy) \\ &= \int_{\Omega_x^*} l(y)\mu_x(dy) \\ &\leq l^*(x) \end{aligned}$$

skoro izvesno, jer je $\mu_x(\Omega_x^*) = 1$ skoro izvesno. Prema McMillanovom teoremu je i

$$H = El = E(E(l\,|\,\varphi)) \leq E l^*$$

Ukoliko bi bilo $\mu(S^*) > 0$, tada je $l^*(x) < H_x$ na skupu S^* pozitivne mere, pa je, uzimajući u obzir i nejednakost /1/,

$$El^* < EH_x \leq H$$

Prema tome je $H \leq El^* < H$, što je nemoguće. Dakle je $\mu(S^*) = 0$, te je

/6/ $\quad l^*(x) = H_x \qquad$ skoro izvesno.

Sada je potrebno dokazati da je $l(y)$ μ_x-skoro izvesno konstanta, naravno, jednaka $l^*(x)$. U tom cilju, neka je

$$S_x^* = \{y;\ l(y) < l^*(x)\}$$

$$S^{**} = \{x;\ \mu_x(S_x^*) > 0\}$$

Kao i u dokazu leme 2, dokazuje se da je S^{**} merljiv skup. Ako je $\mu(S^{**}) = 0$, tada je $\mu_x(S_x^*) = 0$ skoro izvesno, što povlači da je

/7/ $\quad l(y) = l^*(x) \qquad \mu$-skoro izvesno po x,

μ_x-skoro izvesno po y,

što i jeste željeno tvrđenje. Ukoliko je $\mu(S^{**}) > 0$, tada je

$$\begin{aligned} E(h\,|\,\varphi) &= \int_{\Omega} h(y)\,\mu_x(dy) \\ &= \int_{\Omega_x} h(y)\,\mu_x(dy) \\ &< h^*(x) \end{aligned}$$

za $x \in S^{**}$, jer je $h(y) < h^*(x)$ za y iz skupa S_x^* čija je μ_x-mora pozitivna, prema definiciji skupa S^{**}. Na osnovu toga, pretpostavke $\mu(S^{**}) > 0$, jednakosti /6/, nejednakosti /1/ i McMillanovog teorema, sledi da je

$$H = Eh = E(E(h\,|\,\varphi)) < Eh^* = EH_x \le H$$

što je nemoguće. Time je dokazana jednakost /7/.

Konačno, iz $y \in \Omega_x$ sledi $\mu_y = \mu_x$ μ_x-skoro izvesno po y, te je zato za svako $x \in \Omega$

/8/ $H_y = H_x$ μ_x-skoro izvesno po y

Jednakosti /6/, /7/ i /8/ zajedno daju

$h(y) = H_y$ μ-skoro izvesno po x,
μ_x-skoro izvesno po y.

Drugim rečima, za skup

$$S^{***} = \{y;\ h(y) = H_y\}$$

važi

$$\mu_x(S^{***}) = 1 \quad \mu\text{-skoro izvesno po } x.$$

Stoga je i

$$\mu(S^{***}) = \int_{\Omega} \mu_x(S^{***})\,\mu(dx) = 1$$

čime je teorema u potpunosti dokazana.

Napomena : Nije moguće, sledeći postupak iz tačke 4, pooštriti iskaz teorema tako da $H_x = h(x)$ važi za svako $x \in \Omega$. Razlog tome je što funkcija $h(x)$ ne mora biti ograničena / na primer, neka reč u_n može biti mere nula, te je $\ell(x)$ beskonačno za $x \in u_n$/, dok su entropije H_x ograničene / sa $\log N$, gde je N broj slova alfabeta/.

Neposredan značaj ovog teorema je u povezivanju entropijâ članova dekompozicije - izvora $[\alpha, \mu_x]$ - sa funkcijom $\ell(x)$ čija je definicija direktna, za razliku od pomenutih entropijâ. Međutim, više značaja će imati posledice koje proističu iz ovog teorema.

6. POSLEDICE TEOREMA O ENTROPIJAMA

a. Lema

Prva, neposredna posledica teorema o entropijama je poboljšanje relacije /1/ iz tačke 5. Naime, kako je $H_x = \ell(x)$ s. i. i $E\ell = H$ to je

/1/ $$EH_x = H$$

to jest $H(X|\mathcal{Y}) = H(X)$

Da bi se dokazalo opštije tvrđenje, da i u lemi 1 iz prehodne tačke važi znak jednakosti, predhodno je potrebno dokazati jednu lemu, koja i sama predstavlja značajnu posledicu teorema o entropijama.

Lema 1. Ako je za neko $S \in \mathcal{Y}$ $\mu(S) > 0$, tada je

/2/ $$H(X|S) = \frac{1}{\mu(S)} \int_S H_x \, \mu(dx)$$

Dokaz. Ako je $\mu(S) = 1$, iskaz leme važi, jer je tada

$$\int_S H_x \, d\mu = EH_x = H(X) = H(X|S)$$

prema dokazanoj relaciji /1/.

Sada, ukoliko je $0 < \mu(S) < 1$, tada će se najpre dokazati da u relaciji /2/ važi znak $\geq$.

Pre svega, prema definiciji,

$$H(X|S) = -\lim_n \frac{1}{n} \sum_{u_n} \mu(u_n|S) \log \mu(u_n|S)$$

Kako je, prema definiciji mera μ_x,

$$\mu(u_n|S) = \frac{1}{\mu(S)} \int_S \mu_x(u_n) \, \mu(dx)$$

to se, posle kraćeg sređivanja, dobija

$$H(X|S) = -\frac{1}{\mu(S)} \lim_n \sum_{u_n} \left[\int_S \mu_x(u_n)\,\mu(dx)\right] \cdot \log\left[\int_S \mu_x(u_n)\,\mu(dx)\right] + \frac{1}{\mu(S)} \lim_n \frac{1}{n} \log \mu(S)$$

Drugi sabirak je, jasno, jednak nuli, dok se na prvi može primeniti Jensenova nejednakost u obliku

$$-\left[\int_S \mu_x\,\mu(dx)\right] \log\left[\int_S \mu_x\,\mu(dx)\right] \geq -\int_S \mu_x \log \mu_x\,\mu(dx)$$

Na osnovu toga, sledi da je

$$H(X|S) \geq -\frac{1}{\mu(S)} \lim_n \frac{1}{n} \sum_{u_n} \int_S \mu_x(u_n) \log \mu_x(u_n)\,\mu(dx)$$

i kako $-\mu_x \log \mu_x$ ima integrabilnu majorantu $e^{-1}\log e$, to se sa limesom može ući pod integral, pa je

$$H(X|S) \geq -\frac{1}{\mu(S)} \int_S \lim_n \frac{1}{n} \sum_{u_n} \mu_x(u_n) \log \mu_x(u_n)\,\mu(dx)$$

$$= \frac{1}{\mu(S)} \int_S H_x\,\mu(dx)$$

prema definiciji entropija H_x. Time je dobijena tražena nejednakost.

Sada treba dokazati da ni za jedno $S \in \mathcal{S}$, $0 < \mu(S) < 1$ tu ne može stajati efektivno znak $>$. Zaista, ukoliko bi to bio slučaj, za algebru $\mathcal{S}_2$, koja je generisana sa S i S^c, važi

$$\begin{aligned} H(X|\mathcal{S}_2) &= \mu(S) H(X|S) + \mu(S^c) H(X|S^c) \\ &> \int_S H_x\,\mu(dx) + \int_{S^c} H_x\,\mu(dx) \\ &= EH_x \\ &= H(X) \end{aligned}$$

prema /1/. No, prema lemi 1 iz prethodne tačke mora biti

$$H(X|\mathcal{S}_2) \leq H(X|\mathcal{S}_0) = H(X)$$

čime se dobija kontradikcija. Dakle, u /2/ mora stajati znak jednakosti za svako netrivijalno $S \in \mathcal{S}$, čime je lema u potpunosti dokazana.

Interesantno je da se ova lema može napisati u obliku koji uključuje slučaj $\mu(S)=0$. Naime, ako je

$$H(X \cap S) = -\lim_n \frac{1}{n} \sum_{u_n} \mu(u_n \cap S) \log \mu(u_n \cap S)$$

takozvana "nekompletna entropija" izvora $X=[\alpha,\mu]$, tada je lako proveriti da je iskaz leme ekvivalentan sa

$$H(X \cap S) = \int_S H(X \| \mathcal{S})\, d\mu, \quad S \in \mathcal{S}$$

što baca novo svetlo na uslovnu entropiju $H(X\|\mathcal{S}) = H_x$.

b. Teoremi

Sada je moguće, kao posledicu teorema o entropijama, dokazati da u lemi 1 iz tačke 5 mora stajati znak jednakosti.

Teorem 1. Za svaki stacionaran izvor $X=[\alpha,\mu]$ i za svaku σ-podalgebru invarijantnih skupova $\mathcal{S}'$ je

$$H(X|\mathcal{S}') = H(X)$$

Dokaz. Kako je $\mathcal{S}_0 \subset \mathcal{S}' \subset \mathcal{S}$, to je, na osnovu leme 1 iz tačke 5,

$$H(X|\mathcal{S}_0) \geq H(X|\mathcal{S}') \geq H(X|\mathcal{S})$$

No, $H(X|\mathcal{S}_0) = H(X)$, a na osnovu /1/ je i $H(X|\mathcal{S}) = EH_x = H(X)$, čime de dobija tražena jednakost.

Primedba: U opštem slučaju, $H(X|\mathcal{B})$, za $\mathcal{B} \subset \mathcal{A}$, može biti i manje od $H(X)$. Na primer, $H(X|\mathcal{A})=0$, jer je $\mu(u_n|\mathcal{A}) = I_{u_n} = 0$ il i 1. Prema tome, značajna osobima invarijantnih skupova je u tome da se uslovna entropija stacionarnog izvora ne smanjuje ako se realizovao neki skup iz bilo koje σ-podalgebre invarijantnih skupova.

To se može iskazati i u terminima slučajnih promenljivih. Naime, ako se za slučajnu promenljivu ξ definiše

$$H(X|\xi) = H(X \mid \sigma\text{-algebra generisana sa } \xi)$$

tada, za invarijantno ξ, teorem 1 daje

$$H(X|\xi) = H(X)$$

što znači da znanje o invarijantnom ξ, to jest, informacija koju ono nosi, ne umanjuje entropiju stacionarnog izvora, kao što bi se moglo očekivati.

Sledeći teorem, čiji se dokaz oslanja na prethodnu lemu, daje potrebne i dovoljne uslove za osobinu $\mathcal{E}$.

<u>Teorem 2</u>. Za svaki stacionaran izvor X, sledeća tri iskaza su ekvivalentna:

(i) X ima osobinu $\mathcal{E}$

(ii) H_x je konstanta skoro izvesno /naravno, jednaka H /

(iii) za svaki invarijantan skup S pozitivne mere je $H(X|S) = H(X)$

Dokaz. Prema McMillanovom teoremu, (i) znači da je $h(x) = H$ skoro izvesno. Prema teoremu o entropijama, sledi da je ii. ekvivalentno sa $H_x = H$ skoro izvesno. Kako je $EH_x = Eh = H$, sledi da svaka ~~konstantna~~ vrednost za H_x mora skoro izvesno biti jednaka H. Time je dokazano da je (i) $\Leftrightarrow$ (ii).

Iz (ii) sledi, prema prethodnoj lemi,

$$H(X|S) = \frac{1}{\mu(S)} \int_S H \, d\mu$$

za svako invarijantno S pozitivne mere. Dakle, (ii) $\Rightarrow$ (iii).

Obratno, iz (iii) sledi da invarijantni skup

$$S_1 = \{x;\ H_x > H\}$$

ima meru nula. Naime, ako bi bilo $\mu(S) > 0$, tada je

$$H(X|S_1) = \frac{1}{\mu(S_1)} \int_{S_1} H_x \, d\mu > H$$

što je u suprotnosti sa pretpostavkom iii. Na analogan način se dokazuje da i skup

$$S_2 = \{x;\ H_x < H\}$$

ima meru nula, te je $H_x = H$ skoro izvesno. Dakle, (iii) $\Rightarrow$ (ii), čime je teorem u potpunosti dokazan.

Prema izloženom, i neergodički stacionarni izvori mogu imati osobinu $\mathcal{E}$. Držeći se osnovne konstrukcije iz tačke 4.c, lako je dati primer takvog izvora. Naime, dovoljno je vektore verovatnoće $p^u = (p_a^u)_{a\in\alpha}$ birati tako da entropije

$$H^u = -\sum_{a\in\alpha} p_a^u \log p_a^u$$

budu iste. To se može postići ukoliko alfabet α ima bar tri slova. Dobija se neergodički stacionaran izvor takav da entropije članova dekompozicije $H_x = H^{\phi x}$ budu skoro izvesno jednake konstanti, te će dobijeni izvor imati osobinu $\mathcal{E}$.

Naravno, u opštem slučaju, stacionaran izvor ne mora imati osobinu $\mathcal{E}$. Tada, kako god bila izabrana nisko verovatna grupa /male ukupne verovatnoće/, reči u preostaloj, visoko verovatnoj grupi, neće zadovoljavati osnovnu relaciju /7/ iz tačke 3.

Ipak, kao posledica teorema o entropijama, može se pokazati da važi nešto opštija osobina.

<u>Teorem 3</u>. Za svaki stacionaran izvor $X = [\alpha, \mu]$ i proizvoljne $\varepsilon, \delta > 0$, postoji prirodan broj n, tako da se sve reči u_n mogu podeliti u dve grupe, tako da

- ukupna verovatnoća reči u nisko verovatnoj grupi bude manja od δ, i

- za reči u visoko verovatnoj grupi važi

$$2^{-n(H^*+\varepsilon)} < \mu(u_n) < 2^{-n(H_*-\varepsilon)}$$

gde su

$$H_* = \operatorname{ess\,inf} h(x)$$
$$H^* = \operatorname{ess\,sup} h(x)$$

nenegativni konačni brojevi.

Dokaz. Ako je N broj slova u alfabetu α, sledi da je

$$0 \leq h(x) = H_x \leq \log N \quad \text{skoro izvesno}$$

te je

$$0 \leq H_*, H^* \leq \log N$$

Iz $H_* \leq h(x) \leq H^*$ skoro izvesno i

$$-\frac{1}{n}\log \mu(u_n(x)) \xrightarrow[n\to\infty]{} h(x) \quad \text{u verovatnoći}$$

sledi preostali deo tvrđenja, čime je i ovaj teorem dokazan.

Brojevi H_* i H^* mogu se nazvati donja i gornja entropija stacionarnog izvora X. Na osnovu gornjeg teorema, broj reči u visoko verovatnoj grupi će sada biti manji od $2^{n(H^*+\varepsilon)}$. Prema tome, za $H^* < \log N$, taj je broj vrlo mali u odnosu na broj svih reči u_n, pri dovoljno velikom n. Na taj način, i izvori bez osobine $\mathcal{E}$ mogu imati osobine analogne onima koje imaju izvori sa osobinom $\mathcal{E}$.

Tako, na primer, u Shannonovim teoremima o kodiranju i prenosu informacija kroz kanal sa šumovima /Guiașu, str. 159 i 163/, osnovni uslov je da je izvor ergodičan sa entropijom $H < C$, gde je C kapacitet kanala. Pri tome se kodiranje zasniva na rečima iz visoko verovatne grupe i osnovnu ulogu igra njihov broj.

Stoga će, prema gore dokazanom teoremu, Shannon-ove teoreme važiti i za proizvoljne stacionarne izvore, bez obzira na ergodičnost ili osobinu $\mathcal{E}$, pod uslovom da je $H^{*} < C$. Na taj se način znatno proširuje područje njihove primene, što jeste od značaja.

7. SLUČAJ IZVORA MARKOVA

a. Rastavljivost

Sada će se dosadašnja razmatranja primeniti na stacionarne izvore Markova. Zato, neka u ovoj tački $X = [\alpha, \mu]$ označava izvor Markova definisan matricom verovatnoće Π i stacionarnim vektorom pozitivnih verovatnoća p. Prema definiciji /Feller, str. 384/, matrica Π je rastavljiva ako postoji $\alpha_0 \subset \alpha$, $\alpha_0 \neq \emptyset, \alpha$ tako da je

$$\sum_{b \in \alpha_0} p_{ab} = 1 \qquad \text{za svako } a \in \alpha_0$$

Drugim rečima, tada je $\Pi_0 = [p_{ab}]_{a, b \in \alpha_0}$ prava podmatrica verovatnoće od Π. Kako je tada

/1/
$$\begin{aligned} \mu(\alpha_0^I) &= \lim_n \mu\{x;\ x_{-n} \in \alpha_0, \ldots x_n \in \alpha_0\} \\ &= \lim_n \sum_{\alpha_0} p_{a_{-n}} p_{a_{-n}a_{-n+1}} \cdots p_{a_{n-1}a_n} \\ &= \sum_{a \in \alpha} p_a \\ &\neq 0, 1 \end{aligned}$$

jer su svi p_a pozitivni, takav izvor Markova nije ergodičan.

Obratno, ako stacionaran izvor Markova X nije ergodičan, može se pokazati da je matrica Π rastavljiva /Billingsley, str. 31/. Prema tome, neergodičnost stacionarnog izvora Markova jeste ekvivalentna rastavljivosti njegove matrice.

Sada, neka stacionaran izvor Markova X nije ergodičan. Ako je α_0 odgovarajući pravi podalfabet, uslov stacionarnosti $p\Pi = p$ se može napisati u obliku

$$\sum_{a\in\alpha_0} p_a p_{ab} + \sum_{a\notin\alpha_0} p_a p_{ab} = p_b$$

Sumirajući tu jednakost po svim $b\in\alpha_0$ i uzimajući u obzir da je matrica Π rastavljiva, dobija se

$$\sum_{b\in\alpha_0} p_{ab} = 0 \qquad \text{za svako } a\notin\alpha_0$$

te je i

$$\sum_{b\in\alpha_1} p_{ab} = 1 \qquad \text{za svako } a\in\alpha_1$$

gde je $\alpha_1 = \alpha - \alpha_0$.

Drugim rečima, tada je i $\Pi_1 = [p_{ab}]_{a,b\in\alpha_1}$, prava podmatrica verovatnoće od Π.

Iz toga sledi, prema /1/, da invarijantni skupovi

$$\Omega_0 = \alpha_0^I, \quad \Omega_1 = \alpha_1^I$$

zadovoljavaju

/2/
$$\mu(\Omega_0) \neq 0,1, \qquad \mu(\Omega_1) \neq 0,1$$
$$\mu(\Omega_0) + \mu(\Omega_1) = 1$$

Na taj način, dobija se sledeće tvrđenje:

Lema 1. Ako stacionaran izvor Markova X nije ergodičan, njegov se alfabet α može razbiti na disjunktne delove α_0 i α_1, tako da važi /2/.

b. Dekompozicija izvora Markova

Prema predhodnom, za stacionarni neergodički izvor Markovâ X, odgovarajuća matrica verovatnoće Π se može rastaviti na dve podmatrice verovatnoće Π_0 i Π_1. Ako je neka od matricâ Π_0, Π_1 takođe rastavljiva / odnosno odgovarajući izvor neergodičan /, isti se postupak može primeniti i na nju. Nastavljajući tako i dalje, na kraju se dolazi ili do nerastavljivih podmatrica verovatnoće, ili, u krajnjem slučaju, do matrica verovatnoće dimenzija $(1,1)$ koje su, očigledno, opet nerastavljive.

Pri tome se alfabet α razbija na disjunktne delove $\alpha_1, \dots \alpha_k$ tako da su odgovarajuće matrice verovatnoće $\Pi_K = [p_{a b}]_{a, b \in \alpha_K}$, $K = 1, \dots k$ nerastavljive.

Prema gornjoj lemi, za $\Omega_K = \alpha_K^I$, važi

$$p_K = \mu(\Omega_K) \neq 0, 1$$
$$\sum_{K=1}^{k} p_K = \sum_{K=1}^{k} \mu(\Omega_K) = 1$$

Naravno, $k \leq N$, gde je N broj slova u alfabetu α / jer svako α_K sadrži bar po jedno slovo /.

Za nove mere verovatnoće

$$\mu_K(\cdot) = \mu(\cdot \mid \Omega_K)$$

i za $a_0, \dots a_{n-1} \in \alpha_K$ važi

$$\mu_K\{x;\ x_t = a_0, \dots x_{t+n-1} = a_{n-1}\} = p_{a_0}^K \, p_{a_0 a_1} \cdots p_{a_{n-2} a_{n-1}}$$

gde je

$$p_a^K = \frac{p_a}{p_K}, \quad \sum_{a \in \alpha_K} p_a^K = 1$$

Kako je i

$$\sum_{a \in \alpha_K} p_a^K p_{a b} = \frac{1}{p_K} \sum_{a \in \alpha} p_a p_{a b} = p_b^K$$

za $b \in \alpha_K$, to su $[\alpha_K, \mu_K]$ stacionarni izvori Markova sa nerastavljivim matricama Π_K. Znači, mere μ_K su ergodičke i

$$\mu = \sum_{K=1}^{k} p_K \mu_K$$

predstavlja dekompoziciju mere μ na $k \leq N$ ergodičkih merâ μ_K.

Na osnovu rezultata iznetih u tačkama 5 i 6, entropije izvorâ $[\alpha, \mu_K]$ zadovoljavaju

$$H_k = h(x) \quad \text{skoro izvesno za } x \in \Omega_k$$

Dalje, stacionaran izvor Markova $X = [\alpha, \mu]$ ima osobinu $\mathcal{E}$ ako i samo ako su entropije H_k jednake međusobno. Ukoliko X nema osobinu $\mathcal{E}$, tada ipak postoje visoko verovatna i nisko verovatna grupa u smislu iznetom u teoremu 3 iz predhodne tačke. Specijalno, tada je

$$H_* = \min_{1 \le k \le k} H_k$$

$$H^* = \max_{1 \le k \le k} H_k$$

LITERATURA

Ash R. B. Information Theory, Wiley, New York, 1965.

Billingsley P. Ergodic Theory and Information, Wiley, New York, 1965.

Breiman L. The individual ergodic theorem of information theory, Ann. Math. Statist., 28, 809-811, 1958.

Breiman L. Corection to "The individual ergodic theorem of information theory", Ann. Math. Statist., 31, 809-810, 1960.

Dąbrowski A. O teorii informacji, Bibl. mat., WSiP, Warszawa, 1974.

Feinstein A. Foundations of Information Theory, McGraw-Hill, New York, 1958.

Feller W. An Introduction to Probability Theory and Its Application, Vol. I, Chap. XV, Wiley, New York, 1968.

Guiaşu S. Information Theory with Applications, McGraw-Hill, New York, 1977.

Ivković Z. Teorija verovatnoće sa matematičkom statistikom, Gl. VI, Građevinska knjiga, Beograd, 1976.

Jaglom A. M., Jaglom I. M. Verojatnost' i informacija, Nauka, Moskva, 1973.

Janković S. Ergodičnost izvora informacija i osobina E, magistarski rad, Prir. mat. fakultet, Beograd, 1976.

Janković S. Neergodički izvori Markova, biće publikovano u Matematičkom vesniku, Beograd.

Khinchin A. I. Mathematical Foundations of Information Theory, Dover, New York, 1957.

McMillan B. The basic theorems of information theory, Ann. Math. Statist., 24, 196-219, 1953.

Neveu J. Mathematical Foundations of the Calculus of Probability, Holden-Day, San Francisco, 1965.

Ornstein D. Ergodic Theory, Randomness and Dinamical Systems, Yale Univ. Press, New Haven and London, 1974.

Shannon C. E. A mathematical theory of communication, Bell Syst. Techn. Journ., 27, 379-423, 625-656, 1948.

Shannon C. E., Weaver W. The Mathematical Theory of Communication, Univ. of Ill., Illinois, 1949.

Sinai Ya. G. Introduction to Ergodic Theory, Princeton Univ. Press, Princeton, 1976.

Thomasian A. J. Elementary proof of the AEP of information theory, Ann. Math. Statist., 31, 452-456, 1960.

INDEKS POJMOVA