že platí

$$\lim_{N \to \infty} P_e(R, T_N) = 0$$

Dôkaz:

Nech $R > H(X)$ a $\varepsilon = R - H(X) > 0$. Pre každé $N$ zostrojíme $(R, T_N)$ kód, pre ktorý $T_N = T_N(\varepsilon)$. Potom podľa (5) platí uvedená veta.

Veta (obrátená veta o kódovaní zdroja bez pamäte)

Pre každý stacionárny zdroj bez pamäti s abecedou $X$ existuje pri $R < H(X)$ také číslo $\delta > 0$, že pre každé $N$ a každý kód $(R, T_N)$ je

$$P_e(R, T_N) > \delta$$

Okrem toho pre každú postupnosť kódov $(R, T_N)$, $N = 1, 2, \ldots$ platí

$$\lim_{N \to \infty} P_e(R, T_N) = 1$$

Dôkaz neuvádzame, čitateľ ho môže nájsť v [10] . V tejto literatúre je aj návod dôkazu, že priama a inverzná veta o kódovaní stacionárneho zdroja platí aj všeobecnejšie - pre ergodické zdroje informácie, ak entropiu $H(X)$ nahradíme entropiou $H(X/X^{\infty})$.

Poznámka: Diskrétny stacionárny zdroj voláme ergodickým, ak pre ľubovoľné $k$, ľubovoľnú reálnu funkciu $\varphi(x_1, \ldots, x_k)$, $x_i \in X$, $i = 1, 2, \ldots, k$ a ľubovoľné kladné čísla $\varepsilon$, $\delta$ existuje také $n$, že pre všetky $N > n$ platí

$$p\left( \left| \frac{1}{N} \sum_{i=1}^{N} \varphi(x_1, \ldots, x_k) - E[\varphi(x)] \right| > \varepsilon \right) \leq \delta$$

## 8.5 KANÁL PRENOSU INFORMÁCIE

Nech je daná vstupná abeceda kanála

$$X = \{x_1, x_2, \ldots, x_{Mx}\}$$

výstupná abeceda kanála

$$Y = \{y_1, y_2, \ldots, y_{My}\}$$

vzdialenosť

$$d(\mathbf{x}, \mathbf{y}) \geq 0, \quad \mathbf{x} \in X^N, \quad \mathbf{y} \in Y^N$$

(kde $N$ je prirodzené číslo) a rozdelenie pravdepodobnosti $p(\mathbf{x}, \mathbf{y})$, $\mathbf{x} \in X^N$, $\mathbf{y} \in Y^N$, kde

$$p(\mathbf{x}, \mathbf{y}) = p(\mathbf{x}) \cdot p(\mathbf{y}/\mathbf{x})$$

Trojicu $\mathcal{K} = (X^N, Y^N, p(\mathbf{y}/\mathbf{x}))$ nazývame kanálom prenosu informácie.

Pokiaľ pri pripojení stacionárneho zdroja na vstup kanála aj výstup kanála je stacionárnym zdrojom, budeme kanál nazývať stacionárnym. Pokiaľ pri pripojení Markovovho zdroja, resp. zdroja bez pamäti na vstup kanála, bude aj výstup kanála Markovovým zdrojom, resp. zdrojom bez pamäti, budeme kanál nazývať Markovovým (s konečnou pamäťou), resp. bez pamäti.

Budeme hovoriť, že chyba prenosu je menšia než $\varepsilon$, ak

$$\frac{1}{N} \sum_{X^N x Y^N} p(\mathbf{x}\mathbf{y}) d(\mathbf{x}, \mathbf{y}) < \varepsilon \tag{1}$$

Je zrejmé, že chyba prenosu podľa tejto definície nezáleží len na vlastnostiach kanála, ktoré sú vyjadrené rozdelením $p(\mathbf{y}/\mathbf{x})$, ale aj na použitej vzdialenosti $d(\mathbf{x}, \mathbf{y})$.

Definícia:

$\varepsilon$ - entropiou $H_x(\varepsilon)$ vzhľadom na vzdialenosť $d(x,y)$ voláme minimálne množstvo strednej vzájomnej informácie $\mathcal{J}(X,Y)$

$$H_X(\varepsilon) = \min_{p(y/x) \in P} \mathcal{J}(X,Y)$$

pre všetky rozdelenia

$$P = \left\{ p(y/x);\ \sum_{x} \sum_{Y} p(x)p(y/x)\ d(x,y) \leq \varepsilon \right\}$$

Vlastnosti $\varepsilon$ - entropie

1. $H(\varepsilon) \geq 0$
   Nezápornosť $\varepsilon$ - entropie vyplýva z nezápornosti strednej vzájomnej informácie
2. $\varepsilon$ - entropia je nerastúca funkcia premennej $\varepsilon$, t.j. ak $\varepsilon_1 > \varepsilon_2$, potom $H(\varepsilon_1) \leq H(\varepsilon_2)$

Skutočne, nech $p''(y/x)$ je rozdelenie, pri ktorom sa dosiahne $H(\varepsilon_2)$. Pretože $\varepsilon_1 > \varepsilon_2$, je zrejmé, že toto rozdelenie vyhovuje nerovnosti (1) pri $\varepsilon = \varepsilon_1$. Nech $p'(y/x)$ je rozdelenie, pri ktorom sa dosiahne $H(\varepsilon_1)$. Pretože $p'(y/x)$ aj $p''(y/x)$ splňuje nerovnosť (1) potom z definície $\varepsilon$ - entropie ako minima strednej vzájomnej informácie

$$H(\varepsilon_1) = \sum_{XY} p(x)\ p'(y/x)\ \mathcal{J}(x,y) \leq \sum_{XY} p(x) p''(y/x)\ \mathcal{J}(x,y) = H(\varepsilon_2)$$

3. Nech $y_o \in Y$ je taký prvok výstupnej abecedy kanála, že

$$\mathcal{E}_o = \min_y \sum_X p(x)d(x,y) = \sum_X p(x)d(x,y_o)$$

Inými slovami $y_o$ je univerzálny výstupný symbol kanála. Potom $H(\mathcal{E}) = 0$ pre všetky $\mathcal{E} > \mathcal{E}_o$.

Skutočne, pri nájdení hodnoty $y_o$, hodnota chyby pre všetky $x$ nebude väčšia než $\mathcal{E}_o$. Potom kanál má tvar

$$P(y/x) = \begin{cases} 1, & y = y_o \\ 0, & y \neq y_o \end{cases}$$

odkiaľ vyplýva, že $X, Y$ sú nezávislé, teda $\mathcal{J}(X,Y) = 0$. Z podmienky (2) vyplýva, že uvedené rozdelenie $P(y/x)$ vyhovuje nerovnici (1) pre všetky $\mathcal{E} > \mathcal{E}_o$.

4. $\mathcal{E}$ - entropia je konvexná funkcia, t.j.

$$\sum_{i=1}^{k} \lambda_i H(\mathcal{E}_i) \geq H\left( \sum_{i=1}^{k} \lambda_i \mathcal{E}_i \right)$$

Dôkaz nájde čitateľ v [10].

Tak ako pri zdroji informácie, aj na výstupe kanála môžeme množinu $Y^N$ rozdeliť na podmnožinu $\overline{T}_N$ nejednoznačne dekódovateľných slov a na množinu

$$T_N = \{ \mathbf{u}_1, \ldots, \mathbf{u}_L \}$$

jednoznačne dekódovateľných slov.

Chybou prenosu voláme veličinu

$$d = \mathcal{E} \{ d(\mathbf{x}, u(\mathbf{x})) \} = \sum_{X^N} p(\mathbf{x})d(\mathbf{x}, u(\mathbf{x}))$$

Rýchlosťou prenosu kanála $(X^N, Y^N, p(\mathbf{y}/\mathbf{x}))$ voláme veličinu

$$R = \frac{\log_2 L}{N} \qquad [\text{bit/správa}]$$

ktorá udáva počet dvojkových symbolov potrebných na prenos jednej správy zdroja. Z dvoch kanálov s rovnakou chybou prenosu je lepší ten, ktorý má menšiu rýchlosť prenosu. Pretože z hľadiska prenosu informácie nezáleží na tom aké symboly tvoria vstupnú a výstupnú abecedu, budeme kanál popisovať aj trojicou $(R, T_N, d)$, kde $R$ a $N$ udávajú počet jednoznačne dekódovateľných slov $L = 2^{NR}$, $T_N$ je množinou týchto slov a $d$ je chyba prenosu.

V ďalšej časti ukážeme, akú rýchlosť prenosu má mať kanál, aby preniesol informáciu generovanú zdrojom so zadanou chybou prenosu. Skôr však dokážeme nasledujúcu pomocnú vetu.

Veta:

Pre každý diskrétny zdroj bez pamäti s abecedou X a kanál $(R, T_N, d)$ s chybou

$$d = \mathcal{E}\left\{d(\mathbf{x}, \mathbf{y})\right\} = \frac{1}{N}\sum_{i=1}^{N} \mathcal{E}\left\{d(x_i, y_i) \leq \varepsilon\right\}$$

kde $\mathbf{x} = (x_1, \ldots, x_N)$, $\mathbf{y} = (y_1, \ldots, y_N)$ platí

$$H_N(Y) = \frac{1}{N} H(Y^N) \geq H_x(\varepsilon)$$

Dôkaz:

Z vlastností strednej vzájomnej informácie, keď berieme do úvahy zdroj bez pamäti dostávame

$$\frac{1}{N} H(Y^N) \geq \frac{1}{N} \mathcal{J}(X^N, Y^N) = \frac{1}{N} H(X^N) - \frac{1}{N} H(X^N/Y^N) =$$

$$= \frac{1}{N}\sum_{i=1}^{N} H(X_i) - \frac{1}{N}\sum_{i=1}^{N} H(X_i/X^{i-1}\, Y^N) \geq \frac{1}{N}\sum_{i=1}^{N} H(X_i/Y_i) =$$

$$= \frac{1}{N}\sum_{i=1}^{N} \mathcal{J}(X_i, Y_i)$$

Nech $\varepsilon_i$ je chyba prenosu i-tej zložky slova $\mathbf{x}$.

$$\varepsilon_i = \mathcal{E}\left\{d(x_i, y_i)\right\} = \sum_{x_i}\sum_{y_i} p(x)p_i(y/x)\, d(x,y)$$

kde

$$p(x)p_i(y/x) = p_i(xy)$$

je pravdepodobnosť toho, že i-ty symbol slova $\mathbf{x}$ sa rovná $x$ a i-ty symbol postupnosti $\mathbf{y}$ sa rovná $y$. Potom z (3) a definície $\varepsilon$ - entropie dostávame

$$H_N(Y) = \frac{1}{N} H(Y^N) \geq \frac{1}{N}\sum_{i=1}^{N} H(\varepsilon_i) \geq H(\varepsilon)$$

kde posledná nerovnosť vyplýva z konvexnosti $\varepsilon$ - entropie (tu $\lambda_1 = \dots = \lambda_N = 1/N$).

Veta (obrátená veta o kódovaní zdroja pri prenose kanálom)

Pre každý zdroj bez pamäti s abecedou $X$ a každý kanál $(R, T_N, d)$ s rýchlosťou

$$R < H_X(\varepsilon)$$

platí

$$d > \varepsilon$$

Dôkaz:

Predpokladajme, že existuje taký kód $(R, T_N, d)$, že $R < H_X(\varepsilon)$ a $d \leq \varepsilon$. Ak dekódujeme podľa pravidla $\mathbf{y} = u(\mathbf{x})$, potom

$$H(Y^N) = H(T_N) = \sum_{T_N} p(\mathbf{y}) \log \frac{1}{p(y)} \leq \log 2^{NR} = NR < NH_X(\varepsilon)$$

a to je spor s predchádzajúcou vetou.

Bez dôkazu uvádzame nasledujúcu vetu.

Veta (priama veta o kódovaní zdroja pri prenose kanálom)

Nech je daný diskrétny zdroj bez pamäti s abecedou $X$ a $\varepsilon$ - entropiou $H_X(\varepsilon)$ vzhľadom na metriku $d(x,y) < \infty$ pre každé $(x,y)$. Pre ľubovoľné $\delta > 0$ existuje také $n$, že pre $N > n$ existuje kanál $(R, T_N, d)$ s rýchlosťou prenosu $R = H_X(\varepsilon) + \delta$, tak, že

$$d = \mathcal{E}\left\{ d(\mathbf{x}, u(\mathbf{x})) \right\} \leq \max_{X^N} d(\mathbf{x}, u(\mathbf{x})) \leq \varepsilon + \delta .$$

## 8.6 KANÁL A HAMMINGOVOU METRIKOU

V tejto kapitole zkonkretizujeme výsledky predchádzajúcej kapitoly pre prípad, kedy vzdialenosť $d(\mathbf{x}, \mathbf{y})$ je Hammingovou vzdialenosťou vztiahnutou na jeden symbol

$$d(\mathbf{x}, \mathbf{y}) = \frac{1}{N} \sum_{i=1}^{N} d(x_i, y_i)$$

kde

$$d(x_i, y_i) = \begin{cases} 0, & x_i = y_i \\ 1, & x_i \neq y_i \end{cases} \qquad i = 1, 2, \dots, N$$

a $x$, $y$ sú číslicové signály s $M$ prvkovou abecedou $\{0,1, \ldots, M-1\}$.

Pre $N = 1$ je chyba prenosu

$$\mathcal{E}\{d(x,y)\} = P(x \neq y)$$

čo vyplýva z definície strednej hodnoty náhodnej veličiny $d(x,y)$. Množina rozdelení $p(y/x)$ v definícii $\mathcal{E}$ - entropie potom je

$$P = \{p(y/x);\ p(x \neq y)\}$$

Ak $\mathcal{E} \geq 1 - \max p(x)$, potom existuje univerzálny prvok $y$ a $H(\mathcal{E}) = 0$. Preto budeme ďalej predpokladať, že

$$\mathcal{E} < 1 - \max_{X} p(x)$$

Pripomeňme, že ak

$$y = x \underset{M}{\oplus} z$$

potom rozdelenia pravdepodobnosti splňujú vzťahy

$$p(x,y) = p(x,z) = p(y,z)$$

a odtiaľ

$$p(x/y) = p(z/y),\ p(y(x)) = p(z/x)$$

Pre určenie $\mathcal{E}$ - entropie $H_X(\mathcal{E})$ máme nájsť také $p(y/x) \in P$, pri ktorom $\mathcal{J}(X,Y)$ bude minimálna. Pretože

$$\mathcal{J}(X,Y) = H(x) - H(X/Y)$$

a $H(X)$ nezávisí na $p(y/x)$, prejde úloha na hľadanie takéhoto $p(y/x)$, pri ktorom $H(X/Y)$ dosahuje maximum, t.j.

$$H(\mathcal{E}) = \min_{p(y/x) \in P} \mathcal{J}(X,Y) = H(X) - \max_{p(y/x) \in P} H(X/Y)$$

Z definície $H(X/Y)$ s použitím vzťahov pre pravdepodobnosti

$$H(X/Y) = \sum_{X} \sum_{Y} p(y)p(x/y) \log \frac{1}{p(x/y)} =$$

$$= \sum_{Z} \sum_{Y} p(y)p(z/y) \log \frac{1}{p(z/y)} = H(Z/Y) \leq H(Z)$$

Teda

$$H(\mathcal{E}) \geq H(X) - \max_{P} H(Z) \tag{5}$$

Rozdelenie $p(y/x)$, pri ktorom nastane rovnosť a tá nastane, ak $y, z$ sú nezávislé, t.j. $p(z/y) = p(z)$, určíme riešením sústavy

$$p(x_i)p(y_j/x_i) = p(x_i/y_j) \sum_{k=1}^{M} p(x_k)p(y_j/x_k)$$

$$p(x_i/y_j) = p(z_\ell), \quad z_\ell \underset{M}{\oplus} x_i = y_j \tag{6}$$

Ďalší krok spočíva v nájdení maxima $H(Z)$ pre všetky $p(z)$.
Ohraničenie

$$p(x \neq y) \leq \mathcal{E}$$

dáva pre $p(z)$ podmienku

$$1 - p(z=0) \leq \mathcal{E} \tag{7}$$

pretože $x=y$ práve ak $z=0$. Potom

$$H(\mathcal{E}) = H(X) - \max_{p(z)} H(z)$$

kde maximum hľadáme pre všetky $p(z)$, ktoré splňujú (7). Zafixujeme veličinu $p(z=0) = 1 - \alpha$ a nájdime pravdepodobnosť $p(z)$, $z \neq 0$ maximalizujúce $H(Z)$ pri tejto podmienke. Podľa definície entropie

$$H(Z) = - \sum_{Z} p(z) \log p(z)$$

$$= -\alpha \log \alpha - (1-\alpha) \log(1-\alpha) + \alpha \log \alpha - \sum_{z \neq 0} p(z) \log p(z) =$$

$$= h(\alpha) - \alpha \sum_{z \neq 0} \frac{p(z)}{\alpha} \log \frac{p(z)}{\alpha} \tag{8}$$

kde

$$h(\alpha) = -\alpha \log \alpha - (1-\alpha) \log(1-\alpha)$$

a

$$\sum_{z \neq 0} p(z) = \alpha$$

Ďalej poznamenajme, že

$$\alpha^{-1} \sum_{z \neq 0} p(z) = 1$$

a preto suma so záporným znamienkom vo výraze (8) je entropiou množiny M-1 symbolov.

$$H(Z) \leq h(\alpha) + \log(M-1) \tag{9}$$

pričom rovnosť nastáva pri

$$p(z) = \frac{\alpha}{M-1} \qquad \text{pre všetky } z \neq 0$$

a $\alpha$ je ľubovoľné číslo nie väčšie než $\mathcal{E}$. Maximum pravej časti poslednej nerovnosti získame deriváciou podľa $\alpha$ a položením do rovnosti s nulou:

$$- \log \alpha - \log e + \log(1 - \alpha) + \log e + \log(M-1) = 0$$

odkiaľ

$$\alpha = \frac{M-1}{M}$$

Na druhej strane, pretože $\mathcal{E} < 1 - \max_X p(x)$ a $\max_X p(x) \geq \frac{1}{M}$, bude

$$\mathcal{E} < 1 - \frac{1}{L}$$

takže optimálnu hodnotu nedosahuje. Pretože pravá časť nerovnice (9) monotónne rastie od 0 do (M-1)/M, maximálnu hodnotu nadobúda pri $\alpha = \mathcal{E}$.

Potom

$$\max_{p(z)} H(Z) \leq h(\mathcal{E}) + \log(M-1)$$

a odtiaľ

$$H(\mathcal{E}) \geq H(X) - h(\mathcal{E}) - \mathcal{E} \log(M-1) \tag{10}$$

pričom rovnosť nastáva pri

$$p(z) = \begin{cases} 1 - \mathcal{E} & \text{pre } z = 0 \\ \frac{\mathcal{E}}{M} - 1 & \text{pre } z \neq 0 \end{cases} \tag{11}$$

a pri podmienke y,z sú nezávislé.

Možnosť alebo nemožnosť splnenia týchto podmienok závisí na tom, či je možné riešiť sústavu (6) pri podmienke (11). Ukážeme jej riešenie. Označme

$p_x(\cdot) = p(x)$, $p_y(\cdot) = p(y)$, $p_z(\cdot) = p(z)$. Potom je sústava (6) ekvivalentná sústave

$$P(y=b/x=a) = \frac{p_z(b-a)\ p_y(b)}{p_x(a)} \tag{12}$$

$$p_x(a) = \sum_{b=0}^{M-1} p_y(b)\ p_z(b-a)\ ; \quad a,b \in \{0,1,\ \dots,\ M-1\}$$

kde prvú rovnicu sme dostali pomocou

$$\sum_{k=1}^{M} p(x_k)\ p(y=b/x_k) = p_y(b)$$

a v prvej a druhej rovnici je použitý vzťah

$$p(x=a/y=b) = p_z(b \underset{M}{\ominus} a)$$

Potom, ak pri danom rozdelení pravdepodobností $p_x(a)$ a $p_z(0) = 1-\varepsilon$, $p_z(a) = \varepsilon/M-1$, $a\neq 0$ riešime sústavu

$$p_x(a) = \sum_{b=0}^{M-1} p_y(b)\ p_z(b-a)\ , \quad a = 0,1,\ \dots,\ M-1$$

táto sústava má riešenie aj pri podmienke (11), takže nerovnosť (10) prejde na rovnosť a s použitím (11) ju môžeme zapísať

$$p_x(a) = p_y(a)(1-\varepsilon) + [1-p_y(a)]\ \frac{\varepsilon}{M-1}$$

Jej riešením je

$$p_y(a) = \frac{(M-1)\ p_x(a) - \varepsilon}{(1-\varepsilon)\ L-1}\ ; \quad a = 0,1,\ \dots,\ M-1$$

Aby $\{p_y(a),\ a = 0,1,\ \dots,\ M-1\}$ bolo rozdelením pravdepodobnosti, musí byť

$$p_y(a) \geq 0\ , \quad a = 0,1,\ \dots,\ M-1$$

$$\sum_{a=0}^{M-1} p_y(a) = 1$$

Normujúca podmienka je splnená, pretože $p_x(a)$ je rozdelením pravdepodobnosti. Podmienka nezápornosti je splnená práve, keď

$$\min_X p_x(a) \geq \frac{\varepsilon}{M-1}$$

Z toho, čo sme v tejto kapitole uviedli vyplýva nasledujúca veta.

Veta:

Nech X,Y je vstupná a výstupná abeceda kanála, každá s M prvkami, ktoré nastávajú s pravdepodobnosťami p(x), p(y), vzdialenosť d(x,y) je Hammingovou vzdialenosťou a nezáporné číslo $\varepsilon$ vyhovuje nerovnosti

$$\varepsilon \leq (M-1) \min_X p(x)$$

Potom $\varepsilon$ - entropia vstupných symbolov vzhľadom na metriku d(x,y) je

$$H_X(\varepsilon) = \begin{cases} 0\ ; & \varepsilon \geq 1 - \max\limits_X p(x) \\ H(X) - h(\varepsilon) - \varepsilon \log(M-1); & \varepsilon < 1 - \max\limits_X p(x) \end{cases}$$

kde H(X) je entropia vstupných symbolov a

$$h(\varepsilon) = -\varepsilon \log \varepsilon - (1-\varepsilon)\log(1-\varepsilon)\ .$$

## 8.7 KAPACITA KANÁLA

Priama veta o kódovaní zdroja pri prenose kanálom tvrdila, že pre diskrétny zdroj bez pamäte s $\varepsilon$ - entropiou $H_X(\varepsilon)$ existuje kanál s rýchlosťou $R < H_X(\varepsilon)$ tak, že chyba prenosu nebude väčšia než $\varepsilon$. V tejto kapitole budeme naopak študovať, akou rýchlosťou môže daný kanál prenášať informáciu, pričom kvôli jednoduchosti sa budeme zaoberať len prípadom, kedy $\varepsilon \to 0^+$, t.j. chyba prenosu je ľubovoľne malá.

Definícia:

Kapacitou kanála $(X^N, Y^N, p(\mathbf{y}/\mathbf{x}))$ voláme číslo

$$C = \sup_{p(\mathbf{x}),n} \frac{1}{n} \mathcal{I}(X^n, Y^n)$$

kde supremum hľadáme cez všetky rozdelenia vstupných slov $\mathbf{x} = (x_0, x_1, \ldots, \ldots, x_{n-1})$ a všetky $n \leq N$.

Pre stacionárny kanál bez pamäte platí, že najväčšie množstvo informácií pri danom $n$ prenesieme pri použití zdroja bez pamäte, takže

$$\mathcal{J}(X^n, Y^n) = n\, \mathcal{J}(X,Y)$$

a teda jeho kapacita bude

$$C = \sup_{p(x)} \mathcal{J}(X,Y)$$

kde supremum hľadáme cez všetky rozdelenia vstupných symbolov. Kapacitu diskrétneho kanála $(X, Y, p(y/x))$ určíme teda tak, že nájdeme rozdelenie $p(x)$, pri ktorom stredná hodnota vzájomnej informácie medzi symbolom na vstupe kanála a výstupe kanála dosahuje maximum. Hľadanie takéhoto rozdelenia podstatne zjednodušuje nasledujúca veta.

Veta:

Pre rozdelenie pravdepodobnosti vstupných symbolov, ktoré maximalizuje $\mathcal{J}(X,Y)$, podmienená stredná hodnota vzájomnej informácie $\mathcal{J}(x_k, Y)$ nezávisí na vstupnom symbole $x_k$, ak $p(x_k) \neq 0$ a platí

$$\mathcal{J}(x_k, Y) = \sum_{Y} p(y/x_k) \log \frac{P(y/x_k)}{P(y)} = C$$

kde $C$ je kapacita kanála.

Dôkaz:

Podľa metódy Lagrangeových multiplikátorov bude mať funkcia $\mathcal{J}(X,Y)$ stacionárny bod na oblasti

$$P = \left\{ p(x);\ \sum_{X} p(x) = 1 \right\}$$

práve vtedy, keď

$$\frac{\partial}{\partial P(x_k)} \left[ \mathcal{J}(X,Y) + \lambda \sum_{i=1}^{M_x} P(x_i) \right] = 0 \quad , \quad k = 1, 2, \ldots, M_x$$

kde $\lambda$ je Lagrangeov multiplikátor. Po rozpísaní

$$\mathcal{J}(X,y) = H(Y) - H(Y/X)$$

a derivovaní podľa $P(x_k)$

$$\frac{\partial H(Y)}{\partial P(x_k)} = \sum_{i=1}^{M_y} \frac{\partial H(Y)}{\partial P(y_j)} \cdot \frac{\partial P(y_j)}{\partial P(x_k)} = - \sum_{j=1}^{M_y} \left[ 1 + \ln P(y_j) \right] P(y_j/x_k) =$$

$$= - \left[ 1 + \sum_{j=1}^{M_y} P(y_j/x_k) \, \ell n \, P(y_j) \right]$$

$$\frac{\partial H(Y/X)}{\partial P(x_k)} = - \sum_{j=1}^{M_y} P(y_j/x_k) \, \ell n \, P(y_j/x_k)$$

$$\frac{\partial}{\partial P(x_k)} \sum_{i=1}^{M_x} P(x_i) = 1$$

dostávame

$$\sum_{j=1}^{M_y} P(y_j/x_k) \, \ell n \, \frac{P(y_j/x_k)}{P(y_j)} - 1 + \lambda = 0$$

čo môžeme napísať

$$\mathcal{I}(x_k, Y) = K \, , \quad k = 1, 2, \ldots, M_x$$

kde

$$K = 1 - \lambda$$

Vynásobením k-tej rovnice $p(x_k)$ a sčítaním všetkých rovníc dostávame

$$\sum_{k=1}^{M_x} \mathcal{I}(x_k, Y) \, p(x_k) = K \sum_{k=1}^{M_x} p(x_k)$$

t.j.

$$\mathcal{I}(X, Y) = K$$

Vzhľadom na konkávnosť funkcie (dokážte)

$$\mathcal{I}(X, Y) = f(p(x_1), \ldots, p(x_{M_x}))$$

na oblasti

$$\sum_{k=1}^{M_x} p(x_k) = 1 \, , \quad p(x_k) \neq 0 \, , \quad k = 1, 2, \ldots, M_x$$

je stacionárny bod maximom funkcie $\mathcal{I}(X, Y)$ a teda konštanta $K$ udáva kapacitu kanála

$$C = 1 - \lambda$$

Rovnica

$$\mathcal{J}(x_k, Y) = C$$

teda platí pre všetky $k = 1,2, \ldots, M_x$.

Na výpočet rozdelenia pravdepodobnosti $p(x_k)$, ktoré maximalizuje $\mathcal{J}(X,Y)$, prepíšeme rovnice

$$\sum_Y p(y/x_k) \log \frac{P(y/x_k)}{P(y)} = C \, , \quad k = 1,2, \ldots, M_x$$

do tvaru

$$- \sum_Y p(y/x_k) \log \frac{1}{P(y/x_k)} + \sum_Y p(y/x_k) \log \frac{1}{P(y)} = C$$

resp.

$$\sum_{j=1}^{M_y} P(y_j/x_k) \left[ C + \ln P(y_j) \right] = - H(Y/x_k)$$

pre $k = 1,2, \ldots, M_x$, kde

$$P(y_j) = \sum_{k=1}^{M_x} P(x_k) P(y_j/x_k) \, , \quad j = 1,2, \ldots, M_y$$

za normujúcej podmienky

$$\sum_{j=1}^{M_y} P(y_j) = 1$$

Riešením týchto $M_x + M_y + 1$ rovníc (aj nelineárnych) dostaneme neznáme $p(x_k)$, $k = 1,2, \ldots, M_x$ ; $p(y_j)$, $j = 1,2, \ldots, M_y$ a $C$.

V prípustnej množine riešení $p(x_k)$

$$\left\{ p(x_k) \, , \quad \sum_{k=1}^{M_x} p(x_k) = 1 \right\}$$

predchádzajúcej vety sme nepredpokladali nezápornosť pravdepodobnosti $p(x_k)$, pretože by tieto podmienky vniesli ďalšie Lagrangeove multiplikátory a veta by neplatila. Z toho dôvodu riešením uvedenej sústavy môžu byť aj záporné čísla, ktoré nemôžeme interpretovať ako pravdepodobnosti. V tomto prípade bude funkcia $\mathcal{J}(X,Y)$ nadobúdať maximum pri rešpektovaní nezápornosti riešenia na hranici oblasti danej normujúcou podmienkou, to znamená, že aspoň jedna hodnota $p(x_k)$ sa rovná nule. Nie je známy priamy spôsob určenia týchto pravdepo-

dobností. Preto budeme postupne vynechávať jednotlivé vstupné symboly a riešiť uvedenú sústavu rovníc. Ak v niektorých prípadoch dostaneme nezáporné riešenie, potom riešením úlohy o maximalizácii $J(X,Y)$ je to, u ktorého je hodnota $C$ najväčšia. V prípade, že sme nezáporné riešenie nedostali, budeme vynechávať dvojice symbolov, trojice symbolov atď., a to dovtedy, pokiaľ nedostaneme aspoň jedno nezáporné riešenie. Pre $n$ vypustených vstupných symbolov však vždy urobíme všetky kombinácie spôsobu výberu.
Na záver uvedieme bez dôkazu dve dôležité vety.

Veta(priama Shannonova veta)

Ak entropia $H(X)$ ergodického zdroja je menšia než kapacita $C$ kanála s konečnou pamäťou, potom existuje taký kód, že chyba prenosu je ľubovoľne malá.

Veta (nepriama Shannonova veta)

Ak entropia $H(X)$ ergodického zdroja je väčšia než kapacita $C$ kanála s konečnou pamäťou, potom neexistuje kód, že chyba prenosu je ľubovoľne malá.