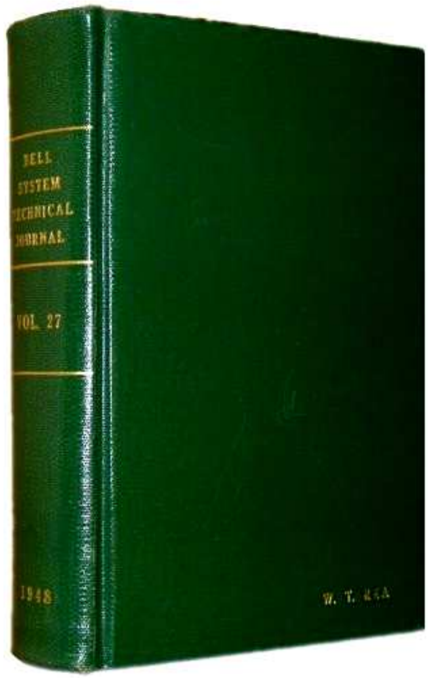
A HÍRKÖZLÉS MATEMATIKAI ELMÉLETE

Claude E. Shannon

The Bell System Technical Journal,[1]
Vol. 27, pp. 379–423, 623–656, July, October, 1948.



Bevezetés

A fejlődés, amely a közelmúltban egyes modulációs módok terén (mint pl. a PCM és a PPM-moduláció, amelyeknél a sávszélesség helyett a jel/zaj-viszony kerül előtérbe) végbement, megnövelte az érdeklődést a hírközlés általánosabb elmélete iránt. Ennek az elméletnek az alapjait Nyquist[2] és Hartley[3] e témában közzétett fontos cikkeikben rakták le. A jelen tanulmányban kiterjesztjük ezt az elméletet és ennek során néhány új tényezőt veszünk figyelembe, mint pl.: a zaj hatását a csatornában és az eredeti üzenet statisztikus szerkezete, valamint az információ végső rendeltetési helyének természete folytán elérhető megtakarítást.

A hírközlés alapproblémája: *egy adott helyen kiválasztott üzenet pontos vagy megközelítő visszaállítása egy másik helyen.*
Az üzeneteknek gyakran *jelentésük* van; ez azt jelenti, hogy valamely – bizonyos fizikai vagy fogalmi dolgokkal jellemzett – rendszerre vonatkoznak, illetőleg aszerint korreláltak. **A hírközlés elméletének e szemantikai vonatkozásai közömbösek a műszaki probléma szempontjából. A lényegi kérdés az, hogy a tényleges üzenet, *egy sor lehetséges közül kiválasztott* egyetlen üzenet.** A rendszert úgy kell megtervezni, hogy valamennyi választási lehetőséggel együtt tudjon működni, ne csak azzal az eggyel amely végül kiválasztásra kerül, hiszen azt, hogy ez melyik lesz, a tervezés szakaszában még nem tudjuk.
Ha a sorozatot alkotó üzenetek száma véges, úgy ez a szám, vagy ennek bármely monoton függvénye úgy tekinthető, mint annak az információnak a mértéke, amelyet e készletből egyforma valószínűséggel kiválasztott bármelyik üzenet hordoz. Ahogyan azt Hartley kimutatta, legtermészetesebb a logaritmusfüggvény választása, és amennyiben folytonos üzenetsorral rendelkezünk, ez a meghatározás jelentős általánosításra szorul, mégis, lényegileg minden esetben valamilyen logaritmikus mértéket fogunk alkalmazni.

A logaritmikus mérték több okból kényelmesebb:

1. A gyakorlatban hasznosabb, mivel a műszaki szempontból lényeges paraméterek, mint: idő, sávszélesség, jelfogók száma stb. a lehetőségek számának logaritmusával lineárisan változnak. Pl.: egy csoport jelfogóhoz egy újabbat hozzátéve megkétszerezzük a jelfogók lehetséges állapotainak számát, ui. ekkor a szám 2-es

---

[1] A jelen cikkbeli ábrák nem az eredeti cikkből kerültek kimásolásra, de tartalmilag azok interpretációinak felelnek meg. (szerk. T.Dénes T.)

[2] Nyquist, H.: *Certain Factors Affecting Telegraph Spedd* („A táviratozási sebességet befolyásoló néhány tényezőről"), Bell System Technical Journal. 1924. április, 324. old.: „Certain Topics in Telegraph Transmission Theory" („A távíró-átviteli elmélet néhány kérdése").A.I.E.E. Trans., 47.k., 1928. április, 617. old.

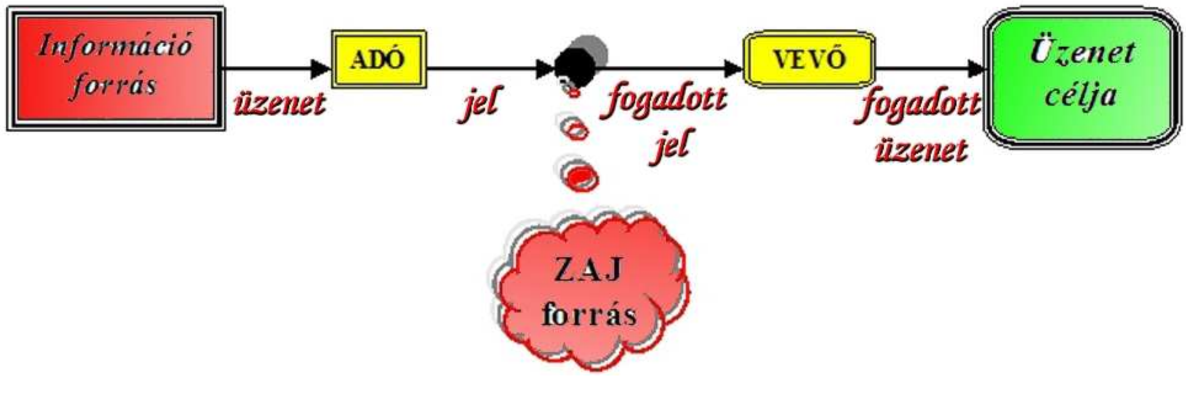
[3] Hartley, R.V.L.: *Transmission of Information* („Az információ átvitele"). Bell System Technical Journal, 1928. július, 535. old.

alapú logaritmusát 1-gyel megnöveljük. Megkétszerezve az időt, durván négyzetesen változik a lehetséges üzenetek száma, vagy megduplázódik a logaritmus stb.
2. Közelebb áll az intuitív érzésünkhöz, mint a megfelelő mértékhez. Ez szorosan összefügg az 1. ponttal, mivel a dolgokat intuitív módon úgy mérjük meg, hogy általánosan használt mértékegységek segítségével lineáris összehasonlítást végzünk. Az ember pl. úgy érzi, hogy két lyukkártyának kétszer annyi információ tárolására kellene alkalmasnak lennie mint egyetlennek, és két, teljesen azonos csatorna kétszeres információátviteli kapacitást képvisel az egyikhez képest.
3. Matematikai szempontból alkalmasabb, ui. a határérték számítási műveletek logaritmikusan könnyen elvégezhetők, míg a lehetőségek száma miatt egyébként nehézkezesen kezelhető formát öltenének.

A logaritmusalapot a mérendő információ egységének megfelelően választjuk meg. Ha 2-es alapú logaritmust választunk, az így kapott egységet bináris digitnek, vagy röviden bit-nek nevezhetjük, mely elnevezést J. W. Tukey javasolta. Egy, két stabil állapottal rendelkező eszköz, mint e jelfogó, vagy a multivibrátor-áramkör, egy bitnyi információ tárolására alkalmas. *N* számú ilyen eszköz *N* bitet tud elraktározni, minthogy a lehetséges állapotok száma 2^N, és $\log_2 2^N = N$. 10-es alapú logaritmus használata esetén az egységet decimális digitnek nevezhetjük. Mivel

$$log_2\, M = \frac{\log_{10} M}{\log_{10} 2} = 3{,}32\, log_{10}\, M,$$

Egy decimális digit mintegy $3\frac{1}{3}$ bitnek felel meg. Egy asztali számológép számkerekének tíz stabil állapota van, ennélfogva egy decimális digit tárolókapacitással rendelkezik. A műszaki-tudományos munkában, ahol differenciál- és integrálszámítást alkalmazunk, gyakran hasznos az *e*-alapú logaritmusról a *b*-alapúra történő átszámítás során csupán a $\log_b a$ tényezővel való szorzást kell elvégezni.



1. ábra *Általános hírközlési rendszer vázlata*

Hírközlési rendszerben az 1. ábrán vázlatosan bemutatott típusú rendszert fogjuk érteni. Ez lényegében öt részből áll:
1. Az információforrás üzenet, vagy üzenetek sorát állítja elő, amelyeket a vételi végállomáshoz kívánunk eljuttatni. Az üzenet többfajta lehet, úgymint:
   (a) betűk sorozata, mint pl. a vezetékes- vagy rádiótávírónál,
   (b) az időnek egyszerű *f(t)* függvénye, mint a rádió vagy a távbeszélő esetében,
   (c) az időnek és más változóknak a függvénye, mint a fekete-fehér televíziós átvitelnél, ahol az üzenet felfogható mint *f(x,y,t)*, azaz két tér- és egy

időkoordináta függvényeként. Itt az üzenet a képernyő egy adott *(x,y)* pontján, egy adott *t* időben észlelhető fényerősséget jelenti,

(d) két vagy több időfüggvény, mondjuk *f(t), g(t), h(t)* – ez az eset áll fenn a „háromdimenziós" hangátvitelnél, vagy ha a rendszerrel több, egyedi csatornát kívánunk nyalábolni (multiplexelni);

(e) különféle változók különféle függvénye; - a színes televíziónál az üzenet három függvényből – *f(x,y,t), g(x,y,t), h(x,y,t)* áll, amelyek egy háromdimenziós kontinuumban vannak definiálva; (ezeket a függvényeket úgy is felfoghatjuk, mint egy, ebben a tartományban definiált vektortér elemeit), - hasonlóképpen, több fekete-fehér televíziós forrás produkálhat olyan üzenetet, amely három változó több függvényéből áll;

(f) a fenti esetek különféle kombinációi is előfordulnak, pl. a televíziós átvitelnél a képhez tartozó hangcsatorna.

2. Az adó úgy módosítja az üzenetet, hogy abból a csatornán történő átvitelre alkalmas jelet állít elő. A távbeszélő technikában ez a művelet egyszerűen abból áll, hogy a hangnyomást azzal arányos elektromos árammá alakítjuk át. A távírótechnikában olyan kódolási műveletet végzünk, amelynek során az üzenetnek megfelelő pontok, vonások és szünetek sorozatát állítjuk elő az átviteli csatorna számára. Az impulzus-kódmodulált (PCM) multiplex-rendszereknél a különböző beszéd-függvényekből mintát veszünk, ezeket komprimáljuk, kvantáljuk, kódoljuk és végül – a jel összeállítása érdekében – megfelelően multiplexeljük. A vokóder-rendszerek, a tv és a frekvenciamoduláció, a jelnek üzenetből történő összeállítása összetett műveletére mutatnak más és más példát.

3. Csatornán pusztán az átvivő közeget értjük, amelyen keresztül a jel az adótól a vevőhöz eljut. Ez lehet kéterű vezeték, koaxiális kábel, rádiófrekvenciás sáv, fénysugár stb. A jelet az átvitel során, vagy valamelyik végponton zaj zavarhatja meg. Ezt az 1. ábrán sematikusan egy zajforrással jelöltük, amely az adó jelét vétel közben befolyásolja.

4. A vevő rendesen az adás műveletének fordítottját végzi, azzal, hogy a jelből visszaállítja az üzenetet.

5. Az üzenet rendeltetésén azt a személyt (vagy dolgot) értjük, akinek (amelynek) a részére az üzenet szól.

Meg fogjuk vizsgálni a hírközlési rendszereknél felmerülő bizonyos általános problémákat. Ennek érdekében először az egyes előforduló elemeket, mint matematikai mennyiségeket kell előállítani, fizikai megfelelőjüktől kellően elvonatkoztatva (idealizálva). Első közelítésben a kommunikációs rendszereket három fő csoportba sorolhatjuk: *diszkrét, folytonos és vegyes rendszereket különböztetünk meg*. Az elsőn olyan rendszert értünk, amelynél mind az üzenet, mind a jel diszkrét szimbólumok sorozatából áll. Tipikus példa erre a távírótechnika, ahol az üzenet betűk, a jel pedig pontok, vonások és szünetek sorozatából áll. Folytonos rendszer az, amelyben az üzenetet és a jelet egyaránt mint folytonos függvényeket kezeljük, pl. a rádió- és a tv műsorátvitelnél. Vegyes típusú rendszereknél mind a diszkrét, mind a folytonos változók előfordulnak (pl. PCM beszédátvitelnél).
Először a diszkrét rendszerek esetét vizsgáljuk meg, amelyeket nemcsak a hírközléselméletben alkalmaznak, hanem a számítógépek, a távbeszélőközpontok tervezésében és egyéb területeken is. Ezen túlmenően a diszkrét eset képezi az alapját a

folytonos és a vegyes eseteknek, amelyeket jelen tanulmány második részében fogunk tárgyalni.

I. Diszkrét zajmentes rendszerek

1. A diszkrét zajmentes csatorna

A Morse-típusú távíró és a géptávíró két egyszerű példa a diszkrét információátviteli csatornára. Általánosságban diszkrét csatornán olyan rendszert fogunk érteni, amelyben egy $S_1,\ldots,S_n$ elemi szimbólumokból álló véges készletből kiválasztunk egy jelsorozatot, hogy azt az egyik ponttól a másikig átvigyük. Valamennyi S_i szimbólum t_i időtartama azonos legyen, (pl. Morse-táviratozásnál a pontok és a vonások hossza különbözik). Nem feltétel az, hogy a rendszer az S_i szimbólumkészletből alkotható valamennyi lehetséges sorozat átvitelére alkalmas legyen, megengedhetünk csupán bizonyos sorozatokat, amelyek a csatorna számára lehetséges jeleket fogják képviselni. Így a Morse-táviratozásnál a következő szimbólumokat tételezzük fel: (1) a pontot, amely az áramkör egységnyi idejű zárásából, majd ugyanolyan nyitásából áll, (2) a vonást, amely 3 időegységnyi zárást és egy egységnyi nyitást tartalmaz, (3) a betűköz, amely – mondjuk – 3 egységnyi ideig nyitott vonalat jelent, és (4) a szóközt, amely 6 időegységnyi ideig nyitott vonalnak felel meg. A megengedhető sorozatokra nézve azt a megszorítást tehetjük, hogy közök nem követhetik egymást (mivel, ha két betűköz kerül egymás mellé, az egyenértékű egy szóközzel). Felmerül a kérdés most már, hogyan mérhetjük meg egy ilyen információátviteli csatorna kapacitását?

A távíró esetében, ahol valamennyi szimbólum azonos időtartamú, és a 32 szimbólum bármilyen sorozatának előfordulása megengedett, a válasz egyszerű. Minden szimbólum 5 bit információt hordoz. Ha a rendszer másodpercenként n számú szimbólumot bocsát ki, nyilvánvalóan azt mondjuk, hogy a csatorna kapacitása *5n bit/s*. Ez nem jelenti azt, hogy a távírócsatorna mindig ilyen sebességgel adja az információt, ez a maximálisan elérhető érték, és az, hogy ezt valóban elérjük-e, vagy sem, - amint később látni fogjuk – a csatornát tápláló információforrástól függ.

Abban az általánosabb esetben, ha a szimbólumok időtartama nem egyforma és a megengedett sorozatok számát korlátozzuk, a következő definíciót adjuk meg:

Definíció: Egy diszkrét csatorna C kapacitása

$$C = \lim_{T\to\infty} \frac{\log N(T)}{T}$$

ahol $N(T)$ a megengedett T időtartamú jelek száma.

Könnyen belátható, hogy a távíró esetében a képlettel a fentebb megadott eredményt kapjuk. Meg lehet mutatni, hogy a szóban forgó határérték az érdeklődésre számot tartó esetek legnagyobb részében véges számnak adódik. Tételezzük fel, hogy az $S_1,\ldots,S_n$ szimbólumkészlet bármiféle sorozatának előfordulása megengedett és hogy az azt alkotó szimbólumok időtartama rendre $t_1,\ldots,t_n$. Mi lesz ezek után a csatorna kapacitása?

Ha $N(t)$-vel jelöljük a t időtartamú sorozatok számát, akkor

$$N(t) = N(t-t_1) + N(t-t_2) + \ldots + N(t-t_n)$$

A teljes szám egyenlő az $S_1, S_2, \ldots, S_n$-nel végződő sorozatok számának összegével, és ezek rendre: $N(t-t_1)$, $N(t-t_2)$, …, $N(t-t_n)$. A véges különbségeknél jól ismert eredmény szerint $N(t)$ nagy t értékeknél aszimptotikusan tart egy X_0^t értékhez, ahol X_0 pedig a legnagyobb valós megoldása a következő karakterisztikus egyenletnek:

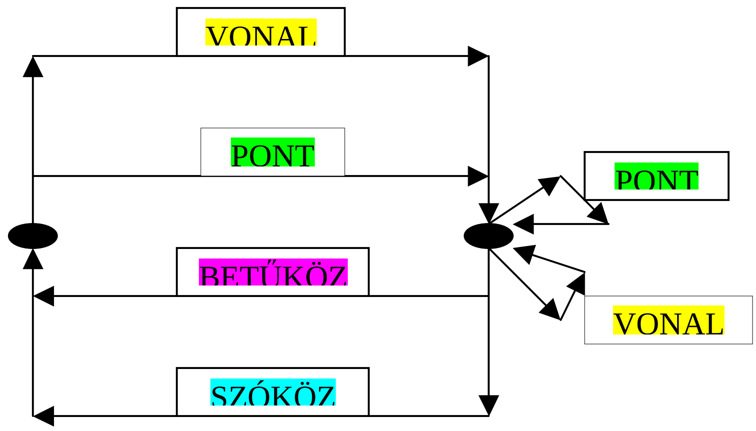
$$X^{-t_1} + X^{-t_2} + \ldots + X^{-t_n} = 1 \qquad \text{továbbá} \qquad C = \log X_0$$

Abban az esetben, ha a megengedett sorozatokban megszorításokat vezetünk be, gyakran ilyen típusú differenciálegyenletekhez juthatunk és C értékét a karakterisztikus egyenletből határozhatjuk meg. A fent említett távírócsatorna esetén

$$N(t) = N(t-2) + N(t-4) + N(t-5) + N(t-7) + N(t-8) + N(t-10)$$

amint ez a szimbólumsorozatok összeállításakor látható, az előforduló utolsó, vagy utolsó előtti szimbólumnak megfelelően. Tehát $C = -\log \mu_0$, ahol μ_0 az $1=\mu^2+\mu^4+\mu^5+\mu^7+\mu^8+\mu^{10}$ egyenlet pozitív gyöke. Az egyenlet megoldva $C=0.539$.

A megengedett sorozatoknál nagyon gyakran a következő megszorítással élünk: $a_1, a_2 \ldots, a_n$ olyan állapotok, amelyek mindegyikéhez csupán egyetlen elem tartozhat az $S_1, \ldots, S_n$ szimbólumkészletből az adás során (különböző készletek a különböző állapotokhoz). Amikor ezek közül egyet adásra kiválasztunk, az állapot új állapotra változik, s ez a változás a régi állapot, valamint az átvitt, szóban forgó szimbólum sajátosságától függ. Ennek egyszerű példája a távíró esete, ahol két állapot létezik, attól függően, hogy szünet volt-e az adott utolsó szimbólum vagy sem. Ha nem szünet volt az utolsó szimbólum, bármilyen jel adásra kerülhet, és ha szünet következik, az állapot megváltozik, egyébként pedig változatlan marad. A viszonyokat lineáris gráfon lehet szemléltetni, ahogyan az a 2. ábrán látható. Itt a csomópontok felelnek meg az állapotoknak, míg a vonalak az egy adott állapotban lehetséges szimbólumokat és az eredményül kapott állapotot jelképezik.



2. ábra *Távírójelekre alkalmazott megszorítások grafikus ábrázolása*

Az 1.sz. Függelékben megmutattuk, hogy amennyiben a megengedett sorozatokra vonatkozó feltételek ebben a formában leírhatók, úgy C létezik, és a következő tételnek megfelelően számítható:

1. TÉTEL:

Legyen $b_{ij}^{(s)}$ az *s*-edik szimbólum időtartama, amelynek előfordulása az *i* állapotban megengedett, és amely a *j* állapothoz vezet. Ekkor a *C* csatornakapacitás *logW*-vel egyenlő, ahol W a következő determináns egyenlet legnagyobb valós gyöke:

$$\left| \sum_s W^{-b_{ij}^{(s)}} - \delta_{ij} \right| = 0 \quad \text{ahol } \delta_{ij} = 1 \text{ ha } i{=}j \text{, egyébként } \delta_{ij} = 0$$

Például a 2. ábra távírója esetén a determináns a következő:

$$\begin{vmatrix} -1 & \left(W^{-2} + W^{-4}\right) \\ \left(W^{-3} + W^{-6}\right) & \left(W^{-2} + W^{-4} - 1\right) \end{vmatrix} = 0$$

amelyet kifejtve a fentebb megadott egyenlőséget kapjuk.

2. A diszkrét információforrás

Láttuk, hogy – nagyon általános feltételek mellett – egy diszkrét csatorna lehetséges jelei számának logaritmusa az idővel lineárisan változik. Az információadás kapacitását úgy definiálhatjuk, hogy megadjuk a felhasznált jel jellemzéséhez szükséges másodpercenkénti bitszámot.

Most tekintsük az információforrást. Hogyan lehet matematikailag leírni egy információforrást és másodpercenként hány bit információt állít elő egy adott forrás?

A legfontosabb dolog a kiindulásnál az, hogy statisztikusan ismerjük a forrást, s így az információ helyes kódolásával csökkentsük a szükséges csatornakapacitást. A táviratozásnál például a közvetítendő üzenetek betűk sorozatából állnak. Ezek a sorozatok azonban nem teljesen véletlenszerűek, hanem mondatokat alkotnak és a nyelv – mondjuk az angol – statisztikus szerkezetével rendelkeznek. Így az E betű gyakrabban fordul elő mint a Q, a TH sorozat gyakoribb az XP-nél, stb. Ennek a szerkezetnek a létezése lehetővé teszi számunkra, hogy az időben (vagy a csatorna kapacitásában) megtakarítást érjünk el azáltal, hogy az üzenetsorozatot megfelelő módon kódoljuk jelsorozattá. Ezt bizonyos mértékig meg is valósították a Morse-távírónál, ahol a leggyakrabban előforduló angol betű az E csatornaszimbólumát egy pont jelzi, míg a kevésbé gyakori betűket, pl. a Q, X, Z-t pontok és vonalak hosszabb sorozata jelképezi. Ezt az elvet bizonyos kereskedelmi kódoknál még tovább fejlesztették és itt gyakori szavakat és kifejezéseket 4-5 betűs kódcsoportokkal jelölnek, ezáltal jelentősen lerövidítve az átlagos átviteli időtartamot. A manapság szabványosított üdvözlő és évfordulói táviratoknál ezt az elvet odáig fejlesztették, hogy egy vagy két mondatot viszonylag rövid számsorban kódolva visznek át.

Elképzelhetünk egy diszkrét információforrást, amely szimbólumonként állítja elő az üzenetet. Ez a forrás egymás után fogja a szimbólumokat kiválasztani, és azok előfordulási valószínűsége általában az előzőleg kiválasztott szimbólumoktól, valamint az adott szimbólumtól függ. Az olyan fizikai rendszer, vagy annak matematikai modellje, amely egy valószínűségi sorozat által szabályozott szimbólumsorozatot produkál, sztochasztikus folyamatnak[4] nevezzük. Ennélfogva egy diszkrét forrást úgy foghatunk fel, hogy azt egy

[4] L. pl. S. Chandrasekhar: „Stochastic Problems in Physics and Astronomy". („Sztochasztikus problémák a fizikában és a csillagászatban"), Reviews of Modern Physics, 15. k., 1943. január, 1. sz., 1. old.

sztochasztikus folyamat reprezentálja és fordítva, bármely sztochasztikus folyamat, amely egy véges készletből kiválasztott szimbólumok diszkrét sorozatát állítja elő, diszkrét információforrásnak tekinthető. Ilyenek a következők:

1. Természetes írott nyelvek, pl. az angol, a német, a kínai.
2. Folyamatos információforrások, amelyeket valamilyen kvantálási folyamat segítségével diszkrétté teszünk. Pl.: PCM adóból származó kvantált beszéd, vagy kvantált televíziós jel.
3. Olyan matematikai esetek, amelyeknél absztrakt módon definiálunk egy szimbólumsorozatot előállító sztochasztikus folyamatot. A következőkben erre az utóbbi fajta forrásra sorolunk fel példákat.

(A) Tételezzünk fel öt betűt, ezek: A, B, C, D, E, amelyeket egyaránt 0.2-es valószínűséggel választunk ki úgy, hogy az egymást követő kiválasztások egymástól függetlenek. Ez egyfajta sorozathoz vezet, amelyre tipikus példa az alábbi:
BDCBCECCCADCBDDAAECEEAABBDAEECACEEBAEECBCEAD
Ezt a sorrendet egy véletlenszám táblázat[5] segítségével szerkesztettük.

(B) Ugyanezen betűket használva legyenek a kiválasztási valószínűségek rendre: 0.4, 0.1, 0.2, 0.1 és az egymást követő választások egymástól függetlenek. Ekkor ennek a forrásnak egy tipikus üzenete az alábbi lesz: AAACDCBDCEAADADACEDAEADCABEDADDCECAAAAAD.

(C) A fentieknél bonyolultabb szerkezetet kapunk, ha az egymásután következő szimbólumokat nem egymástól függetlenül választjuk ki, hanem azok kiválasztási valószínűsége függ a megelőző betűktől. Ennek a legegyszerűbb esete az, amikor a választás csupán a megelőző betűtől függ, a korábbiaktól már nem. A statisztikus szerkezetet ezek után a $p_i(j)$ átmenet-valószínűségi sorozattal írhatjuk le, ahol p annak a valószínűsége, hogy az i betűt a j fogja követni. Az i, j indexek az összes lehetséges szimbólumot felölelik. Egy más, de ezzel egyenértékű módja a szerkezet jellemzésének, ha a $p(i,j)$ „digram" („kétbetűs") valószínűségeket adjuk meg, amely az i,j betűpár együttes előfordulásának relatív gyakoriságát jelzi. A $p(i)$ betűgyakoriságok (az i betű valószínűsége), a $p_i(j)$ átmenet-valószínűség és a $p(i,j)$ digram-valószínűségek a következő képlettel kapcsolódnak egymáshoz:

$$p(i) = \sum_j p(i,j) = \sum_j p(j,i) = \sum_j p(j) p_j(i)$$
$$p(i,j) = p(i) p_i(j)$$
$$\sum_j p_i(j) = \sum_i p(i) = \sum_{i,j} p(i,j) = 1$$

Konkrét példaként tételezzünk fel három betűt, A, B, C-t, a következő valószínűségi táblázatokkal:

[5] Kendall and Smith: „Tables of Random Sampling Numbers". („Véletlen mintavételezési számok táblázata"). Cambridge, 1939.

$p_i(j)$		j		
		A	B	C
i	A	0	4/5	1/5
	B	1/2	1/2	0
	C	1/2	2/5	1/10

i	$p(i)$
A	9/27
B	16/27
C	2/27

$p(i,j)$		j		
		A	B	C
i	A	0	4/15	1/15
	B	8/27	8/27	0
	C	1/27	4/135	1/135

Ebből a forrásból származó tipikus üzenet a következő: ABBABABABABABABBBABBBBBABABABABABBBACACABBABB BBABBABACBBBABA

A komplexitás további fokozásával a „trigram” (hárombetűs csoportok) előfordulási gyakoriságaihoz jutunk. Itt egy betű kiválasztása az előző kettőtől függ, de már független az ezt megelőző üzenetrésztől. A *p(i,j,k)* trigram-előfordulási gyakoriságsorozatot – vagy ami ezzel egyenértékű, egy $p_{ij}(k)$ átmeneti valószínűségsorozatot – kellene létrehozni. Ezen az úton továbbmenve egyre bonyolultabb sztochasztikus folyamatokat kapunk. Az általános *n-edrendű* esetben egy $p(i_1, i_2,\ldots,i_n)$ *n-gram* valószínűségre, ill. a $p_{i_1,i_2,\ldots,i_{n-1}}(i_n)$ átmenet valószínűségre van szükség a statisztikus szerkezet jellemzéséhez.

(D) Sztochasztikus folyamatot definiálhatunk úgy is, hogy „szavak” sorozatából álló szöveget állítunk elő. Tételezzünk fel öt betűt, ezek: A, B, C, D, E és 16 „szót” ebben a nyelvben, melyekhez tartozó valószínűségek:

0.10 A	0.16 BEBE	0.11 CABED	0.04 DEB
0.04 ADEB	0.04 BED	0.05 CEED	0.15 DEED
0.05 ADEE	0.02 BEED	0.08 DAB	0.01 EAB
0.01 BADD	0.05 CA	0.04 DAD	0.05 EE

Tételezzük fel továbbá, hogy az egymást követő „szavakat” egymástól függetlenül választottuk ki, s azok között szünet van. Ekkor egy tipikus üzenet a következő lehet:
DAB EE A BEBE DEED DEB ADEE ADEE EE DEB BEBE BEBE BEBE ADEE BED DEED DEED CEED ADEE A DEED DEED BEBE CABED BEBE BED DAB DEED ADEB

Ha valamennyi szó véges hosszúságú, úgy ez a folyamat az előzőekkel egyenértékű, de a leírás a szószerkezet és a valószínűségek szempontjából egyszerűbb lehet. Itt ismét általánosíthatunk és bevezethetjük a szavak közötti átmenet valószínűséget stb.

Ezek a mesterséges nyelvek hasznosak arra, hogy segítségükkel egyszerű problémákat, példákat állítsunk elő, annak érdekében, hogy illusztráljuk a különféle valószínűségeket. Egy természetes nyelvet is megközelíthetünk egy sor egyszerű mesterséges nyelv segítségével. A nulladrendű megközelítést úgy végezzük, hogy valamennyi betűt egymástól függetlenül, azonos valószínűséggel választjuk ki. Az elsőrendű közelítés során az egymást követő betűket egymástól függetlenül választjuk ki, de minden betűnek a kiválasztási valószínűsége megegyezik a természetes nyelvben képviselt értékkel.[6] Így az angol nyelv elsőrendű

[6] Betű-, kétbetűs és hárombetűs csoportok előfordulási gyakoriságai Fletcher Pratt: *Secret and Urgent* („Titkos és sürgős”) című, Blue Ribbon Books, 1939-ben megjelent könyvében találhatók. Szó- előfordulási gyakoriságokat tartalmaz táblázatos formában G. Dewey: *Relative Frequency of English Speech Sounds* („Az angol nyelv beszédhangjainak relatív előfordulási gyakorisága”) c. műve, Harvard University Press, 1923.

közelítésében az E betűt 0.12-es valószínűséggel választjuk ki (ez az E hang előfordulási gyakorisága a hétköznapi angol nyelvben), míg a W valószínűsége 0.02, azonban a szomszédos betűk között nincs kölcsönhatás, és nem kívánunk TH, ED és ehhez hasonló digramokat képezni. A másodrendű approximáció során digramszerkezeteket vezetünk be. Miután egy betűt már kiválasztottunk, a következőt annak a valószínűségnek megfelelően választjuk ki, amellyel a különböző betűk az elsőt követik. Ez szükségessé teszi a $p_i(j)$ digram-előfordulási gyakoriságokat tartalmazó táblázat alkalmazását. A harmadrendű közelítésben trigramszerkezetek jelennek meg, amelyeknél minden betűt az előző két betűtől függő valószínűséggel választunk ki.

4. Az angol nyelv megközelítésének folyamata

Hogy képet alkossunk, hogyan lehet ezekkel az eljárásokkal egy nyelvet jellemezni, az alábbiakban megadjuk az angol nyelv megközelítésének tipikus lépéseit. Minden esetben 27 szimbólumból álló ABC-t tételezzünk fel, 26 betűt és 1 közt.

1. Nulladrendű megközelítés (a szimbólumok egymástól függetlenek és előfordulási valószínűségük azonos).
   XFOML RXKHRJFFJUJ ZLPWCFWKCYJ FFJEYVKCQSGHYD QPAAMKBZAACIBZLHJQD.

2. Elsőrendű megközelítés (a szimbólumok egymástól függetlenek, de az előfordulási gyakoriságuk az angol nyelvű szövegekének megfelelő).
   OCRO HLI RGWR NMIELWIS EU LL NBNESEBYA THE EEI ALHENHTTPA OOBTTVA NAH BRL.

3. Másodrendű megközelítés (az angol nyelvben előforduló digram szerkezetek).
   ON IE ANTSOUTINYS ARE TINCTORE ST BE S DEAMY ACHIN D ILONASIVE TOCOOWE AT TEASONARE FUSO TIZIN ANDY TOBE SEACE CTISBE.

4. Harmadrendű megközelítés (az angol nyelvben előforduló trigram szerkezetek).
   IN NO IST LAT WHEY CRATICT FROURE BIRS GROCID PONDENOME OF DEMONSTURES OF THE REPTAGIN IS REGOACTIONA OF CRE.

5. Elsőrendű szó-approximáció. Ahelyett, hogy *tetragram,…,n-gram* szerkezetekkel folytatnánk, egyszerűbb és jobb, ha ezen a ponton szóegységekre váltunk át. Itt az egyes szavakat függetlenül választjuk meg, de figyelembe vesszük a megfelelő előfordulási gyakoriságukat.
   REPRESENTING AND SPEEDILY IS AN GOOD APT OR COME CAN DIFFERENT NATURAL HERE HE THE A IN CAME THE TO OT TO EXPERT GRAY COME TO FURNISHES THE LINE MESSAGE HAD BE THESE.
   (magyar fordítás szavanként: KÉPVISELVE ÉS GYORSAN VAN EGY JÓ FOGÉKONY VAGY JÖNNI LEHET KÜLÖNBÖZŐ TERMÉSZETES ITT Ő A –BAN JÖTT AZ –HOZ –NAK –HOZ SZAKÉRTŐ SZÜRKE JÖNNI –HOZ SZÁLLÍT A VONAL ÜZENET VOLT LENNI EZEK.)

6. Másodrendű szó-approximáció. A szó átmeneti valószínűség-értékek valóságosak, azonban további szerkezetet már nem veszünk figyelembe.

THE HEAD AND IN FRONTAL ATTACK ON AN ENGLISH WRITER THAT CHARACTER OF THIS POINT IS THEREFORE ANOTHER METHOD FOR THE LETTERS THAT THE TIME OF WHO EVER TOLD THE PROBLEM FOR AN UNEXPECTED.
(A szavak, ill. a szöveg magyar fordítása: A FEJ ÉS FRONTÁLIS TÁMADÁS EGY ANGOL ÍRÓVAL SZEMBEN HOGY E PONT JELLEGE ENNÉLFOGVA EGY MÁSIK MÓDSZER A BETŰKRE NÉZVE, HOGY AZ IDEJE? AKI SOHA NEM MONDTA A PROBLÉMA EGY VÁRATLAN RÉSZÉRE.
A két nyelv eltérő sajátosságai, ill. a szövegösszefüggés hiányos volta miatt a szövegre a fentitől kissé eltérő értelmezés is megadható, azaz nem létezik teljesen adekvát fordítás.)

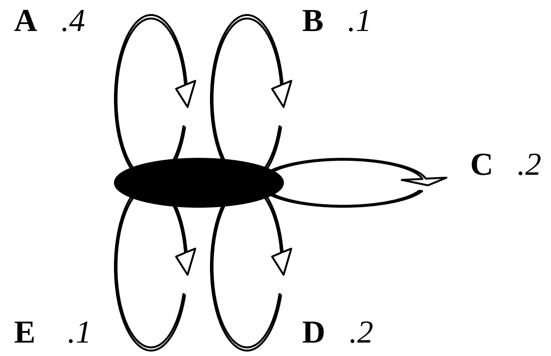
A szokványos angol nyelvű szövegekhez való hasonlóság minden egyes fenti lépéssel jelentősen növekedett. Figyeljük meg, hogy ezek a minták elfogadhatóan jó szerkezetűek, körülbelül kétszer annyi szerkezettel, mint amennyit megalkotásukkor figyelembe vettünk. Így a (3)-nál a statisztikus folyamat kétbetűs sorozatból álló, ésszerű szöveget ad, míg a minta 4-betűs sorrendjeit jó mondatokba lehet illeszteni. A (6)-nál a négy-, vagy ennél több szóból álló sorozatok minden különösebb szokatlanság vagy erőltetettség nélkül mondatokba illeszthetők. A kilenc szóból álló, előbb vizsgált sorozat: *„támadás egy angol íróval szemben, hogy e pont jellege”*… egyáltalán nem elképzelhetetlen. Úgy tűnik ezek után, hogy egy elegendően komplex sztochasztikus folyamat jól jellemez egy diszkrét forrást.
Az első két mintát egy véletlen számokkal foglalkozó könyv segítségével szerkesztettük, (pl. a 2.-nál) egy betű-előfordulási gyakoriságot tartalmazó táblázatot használva. Ezt a módszert a 3., 4. és 5. esetekre is kiterjeszthettük volna, mivel digram, trigram és szó-előfordulási gyakoriságokat tartalmazó táblázatok léteznek, azonban mi egy egyszerűbb, de ezzel egyenértékű módszert használtunk. Pl. a 3. megalkotásához az ember kinyit egy könyvet valahol, és ezen az oldalon találomra kiválaszt egy betűt, s azt feljegyzi. Ezután a könyvet egy másik oldalon nyitja ki, és addig olvas, míg ezt a betűt meg nem találja. Ekkor az ezt követő betűt jegyzi fel. Más oldalra lapozva ezt a második betűt keresi meg, és az azt követőt jegyzi fel, és így tovább. Hasonló módszert alkalmaztunk 4., 5. és 6. esetében. Érdekes lenne további approximációkat szerkeszteni, azonban az ezzel járó munka a következő lépésnél hihetetlenül megnövekszik.

4. Egy Markov-folyamat grafikus ábrázolása

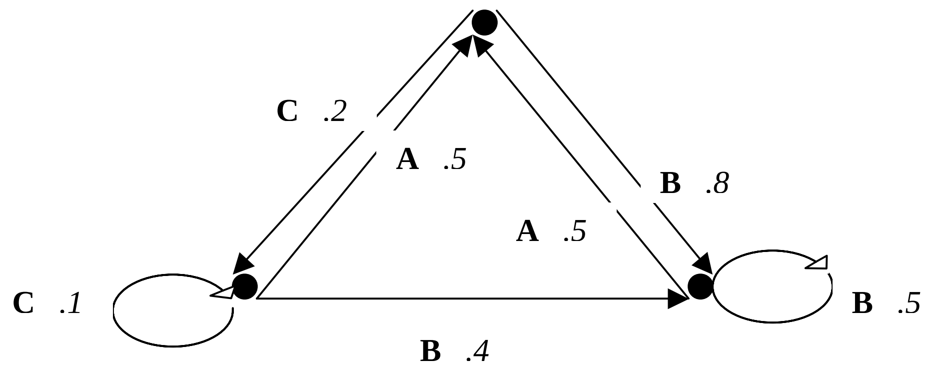
A fentebb leírt típusú sztochasztikus folyamatok a matematikában diszkrét Markov-folyamatokként ismeretesek, és azokkal a szakirodalom igen behatóan foglalkozik.[7] Az általános eset a következőképpen írható le: létezik egy rendszer véges számú lehetséges állapota: $S_1, S_2, \ldots, S_n$. Ezenkívül létezik a $p_i(j)$ átmeneti valószínűségsorozat, ami azt jelenti, hogy ha a rendszer az S_i állapotból mekkora valószínűséggel megy át következő lépésben S_j állapotba. Hogy ezt a Markov-folyamatot információforrásá tegyük, csupán azt kell feltételeznünk, hogy minden egyes állapotból a másikba való átmenet során egy betűt állít elő a rendszer. Az állapotok meg fognak felelni az előző betűkből „visszamaradó hatás”-nak.
A helyzetet rajzban a 3. 4. és 5. ábrákon látható módon vázolhatjuk. Az „állapotok” a gráf csomópontjai, míg az átmenethez létrehozott betűket és valószínűségeket a megfelelő vonal (él) mellett tüntettük fel. A 3. ábra a 2. fejezetben ismertetett B-példát mutatja, míg a 4. ábra az ugyanott szereplő C példára vonatkozik. A 3. ábrán csupán egy állapot van, mivel az

[7] Részletes tárgyalás található M. Frechet: *Methods des fonctions arbitraires. Theorie des événements en chaine dans le cas d'un nombre fini d'états possibles* (Tetszés szerinti függvények módszerei. A lác-események elmélete véges számú lehetséges állapot esetében) c. művében, (Paris, Gauthier Villars, 1938.)

egymást követő betűk függetlenek. A 4. ábrán annyi állapot van, ahány betű. Ha egy trigram-példát szerkesztenénk, a kiválasztott betűt megelőző lehetséges betűpárokra vonatkozólag legfeljebb n^2 számú állapot létezhet. Az 5. ábra a D példában bemutatott szószerkezetre mutat gráfot. Itt *S* a „betű”- vagy „szóközt” jelzi.



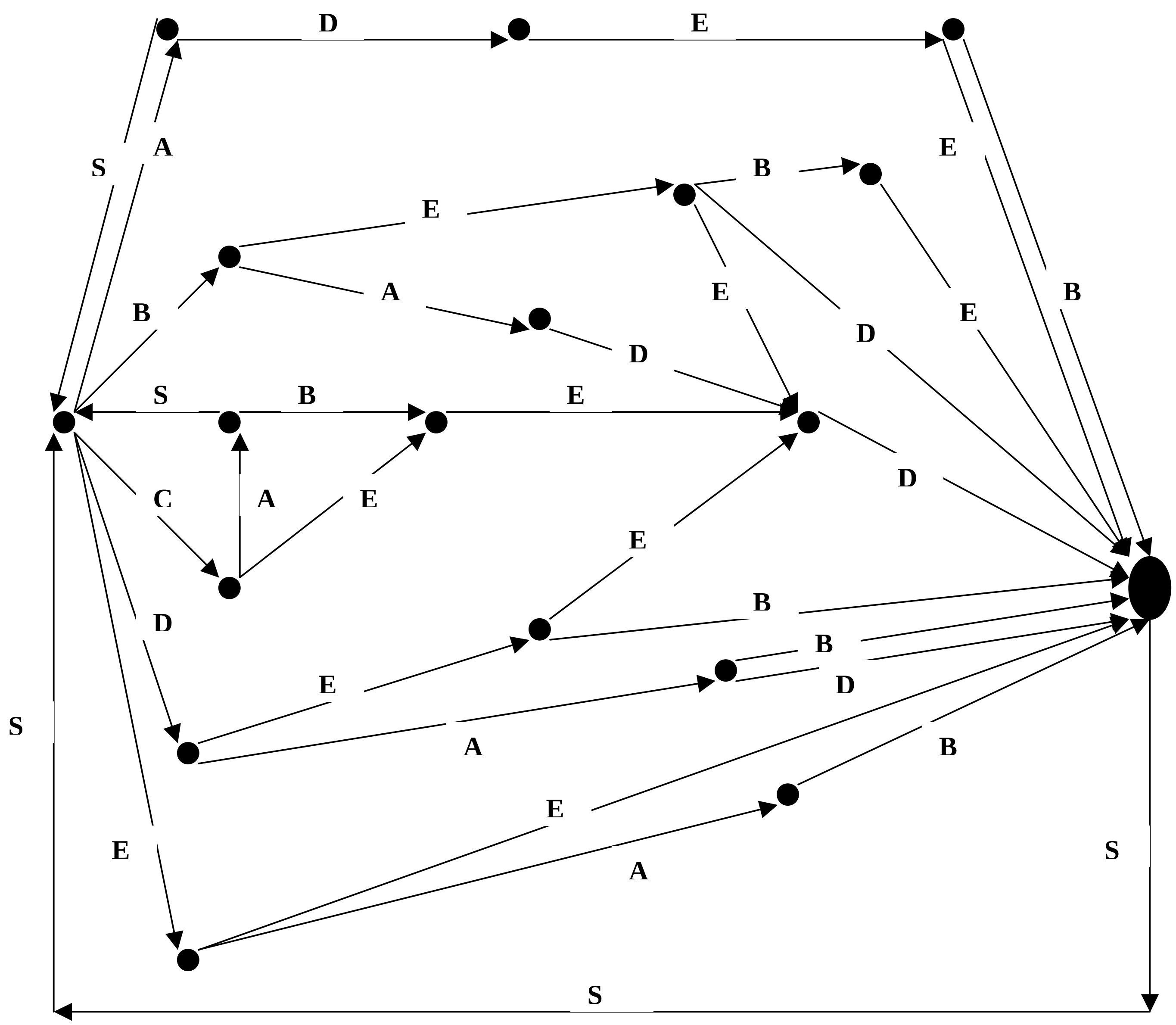
3.ábra *A B példabeli forrásnak megfelelő gráf*



4.ábra *A C példabeli forrásnak megfelelő gráf*

5. Ergodikus és vegyes források

Amint azt fentebb jeleztük, egy a céljainknak megfelelő diszkrét forrást úgy tekinthetünk, hogy az egy Markov-folyamattal reprezentálható. A lehetséges diszkrét Markov-folyamatok egy csoportja sajátos tulajdonságai révén jelentős szerephez jut a hírközléselméletben. Ez a különleges csoport az *ergodikus* folyamatokat foglalja magában és így a megfelelő forrásokat ergodikus forrásoknak fogjuk hívni. Bár egy ergodikus folyamat pontos definícióját elég nehéz megadni, az általános elv egyszerű. Egy ergodikus folyamat által létrehozott sorozatok statisztikus tulajdonságok szempontjából azonosak. Így az egyes sorozatokból kapott betűk, ill. kétbetűs összetételek (digramok) stb. előfordulási gyakoriságai a sorozatok hosszának növekedésével meghatározott, az adott sorozattól független korláthoz fognak tartani. Valójában ez nem igaz minden sorozatra, de az a készlet, amelyre ez hamis, zérus valószínűségű. Nagyvonalúan: az ergodikus tulajdonság statisztikus homogenitást jelent.



5.ábra *A D példabeli forrásnak megfelelő gráf*

Valamennyi, fentebb megadott, mesterséges nyelvre vonatkozó példa ergodikus. Ezt a sajátosságot a megfelelő gráfszerkezethez kapcsoljuk. Ha a gráf rendelkezik a következő két tulajdonsággal[8], akkor az adott folyamat ergodikus.

1. A gráf nem áll két olyan különálló részből (A és B), amelyekben a nyilak irányában a gráf vonalai mentén ne lehetne az A részben lévő csomópontokból a B-ben fekvő csomópontokhoz eljutni és viszont. (A gráf összefüggő! Szerk.)
2. Az azonos irányba mutató, gráfban záródó élsorozatot „körnek” nevezzük. A kör „hosszát” a benne lévő élek száma adja. Így az 5. ábrán a BEBES sorozat 5 egységnyi hosszúságú kört jelent. A második szükséges jellemző az, hogy a gráfban lévő valamennyi kör hosszúságának legnagyobb közös osztója 1 legyen.

[8] Ezek a Frechet idézett művében szereplő feltétel gráfok szempontjából ismételt állítások.

Ha az első feltétel teljesül, de a második nem (azáltal, hogy a legnagyobb közös osztó $d>1$), az élsorozatok bizonyos ismétlődő szerkezettel rendelkeznek. A különböző sorozatok d különböző csoportba oszthatók, amelyek statisztikusan azonosak, eltekintve attól, hogy kezdetük egymáshoz képest eltolódott (az a betű, amelyet a sorozatban 1-nek nevezünk). Egy nullától $d-1$-ig terjedő eltolással statisztikusan bármely sorozat bármely másikkal egyenértékűvé tehető. A $d=2$ esetben erre egy egyszerű példa a következő: három lehetséges betűnk van, ezek: a, b és c. Az a-t akár a b, akár a c követheti, rendre $1/3$ ill. $2/3$ valószínűséggel, míg a b és c betűt mindig az a követi. Ekkor egy tipikus sorozat:

$$a\ b\ a\ c\ a\ c\ a\ c\ a\ b\ a\ c\ a\ b\ a\ b\ a\ c\ a\ c.$$

Az ilyenfajta helyzeteknek a mi szempontunkból nincs sok jelentősége.
Ha az első feltételt megsértjük, úgy a gráfot feloszthatjuk egy sor olyan részgráfra, amelyek mindegyike teljesíti azt. Feltételezzük, hogy a második feltétel is teljesüljön minden részgráfra. Ez esetben olyan, ún. „vegyes” forrásunk van, amelyet egy sor tiszta összetevő alkot. Az elemek a különböző részgráfoknak felel meg. Ha L_1, L_2, L_3,... a komponens-források, felírhatjuk, hogy

$$L = p_1L_1 + p_2L_2 + p_3L_3 + \ldots.$$

Ahol p_i az L_i komponens-forrás valószínűsége.

Fizikailag a vázolt helyzet a következő: különböző L_1. L_2, L_3,... források léteznek, amelyek mind homogén statisztikai szerkezettel rendelkeznek (azaz ergodikusak). A priori nem tudjuk, melyik kerül sorra, de ha egy adott L_i tiszta komponensből egy sorozat elkezdődik, annak statisztikus szerkezetének megfelelően határozatlanul folytatódik.
Példaként a fent meghatározott folyamatok közül kettőt kiválasztva tételezzük fel, hogy $p_1=0.2$ és $p_2=0.8$. A vegyes forrás egy sorozatára ekkor azt kapjuk, hogy

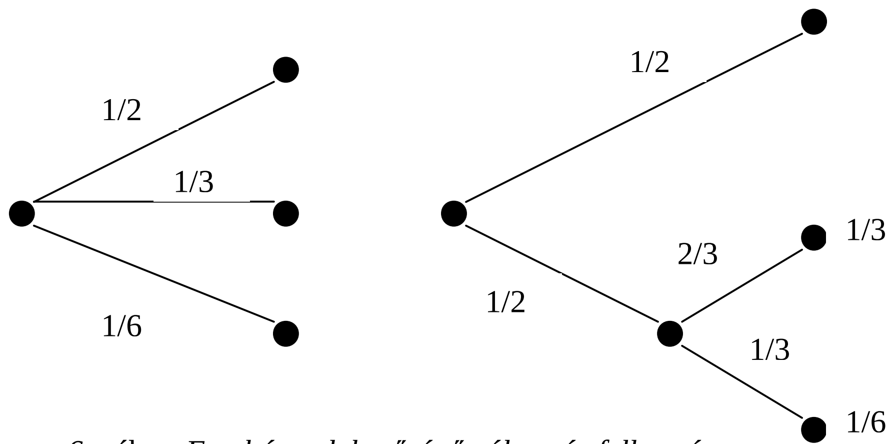
$$L = 0.2L_1 + 0.8L_2$$

Ha L_1-re 0.2, L_2-re 0.8-as valószínűséget választunk, és bármelyiket is választottuk, abból ezután egy sorozatot állítunk elő.
Hacsak az ellenkezőjét nem rögzítjük, a forrást ergodikusnak fogjuk feltételezni. Ez a feltételezés lehetővé teszi számunkra, hogy egy sorozat tagjaiból képezett átlagokat a lehetséges sorozatok összességének átlagával azonosítsuk (az eltérés valószínűsége eközben zérus). Pl. az A betű relatív gyakorisága egy adott végtelen sorozatban bizonyos, hogy a sorozatok összességében található relatív gyakoriságával lesz egyenlő.
Ha P_i az i állapot valószínűsége és $p_i(j)$ a j állapotba való átmenet valószínűsége akkor, hogy a folyamat stacionárius legyen világos, hogy P_i-nek ki kell elégitenie az alábbi egyensúlyi feltételeket: $P_j = \sum_i P_i P_i(j)$.
Az ergodikus esetben meg lehet mutatni, hogy bármilyen kezdeti feltételek mellett N szimbólum után a j állapotban tartózkodás $P_j(N)$ valószínűségei az egyenlőségi feltételben szereplő értékhez tartanak, ha $N \to \infty$.

6. Választás, bizonytalanság és entrópia

Egy diszkrét információforrást mint Markov-folyamatot reprezentáltunk. A kérdés az, definiálhatunk-e egy mérhető mennyiséget, amely bizonyos tekintetben megmutatja, mennyi információt állítottunk elő ezzel a folyamattal, vagy helyesebben szólva, milyen sebességgel állítjuk elő az információt?
Tételezzük fel, hogy egy sor olyan lehetséges eseményünk van, amelyek előfordulási valószínűségei: p_1, p_2, …, p_n. Ezek a valószínűségek ismertek, azonban ez minden, amit arról tudunk, hogy melyik esemény fog előfordulni. Találhatunk-e olyan mennyiséget, ami azt jellemzi, mennyi választási lehetőségünk volt a kiválogatás során, illetve mennyire vagyunk bizonytalanok a kimenetelt illetően?
Ha van ilyen mennyiség, mondjuk $H(p_1,p_2,\ldots,p_n)$, célszerű kikötni, hogy a következő tulajdonságokkal rendelkezzék:

1. H legyen folytonos a p_i tartományban.
2. Ha minden p_i azonos, azaz $p_i = \frac{1}{n}$, akkor H monoton növekvő függvénye n-nek. Egyformán valószínű eseményeknél több választási lehetőség, ill. bizonytalanság van, mint amikor valószínűbb események is előfordulnak.
3. Ha egy választási lehetőséget két, egymás utáni választási lehetőségbe ágaztatunk el, az eredeti H értéket az egyes egyedi H értékek súlyozott összegeként kell számítani. Ennek a jelentését a 6. ábrán ábrázoltuk.



6. ábra *Egy háromlehetőségű választás felbontása*

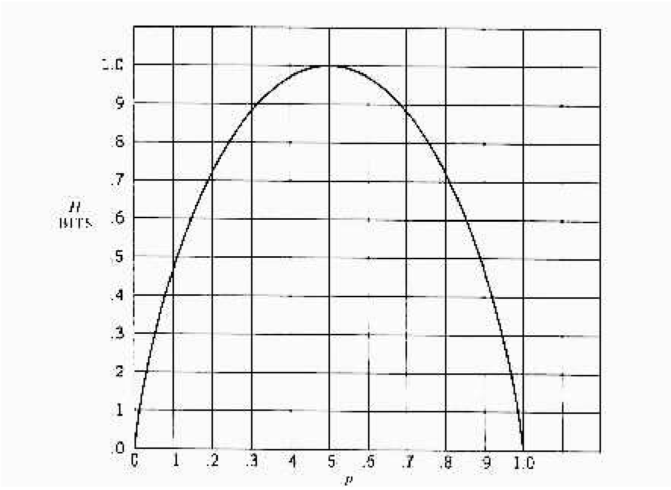
Baloldalt három valószínűségünk van, $p_1 = \frac{1}{2}$, $p_2 = \frac{1}{3}$, $p_3 = \frac{1}{6}$. Jobbra először két, egyaránt $\frac{1}{2}$ valószínűségű esemény közül választunk és ha ezek közül a második fordul elő, ez egy további választást ad, $\frac{2}{3}$ *és* $\frac{1}{3}$ valószínűségekkel. A végeredmények ugyanazt a valószínűségértéket adják, mint az előbb. Ebben a speciális esetben megköveteljük, hogy

$$H\left(\frac{1}{2},\frac{1}{3},\frac{1}{6}\right) = H\left(\frac{1}{2},\frac{1}{2}\right) + \frac{1}{2}H\left(\frac{2}{3},\frac{1}{3}\right)$$

Az $\frac{1}{2}$-es együtthatót mint súlytényezőt azért vezettük be, mivel ez a második választási lehetőség csak fele annyi ideig áll fenn. A 2. sz. Függelékben a következő eredményt állapítjuk meg:

2. TÉTEL: A *H* értékének egyetlen olyan kifejezése, amely a fenti három feltevést kielégíti, a következő: $H = -K\sum_{i=1}^{n} p_i \log p_i$, ahol *K* pozitív konstans szám.

Ez a tétel és a bizonyításához szükséges feltételezések semmiképpen nem szükségesek jelen elméletünkhöz, csupán azért közöltük röviden, hogy néhány későbbi definíciót ezáltal kézenfekvővé tegyünk. Ezeknek a definícióknak az igazi jelentősége hatásaiban rejlik.
A $H = -\sum p_i \log p_i$ jellegű mennyiségek (a *K* konstans csupán a mértékegység megválasztásától függ) központi szerepet játszanak az információelméletben, hiszen az információ mennyiségének, a választási lehetőségeknek és a bizonytalanságnak a mérésére szolgálnak. *H* ilyen alakjában felismerhetjük a statisztikus mechanika[9] egyes képleteiben meghatározott entrópia fogalmát, ahol p_i annak a valószínűségét jelenti, hogy egy rendszer fázisterének éppen az *i cellájában* van.
Ekkor *H* pl. a Boltzmann híres elméletéből ismert *H* lesz. A $H = -\sum p_i \log p_i$ értéket a p_1, p_2, …, p_n valószínűség-halmaz entrópiájának nevezzük. Az *x* valószínűségi változó entrópiáját *H(x)*-el jelöljük. Így *x* nem egy függvény változóját jelenti, hanem egy szám címkéjét jelöli, amely megkülönbözteti az *y* valószínűségi változó *H(y)* entrópiájától.
Az entrópia két *p* és *q=1-p* valószínűségű lehetőség esetében: $H = -(p\log p + q\log q)$, amelyet a 7. ábrán *p* függvényében ábrázoltunk.



7. ábra *Két p és (1-p) valószínűségű lehetőség entrópiája*

A *H* mennyiség egy sor érdekes tulajdonsággal rendelkezik, amelyek még inkább alkalmassá teszik arra, hogy az információ mértéke legyen.

1. *H=0* eset akkor, és csak akkor áll fenn, ha egyetlen kivételével valamennyi p_i érték zérus, ez az egy pedig éppen *1*-gyel egyenlő. Így csupán akkor tűnik el *H*, ha biztosak vagyunk a kimenetelben. Egyébként *H* pozitív.
2. Adott *n* értékre *H* maximális és *log n*-nel egyenlő, ha minden p_i egyenlő (azaz $p_i = \frac{1}{n}$)

   Ez intuitíve belátható, hogy a legbizonytalanabb helyzet.

[9] L. Pl.: R.C. Tolman: „Principles of Statistical Mechanics” (A statisztikus mechanika elvei), Oxford, Clarendon, 1938.

3. Tételezzünk fel két eseményt, *x*-et és *y*-t, rendre *m* illetve *n* lehetőséggel. Legyen *p(i,j)* a két esemény együttes előfordulásának valószínűsége, ahol *i* az első, *j* a második eseményre vonatkozik. Az együttes esemény entrópiája:

$$H(x,y) = -\sum_{i,j} p(i,j)\log p(i,j)$$

ahol $H(x) = -\sum_{i,j} p(i,j)\log\sum_{j} p(i,j)$ és $H(y) = -\sum_{i,j} p(i,j)\log\sum_{i} p(i,j)$

Könnyen megmutatható, hogy

$$H(x,y) \le H(x) + H(y)$$

és egyenlőség csak akkor áll fenn, ha az események függetlenek egymástól (azaz *p(i,j)=p(i)p(j)).* Azaz, az együttes esemény bizonytalansága kisebb, vagy egyenlő az egyedi események bizonytalanságainak összegével.

4. Bármiféle változás, amelynek során a p_1, p_2, …, p_n valószínűségek egymáshoz közelednek, *H* értékét növeli. Így ha $p_1 \langle p_2$ és növeljük p_1-et miközben p_2-t ugyanolyan mértékben csökkentjük, úgy hogy p_1 *és* p_2 csaknem teljesen azonos lesz, *H* értéke növekszik. Általánosabban: ha bármiféle „átlagoló” műveletet hajtunk végre

$$p_i^{'} = \sum_j a_{ij} p_j \quad \text{ahol} \quad \sum_i a_{ij} = \sum_j a_{ij} = 1 \quad \textit{és} \quad a_{ij} \ge 0$$

ekkor *H* növekszik (kivéve azt a speciális esetet, amelyben ez az átalakítás nem terjed tovább p_j permutációjánál, amikor is *H* természetesen változatlan marad).

5. Tételezzük fel, hogy van két véletlen eseményünk, *x* és *y* mint a 3. pontban, s ezek nem feltétlenül függetlenek egymástól. Bármely *i*-re, amelyet *x* felvesz, fennáll egy $p_i(j)$ feltételes valószínűség, hogy *y* a *j* értéket veszi fel. Ezt a következő képlet fejezi ki:

$$p_i(j) = \frac{p(i,j)}{\sum_j p(i,j)}$$

Az *y* feltételes entrópiáját $H_x(y)$-t olyan átlagos entrópiájaként definiáljuk, amely *x* minden értékét figyelembe veszi és amelyet az adott *x* érték valószínűségének megfelelően súlyozunk. Ezek szerint

$$H_x(y) = -\sum_{i,j} p(i,j)\log p_i(j)$$

Ez a mennyiség azt mutatja, mennyire vagyunk bizonytalanok *y* átlagos értékében, ha *x*-et ismerjük. Helyettesítve $p_i(j)$ értékét, kapjuk:

$$H_x(y) = -\sum_{i,j} p(i,j)\log p(i,j) + \sum_{i,j} p(i,j)\log\sum_j p(i,j) = H(x,y) - H(x)$$

$$\text{vagy} \quad H(x,y) = H(x) + H_x(y)$$

Az *x, y* együttes esemény bizonytalansága (vagy entrópiája) az *x* és *y* bizonytalanságának összegével egyenlő, ha *x*-et ismerjük.

6. A 3. és az 5. pontból a következik: $H(x) + H(y) \ge H(x,y) = H(x) + H_x(y)$

$$\text{Tehát } H(y) \geq H_x(y)$$

y bizonytalansága sosem növekszik azáltal, hogy *x*-et ismerjük, sőt csökkenni fog, hacsak *x* és *y* nem független események, míg ez utóbbi esetben nem változik.

7. Egy információforrás entrópiája

Tekintsünk egy fent leírt fajtájú, véges állapotú diszkrét forrást. Valamennyi lehetséges *i* állapothoz tartozni fog egy, a különböző lehetséges *j* szimbólumok $p_i(j)$ valószínűségeiből álló sorozat. Így minden állapothoz tartozik egy H_i entrópia. A forrás entrópiáját ezek után mint ezen H_i értékeknek a szóban forgó állapotok előfordulási valószínűségei szerint súlyozott átlagát definiálhatjuk: $H = \sum_i P_i H_i = -\sum_{i,j} P_i p_i(j) \log p_i(j)$

Ez a forrásnak a szöveg egy szimbólumára vonatkoztatott entrópiája. Ha a Markov-folyamat időben meghatározott sebességgel zajlik, arra meghatározható a másodpercenkénti entrópia is.

$$H' = \sum_i f_i H_i$$

ahol f_i az *i* állapot átlagos frekvenciája (azaz az *i* állapot másodpercenkénti előfordulásainak száma). Világos, hogy $H' = mH$,

ahol *m* a másodpercenként előállított szimbólumok átlagos száma. *H* vagy *H'* a forrás által szimbólumonként ill. másodpercenként produkált információ mennyiségét méri. Ha a logaritmus-alap 2, ezek a mennyiségek bit/szimbólum, ill. bit/s mértékegységet képviselnek. Amennyiben az egymást követő szimbólumok függetlenek, úgy *H* egyszerűen a - $\sum p_i \log p_1$-ből számítható, ahol p_i az *i* szimbólum valószínűsége. Tételezzük fel jelen esetben, hogy egy *N* szimbólumból álló hosszúságú üzenetünk van. Ez nagy valószínűséggel $p_1 N$-szeresen fogja tartalmazni az első szimbólum előfordulását, $p_2 N$-szer a másodikét stb. Ennélfogva ennek a sajátos üzenetnek a valószínűsége első közelítésben

$$p = p_1^{p_1 N} p_2^{p_2 N} \cdots p_n^{p_n N}$$

$$\text{vagy} \quad \log p = N \sum_i p_i \log p_i = -NH \quad \textit{ahol} \quad H = \frac{\log \frac{1}{p}}{N}$$

H tehát közelítőleg egyenlő egy tipikus hosszúságú sorozat reciprok valószínűsége logaritmusának és a sorozatban lévő szimbólumok számának a hányadosával. Bármilyen forrásra ugyanezt az eredményt kapjuk. Pontosabban fogalmazva (ld. A 3. Függeléket) a következőket mondhatjuk:

3. TÉTEL: Bármely adott $\varepsilon \rangle 0$ *és* $\delta \rangle 0$ esetében találhatunk olyan N_0-t, hogy a tetszés szerinti $N \geq N_0$ hosszúságú sorozatok két csoportba oszthatók.

1. Egy olyan készlet, melynek teljes valószínűsége kisebb mint ε.

2. A fennmaradó valamennyi többi tag valószínűsége kielégíti a következő egyenlőtlenséget: $\left| \frac{\log p^{-1}}{N} - H \right| \langle \delta$

Más szavakkal: csaknem bizonyosak lehetünk abban, hogy $\frac{\log p^{-1}}{N}$ nagyon megközelíti H-t, ha N értéke nagy.
Egy másik ezzel szorosan összefüggő eredmény adódik a különböző valószínűségek sorozataira. Tekintsünk ismét N hosszúságú sorozatokat és rendezzük azokat csökkenő valószínűségi sorrendbe. Definiáljuk az *n(q)* számot, amelyet a legvalószínűbb taggal kezdve veszünk a sorozatból, annak érdekében, hogy a figyelembevett sorozatokra a teljes q valószínűséget meghatározzuk.

4. TÉTEL: $\lim_{N\to\infty} \frac{\log n(q)}{N} = H \quad ahol \quad q \neq 0 \quad és \quad q \neq 1$

log n(q)-t mint a sorozat meghatározásához szükséges bitek számát értelmezhetjük, abban az esetben, ha csak a legvalószínűbb sorozatokat vesszük figyelembe, q teljes valószínűséggel. Ekkor $\frac{\log n(q)}{N}$ a bitek száma szimbólumonként.
A tétel azt mondja ki, hogy nagy N értékekre ez a q-tól független lesz és értéke H-val egyenlő. Az ésszerű valószínűséggel bekövetkező sorozatok száma logaritmusának növekedési sebességét H adja meg, tekintet nélkül arra, hogyan értelmezzük az „ésszerűen valószínű” kifejezést. Ezen eredmények alapján (amelyeket a 3. Függelékben bizonyítunk) a legtöbb célra hosszú sorozatokat úgy kezelhetünk, mintha azok csak 2^{HN} sorozatból állnának, melyek mindegyike 2^{-HN} valószínűséggel rendelkezik.
A következő két tétel azt mutatja meg, hogy H-t és H'-t határérték számítással közvetlenül az üzenetsorozatok statisztikájából határozhatjuk meg anélkül, hogy tekintettel lennének az egyes állapotokra és az azok közötti átmeneti valószínűségekre.

5. TÉTEL:
Legyen $p(B_i)$ a forrásból származó B_i szimbólumok egy sorozatának valószínűsége. Legyen továbbá $G_N = -\frac{1}{N}\sum_i p(B_i)\log p(B_i)$

ahol az összeg minden n szimbólumból álló B_i sorozatot figyelembe vesz. Ezek után G_N N-nek monoton csökkenő függvénye és $\lim_{N\to\infty} G_N = H$

6. TÉTEL:
Legyen $p(B_i, S_j)$ a B_i sorozat valószínűsége, amelyet az S_j szimbólum követ és $p_{Bi}(S_j)=p(B_i, S_j)/p(B_i)$ annak a feltételes valószínűsége, hogy B_i-t S_j követi.
Legyen $F_N = -\sum_{i,j} p(B_i, S_j)\log p_{B_i}(S_j)$

ahol az összeg az $N\!-\!1$ szimbólum valamennyi B_i tömbjére és valamennyi S_j szimbólumra vonatkozik. Ezután F_N az N-nek monoton csökkenő függvénye:

$$F_N = NG_N - (N-1)G_{N-1}, \quad G_N = \frac{1}{N}\sum_{n=1}^{N} F_n, \quad F_N \leq G_N, \quad \lim_{N\to\infty} F_N H$$

Ezeket az eredményeket a 3. Függelékben levezettük. Az eredmények azt mutatják, hogy *H*-ra közelítések sorát lehet megadni úgy, hogy az *1,2,…,N* szimbólumokra kiterjedő sorozatoknak csupán a statisztikus szerkezetét vesszük figyelembe. F_N a jobb közelítés. Valójában F_N a fentebb leírt fajtájú forrás *N*-edrendű közelítésének entrópiája. Ha a statisztikai befolyás nem terjed ki *N*-nél több szimbólumra, azaz ha az előző *(N–1)* szimbólum ismeretében a következő szimbólum feltételes valószínűsége nem változik attól, hogy bármely előző szimbólumét ismerjük, akkor $F_N = H$. Természetesen F_N a következő szimbólum feltételes entrópiája, ha az *(N–1)* megelőző szimbólumok ismertek, míg G_N az *N* szimbólumból álló tömbök szimbólumonkénti entrópiája.

Egy forrás entrópiájának a felvehető maximális értékhez való viszonyát (amely még mindig ugyanazon szimbólumokra korlátozódik) relatív entrópiának nevezzük. Ez, amint később látni fogjuk, a maximális lehetséges kompressziót jelenti, ha ugyanazon ABC-ben kódolunk.

1-ből kivonva a relatív entrópiát a *redundanciát* kapjuk. A köznapi angol nyelv redundanciája, a kb. 8 betűnél nagyobb távolságokra nem véve figyelembe a statisztikus szerkezetet, durván 50 %. Ez azt jelenti, hogy amikor angol nyelven írunk, az írott szöveg felét a nyelv szerkezete határozza meg, míg a másik felét szabadon választjuk. Az 50 %-os érték különböző módszerekkel nagyjából azonosnak adódik. E módszerek közül az egyik az angol nyelv approximációi entrópiájának számításán alapszik. Egy másik módszer szerint egy angol szövegből vett minta betűinek egy bizonyos részét töröljük és azután megkíséreltetjük valakivel visszaállítani azokat. Ha 50 %-os elhagyás esetén vissza lehet állítani az eredeti szövegrészt, akkor a redundanciának nagyobbnak kell lennie 50 %-nál. Egy harmadik módszer a titkosírás bizonyos ismert eredményein alapszik.

A redundancia két szélsőséges példája az angol prózában a Basic English és James Joyce *Finnegan ébredése* c. könyve. A *Basic English* nyelv szókészlete 850 szóra korlátozódik és redundanciája igen nagy. Ez tükröződik abban a tényben, hogy egy bekezdést Basic English-re lefordítva az meghosszabbodik. Másfelől Joyce megnövelte a szókészletet és – úgy tartjuk – a szemantikai tartalom tömörítését érte el.

Egy nyelv redundanciája a keresztrejtvények létezésével van összefüggésben. Ha a redundancia zérus, úgy bármely betűsorozat a nyelvben előforduló, értelmes szöveget ad és betűk bármely kétdimenziós elrendezése keresztrejtvényt képez. Ha a redundancia túl nagy, a nyelv túl sok megszorítást tartalmaz ahhoz, hogy nagy keresztrejtvényeket szerkeszthessünk. Részletesebb vizsgálat megmutatja, hogy amennyiben a nyelvre vonatkozó megszorítások jobbára kaotikus és véletlen természetűek, nagyméretű keresztrejtvények éppen 50 %-os redundancia esetén készíthetők. Ha a redundancia 33 % háromdimenziós keresztrejtvények lehetségesek stb.

8. A kódolási és dekódolási műveletek leírása

Le kell még írnunk matematikailag az adó és a vevő segítségével az információn végzett kódolási és dekódolási műveleteket. Mindkét eszközt diszkrét átalakítónak fogjuk nevezni. Az átalakító bemenetére jutó jelsorozatot a bemenő szimbólumok, a kimeneti jelet a kimeneti szimbólumok sorozata alkotja. Az átalakítónak lehet olyan belső tárolóegysége is, hogy a kimeneti jel nemcsak az éppen a bemeneten lévő jeltől, hanem az elmúlt időben történtektől is függ. Feltételezzük, hogy ez a belső memória véges, ami azt jelenti, hogy az átalakítónak véges számú (*m*) állapota létezik, és hogy kimeneti jele a jelenlegi állapotnak és a jelenlegi bemeneti szimbólumnak a függvénye. A következő állapot e két mennyiség egy újabb függvénye lesz. Így egy átalakító két függvény segítségével írható le:

$$y_n = f(x_n, \alpha_n) \qquad \alpha_{n+1} = g(x_n, \alpha_n)$$

ahol:
x_n az *n*-edik bemeneti jel,
α_n az átalakító állapota, ha bemenetén az n-edik bemeneti szimbólum van jelen,
y_n a kimeneti szimbólum (vagy kimeneti szimbólumok sorozta), amely x_n esetén áll elő, ha az átalakító állapota eközben α_n .

Ha egy átalakító kimeneti szimbólumait egy második átalakító bemeneti szimbólumaként azonosíthatjuk, úgy a két átalakító sorba kapcsolható, és eredményül ismét egy átalakítót kapunk. Ha a második átalakító az első kimenetéről működik és az eredeti bemeneti jeleket állítja vissza, az első átalakítót nem-szingulárisnak, míg a másodikat annak inverzének nevezzük.

7. TÉTEL:
Egy véges állapotú statisztikus jelforrás által meghajtott véges állapotú átalakító kimenete egy véges állapotú statisztikus forrás, melynek (egységnyi időre jutó) entrópiája kisebb, vagy egyenlő a bemenet entrópiájával. Ha az átalakító nem-szinguláris, a két mennyiség azonos.

Jellemezze α a forrás állapotát, amely x_i szimbólumsorozatot állít elő, legyen β annak az átalakítónak az állapota, amely kimenetén y_j szimbólum-tömböket ad. Ez a kombinált rendszer az (α,β) értékpár „szorzat állapot-terével" jellemezhető. Két pont a térben (α_1, β_1) és (α_2, β_2) melyeket egy vonal köt össze, s ha α_1 elő tud állítani olyan *x*-et, amely β_1 -et β_2 –be viszi át, úgy ez a vonal jelen esetben ennek az *x*-nek a valószínűségét jelenti. A vonalat az átalakító által előállított y_1 szimbólumok tömbjével jelöljük meg. A kimenet entrópiáját az állapotok súlyozott összegeként számíthatjuk. Ha először β-ra végezzük el az összegezést, az eredményül kapott tagok kisebbek, vagy egyenlők lesznek, mint α, tehát az entrópia nem növekszik. Ha az átalakító nem-szinguláris, kapcsoljuk kimenetét az inverzátalakító bemenetére. Ha H_1', H_2' és H_3' rendre a forrás, az első és a második átalakító kimeneti entrópiái, akkor $H_1^{'} \geq H_2^{'} \geq H_3^{'} = H_1^{'}$ és így H_1'=H_2'.
Tételezzük fel, hogy olyan rendszerrel rendelkezünk, amely a lehetséges sorozatokra nézve olyan megszorításokat tartalmaz, mint amilyeneket a 2. ábrán bemutatott lineáris gráf képvisel. Ha az *i* és *j* állapotok közötti különböző vonalakhoz $p_{ij}^{(s)}$ valószínűséget rendelünk hozzá, úgy a gráfból forrás lesz. Létezik egy olyan sajátos hozzárendelés, amely maximálja az eredő entrópiát (ld. A 4. sz. Függeléket).

8. TÉTEL:
Legyen a figyelembe vett megszorításokkal rendelkező rendszer olyan csatorna, amelynek kapacitása *C=log W* . Ha teljesül

$$p_{ij}^{(s)} = \frac{B_j}{B_i} W^{-l_{ij}^{(s)}}$$

ahol $l_{ij}^{(s)}$ az *i*-től *j* állapotig tartó *s*-edik szimbólum időtartama, és B_i –re fennáll:

$$B_i = \sum_{s,j} B_j W^{-l_{ij}^{(s)}}$$

akkor *H* maximális értékű és *C*-vel egyenlő.

Az átmeneti valószínűségek helyes megválasztásával egy csatorna szimbólumainak entrópiája maximálisan a csatornakapacitás értékéig növelhető.

9. A zajmentes csatorna alaptétele

A következőkben *H* értelmezését mint az információ előállításának sebességét fogjuk igazolni azzal, hogy bebizonyítjuk, *H* meghatározza a legjobb hatásfokú kódoláshoz szükséges csatornakapacitást.

9. TÉTEL:
Legyen egy forrás *H* (bit/szimbólum) entrópiájú, és egy csatorna *C* (bit/s) kapacitású. Ezek után lehetséges a forrás kimenetét olyan módon kódolni, hogy átlagosan $\frac{C}{H}-\varepsilon$ szimbólum/s sebességgel adjon a csatornán, ahol ε tetszés szerinti kis érték. Nem lehetséges *C/H*-nál nagyobb átlagos sebességgel adni.

A tétel második részét, hogy *C/H*-t nem lehet meghaladni, úgy bizonyíthatjuk, hogy észrevesszük: a csatorna másodpercenkénti bemeneti entrópiája a forrásével egyenlő, minthogy az adónak nem-szingulárisnak kell lennie, s ez az entrópia nem haladhatja meg a csatornakapacitást. Ennélfogva $H^{'} \le C$ és a másodpercenkénti szimbólumok száma = $H^{'}/H \le C/H$.
A tétel első részét kétféleképpen bizonyítjuk. Az első módszer az, hogy a forrás által előállított *N* szimbólum valamennyi sorozatának készletét figyelembe vesszük. Mivel *N* nagy, ezeket két csoportba osztjuk: az egyik $2^{(H+\eta)N}$-nél kevesebb tagból, a másik 2^{RN}-nél kevesebb tagból áll (ahol *R* a különböző szimbólumok számának logaritmusa) és a teljes valószínűsége μ-nél kisebb. *N* növekedésével η és μ zérushoz tart. A csatornában előforduló *T* időtartamú jelek száma nagyobb, mint $2^{(C-\theta)T}$, miközben θ kicsiny, ha *T* nagy. Ha

$$T=\left(\frac{H}{C}+\lambda\right)N$$

akkor elegendő számú csatornaszimbólum-sorozat adódik a nagy valószínűségű csoport esetében, ha *N* és *T* elegendően nagyok (bármilyen kicsiny is λ). A nagy valószínűségű csoportot egy tetszés szerinti, kölcsönösen egyértelmű módszerrel kódoljuk be a készletbe. A fennmaradó sorozatokat nagyobb sorozatok képviselik, amelyek kezdetén és végén egy, a nagy valószínűségű csoportban nem alkalmazott sorozat áll. Ez a speciális sorozat start- és stopjelként szolgál a különböző kódok számára. E kettő között elegendő időt engedünk meg ahhoz, hogy valamennyi kis valószínűségű üzenethez elegendő sorozatot adhassunk. Ehhez

$$T_1=\left(\frac{R}{C}+\varphi\right)N$$

ahol φ kicsiny. Az átvitel közepes sebessége szimbólum/s-ban kifejezve ekkor nagyobb lesz, mint

$$\left[(1-\delta)\frac{T}{N}+\delta\frac{T_1}{N}\right]^{-1}=\left[(1-\delta)\left(\frac{H}{C}+\lambda\right)+\delta\left(\frac{R}{C}+\varphi\right)\right]^{-1}$$

amint *N* növekszik, δ, λ és φ zérushoz és a sebesség *C/H*-hoz tart.

Ennek a kódolásnak – és ezáltal a tétel bizonyításának – egy másik módja a következőképpen adható meg: az N hosszúságú üzeneteket csökkenő valószínűségű sorrendben rendezzük és feltételezzük, hogy a valószínűségek $p_1 \geq p_2 \geq ... \geq p_n$. Legyen $P_s = \sum_1^{s-1} p_i$ azaz P_s a p_s-ig terjedő, de azt már nem tartalmazó valószínűségek összege. Először bináris rendszerbe kódolunk. Az s üzenetre a bináris kódot úgy kapjuk, hogy P_s-t mint bináris számot fejtjük ki. A kifejtés m_s helyiértékre történik, ahol m_s az egészrész, melyre igaz, hogy

$$\log_2 \frac{1}{p_s} \leq m_s \langle 1 + \log_2 \frac{1}{p_s}$$

Így a nagy valószínűségű üzeneteket rövid kódokkal, míg a kis valószínűségűeket hosszú kódokkal reprezentáljuk. Ezekből az egyenlőtlenségekből kapjuk, hogy

$$\frac{1}{2^{m_s}} \leq p_s \langle \frac{1}{2^{m_s-1}}$$

A P_s-re kapott kód egy vagy több m_s helyiértékében el fog térni a következőktől, minthogy valamennyi fennmaradó P_i legalább $\frac{1}{2^{m_s}}$-nel nagyobb és bináris kifejtésük ennélfogva az első m_s helyiértékben különbözni fog. Ezért valamennyi kód más és más, és így az üzenetet kódolt formájából vissza lehet állítani. Ha a csatorna-sorozatok már nem eleve bináris digitek sorozataiból állnak, úgy azok tetszés szerinti módon bináris számokká írhatók át, s így a bináris kódot a csatorna számára alkalmas jelekké tesszük át.
Az eredeti üzenet szimbólumonként felhasznált bináris digitjeinek átlagos H' számát könnyen meghatározhatjuk. Azt kapjuk, hogy

$$H' = \frac{1}{N} \sum m_s p_s$$

De $\frac{1}{N} \sum \left(\log_2 \frac{1}{p_s} \right) p_s \leq \frac{1}{N} \sum m_s p_s \langle \frac{1}{N} \sum \left(1 + \log_2 \frac{1}{p_s} \right) p_s$ és így $G_N \leq H' \langle G_N + \frac{1}{N}$

N növekedésével G_N tart H-hoz, amely a forrás entrópiája, és H' tart H-hoz.
Ebből azt látjuk, hogy – ha az N szimbólumoknak csupán egy véges eltolását alkalmazzuk – a kódolás vesztesége nem lehet nagyobb mint $1/N$ plussz a valóságos H entrópia és az N hosszúságú sorozatokra kiszámított G_N entrópia közötti különbség. Az ideálison túl szükséges többletidőtartam százalékosan ennélfogva kisebb, mint

$$\frac{G_N}{H} + \frac{1}{HN} - 1$$

Ez a kódolási mód lényegében azonos azzal, amelyet R.M. Fano[10] dolgozott ki, tőlünk függetlenül. Módszere az volt, hogy az N hosszúságú üzeneteket csökkenő valószínűségi sorrendbe rendezte. Ezeket a sorozatokat azután két, lehetőség szerint legjobban egyező valószínűségű csoportba osztotta. Ha az üzenet az első csoportba tartozik, első bináris digitje 0 lesz, egyébként pedig 1. A csoportok ezután hasonlóan közel azonos valószínűségű részsorozatokra oszlanak, és a megfelelő részsorozat határozza meg a második bináris digitet. Ezt a folyamatot addig ismételjük, míg mindegyik részsorozat csupán egyetlen üzenetet

[10] *Technical Report No. 65.* The Research Laboratory of Electronics, M. I. T. (65.sz. Kutatási jelentés, Massaxhusetts' Institute of Technology Elektronikai Kutató Laboratóriuma, 1949. március 17.)

tartalmaz. Könnyen belátható, hogy kisebb különbségektől eltekintve (amelyek általában az utolsó digitben fordulnak elő) ez a módszer ugyanazt az eredményt adja, mint a fentebb leírt aritmetikai eljárás.

10. Értékelés és példák

Annak érdekében, hogy a generátortól a terhelésig a maximális teljesítmény átvitelt érjük el, általában olyan transzformátort kell alkalmazni, amellyel a generátor a terhelés felől nézve a terhelés ellenállásával rendelkezik. A helyzet nagyjából analóg ezzel. A kódolást végző átalakítónak statisztikai szempontból illesztenie kell a forrást a csatornához. A forrás a csatorna felől, az átalakítón keresztül nézve azon forrásével azonos statisztikai szerkezettel kell, hogy rendelkezzen, amely a csatornában a maximális entrópiát adja. A 9. tétel tartalma az, hogy – bár elérni általában nem lehetséges – az egzakt illesztést a kívánt mértékig megközelíthetjük. A tényleges átviteli sebesség és a C csatornakapacitás arányát a kódoló rendszer hatásfokának nevezhetjük, s ez természetesen megegyezik a csatornaszimbólumok tényleges entrópiájának és a maximális lehetséges entrópiának a hányadosával.
Az ideális, vagy közel ideális kódolási folyamat az adóban és a vevőben általában nagy késleltetést tesz szükségessé. Az eddigiekben tekintett zajmentes esetben e késleltetés fő szerepe abban állt, hogy a valószínűségek és a megfelelő sorozathosszak között elfogadható illesztést hozzon létre.
Jó kód alkalmazásával egy hosszú üzenet reciprok-valószínűségének logaritmusa arányos kell legyen a megfelelő jel időtartamával, azaz a $\left|\frac{\log p^{-1}}{T} - C\right|$ kifejezésnek mindig kis értékűnek kell lennie, kivéve a hosszú üzenetek kis töredékét.
Ha egy forrás csupán egyetlen egyedi üzenetet képes előállítani, entrópiája zérus és csatornára nincs szükség. Például, egy számológép, amelyet a π egymás utáni jegyeinek kiszámítására építettek meg, meghatározott sorozatot állít elő, amelyben nincs véletlen elem. Nincs tehát csatornára szükség, hogy ezt „átvigyük” egy más pontba, mivel konstruálhatunk egy másik gépet, amely ugyanezt a sorozatot az adott pontban kiszámítja, jóllehet ez esetleg nem célszerű. Ilyen esetekben úgy dönthetünk, hogy teljesen vagy részlegesen figyelmen kívül hagyjuk a forrásra vonatkozó statisztikai ismeretünket. A π számjegyeit mint véletlen sorozat elemeit tekintjük, amennyiben bármilyen számjegysorozatok adására képes rendszert konstruálhatunk. Hasonlóképpen úgy is dönthetünk, hogy az angol nyelvre vonatkozó statisztikai ismereteink egy részét – de nem az egészet – felhasználhatjuk egy kód megalkotására. Ilyen esetben a forrást a megtartani kívánt statisztikus feltételeknek alávetett, maximális entrópiával rendelkezőnek tekintjük. Ennek a forrásnak az entrópiája meghatározza a szükséges és elégséges csatornakapacitást. A π példájában az egyetlen, még figyelembe veendő információ az, hogy valamennyi számjegyet a 0, 1, …, 9 készletből választjuk. Az angol nyelv esetén felhasználhatjuk a betűk előfordulási gyakoriságából adódó lehetséges statisztikai megtakarítást, azonban mást már nem. A maximális entrópiájú forrás ezek után az angol nyelv első közelítésének felel meg, s entrópiája meghatározza a szükséges csatornakapacitást.
Ezek közül az eredmények közül néhánynak egyszerű példáját adva, tekintsünk egy forrást, amely az A, B, C, D betűk közül választva egy betűsorozatot állít elő. Az egyes betűk kiválasztási valószínűsége rendre $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}$ és az egymást követő betűket függetlenül választjuk ki. Ekkor

$$H = \left(\frac{1}{2}\log\frac{1}{2} + \frac{1}{4}\log\frac{1}{4} + \frac{2}{8}\log\frac{1}{8}\right) = \frac{7}{4}\,bit\,/\,jel$$

Így megközelíthetünk egy olyan kódolási rendszert, amelyben ennek a forrásnak az üzeneteit szimbólumonként átlagosan $\frac{7}{4}$ bit/jel felhasználásával bináris számjegyek formájában kódolhatjuk. Ez esetben a következő kóddal ténylegesen elérhetjük a határértéket (az eredményt a 9. tétel második fajta bizonyítási eljárása segítségével kaptuk):

A	0
B	10
C	110
D	111

Az *N* szimbólumból álló sorozat kódolásához felhasznált bináris számjegyek átlagos száma:

$$N\left(\frac{1}{2}\cdot 1 + \frac{1}{4}\cdot 2 + \frac{2}{8}\cdot 3\right) = \frac{7}{4}N$$

Könnyen belátható, hogy a *0, 1* bináris egységek ½, ½ valószínűséggel rendelkeznek, s így *H* értéke a kódolt sorozatok esetén szimbólumonként *1* bit. Ennélfogva az eredeti betűkre átlagosan 7/4 bit jut, s az entrópiák az időben azonosak. Az eredeti készlet maximális lehetséges entrópiája *log 4 = 2*, amely akkor áll fenn, ha A, B, C és D egyaránt ¼-¼ valószínűségűek. Ekkor a relatív entrópia 7/8. A bináris sorozatokat kettő az egyhez alapon lefordíthatjuk az eredeti szimbólumkészletre, a következő táblázat segítségével:

00	A’
01	B’
10	C’
11	D’

Ez a kettős eljárás ezután az eredeti üzenetet ugyanazokba a szimbólumokba kódolja, de az átlagos kompresszió-viszony most már 7/8.
Második példaként tekintsünk egy forrást, amely egy *A’* és *B’*-ből álló sorozatot produkál, miközben A-nak *p*, B-nek pedig *q* a valószínűsége. Ha $p \langle\langle q$, akkor

$$H = -\log p^{p}(1-p)^{1-p} = -p\log p(1-p)^{(1-p)/p} = p\log\frac{e}{p}$$

Ilyenkor egy *0, 1* állapotú csatornán az üzenetet elég jól kódolhatjuk úgy, hogy egy speciális sorozatot, mondjuk a *0000*-t adjuk a kevéssé gyakori A szimbólumra, és utána egy olyan sorozatot adunk, amely a következő B számát jelzi. Ezt olyan bináris írásmóddal lehet jelezni, amelyben a speciális eltörölt sorozatot tartalmazó valamennyi szám szerepel. *16*-ig minden szám szokás szerint szerepel, a *16*-ot az utána következő bináris szám reprezentálja, amely nem tartalmaz négy zérust, mégpedig *17 = 10001* stb.
Kimutatható, hogy amint *p→0*, a kódolás az ideálishoz tart, feltéve, hogy a különleges sorozat hosszát megfelelően választjuk meg.

II. Diszkrét zajos csatorna

11. A diszkrét zajos csatorna leírása

A következőkben azt az esetet vizsgáljuk, amikor a jelet az átvitel során, illetve az egyik vagy a másik végponton zaj zavarja. Ez azt jelenti, hogy a vett jel nem szükségképpen azonos azzal, amelyet az adó adott. Két esetet különböztetünk meg. Ha egy bizonyos leadott jelnek határozott függvénye, akkor a hatást torzításnak nevezhetjük. Ha ennek a függvénynek van inverze – nincs két leadott jel, amely ugyanazt a vett jelet hozná létre – a torzítás korrigálható, legalábbis elvben, egyszerűen a vett jelen végzett ellentétes művelet végrehajtásával.
Az érdeklődésre számot tartó eset az, amelyben a jel nem mindig ugyanazt a változást szenvedi el az átvitel során. Ekkor az *E* vett jelet úgy foghatjuk fel, mint amely a leadott *S* jel és az *N* zaj változók függvénye:

$$E = f(S,N)$$

A zajt éppolyan véletlen változónak tekinthetjük, mint az előzőekben az üzenetet, s azt általánosságban egy megfelelő sztochasztikus folyamattal reprezentálhatjuk. A diszkrét zajos csatorna legáltalánosabb típusa, amelyet szemügyre veszünk, általánosítása a korábban leírt zajmentes, véges állapotú csatornának. Tételezzünk fel véges számú állapotot és egy valószínűségi sorozatot:

$$p_{\alpha,i}(\beta, j)$$

Ez annak a valószínűsége, hogy – amennyiben a csatorna α állapotban van és az *i* szimbólum került leadásra – a *j* lesz a következő vett szimbólum és a csatorna β állapotba kerül. Így α és β átfogja a lehetséges állapotokat, míg *i* a lehetséges adott, *j* pedig a lehetséges vett jeleket öleli fel. Abban az esetben, ha az egymást követő szimbólumokat a zaj függetlenül zavarja, csak egyetlen állapot létezik, s a csatorna a $p_i(j)$ átmenet valószínűségekkel írható le, amelyek annak a valószínűségét jelentik, hogy *i* leadott szimbólumra a vételi oldalon *j* szimbólumot kapunk.
Ha egy forrás egy zajos csatornát táplál, két statisztikus folyamat működik: a forrás és a zaj. Így több entrópiát lehet számítani. Elsőként létezik egy $H(x)$ entrópia, amely a forrás, ill. a csatorna bemenet entrópiája (a kettő egyenlő, ha az adó nem szinguláris). A csatornakimenet, azaz a vett jel entrópiáját $H(y)$-nal fogjuk jelölni. Zajmentes esetben $H(y) = H(x)$. A ki- és bemenet közös entrópiája $H(x,y)$ lesz. Végül két feltételes entrópia is létezik: $H_x(y)$ és $H_y(x)$, amelyek ismert bemenet esetén a kimenetre vonatkoznak és fordítva. Ezen mennyiségek között a következő kapcsolat áll fenn: $H(x, y) = H(x) + H_x(y) = H(y) + H_y(x)$
Mindezeket az entrópiákat másodpercenkénti, ill. szimbólumonkénti mértékegységekben is mérhetjük.

12. Ekvivokáció és csatornakapacitás

Zajos csatorna esetén a vett jelen végrehajtott semmilyen művelet segítségével nem lehetséges – általában – az eredeti üzenet ill. az adott jelet teljes bizonyossággal visszaállítani. Léteznek azonban olyan információadási módok, amelyek a zaj leküzdése szempontjából optimálisak. Ezt a problémát vizsgáljuk meg most.

Tételezzünk fel két lehetséges szimbólumot *0*-át és *1*-et, melyeket 1000 szimbólum/s sebességgel, $p_0 = p_1 = ½$ valószínűséggel viszünk át. Így információforrásunk 1000 bit/s sebességgel ad. Az adatátvitel során a zaj olyan hibát okoz, hogy átlagosan minden 100 vett jelből 1 lesz hibás (a 0 helyett 1-et, 1 helyett 0-t kapunk). Mennyi lesz most már az információátvitel sebessége? Biztos, hogy kevesebb, mint 1000 bit/s, minthogy a vett jelek mintegy 1%-a helytelen. Első látásra azt mondhatnánk, hogy 990 bit/s, egyszerűen levonva a hibás jel darabszámot. Ez azonban nem megfelelő, mivel ebben nem vettük figyelembe ismeretünk hiányát, hogy hol fordul elő ez a hiba? Ezt egészen addig a szélsőséges esetig fokozhatjuk, hogy feltételezhetjük: a zaj olyan nagy, hogy annak következtében a vett jelek teljesen függetlenek az adott jelektől. Annak a valószínűsége, hogy 1-et veszünk ½, bármilyen jelet is adunk, és ez fordítva is áll a 0-ra. Ezek után a vett jelek kb. fele pusztán a véletlen folytán helyes lesz, és hajlamosak lennénk arra, hogy a rendszert 500 bit/s átvitelűnek tekintsük, miközben tulajdonképpen egyáltalán nem is szállít információt. Hasonlóan „jó” átvitelt kaphatnánk ugyanis, ha a csatornát teljesen kikapcsolnánk, és a vételi ponton egy érme feldobásával határoznánk meg a szimbólumokat.
A kibocsátott információ mennyiségét nyilván úgy kell helyesen módosítani, hogy a vett jelből hiányzó információ mennyiségével korrigáljuk azt, vagy másképpen azzal a bizonytalansággal, amely az éppen vett jel esetében a leadott jelre fennáll. Az entrópiáról mint a bizonytalanság mértékéről korábban tett megállapításaink alapján – a vett jel ismeretében – célszerűnek látszik az üzenet feltételes entrópiáját használni a hiányzó információ mértékeként. Ez, ahogyan a későbbiekben látni fogjuk, valóban a helyes definíció. Ezt az elvet követve *R*, a valóságos átvitel sebessége úgy határozható meg, hogy az előállítási sebességből (azaz a forrás entrópiájából) a feltételes entrópia átlagos értékét levonjuk:

$$R = H(x) - H_y(x)$$

A $H_y(x)$ feltételes entrópiát, amely a vett jel átlagos bizonytalanságát méri, célszerűen ekvivokációnak fogjuk nevezni.
A fenti példában, ha a vételi oldalon egy 0-át észlelünk, annak az a posteriori valószínűsége, hogy 0 került adásra 0.99, és annak, hogy 1-est adtunk 0.01. Ezek az értékek felcserélődnek, ha 1-est veszünk. Így

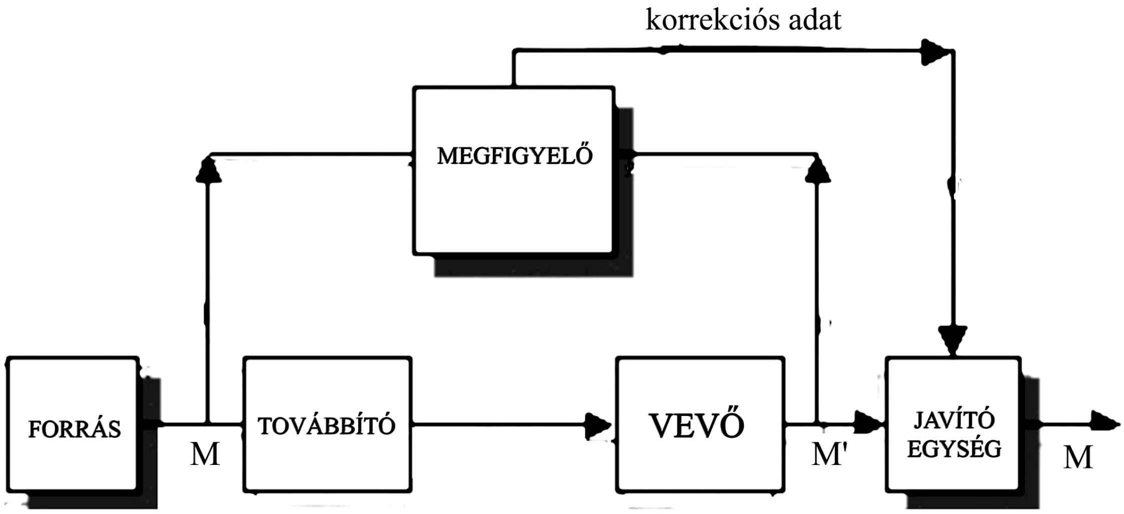
$$H_y(x) = -[.99 \log .99 + 0.01 \log 0.01] = 0.081\, bit/jel \text{ vagy 81 bit/sec.}$$

Azt mondjuk, hogy a rendszer 1000–81=919 bit/sec sebességgel ad. Abban az extrém esetben, amikor a 0-át azonos valószínűséggel vehetjük 0-ként és 1-ként egyaránt, (ill. hasonlóképpen 1-esre is), az a posteriori valószínűségek ½ - ½ értékűek, és

$$H_y(x) = -\left[\frac{1}{2}\log\frac{1}{2} + \frac{1}{2}\log\frac{1}{2}\right] = 1\, bit/jel \text{ vagy 1000 bit/sec.}$$

Az átviteli sebesség ekkor, ahogyan az várható, 0 lesz.

A következő tétel a ekvivokáció közvetlen intuitív értelmezését és annak mint különlegesen alkalmas mértékének az igazolását adja. Tekintsünk egy információátviteli rendszert és egy megfigyelőt (vagy valamilyen segédeszközt), amelyik mind a leadott, mind pedig a vételi oldalon (a zaj miatt bekövetkező hibákkal) visszaállított üzenetet képes észlelni. Ez a megfigyelő észleli a hibákat a vételi oldalon a vett üzenetben és egy „helyesbítő csatornán” keresztül adatokat küld a vételi ponthoz, hogy lehetővé tegye azt, hogy a vevő a hibákat korrigálhassa. A helyzetet vázlatosan a 8. ábrán mutatjuk be.



8 .ábra *Korrekciós rendszer vázlata*

10. TÉTEL:
Ha a korrekciós csatorna kapacitása $H_y(x)$-szel egyenlő, a korrekciós adatok úgy kódolhatók, hogy azokat e csatornán átküldve a hibákat – eltekintve egy tetszőleges kicsiny ε hányaduktól – korrigálhatjuk. Mindez nem lehetséges, ha a csatorna kapacitása kisebb mint $H_y(x)$.

Első közelítésben tehát $H_y(x)$ annak a járulékos információnak a mennyisége, amelyet a vételi ponton másodpercenként be kell táplálni ahhoz, hogy a vett üzenetet kijavítsuk.
Az első rész bizonyítására tekintsünk egy *M'* vett üzenet hosszú sorozatait, és a megfelelő *M* eredeti üzenetet. Ezen *M*-sorzatoknak lesznek olyan $TH_y(x)$ logaritmusai, amelyek célszerűen létrehozhatták valamennyi *M'*-t. Így minden egyes *T* másodperc alat $TH_y(x)$ bitet küldhetünk. Ez $H_y(x)$ csatornakapacitás mellett ε hibagyakorisággal valósítható meg.
A második részt úgy bizonyíthatjuk, hogy észrevesszük, bármilyen *x*, *y*, *z* diszkrét valószínűségi változóra

$$H_y(x,z) \geq H_y(x)$$

A baloldalt kibővítve, kapjuk:

$$H_y(z) + H_{yz}(x) \geq H_y(x)$$

$$H_{yz}(x) \geq H_y(x) - H_y(z) \geq H_y(x) - H(z)$$

Ha *x*-et a forrás kimenetének, *y*-t a vett jelnek és *z*-t a korrekciós csatornán küldött jelnek feleltetjük meg, úgy a jobboldal az ekvivokáció mínusz a korrekciós csatorna átviteli sebességével egyenlő. Ha ennek a csatornának a kapacitása kisebb az ekvivokáció értékénél, a jobboldal nagyobb lesz zérusnál és $H_{yz}(x) > 0$. Ez azonban a küldött jelek bizonytalanságát adja, ha ismerjük a vett és a korrekciós jeleket. Ha ez nagyobb zérusnál, úgy a hibagyakoriság nem lehet tetszés szerinti kicsiny értékű.
Példa:
Tételezzünk fel a bitek sorozatában véletlenszerűen előforduló hibákat: *p* annak a valószínűsége, hogy egy bit téves, és a $q=1-p$ azé, hogy helyes. Ha ezeknek a hibáknak a helye ismert, akkor korrigálhatók. Így a korrekciós csatornán csupán ezekre a pozíciókra vonatkozó információkat kell átvinni. Ehhez tehát olyan forrásra van szükség, amely bináris digiteket ad: az 1 (helytelen) digiteket *p*, a 0 (helyes) digiteket pedig *q* valószínműséggel.

Ehhez $-[p\log p + q\log q]$ értékű csatornakapacitás szükséges, amely az eredeti rendszer ekvivokációjának felel meg.
Az átviteli sebesség R, a fentebb szereplő azonosságok segítségével, két másik formában is írható:

$$R = H(x) - H_y(x) = H(y) - H_x(y) = H(x) + H(y) - H(x,y)$$

Az első definíciós kifejezést korábban mint a küldött információ mennyisége és a küldött bizonytalanság különbségét értelmeztük. A második a vett információ mennyiségének a zaj következtében csökkentett mértékét méri, míg a harmadik e két mennyiség összege, mínusz a kölcsönös entrópia, azaz más szóhasználattal: a két mennyiségre egyaránt vonatkozó, másodpercenkénti bitszám. Így mindhárom kifejezésnek bizonyos intuitív jelentősége van.
Egy zajos csatorna *C* kapacitása felső határt szab az átviteli sebességnek, azaz azt a sebességet adja meg, amely a forrás és a csatorna helyes illesztésekor áll fenn. Így a csatornakapacitást tehát a következőképpen definiálhatjuk:

$$C = \max(H(x) - H_y(x))$$

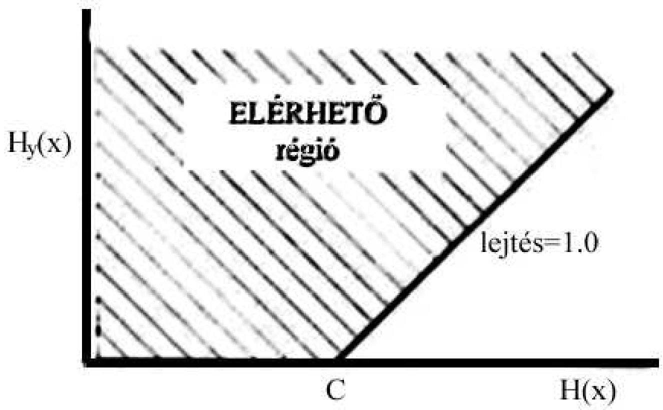
Ahol a maximum a csatornabemeneten szóba jöhető minden lehetséges információforrásra vonatkozik. Ha a csatorna zajmentes, akkor $H_y(x)=0$. Ez esetben a definíció megegyezik azzal, amelyet korábban a zajmentes csatornára adtunk, minthogy a csatorna maximális entrópiája – a 8. tétel értelmében – megegyezik kapacitásával.

13. A diszkrét zajos csatorna alaptétele

Meglepőnek tűnhet, hogy egy zajos csatornára egy meghatározott *C* kapacitást adunk meg, minthogy ilyen esetben sosem küldünk biztos információt. Világos azonban, hogy ha az információt redundáns módon küldjük, a hibák valószínűsége csökkenthető.
Például az üzenetet többször megismételve, annak különböző módon vett változatait statisztikusan vizsgálva, a hibák valószínűsége igen alacsony értékre szorítható. Azt várnánk, hogy ha ezt a hibavalószínűséget a 0-hoz közelítjük, a kódolás redundanciája végtelen nagyra fog növekedni és ennélfogva az átviteli sebesség a zérushoz tart. Ez azonban egyáltalán nem igaz. Ha így volna, nem létezne egy jól meghatározott kapacitás, csupán egy adott hibavalószínűségre vagy egy adott ekvivokációra vonatkozó érték, s ez a kapacitás pedig a hibával szemben szabott követelmények szigorításával csökkenne. Valójában a fentebb definiált *C* kapacitás igen határozott jelentőségű. Lehetséges ugyanis a csatornán keresztül tetszés szerinti kis hibagyakorisággal információt küldeni *C* sebességgel, amennyiben a kódolás megfelelő.
Ez a megállapítás nem érvényes *C*-nél nagyobb sebességekre. Ha kísérletet teszünk arra, hogy *C*-nél nagyobb mondjuk $C+R_1$ sebességgel adjunk, szükségképpen az R_1 növekménnyel egyenlő, vagy annál nagyobb ekvivokáció lép fel. A természet megköveteli az adott mértékű bizonytalanságot, úgyhogy végül is nem tudunk *C*-nél nagyobb értéket korrekt módon elérni.
A helyzet a 9. ábrán látható. A csatorna bemeneti információsebességét a vízszintes, míg az ekvivokációt a függőleges tengelyen ábrázoltuk. A vastag vonal feletti, vonalkázott mezőben bármely pont elérhető, míg a vonal alattiak nem érhetők el. A vonalon lévő pontok általában véve nem érhetők el, de rendszerint lesz két olyan pont, mely elérhető.
Ezek az eredmények a *C*-re vonatkozó definíció legfőbb igazolásai és az alábbiakban közöljük bizonyításukat.

11. TÉTEL:
Rendelkezzen egy diszkrét csatorna C kapacitással és egy diszkrét forrás másodpercenként H entrópiával. Ha $H \leq C$, akkor létezik olyan kódolási rendszer, amelynek esetén a forrás kimenetét tetszés szerinti kis hibagyakorisággal, vagy tetszés szerinti kis ekvivokációval adhatjuk ki a csatornára. Ha $H>C$, akkor a forrás úgy kódolható, hogy az ekvivokáció kisebb, mint $H–C+\varepsilon$, ahol ε tetszés szerinti kicsiny érték. Nem létezik olyan kódolási módszer, amely $H–C$-nél kisebb ekvivokációt adna.

Ennek a tételnek az első részét nem azzal a módszerrel bizonyítjuk, hogy előállítunk egy, a kívánt tulajdonságokkal rendelkező kódolási módszert, hanem úgy, hogy megmutatjuk: a kódok egy bizonyos fajtájú csoportjában ilyen kódnak léteznie kell. Valójában átlagolni fogjuk e csoporton belül a hibagyakoriságokat és kimutatjuk, hogy ezt az átlagértéket ε-nál kisebb értékűvé lehet tenni. Ha számok egy sorozatának átlaga kisebb, mint ε, akkor legalább egy olyan számnak kell léteznie e sorozatban, amely ε-nál kisebb. Ezzel a kívánt eredményt igazoltuk.

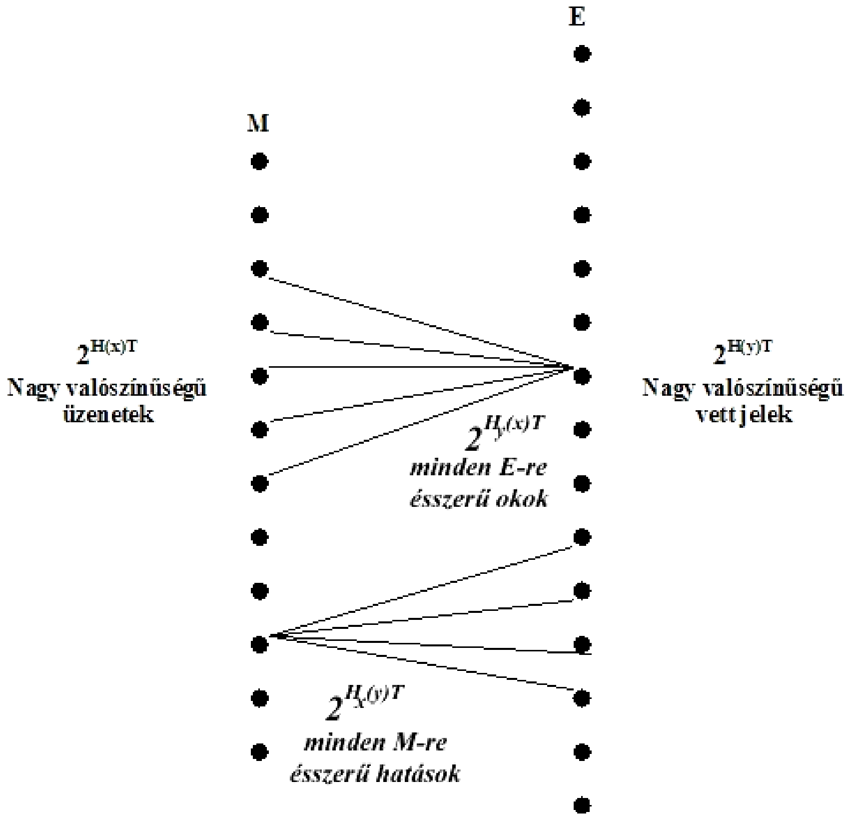


9. ábra *Adott bemenő-entrópiájú csatorna lehetséges ekvivokációja*

Egy zajos csatorna C kapacitását mint $C=Max(H(x) - H_y(x))$ határoztuk meg, ahol x a be-, és y a kimenet. A maximalizálást mindazon forrásokra ki kell terjeszteni, amelyek a csatorna bemeneteként szóba jöhetnek.
Legyen S_0 olyan forrás, amellyel elérhető a maximális C kapacitás. Ha ezt a maximumot nem is éri el egy forrás a valóságban (csupán mint egy határértéket közelíti), legyen S_0 olyan forrás, amely a maximális értékhez közelít. Tételezzük fel, hogy S_0-t a csatorna bemeneteként használjuk. Hosszú T időtartamú adott és vett lehetséges jelsorozatokat tekintünk. Ekkor a következő megállapítások igazak:

1. Az adott sorozatok két csoportba sorolhatók, egy nagy valószínűségű, körülbelül $2^{TH(x)}$ taggal rendelkező csoportba, és a fennmaradó, kis teljes valószínűségű sorozatok csoportjába.
2. Hasonlóképpen, a vett sorozatok egy, $2^{TH(y)}$ tagú, nagy valószínűségű csoportból, és a fennmaradó sorozatok kis valószínűségű készletéből állnak.
3. Minden nagy valószínűségű kimenetet mintegy $2^{TH_y(x)}$ bemenet állíthat elő. Valamennyi más eset teljes valószínűsége kicsiny.

Az ezekben a kifejezésekben használt „kicsiny” és „mintegy” szavakhoz tartozhó ε és δ értékek zérushoz tartanak, ha megengedjük, hogy T növekedjen, és S_0 megközelítse a maximalizált forrás esetét.
A helyzetet a 10. ábrán foglaltuk össze, ahol a bemeneteket a baloldali, a kimeneti sorozatokat pedig a jobboldali pontok jelentik. A felső vonalsereg egy tipikus kimenethez tartozó lehetséges okok sorozatát jelképezi, míg az alsó egy adott, tipikus bemenetből származó lehetséges eredményeket reprezentálja.



10. ábra *Egy csatorna be- és kimenetei közötti kapcsolatok sematikus ábrázolása*

Tételezzük fel most már, hogy egy másik S forrásunk van, amely $R<C$ sebességgel állít elő információt. A T időtartam alatt ez a forrás 2^{TR} számú nagy valószínűségű üzenetet produkál, amelyeket oly módon kívánunk összekapcsolni a lehetséges csatornabemenetekből kiválasztott készlettel, hogy kis hibagyakoriságot érjünk el. Ezt a társítást minden lehetséges módon előállítjuk (eközben azonban csak az S_0 forrás által meghatározott, nagy valószínűségű bemenetek csoportját használjuk), és a hibagyakoriságokat a lehetséges kódolási rendszerek ezen nagy csoportjára átlagoljuk. Ez ugyanaz, mintha a T időtartamú üzenetek és csatornabemenetek véletlen társítása esetére számolnánk ki a hibagyakoriságot. Tételezzük fel, hogy egy adott y_1 kimenetet észlelünk. Mi a valószínűsége annak, hogy az S forrásból származó üzenetek közül egynél több olyat találunk, amelyek y_1 lehetséges okainak készletében előfordulnak?
A $2^{TH(x)}$ pont között véletlenszerűen 2^{TR} számú üzenet oszlik meg. Így annak a valószínűsége, hogy egy adott pont üzenetként szerepeljen: $2^{T(R-H(x))}$

Annak a valószínűsége, hogy egyetlen pont sem képvisel a vonalseregben üzenetet (eltekintve a tényleges kiindulási üzenettől):

$$P=\left[1-2^{T(R-H(x))}\right]^{2^{TH_y(x)}}$$

Mivel $R \langle H(x) - H_y(x)$, ezért $R - H(x) = -H_y(x) - \eta$, ahol η pozitív. Ennek következtében

$$P = \left[1 - 2^{-TH_y(x) - T\eta}\right]^{2^{TH_y(x)}}$$

Az eredmény, ha $T \to \infty$, $\quad akkor \quad 1 - 2^{-T\eta}$.

Ennélfogva egy hiba előfordulási valószínűsége a zérushoz tart, s így a tétel első részét bizonyítottuk.

A tétel második felét egyszerűen bizonyíthatjuk azáltal, hogy észrevesszük: a forrásból csupán *C* bit/sec-os sebességgel tudunk információt adni, teljesen figyelmen kívül hagyva az előállított többi információt. A vételi oldalon a figyelmen kívül hagyott rész *H(x)–C* értékű ekvivokációt ad, az átvitt rész pedig csak ε növekményt igényel. Ezt a határértéket más módon is elérhetjük, amint ezt a folytonos eset vizsgálata során meg fogjuk mutatni.

A tétel utolsó állítása egyszerű következménye a *C*-re adott definíciónak. Tételezzük fel, hogy egy forrást *H(x)=C+a* entrópiával kódolunk, oly módon, hogy $H_y(x) = a - \varepsilon$ értékű ekvivokációt kapjunk (ε pozitív). Ekkor $R = H(x) = C + a \quad és \quad H(x) - H_y(x) = C + \varepsilon$

ahol ε pozitív. Ez ellentmond *C* azon definíciójának, miszerint az a $H(x)$–$H_y(x)$ maximális értéke.

Valójában tehát többet bizonyítottunk, mint amennyit a tételben állítottunk. Ha pozitív számok egy készletének átlaga a zérus ε sugarú környezetébe esik, legfeljebb $\sqrt{\varepsilon}$-ed része lehet $\sqrt{\varepsilon}$-nál nagyobb értékű. Minthogy ε tetszés szerint kicsiny, mondhatjuk, hogy csaknem valamennyi rendszer tetszés szerint megközelítheti az ideálist.

14. Értékelés

A 11. tétel bemutatása, jóllehet nem tiszta egzisztencia bizonyítás, az ilyesfajta bizonyítások néhány hiányosságát viseli magán. Az a törekvés, hogy a bizonyítás módszerét követve az ideális kódolást jól közelítsük, gyakorlati szempontból rendszerint nem célszerű. Valójában – eltekintve néhány egészen triviális esettől és bizonyos korlátozó körülményektől, az ideális eset közelítéseinek explicit leírását mind ez ideig senki sem adta meg. Ez talán nem is véletlen, hanem velejárója annak a nehézségnek, hogy egy véletlen sorozat jó közelítésére kell explicit szerkezetet megadni.

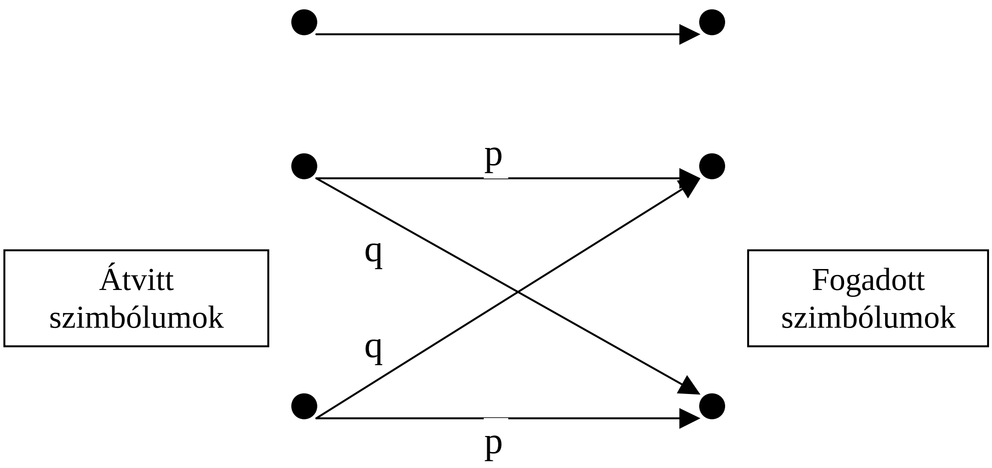
Az ideális eset közelítésének azzal a sajátossággal kellene bírnia, hogy amennyiben a jel a zaj következtében elfogadható módon változik meg, az eredeti jel még mindig visszaállítható legyen. Más szóval ez az eltérés általában nem fogja egy másik, az eredetitől eltérő, ésszerű jelhez közelíteni a jelet. Ez olyan áron teljesíthető, hogy a kódolásnál bizonyos mértékű redundanciát kell alkalmazni. A redundanciát megfelelő módon kell a rendszerbe bevinni, hogy a szóban forgó, sajátos zajszerkezetet leküzdjük, bár általában a forrásban alkalmazott bármely fajta redundancia segít, ha azt a vételi ponton kihasználjuk. Abban a sajátos esetben, ha a forrás már eleve rendelkezik bizonyos redundanciával és nem teszünk rá kísérletet, hogy ezt megszüntessük a csatornához való illesztés során, ez a redundancia segít a zaj hatásának kiküszöbölésében. Például, zajmentes távírócsatorna esetén időben mintegy 50%-os megtakarítást érhetnénk el az üzenet megfelelő kódolása révén. Ez azonban nem teljesül, és a csatornaszimbólumok az angol nyelv redundanciájának zömét megtartják. Ez ugyanakkor azzal az előnnyel jár, hogy a csatornában számottevő zaj engedhető meg. A betűk jelentős hányadát helytelenül véve, az üzenet a szövegösszefüggésből még mindig helyreállítható. Ez

sok esetben valószínűleg nem rossz közelítése az ideális esetnek, minthogy az angol nyelv statisztikus szerkezete meglehetősen bonyolult és az értelmes angol nyelvű betűsorozatok nem esnek távol (a tétel megfogalmazásához szükséges értelemben) egy véletlen kiválasztású sorozattól.
Csakúgy, mint a zajmentes esetben, rendszerint itt is szükség van bizonyos késleltetésre, hogy az ideális kódolást megközelítsük. Ennek most még az a további feladata is van, hogy megengedje azt, hogy a jelre nagy zajminta hasson, mielőtt az eredeti üzenetet illetően bármilyen értékelést végeznénk a vételi ponton. A minta hosszát növelve, a lehetséges statisztikai állítások egyre inkább pontosíthatók.
A 11. tételt és annak bizonyítását kissé eltérő módon is megfogalmazhatjuk, amikor is a zajmentes esettel való összefüggés világosabban kitűnik. Tekintsünk *T* időtartamú lehetséges jeleket és tételezzük fel, hogy ezekből egy másodlagos csoportot választunk ki, amelyet azután használni fogunk. Használjuk ezen alcsoport minden elemét azonos valószínűséggel és tételezzük fel, hogy a vevőt úgy szerkesztették, hogy – eredeti jelként ezen alcsoport elemei közül – a legvalószínűbb okot választja ki, amikor egy zavart jelet vesz. Definiáljuk $N(T,q)$-t mint a jelek olyan maximális számát, amelyet az alcsoport számára választhatunk ki oly módon, hogy a helytelen értelmezés valószínűsége kisebb vagy egyenlő q-val.

12. TÉTEL:
$\lim_{T\to\infty} \frac{\log N(T,q)}{T} = C$, ahol C a csatornakapacitás, feltéve, hogy q nem egyenlő *0*-val vagy *1*-gyel.
Más szavakkal: függetlenül attól, hogy adjuk meg a megbízhatóságra vonatkozó korlátainkat, megbízhatóan meg tudunk különböztetni a *T* időtartam alatt elegendő, körülbelül *CT* bitnek megfelelő üzenetet, ha *T* elegendően nagy. A 12. tételt össze lehet hasonlítani a zajnélküli csatorna kapacitására az 1. fejezetben adott definícióval.

15. Példa diszkrét csatornára és annak kapacitására

Egy diszkrét csatorna egyszerű példáját a 11. ábrán ábrázoltuk. Három lehetséges szimbólum szerepel: ezek közül az elsőt soha nem befolyásolja zaj. A második és harmadik egyaránt p valószínűséggel jut át zavartalanul a csatornán, míg q annak a valószínűsége, hogy az egyik jel a csatornán átjutva a másikká változik át.



11. ábra *Diszkrét csatorna példája*

Legyen $\alpha = -[p\log p + q\log q]$ valamint P és Q annak valószínűsége, hogy az első és második szimbólumot használjuk fel. Ekkor $H(x) = -P\log P - 2Q\log Q \qquad H_y(x) = 2Q\alpha$.

P és *Q* értékét oly módon kívánjuk megválasztani, hogy a $H(x)$–$H_y(x)$-et maximalizálhassuk, a $P+2Q=1$ megszorítás keretén belül. Ennélfogva írhatjuk, hogy

$$U = -P\log P - 2Q\log Q - 2Q\alpha + \lambda(P+2Q)$$

$$\frac{\partial U}{\partial P} = -1 - \log P + \lambda = 0$$

$$\frac{\partial U}{\partial Q} = -2 - 2\log Q - 2\alpha + 2\lambda = 0$$

λ *kiküszöbölése után*:

$$\log P = \log Q + \alpha$$

$$P = Qe^{\alpha} = Q\beta$$

$$P = \frac{\beta}{\beta+2} \qquad Q = \frac{1}{\beta+2}$$

$$a \quad csatorna \quad kapacitás: \quad C = \log\frac{\beta+2}{\beta}$$

Figyeljük meg, hogyan adódnak ebből a $p=1$ és $p=½$ esetben a kézenfekvő értékek. Az első esetben $\beta=1$ és $C=\log 3$ ami helyes, mivel a csatorna a három lehetséges szimbólum esetén zajmentes. Ha $p=½$, $\beta=2$ és $C=\log 2$, ekkor a második és a harmadik szimbólum semmiféleképpen nem különböztethető meg, és együtt úgy szerepelnek, mint egyetlen szimbólum. Az első szimbólumot $P=½$ valószínűséggel használjuk, a második és harmadik együtt ½ valószínűséggel. Ez az érték akárhogyan megoszolhat a két szimbólum között, és mégis elérhető a maximális kapacitás.
A *p* közbenső értékeire a csatornakapacitás *log 2* és *log 3* közé esik. A második és a harmadik szimbólum közötti különbségtétel hordoz ugyan némi információt, de nem annyit, mint a zajmentes esetben. Az első szimbólumot valamivel gyakrabban alkalmazzuk a másik kettőnél, mivel az a zajtól mentes.

16. Csatornakapacitás egyes különleges esetekben

Ha a zaj az egymást követő csatornaszimbólumokra függetlenül hat, azt egy sor p_{ij} átmenet-valószínűséggel írhatjuk le. Ez annak a valószínűsége, hogy ha az *i* szimbólumot küldtük, akkor a *j*-t fogjuk venni. A csatornakapacitást ekkor az alábbi kifejezés maximumaként kapjuk:

$$-\sum_{i,j} P_i p_{ij} \log \sum_i P_i p_{ij} + \sum_{i,j} P_i p_{ij} \log p_{ij}$$

Ahol P_i–t $\sum P_i = 1$ figyelembevételével változtatjuk. Ez a Lagrange-féle módszer segítségével a következő egyenlőséghez vezet:

$$\sum_j p_{sj} \log \frac{p_{sj}}{\sum_i P_i p_{ij}} = \mu \qquad s = 1,2,...$$

P_s-sel szorozva és s tagot összegezve kapjuk, hogy μ=C. Legyen p_{sj} inverze (ha ilyen létezik) h_{st}, úgy hogy $\sum_s h_{st} p_{sj} = \delta_{tj}$, ekkor

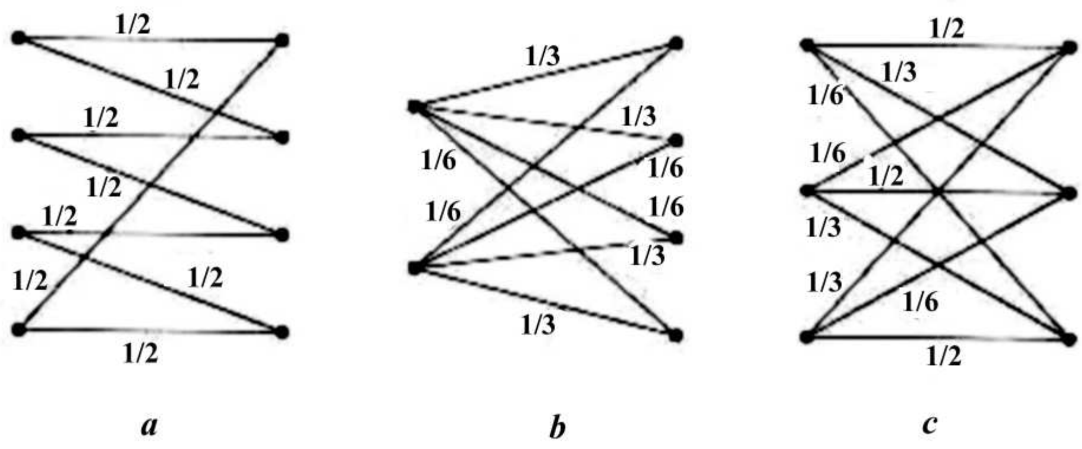
$$\sum_{s,j} h_{st} p_{sj} \log p_{sj} - \log \sum_i P_i p_{it} = C \sum_s h_{st}$$

$$\sum_i P_i p_{it} = \exp\left[-C \sum_s h_{st} + \sum_{s,j} h_{st} p_{sj} \log p_{sj} \right] \quad \textit{vagy}$$

$$P_i = \sum_t h_{it} \exp\left[-C \sum_s h_{st} + \sum_{s,j} h_{st} p_{sj} \log p_{sj} \right]$$

Ez az egyenletrendszer P_i maximális értékeinek meghatározására szolgál, ha C értékét úgy állapítjuk meg, hogy $\sum P_i = 1$. Ha ez teljesül, a csatornakapacitás C lesz, és P_i az ezen kapacitás eléréséhez szükséges csatornaszimbólumokra vonatkozó megfelelő valószínűségeket jelenti.

Ha valamennyi bemeneti szimbólum a belőle kiinduló vonalon felírt, azonos valószínűségi sorral rendelkezik és ugyanez igaz valamennyi kimeneti szimbólumra is, a kapacitás egyszerűen számítható. Erre példákat a 12. ábrán mutatunk. Ilyen esetben $H_x(y)$ független a bemeneti szimbólumok eloszlási valószínűségétől, és értékét $-\sum p_i \log p_i$ adja meg, ahol a p_i a bármely bemeneti szimbólumokból származó átviteli valószínűség.



12. ábra. *Minden be- és kimenetre azonos átmeneti valószínűséggel rendelkező diszkrét csatornák*

A csatornakapacitás így fejezhető ki: $Max[H(y) - H_x(y)] = MaxH(y) + \sum p_i \log p_i$

H(y) maximuma nyilván *log m* lesz, ahol *m* a kimeneti szimbólumok száma, minthogy lehetséges mindegyiket azonos valószínűségűvé tenni azáltal, hogy a bemeneti szimbólumokat azonos valószínűségűvé tesszük. A csatornakapacitás ennélfogva

$$C = \log m + \sum p_i \log p_i$$

A 12/a. ábrán *C=log 4 – log 2 = log 2*
Ez elérhető azáltal, hogy csak az első és a harmadik szimbólumot használjuk fel.

A 12/b. ábrán $C = \log 4 - \frac{2}{3}\log 3 - \frac{1}{3}\log 6 = \log 4 - \log 3 - \frac{1}{3}\log 2 = \log \frac{1}{3} 2^{\frac{5}{3}}$

A 12/c. ábrán $C = \log 3 - \frac{1}{2}\log 2 - \frac{1}{3}\log 3 - \frac{1}{6}\log 6 = \log \frac{3}{2^{\frac{1}{2}} 3^{\frac{1}{3}} 6^{\frac{1}{6}}}$

Tételezzük fel, hogy a szimbólumok több csoportra oszlanak úgy , hogy a zaj egy adott csoportba tartozó szimbólumot sosem befolyásol oly módon, hogy az összetéveszthető lenne egy másik csoportbelivel. Legyen az *n*-edik csoportra a kapacitás C_n (bit/sec-ban), ha csak ebbe a csoportba tartozó szimbólumokat használunk. Ekkor könnyen belátható, hogy – a teljes készlet legjobb kihasználása érdekében – az *n*-edik csoportban lévő összes jel P_n teljes valószínűsége

$$P_n = \frac{2^{C_n}}{\sum 2^{C_n}}$$

Egy csoporton belül a valószínűség úgy oszlik meg, mintha az az érték lenne, amikor csupán ezeket a szimbólumokat használnánk. A csatorna kapacitása $C = \log \sum 2^{C_n}$.

17. Példa hatékony kódolásra

A következő példa, bár némiképp mesterkélt, olyan esetet mutat, amelyben lehetséges egy zajos csatornához való egzakt illesztés. Két csatornaszimbólumunk van: *0* és *1* és a zaj az ezekből kialakított hetes csoportokra hat. Egy ilyen hétjegyű csoportot vagy hiba nélkül adunk ki, vagy egy meghatározott szimbólum nem helyes közülük. Ez a nyolc valószínűségi érték egyforma eséllyel fordulhat elő. Ekkor

$$C = Max\left[H(y) - H_x(y)\right] = \frac{1}{7}\left[7 + \frac{8}{8}\log\frac{1}{8}\right] = \frac{4}{7} bit / jel$$

Egy olyan, jó hatásfokú kód, amely lehetővé teszi a hibák teljes kiküszöbölését és a szimbólumok *C* sebességű adását, a következő formájú lehet (ezt R. Hamming módszere segítségével állítjuk elő):
Legyen egy tömb hét szimbóluma X_1, X_2,…., X_7. Ezek közül az X_3, X_5, X_6 és X_7 üzenetszimbólumok, s ezeket a forrás tetszés szerint választja ki. A fennmaradó három redundáns a következőképpen számítható:

$$X_4 \rightarrow \quad \alpha = X_4 + X_5 + X_6 + X_7$$
$$X_2 \rightarrow \quad \beta = X_2 + X_3 + X_6 + X_7$$
$$X_1 \rightarrow \quad \gamma = X_1 + X_3 + X_5 + X_7$$

Ha egy héttagú csoportot veszünk, az α, β, γ értékek kiszámíthatók, és ha az páros, úgy zérusnak, ha páratlan, akkor egyesnek nevezzük. Az α, β, γ bináris számok ezután megadják a helytelen X_i -t. (Ha zérus, úgy nem volt tévesztés.)[11]

III. Folytonos üzenetek

Most azt az esetet vesszük szemügyre, amikor a jelek, vagy az üzenetek (vagy mindkettő) folytonos változók, szemben az eddig tárgyalt diszkrét jellegű jelenségekkel. A folytonos eset nagyrészt levezethető a diszkrétből határérték képezéssel, úgy hogy az üzenet/jel-kontinuumot nagy, de véges számú kis tartományra osztjuk fel és a szóban forgó különböző paramétereket diszkrét módon számítjuk. Ahogy csökkentjük ezeknek a tartományoknak a méretét, e paraméterek általában a folytonos esethez tartozó megfelelő értékekhez mint határértékhez tartanak. Néhány új jelenség is felbukkan azonban, s a hangsúly az általános eredmények egyes sajátos esetekre történő alkalmazására helyeződik át.
A folytonos esetben nem fogunk arra törekedni, hogy eredményeinket a legnagyobb általánossággal, vagy a tiszta matematika extrém szigorúságával kapjuk meg, minthogy ez nagymértékű absztrakt számelméleti apparátust igényelne, és elhomályosítaná az analízis fő menetét. Előzetes vizsgálatok azonban azt mutatják, hogy az elmélet formába önthető teljesen axiomatikus és szigorú módon, amely magában foglalja mind a diszkrét, mind a folytonos esetet és még másokat is. Azt az esetenkénti szabadságot, amelyet a határérték-meghatározás során megengedünk a következőkben ismertetett analízisben, valamennyi, a gyakorlat számára fontos esetben igazolni lehet.

18. Függvényhalmazok és függvényegyüttesek

A folytonos információ esetén függvény halmazokkal és függvényegyüttesekkel kell foglalkoznunk. Egy függvényhalmaz, ahogy azt a neve is mutatja, csupán egy osztálya vagy gyűjteménye egyváltozós – időfüggvényeknek. A halmazban szereplő különböző függvények explicit – vagy a halmaz egyes függvényei által mutatott, míg másoknál nem jelentkező tulajdonság implicit – kifejezésének megadásával jellemezhetők. Néhány példa erre:

1. A függvények halmaza: $f_\theta(t) = \sin(t + \theta)$
   (Minden egyes értéke meghatároz egy adott függvényt a halmazban.)
2. Minden olyan időfüggvény halmaz, amely nem vesz fel W-nél nagyobb frekvenciákat.
3. A W-frekvenciasávra és az A amplitúdóra korlátozott valamennyi függvény halmaza.
4. Az angol nyelv beszédhangjaiból mint időfüggvényekből álló függvényhalmaz.

---

[11] Néhány további példát az önjavító kódokra lásd M.J.E. Golay: *Notes on Digital Coding* („Néhány megjegyzés a digitális kódolásról”), Proceedings of the Institute of Radio Engineers, 1949. június, 6. sz. 37. k., 637. old.

Függvények együttesén olyan függvényhalmazt értünk, amellyel együtt megadunk egy olyan valószínűségi mértéket is, melynek segítségével meghatározhatjuk egy bizonyos sajátosságokkal rendelkező függvény valószínűségét. Például, az $f_\theta(t) = \sin(t+\theta)$ halmaz esetén θ-ra megadhatunk egy valószínűségi eloszlást, mondjuk $P(\theta)$-t. A halmaz ekkor együttesé válik. Néhány további példa függvényegyüttesre:

1. $f_k(t)$ függvények véges halmaza $(k=1,2,3,\ldots,n)$, ahol az f_k érték p_k valószínűséggel fordul elő.
2. Egy véges dimenziós függvénycsalád $f(\alpha_1, \alpha_2, \ldots, \alpha_n, t)$

   Az α_i paraméterre a következő valószínűségi eloszlással: $p(\alpha_1, \alpha_2, \ldots, \alpha_n)$

   Tekinthetjük például az $f(a_1, \ldots, a_n, \theta_1, \ldots, \theta_n, t) = \sum_{i=1}^{n} a_i \sin i(\omega t + \theta_i)$

   által meghatározott együttest, ahol az a_i amplitúdók eloszlása normális és független, míg a θ_i fázisok ($0 \le \theta_i \le 2\pi$) egyenletes eloszlásúak és egymástól függetlenek.
3. Az $f(a_i, t) = \sum_{n=-\infty}^{+\infty} a_n \dfrac{\sin \pi(2Wt-n)}{\pi(2Wt-n)}$ együttes, ahol a_i normál eloszlású értékei függetlenek és ugyanazon $\sqrt{N}$ szórással rendelkeznek. Ez a „fehér” zajt leíró összefüggés, amely a 0 – W sávban határolt és átlagos teljesítménye N.[12]
4. Válasszunk a t tengelyen pontokat a Poisson-eloszlás szerint. Minden egyes kiválasztott pontban behelyettesítve az $f(t)$ függvényt és a különböző függvényeket összeadva, a következő együttes adódik:

   $$\sum_{k=-\infty}^{+\infty} f(t+t_k)$$

   ahol t_k a Poisson-eloszlás pontjait jelenti. Ezt az együttest az impulzus-, vagy „sörét”-zaj típusának tekintjük, amelynél minden impulzus azonos.
5. Az angol nyelvű beszéd függvényhalmaza, amikor a valószínűségértékeket a mindennapi használatban mutatkozó előfordulási gyakoriság szabja meg.

   Az $f_\alpha(t)$ függvények együttese stacionárius, ha mindegyik függvényt időben bármilyen mértékben eltolva, ugyanazt az együttest kapjuk. Az $f_\theta(t) = \sin(t+\theta)$ együttes stacionárius, ha θ egyenletesen oszlik el 0 és 2π között. Ha mindegyik függvényt t_1-gyel eltoljuk, akkor az $f_\theta(t+t_1) = \sin(t+t_1+\theta) = \sin(t+\varphi)$ kapjuk, ahol φ egyenletesen oszlik el 0 és 2π között. Minden függvény megváltozott, de az együttes egészében véve változatlan maradt az átalakítás során. A többi fenti példa szintén stacionárius.

   Egy együttest akkor nevezünk ergodikusnak, ha az stacionárius és a függvényeknek nincsen 0-tól és 1-től különböző valószínűségű, stacionárius részhalmazuk a halmazon belül.

   A $\sin(t+\theta)$ együttes ergodikus. Ezen 0-tól és 1-től eltérő valószínűségű függvényeknek nincs olyan részhalmazuk, amely minden időbeli eltolás során önmagába

[12] Ezt a meghatározást mint a sávhatárolt fehér-zaj definícióját használhatjuk. Ez bizonyos előnyt jelent abból a szempontból, hogy kevesebb határértékképzéssel jár, mint a korábban használt definíciók. A „fehér-zaj” elnevezés, amely már meglehetősen meghonosodott az irodalomban, talán nem túlságosan szerencsés. Az optikában a fehér fény egy pontszerű fényforrással szemben bármilyen folytonos spektrumot jelenthet, vagy olyan spektrumot jelent, amely a hullámhossz függvényében konstans (ami nem ugyanaz, mintha azt mondanánk, hogy egy spektrum a frekvencia függvényében konstans).

transzformálódik. Másfelől az $a\sin(t+\theta)$ együttes, amelyen a normális eltolású, θ pedig egyenletes, stacionárius, de nem ergodikus. Ezeknek a függvényeknek a részhalmaza, például ha *a* 0 és 1 közé esik, stacionárius és valószínűsége nem egyenlő zérussal, ill. 1-gyel.

A felsorolt példák közül a 3. és 4. szerinti ergodikus, és talán még az 5. is annak tekinthető. Ha egy együttes ergodikus, akkor első közelítésben azt mondhatjuk, hogy a halmazban bármelyik függvény tipikusan jellemző az együttesre. Pontosabban szólva: ismeretes, hogy ergodikus együttes esetén bármely, a teljes együttesre vonatkozó statisztikus átlag (1 valószínűség mellett) egyenlő egy, a halmazban szereplő adott függvény minden időeltolása során kapott átlaggal.[13] Egyszerűen fogalmazva, minden függvényt úgy tekinthetünk, mint amely az idő haladtával, megfelelő frekvenciával a halmazban szereplő bármely függvény valamennyi konvolúcióját felveszi.

Hasonlóképpen ahhoz, amikor számokkal vagy függvényekkel különféle műveleteket hajthatunk végre, hogy új számokat, illetve függvényeket kapjunk, együtteseken is végrehajthatunk műveleteket, hogy új együttesekhez jussunk. Tételezzünk fel például $f_\alpha(t)$ függvényből álló együttest és egy T operátort, amely minden egyes $f_\alpha(t)$ függvényhez egy $g_\alpha(t)$ eredő függvényt rendel hozzá: $g_\alpha(t)=T\,f_\alpha(t)$

A $g_\alpha(t)$ halmazra ugyanúgy definiáljuk a valószínűségi mértéket, mint az $f_\alpha(t)$ halmaz esetében. A $g_\alpha(t)$ függvények egy adott részhalmazának a valószínűsége egyenlő azon $f_\alpha(t)$ függvények részhalmazának valószínűségével, amelyek a T operátor közbejöttével a g függvények adott részhalmazának a tagjait előállítják. Fizikailag ez megfelel egy együttes valamely eszközön, pl. egy szűrőn, egyenirányítón vagy modulátoron történő átbocsátásának. Ekkor a $g_\alpha(t)$ együttest az eszköz kimeneti függvényei adják.

Egy eszközt, vagy a T operátort akkor nevezünk invariánsnak, ha a bemeneten alkalmazott eltolás egyszerű eltolást hoz létre a kimeneten, azaz ha $g_\alpha(t)=T\,f_\alpha(t)$

Ebből következik valamennyi $g_\alpha(t+t_1)=Tf_\alpha(t+t_1)$ értéke. Egyszerűen belátható, (ld. az 5. Függeléket), hogy ha T invariáns és a bemeneti együttes stacionárius, akkor a kimeneti együttes is stacionárius. Hasonlóképpen, ha a bemenet ergodikus, a kimenet is az lesz.

Egy szűrő vagy egy egyenirányító mindenfajta időeltolásban invariáns, a moduláció művelete azonban már nem, mivel a vivőfrekvencia fázisa bizonyos időbeli szerkezetet ad meg. Azonban a moduláció is invariáns mindazon transzformációkban, amelyek a vivőfrekvenciás jel periódusának többszörösei.

Wiener szoros kapcsolatot mutatott ki a fizikai eszközök idő transzformációval szembeni invarianciája és a Fourier-elmélet között.[14] Lényegében azt mutatta meg,

---

[13] Ez a híres ergodikus elmélet, vagy inkább ezen elmélet egyik vonatkozása, amelyet, kissé más megfogalmazásban Birkhoff, Neumann és Koopmann már bizonyított, majd ezt követően Wiener, Hopf, Hurewitz és mások általánosítottak. Az ergodikus folyamatoknak meglehetősen kiterjedt irodalma van, s az olvasónak ezen szerzőktől az idézett dolgozatokat ajánljuk a precíz és általános megfogalmazás megismerése érdekében. Pl.: E. Hopf: *Ergodentheorie* („Ergodikus elmélet"), Ergebnisse der Mathematic und Ihrer Grenzgebiete, 5. kötet, *On Causality Statistics and Probability* („A kauzális statisztikáról és a valószínűségről") Journal of Mathematics and Physics, XIII. kötet, 1934. 1. szám, N. Wiener: *The Ergodic Theoren* („Az ergodikus tétel"), Duke Mathematical Journal, 1939., 5. kötet.

[14] A hírközléselmélet sokat köszönhet Wienernek, aki jelentősen hozzájárult a filozófiai és elméleti alapok lerakásához. Klasszikus NDRC jelentése, a *The Imternpolation, Extraplation and Smoothing of Stationary Time Series*. (Stacionárius idősorok interpolációja, extrapolációja és simítása). Wiley, 1949.) adja meg a hírközléselmélet mint statisztikai probléma első világos megfogalmazását, az idősorokkal végzett műveletekre vonatkozó tanulmányt. Ez a mű, bár főképpen a lineáris előrejelzéssel és szűrési problémákkal foglalkozik,

hogy ha egy rendszer lineáris és emellett invariáns is, akkor a Fourier-analízis az a matematikai eszköz, amellyel a problémát meg kell oldani.
Egy adó által előállított jelek, zavaró zaj, illetve egy folytonos forrás által produkált üzenetek (pl. beszéd) megfelelő matematikai leírása függvények együttesének segítségével történhet. A hírközléselmélet, ahogyan azt Wiener hangsúlyozta, nem egyes függvényekkel végrehajtott műveletekkel foglalkozik. Egy hírközlési rendszert nem egyetlen, sajátos beszédfüggvény, vagy még kevésbé egy színuszos jel átvitelére, hanem beszédfüggvények együttesének átvitelére terveznek.

19. Sávhatárolt függvényegyüttesek

Ha egy *f(t)* időfüggvény a *0–W* frekvenciatartományban sávhatárolt, úgy teljesen meghatározható diszkrét, egymástól $\frac{1}{2W}$-ra lévő pontokban felvett ordinátáival. Ez a következő eredmény szerinti módon történhet.[15]

13. TÉTEL:
Legyen *f(t)* olyan, hogy *W*-nél nagyobb frekvenciát nem tartalmaz. Ekkor

$$f(t) = \sum_{-\infty}^{\infty} X_n \frac{\sin \pi(2Wt - n)}{\pi(2Wt - n)} \quad \text{ahol} \quad X_n = f\left(\frac{n}{2W}\right)$$

Az *f(t)* ezen kifejtésében ortogonális függvények összegeként szerepel. Az egyes tagok X_n együtthatói a végtelen dimenziójú „függvénytérben” foghatók fel mint annak koordinátái. Ebben a térben minden függvény egyetlen pontnak, és minden pont egyetlen függvénynek felel meg.
Egy függvényt lényegében akkor tekintünk *T* időtartamra korlátozottnak, ha valamennyi, ezen időintervallumon kívül eső X_n ordinátája zérus. Ez esetben *2TW* kivételével minden koordináta zérus lesz. Így egy *W* frekvenciasávra és *T* időtartamra korlátozott függvény a térben *2TW* dimenziójú pontoknak felel meg.
A *W* sávszélességű és *T* időtartamú függvények részhalmaza e térben egy tartománynak felel meg. Például azok a függvények, amelyek teljes energiája kisebb, vagy egyenlő *E*-vel, egy *2TW* dimenziójú, $r = \sqrt{2WE}$ sugarú gömbben fekvő pontoknak felelnek meg.
Korlátozott időtartamú és sávszélességű függvényegyüttest egy $p(x_1,\ldots,x_n)$ valószínűségi eloszlás segítségével reprezentálhatjuk a megfelelő *n*-dimenziós térben. Ha az együttes időben nem korlátozott, akkor egy adott *T* intervallumhoz tartozó *2TW* koordinátáit úgy tekinthetjük, mint amelyekkel a függvény *T* időtartam alatti részét és a $p(x_1,\ldots,x_n)$ valószínűségi eloszlást kielégítően jellemezni tudjuk, hogy az ilyen hosszúságú időintervallumokra a függvényegyüttes statisztikus szerkezetét megadjuk.

---

mégis fontos forrásmunka a jelen dolgozat szempontjából. Utalhatunk itt még Wiener *Cybernetics* („Kibernetika”) (Wiley, 1948) c. könyvére, amely a hírközlés és vezérlés általános problémáival foglalkozik.
[15] Ennek a tételnek a bizonyításához és a téma további kifejezéséhez lásd a szerző *Communication in the Presence of Noise* („Hírközlés zaj jelenlétében”) című tanulmányát. Proceedings of the Institute of Radio Engineers, 37.k.1.sz. 1949. jan. 10-21.old.

20. Folytonos eloszlás entrópiája

Egy diszkrét $p_1,\ldots,p_n$ valószínűségekből álló halmaz entrópiáját a következőképpen definiáltuk: $H = -\sum p_i \log p_i$

Hasonló módon definiálhatjuk egy folytonos eloszlás entrópiáját a *p(x)* valószínűség-sűrűség függvény segítségével:

$$H = -\int_{-\infty}^{\infty} p(x)\log p(x)dx$$

n-dimenziós $p(x_1,\ldots,x_n)$ eloszlás esetén a következőt írhatjuk:

$$H = -\int \ldots \int p(x_1,\ldots,x_n)\log p(x_1,\ldots,x_n)dx_1 \ldots dx_n$$

Ha két *x* és *y* változónk van (amelyek maguk is lehetnek többdimenziósak), a *p(x,y)* együttes és feltételes entrópiája:

$$H(x,y) = -\int\int p(x,y)\log p(x,y)dxdy$$

$$H_x(y) = -\int\int p(x,y)\log \frac{p(x,y)}{p(x)}dxdy$$

$$H_y(x) = -\int\int p(x,y)\log \frac{p(x,y)}{p(y)}dxdy$$

$$ahol \quad p(x) = \int p(x,y)dy \quad p(y) = \int p(x,y)dx$$

A folytonos eloszlások entrópiája nagyon sok (bár nem minden) tulajdonság tekintetében azonos a diszkrét esetével. Konkrétan a következőket mondhatjuk el:

1. Ha az *x* terében egy adott *v* értékre korlátozódik, akkor *H(x)* a max. érték és *log v*-vel egyenlő, ha *p(x)=1/v* konstans.
2. Bármely két *x,y* változó esetén azt kapjuk, hogy $H(x,y) \le H(x) + H(y)$ amely akkor és csak akkor megy át egyenlőségbe, ha *x* és *y* egymástól függetlenek, azaz *p(x,y)=p(x)p(y)* (eltekintve egy sorozat, zérus valószínűségű pontjainak halmazától).
3. Tekintsük a következő fajtájú, általánosított átlagolási műveletet:
$$p'(y) = \int a(x,y)p(x)dx$$
$$\int a(x,y)dx = \int a(x,y)dy = 1 \quad ahol \quad a(x,y) \ge 0$$
Ezek után a *p'(y)* átlagolt eloszlás entrópiája egyenlő, vagy nagyobb, mint az eredeti *p(x)* eloszlásé.
4. Írhatjuk, hogy $H(x,y) = H(x) + H_x(y) = H(y) + H_y(x) \quad és \quad H_x(y) \le H(y)$
5. Legyen *p(x)* egydimenziós eloszlás. A maximális entrópiát adó *p(x)* formája (feltéve, hogy *x* szórása σ értékű, és rögzített) Gauss-eloszlású. Hogy ezt kimutassuk, maximálnunk kell a függvényt:
$$H(x) = -\int p(x)\log p(x)dx \quad ahol \quad \sigma^2 = \int p(x)x^2dx \quad és \quad 1 = \int p(x)dx$$
A variációszámítás szerint ehhez maximálni kell:
$$\int \left[-p(x)\log p(x) + \lambda p(x)x^2 + \mu p(x)\right]dx$$
Ennek feltétele: $-1 - \log p(x) + \lambda x^2 + \mu = 0$

   és ebből(a konstanst a megszorításoknak megfelelően megválasztva):

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x^2/2\sigma^2)}$$

A fentiekhez hasonlóan *n*-dimenziós térben, feltételezve, hogy $p(x_1,\ldots,x_n)$ másodrendű momentumát A_{ij}-ben rögzítjük:

$$A_{ij} = \int \ldots \int x_i x_j \, p(x_1,\ldots,x_n) dx_1 \ldots dx_n$$

Ezek után a maximális entrópia akkor áll elő (hasonló számítási eljárással), ha $p(x_1,\ldots,x_n)$ az *n*-dimenziós Gauss-eloszlás, A_{ij} másodrendű momentumokkal.

6. Az egydimenziós Gauss-eloszlás entrópiáját, ha a szórás σ, a következő képlet adja meg: $H(x) = \log \sqrt{2\pi e}\sigma$. Ezt tovább kifejtve:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x^2/2\sigma^2)}$$

$$-\log p(x) = \log \sqrt{2\pi}\sigma + \frac{x^2}{2\sigma^2}$$

$$H(x) = -\int p(x) \log p(x) dx = \int p(x) \log \sqrt{2\pi}\sigma \, dx + \int p(x) \frac{x^2}{2\sigma^2} dx =$$

$$= \log \sqrt{2\pi}\sigma + \frac{\sigma^2}{2\sigma^2} = \log \sqrt{2\pi}\sigma + \log \sqrt{e} = \log \sqrt{2\pi e}\sigma$$

Hasonlóképpen, az n-dimenziós Gauss-eloszlás az a_{ij} kvadratikus alakkal kifejezve a következő lesz:

$$p(x_1,\ldots,x_n) = \frac{|a_{ij}|^{\frac{1}{2}}}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} \sum a_{ij} x_i x_j \right)$$

és az entrópiát a $H = \log (2\pi e)^{n/2} |a_{ij}|^{-\frac{1}{2}}$ kifejezésből számíthatjuk, ahol az $|a_{ij}|$ determináns elemei az a_{ij} elemek.

7. Ha *x* egy $p(x)=0 \; x \leq 0$ félegyenesre korlátozódik és *x* első momentuma *a*-val egyenlő:

$$a = \int_0^\infty p(x) x \, dx$$

akkor a maximális entrópia $p(x) = \frac{1}{a} e^{-(x/a)}$ esetén lép fel, és *log ea*-val egyenlő.

8. Van egy lényeges különbség a folytonos és a diszkrét eset entrópiája között. A diszkrét esetben az entrópia abszolút módon méri a valószínűségi változó véletlenszerűségét, míg a folytonos esetben ez a mérés egy koordinátarendszerre vonatkoztatott relatív mértéket ad. Ha a koordinátákat megváltoztatjuk, az entrópia általában meg fog változni. Konkrétan, ha a koordinátákat $y_1,\ldots,y_n$-re változtatjuk meg, az új entrópia:

$$H(y) = \int \ldots \int p(x_1,\ldots,x_n) J\left(\frac{x}{y}\right) dy_1 \ldots dy_n$$

ahol $J\left(\frac{x}{y}\right)$ Jacobi-féle koordináta-transzformáció. A logaritmus kifejezésével és a változók $x_1,\ldots,x_n$-re történő megváltoztatásával kapjuk:

$$H(y) = H(x) - \int \ldots \int p(x_1,\ldots,x_n)\log J\left(\frac{x}{y}\right)dx_1\ldots dx_n$$

Így az új entrópia egyenlő a régi entrópiának a Jacobi-féle transzformációból adódó logaritmussal csökkentett értékével. Az entrópiát a folytonos esetben a véletlenszerűség, egy felvett egységhez viszonyított mértékének tekintjük, amely egység tulajdonképpen olyan koordinátarendszert jelent, amelynek minden egyes kis $dx_1,\ldots,dx_n$ elemét azonos súllyal vesszük figyelembe. Amikor megváltoztatjuk a koordinátarendszert, az új rendszer entrópiája azt a véletlenszerűséget fogja mérni, amelyben azonos $dy_1,\ldots,dy_n$ elemek azonos súllyal szerepelnek.
A koordinátarendszertől való e függés ellenére is az entrópia fogalma a folytonos esetben éppolyan fontos, mint a diszkrét esetben. Ez amiatt van így, mert az információsebesség és a csatornakapacitás leszármaztatott fogalmai két entrópia különbségétől függenek, míg ez a különbség nem függ a koordinátarendszertől, minthogy mind a két tag azonos mértékben változik.
Egy folytonos eloszlás entrópiája negatív is lehet. A méréshatár egy egységnyi intervallumban való egyenletes eloszlásnak megfelelő, tetszés szerinti zérust állít be. Az ennél jobban határolt eloszlás entrópiája kisebb, negatív lesz. A sebességek és a kapacitások azonban mindig nem-negatívak lesznek.

9. A koordináta transzformáció egy sajátos esete a következő lineáris transzformáció:

$$y_j = \sum_i a_{ij} x_i$$

Ez esetben a Jacobi-transzformáció egyszerűen az $\left|a_{ij}\right|^{-1}$ determináns és $H(y) = H(x) + \log\left|a_{ij}\right|$. A koordináták forgatása (vagy bármely más, a transzformációt megörző módszer) során $J=1$ és $H(y)=H(x)$.

21. Egy függvényegyüttes entrópiája

Tekintsünk egy W sávban határolt ergodikus függvényegyüttest. Legyen

$$p(x_1,\ldots,x_n)$$

az $x_1,\ldots,x_n$ amplitúdók n egymást követő mintavételezési pontban vett értékeinek sűrűségfüggvénye. Az együttes entrópiáját mint szabadságfokot a következőképpen definiáljuk:

$$H' = -\lim_{n\to\infty}\frac{1}{n}\int\ldots\int p(x_1,\ldots,x_n)\log p(x_1,\ldots,x_n)dx_1\ldots dx_n$$

Definiálhatjuk a másodpercenkénti entrópiát is ily módon, hogy nem n-nel, hanem az n mintavételezéshez tartozó T másodpercnyi idővel osztunk. Minthogy $n=2TW$, ezért $H=2WH$.
Fehér termikus zaj esetén p Gauss-eloszlású és

$$H' = \log\sqrt{2\pi eN}, \quad H = W\log 2\pi eN$$

Egy adott átlagos N teljesítményre a fehér-zaj a lehetséges maximális entrópiával bír. Ez következik a fentebb említett Gauss-eloszlás maximálási tulajdonságaiból.
Egy folytonos sztochasztikus folyamat entrópiája több, a diszkrét folyamatokéhoz hasonló tulajdonságot mutat. A diszkrét esetben az entrópia a hosszú sorozatok valószínűségének logaritmusára és az ésszerűen valószínű, hosszú sorozatok számára vonatkozott. A folytonos esetben – hasonló módon – az entrópia egy hosszú mintavételi sorozat valószínűségi sűrűségének logaritmusára és a függvénytér ésszerűen nagy valószínűségű térfogatára vonatkozik.
A fentieket pontosítva, ha feltételezzük, hogy $p(x_1,\ldots,x_n)$ minden x_i értéke valamennyi n-re folytonos, akkor (elegendően nagy n esetében)

$$\left|\frac{\log p}{n} - H'\right| \langle \varepsilon$$

$(x_1,\ldots,x_n)$ bármilyen választása mellett, eltekintve egy olyan sorozattól, amelynek teljes valószínűsége kisebb, mint δ, ha δ és ε tetszés szerinti kis számok. Ez következik az ergodikus tulajdonságból, ha a teret nagyszámú apró sejtre osztjuk fel.
H-nak a térfogathoz való viszonyát a következőképpen adhatjuk meg: ugyanazon feltételezések mellett tekintsük az n-dimenziós teret, amely $p(x_1,\ldots,x_n)$-nek felel meg. Legyen $V_n(q)$ a legkisebb térfogat e térben, amely belsejében q teljes valószínűséget tartalmaz. Ekkor

$$\lim_{n\to\infty} \frac{\log V_n(q)}{n} = H'$$

feltéve, hogy q nem egyenlő 0-val vagy 1-gyel.
Ezek az eredmények azt mutatják, hogy nagy n értékre igen jól meghatározott, nagy valószínűségű térfogat létezik (legalábbis logaritmikus értelemben) és hogy ezen térfogaton belül a valószínűségi sűrűség – ismét logaritmikus értelemben véve – viszonylag egyenletes.
A fehér-zaj esetében az eloszlási függvényt a következő összefüggés adja meg:

$$p(x_1,\ldots,x_n) = \frac{1}{(2\pi N)^{n/2}} \exp - \frac{1}{2N} \sum x_i^2$$

Mivel ez csak $\sum x_i^2$ -től függ, az azonos valószínűség-sűrűségű felületek gömbök és az egész eloszlás gömbszimmetrikus. A nagyvalószínűségű tartomány a $\sqrt{nN}$ sugarú gömb belsejében van. Amint $n \to \infty$, úgy annak a valószínűsége, hogy egy pont a $\sqrt{n(N+\varepsilon)}$ sugarú gömbön kívül esik, a zérushoz közelít, és a gömb térfogata logaritmusának és *1/n*-nek a szorzata $\sqrt{2\pi e N}$ -hez tart.
A folytonos esetben H entrópia helyett kényelmesebb azzal a leszármazott mennyiséggel dolgozni, amelyet entrópia teljesítménynek fogunk nevezni. Ezt mint az eredeti együttes sávjára határolt és azzal azonos entrópiájú fehér-zaj teljesítményét definiálhatjuk. Más szavakkal: ha egy együttes entrópiája H', akkor entrópia teljesítménye:

$$N_1 = \frac{1}{2\pi e} \exp 2H'$$

Geometriai szemléletben ez a mennyiség az azonos térfogatú gömb sugarának négyzetével méri a nagy valószínűségű térfogatot. Minthogy a fehér-zajnak egy adott teljesítményre vonatkoztatva maximális az entrópiája, így bármely zaj entrópia teljesítménye kisebb, vagy egyenlő a tényleges teljesítményével.

22. Entrópia veszteség lineáris szűrőkben

14. TÉTEL:
Ha egy W sávban szabadsági fokonként H_1 entrópiával rendelkező együttest egy $Y(f)$ karakterisztikájú szűrőn bocsátunk át, a kimeneti együttes az alábbi entrópiával rendelkezik:

$$H_2 = H_1 + \frac{1}{W}\int_W \log|Y(f)|^2 df$$

A szűrő működése lényegében egy lineáris koordináta-transzformáció. Ha az eredeti koordinátarendszerben a különböző frekvencia összetevőkre gondolunk, az új frekvencia-komponensek a régieknek egyszerűen többszörösei lesznek. Így a koordináta-transzformáció mátrixának átlóit alapjában véve ezen koordináták adják. A transzformáció Jacobi-féle értéke ezek után (n szinuszos és n koszinuszos összetevő esetére):

$$J = \prod_{i=1}^{n} |Y(f_i)|^2$$

ahol f_i értékei a W sávban egyenlő távolságokban vannak elosztva. Ez határértékben az

$$\exp \frac{1}{W}\int_W \log|Y(f)|^2 df$$

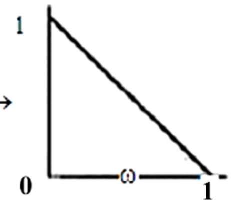
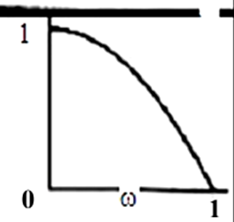
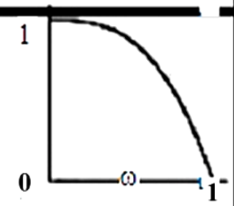
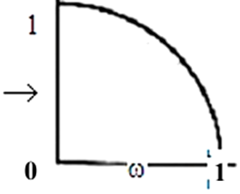
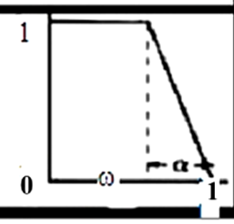
kifejezésbe megy át. Mivel J konstans, átlaga ugyanaz a mennyiség és a tételt a koordinátaváltozással alkalmazva az entrópia változására, ezt az eredményt kapjuk. Ezt megfogalmazhatjuk az entrópia teljesítmény segítségével is. Így, ha az első sorozat entrópia teljesítménye N_1, akkor a másodiké

$$N_1 \exp \frac{1}{W}\int_W \log|Y(f)|^2 df$$

A végső entrópia teljesítményt az eredeti entrópia teljesítménynek és a szűrő-nyereség mértani középarányosának a szorzata adja meg. Ha a nyereséget decibelben mérjük, úgy a kimeneti entrópia teljesítmény növekedni fog a W-sávban, a decibelben kifejezett nyereség számtani középarányosával.

Az I. táblázatban néhány ideális nyereség karakterisztikára adtuk meg az entrópia teljesítmény számított és dB-ben kifejezett értékét. A $W=2\pi$ esetére (miközben a fázist 0-nak tételeztük fel) ugyancsak feltüntettük ezeknek a szűrőknek az impulzus átviteli válaszfüggvényét.
Ezekből az eredményekből számos más esetre levezethető az entrópia veszteség. Például: az első esetre megadott $1/e^2$ entrópia teljesítmény tényező minden olyan nyereség-karakterisztikára is vonatkozik, amelyet l-ω-ból kapunk, oly módon, hogy az ω tengely transzformációját fenntartjuk. Egy lineárisan növekvő $G(\omega)=\omega$ nyereségfüggvény (másként: „fűrészfog” – karakterisztika) 0 és 1 között ugyanezt az entrópia veszteséget mutatja. A reciproknyereség a reciproktényezővel rendelkezik: így $1/\omega$ tényezője e^2. A nyereséget bármilyen hatványra emelve ez a tényezőnek ugyanezen hatványra emelését eredményezi.

I. Táblázat

Entrópia teljesítmény nyereség karakterisztika	**Entrópia teljesít-mény faktor**	**Entrópia teljesítmény nyereség (decibelben)**	**Impulzus válasz**
$1-\omega \rightarrow\rightarrow$ 	$\frac{1}{e^2}$	**-8.69**	$\frac{\sin^2(t/2)}{t^2/2}$
$1-\omega^2$ 	$\left(\frac{2}{e}\right)^4$	**-5.33**	$2\left[\frac{\sin t}{t^3}-\frac{\cos t}{t^2}\right]$
$1-\omega^3$ 	.411	**-3.87**	$6\left[\frac{\cos t-1}{t^4}-\frac{\cos t}{2t^2}+\frac{\sin t}{t^3}\right]$
$\sqrt{1-\omega^2} \rightarrow$ 	$\left(\frac{2}{e}\right)$	**-2.67**	$\frac{\pi}{2}\frac{J_1(t)}{t}$
	$\overline{2\alpha}$	**-8.69α**	$\frac{1}{\alpha t^2}\left[\cos(1-\alpha)t-\cos t\right]$

23. Két függvényegyüttes összegének entrópiája

Ha két függvényegyüttesünk van, $f_\alpha(t)$ és $g_\beta(t)$ "összeadásuk" révén egy új együttest alkothatunk. Tételezzük fel, hogy az első együttes valószínűségi sűrűségfüggvénye $p(x_1,\ldots,x_n)$, míg a másodiké $q(x_1,\ldots,x_n)$. Ezek után az összeg sűrűségfüggvényét a következő konvulúció adja:

$$r(x_1,\ldots,x_n)=\int\ldots\int p(y_1,\ldots,y_n)q(x_1-y_1,\ldots,x_n-y_n)dy_1\ldots dy_n$$

Fizikailag ez az eredeti függvényegyüttesek által képviselt zajok vagy jelek összeadásának felel meg. A következő eredményt a 6. Függelékben vezettük be.

15. TÉTEL:
Legyen két együttes átlagos teljesítménye N_1 , N_2 és legyen entrópia teljesítményük $\overline{N_1}, \overline{N_2}$.
Ekkor az összeg entrópia teljesítményére $\overline{N_3}$ a következő korlátok állnak fenn:

$$\overline{N_1} + \overline{N_2} \leq \overline{N_3} \leq N_1 + N_2$$

A Gauss-eloszlású fehér-zajnak az a sajátságos tulajdonsága, hogy bármely más, hozzáadódó zaj- vagy jel-együttest a fehér-zaj és a jel teljesítmény összegével közel egyenlő eredő entrópia teljesítménnyel tud elnyelni, feltéve, hogy a jel teljesítménye –bizonyos értelemben – kicsiny a zajéhoz viszonyítva. (A jel teljesítményét itt a rendesen zérus értékű átlagos jelértékből mérjük.)
Tekintsük az ezen *n*-dimenziójú együttesekhez tartozó függvényteret. A fehér-zaj e térben a gömbi Gauss-eloszlásnak felel meg. Az együttes egy másik, nem szükségképpen Gauss-féle, ill. gömbi eloszlásnak felel meg. Legyenek ezen eloszlás súlypontja körüli másodrendű momentumai a_{ij} . Ha $p(x_1,\ldots,x_n)$ a sűrűség eloszlás függvény, akkor

$$a_{ij} = \int \ldots \int p(x_i - \alpha_i)(x_j - \alpha_j) dx_1 \ldots dx_n$$

ahol α_i a súlypont koordinátája. Mivel a_{ij} pozitív definit, kvadratikus alak, így koordinátarendszerünket úgy forgathatjuk el, hogy illeszkedjék a fő irányokkal. Ezután b_{ii} diagonál alakra redukáljuk. Megkövetelhetjük, hogy N-hez (a gömbi eloszlás sugarának négyzetéhez) képest minden egyes b_{ii} kicsiny legyen.
Ez esetben a zaj és a jel konvolúciója megközelítőleg olyan Gauss-eloszlást hoz létre, amelynek megfelelő kvadratikus alakja: $N+b_{ii}$

Ennek az eloszlásnak az entrópia teljesítménye: $\left[\prod (N + b_{ii})\right]^{1/n}$
vagy közelítőleg

$$= \left[(N)^n + \sum b_{ii} (N)^{n-1}\right]^{1/n} = N + \frac{1}{n} \sum b_{ii}$$

A második tag a jel, míg az első a zaj teljesítménye.

IV. A folytonos csatorna

24. Egy folytonos csatorna kapacitása

Egy folytonos csatorna bemeneti vagy átvitt jelei az időnek egy meghatározott készlethez tartozó folytonos *f(t)* függvényei, míg a kimeneti- vagy vett jelek ezek zavart változatai lesznek. Mi csak azt az esetet vizsgáljuk, amelyben mind az adott, mind a vett jelek egy adott *W* sávban határoltak. Ennélfogva e jeleket *T* időre vonatkoztatva a *2TW* számmal lehet meghatározni, statisztikus szerkezetük pedig véges dimenziójú eloszlásfüggvények segítségével írható le. Így az átvitt jel statisztikáját a

$$P(x_1,\ldots,x_n)=P(x)$$

Míg a zajét a következő feltételes valószínűség eloszlás határozza meg:

$$P_{x_1,\ldots,x_n}(y_1,\ldots,y_n) = P_x(y)$$

Egy folytonos csatorna esetében az információátvitel sebességét a diszkrét esethez hasonló módon definiáljuk, mégpedig az $R=H(x)-H_y(x)$ kifejezés segítségével, ahol $H(x)$ a bemenet entrópiája és $H_y(x)$ az ekvivokáció. A csatornakapacitást mint R maximumát definiáljuk, miközben a bemenetet valamennyi lehetséges sorozat szerint változtatjuk. Ez azt jelenti, hogy egy véges dimenziójú közelítésben $P(x)=P(x_1,\ldots,x_n)$ kell változtatnunk és maximálni az alábbi kifejezést:

$$-\int P(x)\log P(x)dx+\int\int P(x,y)\log\frac{P(x,y)}{P(y)}dxdy$$

Ezt átírhatjuk a következő alakba

$$\int\int P(x,y)\log\frac{P(x,y)}{P(x)P(y)}dxdy$$

kihasználva azt a tényt, hogy

$$\int\int P(x,y)\log P(x)dxdy=\int P(x)\log P(x)dx$$

Így a csatornakapacitást a következőképpen fejezhetjük ki:

$$C=\lim_{T\to\infty}Max\frac{1}{T}\int\int P(x,y)\log\frac{P(x,y)}{P(x)P(y)}dxdy$$

Ebben az alakban nyilvánvaló, hogy R és C független a koordináta-rendszertől, mivel a $\log\frac{P(x,y)}{P(x)P(y)}$ számlálóját és nevezőjét az x és y bármilyen egy az egyben történő transzformációja során ugyanazokkal a tényezőkkel szorozzuk. C-nek ez az integrál-kifejezése általánosabb, mint $H(x)-H_y(x)$.
Megfelelően értelmezve (ld. 7. függelékben) ez mindig létező, míg $H(x)–H_y(x)$ egyes esetekben felveheti a ∞-∞ határozatlan alakot. Ez például akkor áll elő, ha x-et n-nél kevesebb dimenziójú felületre korlátozzuk az n-dimenziós közelítése során.
Ha $H(x)–H_y(x)$ kiszámításához a kettes alapú logaritmust használjuk, akkor C a másodpercenként a csatornán –tetszés szerinti kis ekvivokáció mellett – átvihető maximális számú bináris digitet adja meg csakúgy, mint a diszkrét esetben. Ezt fizikailag úgy állíthatjuk be, hogy a jelteret nagyszámú kis sejtre osztjuk fel, amelyek elegendően kicsik ahhoz, hogy az x jel y pontba tolódásának $P_x(y)$ valószínűségi sűrűsége lényegében véve konstans a sejten belül (mind x-re, mind y-ra). Ha e sejteket mint különálló pontokat tekintjük, a helyzet lényegileg ugyanaz, mintha diszkrét csatornánk volna, s az ott bemutatott bizonyítások itt is alkalmazhatók. Fizikai szempontból azonban világos, hogy a térfogatnak egyedi pontokká történő kvantálása semmilyen gyakorlati szituációban nem változtathatja meg lényegesen a végső választ, feltéve, hogy a térrészecskék elegendően kicsinyek. Ilyenformán a kapacitás a diszkrét felosztás során kapott kapacitás korlátja lesz és ez éppen a fentebb definiált folytonos kapacitás.

Matematikai oldalról mindenekelőtt meg lehet mutatni (ld. a 7. Függelékben), hogy ha u az üzenet, x a jel, y a vett jel (amelyet zaj zavar) és v a visszaállított üzenet, akkor

$$H(x)-H_y(x)\geq H(u)-H_v(u)$$

Tekintet nélkül arra, milyen műveleteket hajtunk végre u-n, hogy x-et, illetve y-t megkapjuk. Így nem számít az, hogyan kódoljuk a biteket a jel előállítása, illetve az, hogy dekódoljuk a vett jelet az üzenet visszaállítása érdekében, mert a bináris digitek diszkrét sebessége nem haladja meg az általunk definiált csatornakapacitást. Másfelől lehetséges – nagyon általános feltételek közepette – találni olyan kódolási rendszert, amely lehetővé teszi bináris digitek *C*-sebességű adását tetszés szerinti kis ekvivokáció vagy hibagyakoriság mellett. Ez igaz például akkor, ha a jelfüggvények terének véges dimenziójú közelítése során $P(x,y)$ mind x-ben, mind y-ban folytonos, kivéve a zérus valószínűségű pontok készletét.

Egy fontos speciális eset az, amikor a jelhez (attól – valószínűségi értelemben – független) zaj adódik. Ekkor $P_x(y)$ csupán az $n = (y\text{-}x)$ vektor–különbség függvénye $P_x(y) = Q(y - x)$ és a zajhoz egy határozott entrópiát rendelhetünk (ez a jel statisztikájától független lesz), mégpedig a $Q(n)$ eloszlás entrópiáját. Ezt az entrópiát $H(n)$-nel fogjuk jelölni.

16. TÉTEL:
Ha a jel és a zaj függetlenek egymástól és a vett jel a zaj és az adott jel összege, akkor az átvitel sebessége $R = H(y) - H(n)$, azaz a vett jel entrópiája mínusz a zaj entrópiája. A csatornakapacitás

$$C = \underset{P(x)}{Max}\, H(y) - H(n)$$

Ezek után – mivel $y = x + n$ – kapjuk: $H(x,y) = (H(x,n)$,

Bővítve a baloldalt és felhasználva a tényt, hogy x és n függetlenek:

$$H(y) + H_y(x) = H(x) + H(n)$$

Innen: $R = H(x) - H_y(x) = H(y) - H(n)$

Mivel $H(n)$ független $P(x)$-től, R maximálásához $H(y)$-t azaz a vett jel entrópiáját kell maximálni. Ha az adott jelek sorozatára nézve bizonyos megszorítások állnak fenn, úgy a vett jel entrópiáját ezen megszorítások keretei között kell maximálni.

25. A csatornakapacitás az átlagteljesítmény korlátozása esetén

A 16. tétel egyik egyszerű alkalmazásához jutunk, ha a zaj fehér termikus zaj, és az átvitt jelek teljesítménye egy adott P átlagos teljesítményére korlátozódik. Ebben az esetben a vett jelek $P+N$ átlagos teljesítménnyel rendelkeznek, ahol N az átlagos zajteljesítmény. A vett jelekre nézve az entrópia akkor a maximális, amikor azok ugyancsak fehér – zaj sorozatot alkotnak, minthogy a $P+N$ teljesítményre ez a lehetséges legnagyobb entrópia, amelyet az adott jelek sorozatának megfelelő megválasztásával – nevezetesen ha azok P teljesítményű fehér-zaj sorozatot alkotnak – kapunk. A vett sorozat másodpercenkénti entrópiája ezek után $H(y)=W\log 2\pi e(P+N)$ és a zaj entrópiája $H(n)=W\log 2\pi eN$

A csatornakapacitás: $C = H(y) - H(n) = W \log \dfrac{P+N}{N}$

A fentieket összegezve kapjuk:

17. TÉTEL:
Egy N teljesítményű fehér-zaj által zavart, W sávszélességű csatorna kapacitását, ha az átlagos adó-teljesítményt P-re korlátozzuk, a következő kifejezésével kapjuk:

$$C = W \log \frac{P+N}{N}$$

Ez azt jelenti, hogy elegendően fejlett kódolási rendszerek segítségével másodpercenkénti $W\log_2\frac{P+N}{N}$ sebességgel vihetünk át bináris számjegyeket, tetszés szerinti kis hibagyakoriság mellett. Ennél nagyobb sebességű átvitel semmiféle kódolási rendszer segítségével nem lehetséges anélkül, hogy egy meghatározott, pozitív hibaarány ne álljon elő.
Az átviteli sebesség e határértékének megközelítése érdekében – statisztikai tulajdonságok szempontjából – az átvitt jeleknek a fehér-zajhoz kell közelíteniük.[16] Az ideális sebességet megközelítő rendszer a következőképpen írható le: állítsunk elő $M = 2^s$ számú, egyenként T időtartamú fehér-zaj mintát. Ezekhez rendeljünk hozzá 0-tól (M-1)-ig bináris számokat. Az adóban az üzenetsorozatokat s számú csoportra bontjuk és minden egyes csoportra jelként az annak megfelelő zajmintát adjuk be. A vevőben az M mintát ismerjük és a – zaj által zavart – éppen vett jelet ezek mindegyikével összehasonlítjuk. Leadott jelnek azt a mintát választjuk, amely a vett jeltől való legkisebb négyzetes középértékű eltérést mutatja, s ezek után visszaállítjuk a megfelelő bináris számot. Ez az eljárás az (a posteriori) legvalószínűbb jel kiválasztására épül. Az alkalmazott zajminták M száma a megengedett ε hibagyakoriságtól függ, azonban csaknem minden mintavételezésre a

$$\lim_{\varepsilon\to 0}\lim_{T\to\infty}\frac{\log M(\varepsilon,T)}{T}=W\log\frac{P+N}{N}$$

kifejezés érvényes, így közömbös, hogy ε értékét milyen kicsire választjuk, mert T-t elegendően nagynak véve, a T idő alatt leadott bitek számával tetszés szerint megközelíthetjük $TW\log\frac{P+N}{N}$ értékét.

A $C=W\log\frac{P+N}{N}$-hez hasonló képleteket egymástól függetlenül több más szerző is levezetett fehér-zaj esetére, jóllehet az értelmezések némiképp eltérnek egymástól.

Megemlítjük ebben a vonatkozásban N. Wiener[17], V.G. Tuller[18] és H. Sullivan munkásságát.
Egy tetszés szerinti (azaz nem szükségképpen fehér termikus) zavaró zaj esetében a C csatornakapacitás meghatározása során felmerült maximális probléma nem látszik explicite megoldhatónak, azonban C-re az N átlagos zajteljesítmény és az N_1 zaj entrópia-teljesítmény szempontjából megadhatók alsó és felső korlátok. A legtöbb gyakorlati esetben ezek a korlátok elegendően közel vannak egymáshoz ahhoz, hogy a probléma kielégítő megoldását szolgáltassák.

18. TÉTEL:
Egy tetszés szerinti zaj által zavart, W sávszélességű csatorna kapacitása az alábbi egyenlőtlenségek szerint korlátos:

$$W\log\frac{P+N_1}{N_1}\le C\le W\log\frac{P+N}{N_1}$$

---

[16] A fehér-zaj ezen, valamint egyéb tulajdonságait geometriai szempontból a *Communikation in the Presence of Noise* („Hírközlés a zaj jelenlétében") című idézett mű tárgyalja.
[17] „Cybernetics" idézett mű
[18] „Theoretical Limatations on the Rate of Transmission of Information" („Az információátvitel sebességének elméleti korlátai"). Proceding of the Institute of Radio Engineers, 1949. május, 37. k., 5.sz. 468-478. oldal.

ahol
P = átlagos adóteljesítmény,
N = átlagos zajteljesítmény,
N_1 = a zaj entrópia teljesítménye.

A zavart jelek átlagos teljesítménye itt ismét $P+N$ lesz. E teljesítmény entrópiája akkor lenne maximális, ha a vett jel fehér-zaj volna, s ekkor értéke $W\log 2\pi e(P+N)$ venne fel. Ezt nem lehetséges elérni, azaz esetleg nincs egy olyan átviteli jelkészlet sem, amely zavaró zajhoz hozzáadódva a vevőben fehér termikus zajt hozna létre, azonban ez legalább felső határt szab $H(y)$-ra, ennélfogva:

$$C = MaxH(y) - H(n) \leq W\log 2\pi e(P+N) - W\log 2\pi e N_1$$

Ez a tétel által megszabott felső határ. Az alsó határt úgy kaphatjuk meg, hogy az adott jelet P teljesítményű fehér-zajnak választva tekintjük az átviteli sebességet. Ez esetben a vett jel entrópia teljesítménye legalább olyan nagy kell, hogy legyen, mint a $P+N_1$ teljesítményű fehér-zajé, minthogy a 18. Tételben bebizonyítottuk, hogy két sorozat összegének entrópia teljesítménye nagyobb vagy egyenlő az egyes entrópia teljesítmények összegénél, így

$$MaxH(y) \geq W\log 2\pi e(P+N_1)$$

és

$$C \geq W\log 2\pi e(P+N_1) - W\log 2\pi e N_1 = W\log\frac{P+N_1}{N_1}$$

P növelésével a 18. tételben szereplő alsó és felső határ egymáshoz közelít, oly módon, hogy aszimptotikus sebességként a $W\log\frac{P+N}{N_1}$ értéket kapjuk.

Ha maga a zaj fehér, azaz $N=N_1$ és az eredmény a korábban bebizonyított képletre egyszerűsödik

$$C = W\log\left(1+\frac{P}{N}\right)$$

Ha a zaj Gauss-eloszlású, de spektruma nem szükségképpen egyenletes, úgy N_1 a W sávban található különféle frekvenciákra számított zajteljesítmény mértani középértéke. Így

$$N_1 = \exp\frac{1}{W}\int_W \log N(f)df$$

ahol $N(f)$ az f frekvencián mért zajteljesítmény.

19. TÉTEL:
Ha egy adott P adóteljesítményre a kapacitást úgy választjuk meg, hogy

$$C = W\log\frac{P+N-\eta}{N_1}$$

akkor η a P növekedtével monoton csökken és határértékben 0-hoz tart.

Tételezzük fel, hogy egy adott P_1 teljesítményre a csatornakapacitás

$$W\log\frac{P_1+N-\eta_1}{N_1}$$

Ez azt jelenti, hogy a legjobb – mondjuk *p(x)* – jel-eloszlás, a *q(x)* zaj-eloszláshoz adódva $(P_1+N-\eta_1)$ entrópia-teljesítményű *r(y)* vett eloszlást ad. Növeljük a teljesítményt $P_1+\Delta P$ értékre azáltal, hogy ΔP teljesítményű fehér-zajt adunk a jelhez. Ekkor a vett jel entrópiája legalább

$$H(y) = W \log 2\pi e(P_1 + N - \eta_1 + \Delta P)$$

lesz, ha az összeg minimális entrópiájának tételét alkalmazzuk. Ebből – miután elérhetjük a jelzett *H* értéket – a maximáló eloszlás entrópiájának legalább ekkorának, η-nak pedig monoton csökkenőnek kell lennie. Annak bemutatása érdekében, hogy $\eta \to 0 \quad a\min t \quad P \to \infty$, tekintsünk egy jelet, amely *P* nagy teljesítményű fehér-zaj. Ekkor, bármilyen is legyen a zavaró zaj, - ha *P* elegendően nagy – a vett jel közelítőleg fehér-zaj lesz, abban az értelemben, hogy entrópia teljesítménye a *P + N* értéket fogja közelíteni.

26. A csatornakapacitás csúcsteljesítmény korlátozás esetén

Egyes alkalmazásokban az adónak nem az átlagos teljesítménye, hanem a pillanatnyi csúcsteljesítménye van korlátozva. Ez esetben a csatornakapacitás kiszámításának problémája – az átvitt szimbólumok készletének variálásával – a *H(y) – H(n)* maximálását jelenti, azon megszorításnak a figyelembevételével, hogy a készletben lévő valamennyi *f(t)* függvény kisebb, vagy egyenlő kell, hogy legyen $\sqrt{S}$-sel, minden *t*-re. Egy ilyenfajta megszorítás matematikailag nem dolgozható ki olyan jól, mint az átlagteljesítmény korlátozása esetén. A legtöbb, amit erre az esetre kaptunk, egy minden *S/N*-re érvényes alsó korlát, egy „aszimptotikus” felső korlát (amely nagy *S/N*-ekre érvényes) és kis *S/N* értékre *C* egy aszimptotikus értéke.

20. TÉTEL:
Egy *N* teljesítményű termikus fehér-zaj által zavart *W* sávszélességű csatorna *C* kapacitásának korlátját a

$$C \geq W \log \frac{2}{\pi e^3} \frac{S}{N}$$

kifejezés adja meg, ahol *S* az adó megengedett csúcsteljesítménye. Ha *S/N* elegendően nagy

$$C \leq W \log \frac{\frac{2}{\pi e} S + N}{N}(1+\varepsilon)$$

ahol ε tetszés szerint kicsiny. Amint $\frac{S}{N} \to 0$ (és feltéve, hogy a *W* sáv *0*-nál kezdődik)

$$\frac{C}{W \log\left(1 + \frac{S}{N}\right)} \to 1$$

Szeretnénk maximálni a vett jel entrópiáját. Ha *S/N* nagy, ez hamarosan bekövetkezik, amint az átvitt sorozat entrópiáját maximáljuk.
Az aszimptotikus felső korlátot a sorozatra vonatkozó megszorító feltételek lazításával kapjuk. Tételezzük fel, hogy a teljesítményt nem minden pillanatban, csak a mintavételezési

pontokon korlátozzuk *S*-re. Az átvitt együttes maximális entrópiája ezen enyhébb feltételek közepette bizonyosan nagyobb, vagy egyenlő lesz az eredeti feltételek szerintivel. Ez az átalakított probléma könnyen megoldható. A maximális entrópia akkor lép fel, ha a különböző minták függetlenek egymástól és eloszlási függvényük $-\sqrt{S}$ *és* $+\sqrt{S}$ között konstans. Az entrópiát a *Wlog4S* kifejezésből számíthatjuk. Ezek után a vett jel entrópiája kisebb lesz, mint

$$W\log(4S+2\pi eN)(1+\varepsilon)$$

ahol $\varepsilon \to 0$ miközben $\dfrac{S}{N} \to \infty$ és a csatornakapacitást a fehér-zaj entrópiájának $W\log 2\pi eN$ levonásával kapjuk:

$$W\log(4S+2\pi eN)(1+\varepsilon) - W\log(2\pi eN) = W\log\frac{\frac{2}{\pi e}S+N}{N}(1+\varepsilon)$$

Ez a csatornakapacitás keresett felső korlátja.

Az alsó korlát kiszámításához tekintsük ugyanezt a függvényegyüttest. Bocsássuk át ezeket egy háromszög átviteli karakterisztikájú ideális szűrőn. A *0* frekvenciánál mért nyereség legyen egységnyi és csökkenjen lineárisan a *W* frekvenciáig, ahol legyen *0*. Először megmutatjuk, hogy a szűrő kimeneti függvényei mindig (és nemcsak a mintavételezési pontban) korlátozzák az *S* csúcsteljesítményt. Mindenekelőtt megjegyezzük, hogy a szűrűbe beadott $\dfrac{\sin 2\pi Wt}{2\pi Wt}$ impulzus az $\dfrac{1}{2}\dfrac{\sin^2 \pi Wt}{(\pi Wt)^2}$ választ állítja elő a kimeneten. Ez a függvény sosem negatív. Általános esetben a bemeneti függvényt felfoghatjuk az $a\dfrac{\sin 2\pi Wt}{2\pi Wt}$ függvények egy sorának összegeként, ahol a minta amplitúdója (*a*), nem nagyobb mint $\sqrt{S}$. Ennélfogva a kimenet a fenti nem-negatív, eltolt függvények összege, ugyanazokkal az együtthatókkal. Miután ezek a függvények nem-negatívak, a legnagyobb pozitív értéket bármely *t*-re akkor kapjuk, amikor valamennyi *a* együttható maximális pozitív értéket – azaz $\sqrt{S}$ – veszi fel. Ebben az esetben a bemeneti függvény egy $\sqrt{S}$ amplitúdójú konstans volt és mivel a szűrő egyenáramú nyeresége egységnyi, a kimenet ugyanaz. Ennélfogva a kimeneti sorozat csúcsteljesítménye: *S*.

A kimeneti együttes entrópiáját a bementi együttesből számíthatjuk, az ilyen esettel foglalkozó tétel felhasználásával. A kimeneti entrópia egyenlő a bemeneti entrópia és a szűrő nyeresége mértani középarányosának összegével:

$$\int_0^W \log G^2 df = \int_0^W \log\left(\frac{W-f}{W}\right)^2 df = -2W$$

Így a kimenet entrópiája:

$$W\log 4S - 2W = W\log\frac{4S}{e^2}$$

és a csatorna kapacitása nagyobb, mint

$$W\log\frac{2}{\pi e^3}\frac{S}{N}$$

Most szeretnénk megmutatni, hogy kis *S/N*-re (azaz a jel csúcsteljesítménye és a fehér-zaj átlagos teljesítménye közötti arány kis értéke mellett) a csatornakapacitás közelítőleg

$$C = W\log\left(1+\frac{S}{N}\right)$$

Pontosabban fogalmazva $\frac{C}{W\log\left(1+\frac{S}{N}\right)} \to 1$ amint $\frac{S}{N} \to 0$. Mivel a P átlagos jelteljesítmény kisebb, vagy egyenlő az S csúcsértékkel, így minden *S/N*-re következik, hogy

$$C \le W\log\left(1+\frac{P}{N}\right) \le W\log\left(1+\frac{S}{N}\right)$$

Ennélfogva, ha tudunk találni olyan függvényegyüttest, amely megfelel egy, közel $W\log\left(1+\frac{S}{N}\right)$ sebességnek, és amely W – sávszélességre és P csúcsteljesítményre határolt, akkor az eredményt bizonyítottuk. Tekintsük a következő típusú függvényegyüttest. A t minták egy sorozata azonos értékkel, akár $+\sqrt{S}$-sel, akár $-\sqrt{S}$-sel rendelkezik, majd ezt követően a következő t minták értéke azonos, ... Egy sorozatra az értéket véletlenszerűen választjuk, ahol ½ valószínűséggel kerülhet sorra. Ha ezt a sorozatot háromszög nyereségkarakterisztikájú (egyenáramra egységnyi nyereséget mutató) szűrőn bocsátjuk át, a kimenet $\pm S$-re lesz csúcs limitálva. Ezen túlmenően az átlagteljesítmény csaknem S-sel lesz egyenlő és t-t elegendően nagyra választva elérhető, hogy megközelítse ezt az értéket. Ennek, valamint a termikus zajnak az összegére vonatkozó entrópiát úgy kaphatjuk meg, ha alkalmazzuk a zaj és a kis jel összegének tételét. Ez a tétel érvényes, ha $\sqrt{t}\,\frac{S}{N}$ elegendően kicsi. Ez úgy érhető el, ha *S/N*-t elegendően kicsire választjuk (miután t-t már megválasztottuk). Az entrópia teljesítmény tetszés szerint közelíti az $S+N$ értéket és ennélfogva az átviteli sebességet is tetszés szerint szorosan közelíthetjük a $W\log\left(\frac{S+N}{N}\right)$ értékhez.

V. Folytonos forrás sebessége

27. Hűségkiértékelési függvények

Diszkrét információforrás esetén meg tudunk határozni egy határozott információ-előállítási sebességet, azaz az alapul szolgáló sztochasztikus folyamat entrópiáját. Folytonos információforrás esetén a helyzet jelentősen bonyolultabb. Először is egy folytonosan változó mennyiség végtelen számú értéket vehet fel és ennélfogva az egzakt leíráshoz végtelen számú bináris digitet igényel. Ez azt jelenti, hogy egy folytonos forrás kimenetének a vételi ponton egzakt visszaállítással történő átvitele érdekében általában végtelen kapacitású (bit/s) csatornára van szükség. Mivel – normál körülmények között – a csatornáknak bizonyos mértékű zajuk van és ennélfogva kapacitásuk véges, az egzakt átvitel lehetetlen.

A lényegi kérdés azonban nem ez. A gyakorlatban ugyanis, ha folytonos forrással rendelkezünk, nem az egzakt, hanem egy adott tűréshatáron belüli átvitel érdekel bennünket. A kérdés az, vajon hozzá tudunk-e rendelni egy adott átviteli sebességet egy folytonos forráshoz akkor, ha csupán egy adott, megfelelő módon mért visszaállítási hűséget követelünk meg?

Természetes, hogy a hűségi követelmények fokozásával az átviteli sebesség is növekedni fog. Meg fogjuk mutatni, hogy – nagyon általános esetekben – definiálhatunk ilyen sebességet, amely azzal a tulajdonsággal rendelkezik, hogy a megfelelően kódolt információt átviszi egy olyan csatornán, amelynek kapacitása ezzel a sebességgel egyenlő, a hűségi követelmények kielégítése mellett. Ennél kisebb kapacitású csatorna nem felel meg.

Elsőként meg kell adni az átviteli hűség fogalmának általános matematikai megfogalmazását. Tekintsünk egy hosszú –mondjuk T – időtartamú üzenetekből álló készletet. A forrást a hozzárendelt térben azzal a $P(x)$ valószínűségi sűrűséggel írjuk le, amely szerint a forrás a szóban forgó üzenetet kiválasztja. Egy adott hírközlési rendszert (a külső szemlélő szempontjából) az a $P_x(y)$ feltételes valószínűség írja le, hogy ha a forrás az x üzenetet állította elő, akkor a vételi pontban a visszaállított üzenet y lesz. A rendszert egészében véve (beleértve a forrást és az átviteli rendszert is) az a $P(x,y)$ valószínűségi függvény jellemzi, amely szerint az x üzenethez az y végső kimenet tartozik. Ha ezt a függvényt ismerjük, akkor a hűség szempontjából ismerjük a rendszer összes jellemzőjét. A hűség bármilyen értékelése matematikailag egy, ezen $P(x,y)$ függvénnyel végrehajtott műveletnek kell, hogy megfeleljen. Ez a művelet legalábbis a rendszerek egyszerű rendezési tulajdonságaival kell, hogy rendelkezzék, azaz két $P_1(x,y)$ és $P_2(x,y)$ által jellemzett rendszerről meg kell tudnia mondani, hogy – hűségi követelményeinknek megfelelően – akár (1) az elsőnek nagyobb a hűsége, akár (2) a másodiknak nagyobb a hűsége, akár (3) a két rendszer hűsége egyforma. Ez azt jelenti, hogy a hűségkritériumot egy, számszerűen mérhető: $v(P(x,y))$

értékelő függvényével fejezhetjük ki, amelynek argumentuma a $P(x,y)$ lehetséges valószínűségi függvények tartományában változik. A $v(P(x,y))$ függvény a hírközlési rendszereket a hűség szerint rendezi és a könnyebbség kedvéért úgy választjuk meg, hogy a kisebb v értékek felelnek meg a „nagyobb hűségnek”.

Most meg fogjuk mutatni, hogy nagyon általános és ésszerű feltételezések mellett a $v(P(x,y))$ függvényt látszólag sokkal speciálisabb formába lehet átírni, mégpedig egy $p(x,y)$ függvény x és y lehetséges értékkészletében tekintett átlagként:

$$v(P(x,y)) = \int\int P(x,y)\rho(x,y)dxdy$$

Hogy ezt megkapjuk, csupán azt kell feltételeznünk, hogy: (1) a forrás és a rendszer ergodikus, s ilyenformán egy igen hosszú – csaknem 1 valószínűséggel – az együttes jellemzője, és (2) az értékelés „ésszerű”, olyan értelemben, hogy – tipikus x_1 be- és y_1 kimenet mellett – lehetséges ezen minták alapján előzetes értékelést végezni, s ha ezen minták időtartamát növeljük, akkor az előzetes értékelés 1-es valószínűséggel fogja közelíteni a $P(x,y)$ teljes ismeretén alapuló egzakt értékelést. Legyen az előzetes értékelés $\rho(x,y)$. Ekkor a $\rho(x,y)$ függvény (miközben $T \to \infty$) csaknem valamennyi (x,y) értékre, amelyek a rendszernek megfelelő, nagy valószínűségű tartományban találhatók, egy konstanshoz közelít:

$$\rho(x,y) \to v(P(x,y))$$

valamint azt is írhatjuk, hogy $\rho(x,y) \to \int\int P(x,y)\rho(x,y)dxdy$, minthogy $\int\int P(x,y)dxdy = 1$.

Ez a kívánt eredményt adja.

A $\rho(x,y)$ függvény általános természete „távolság” az x és y között.[19] Ez azt méri, hogy – a hűségkritériumunknak megfelelően – milyen mértékben nem kívánatos x adása esetén y vétele. A fentebb megadott általános eredményt a következőképpen is megfogalmazhatjuk: bármely ésszerű értékelés leírható, mint egy olyan üzenetsorozatra és az x és y helyreállított üzenetre vonatkozó távolságfüggvény átlaga, amelyeket az együttes vétel $P(x,y)$ valószínűsége szerint súlyozunk, feltéve, hogy az üzenetek T időtartamát elegendően hosszúra választjuk. A következők egyszerű példák az értékelő függvényekre:

1. Négyzetes középérték-kritérium $v = \overline{(x(t) - y(t))^2}$

A hűség e nagyon általánosan használt mértékénél a $\rho(x,y)$ távolságfüggvény (a konstans tényezőtől eltekintve) a vonatkozó függvénytér x és y pontja közötti közönséges euklideszi távolság négyzete.

$$\rho(x,y) = \frac{1}{T}\int_0^T [x(t) - y(t)]^2\, dt$$

2. Frekvenciában súlyozott négyzetes középérték-kritérium.

Általánosabb esetben a különböző frekvencia összetevőkre különböző súlytényezőket alkalmazhatunk, mielőtt a hűség négyzetes középérték-kritériumát használnánk. Ez egyenértékű azzal, ha az $x(t)$–$y(t)$ különbséget egy alakító (shaping) szűrőn bocsátjuk át, majd meghatározzuk a kimeneten az átlagos teljesítményt. Így legyen

$$e(t)=x(t)-y(t) \quad \text{és} \quad f(t) = \int_{-\infty}^{\infty} e(\tau)k(t-\tau)\,d\tau \text{ , ekkor } \rho(x,y) = \frac{1}{T}\int_0^T f(t)^2\, dt$$

3. Abszolút hibakritérium

$$\rho(x,y) = \frac{1}{T}\int_0^T \left| x(t) - y(t) \right| dt$$

4. A fül és az agy szerkezete implicite egy sor értékelési módszert határoz meg, amelyek beszéd- vagy zeneátvitel esetén alkalmazhatók célszerűen. Van például egyfajta „érthetőségi” kritérium, amely szerint $\rho(x,y)$ egyenlő a helytelenül értelmezett szavak előfordulási gyakoriságával, abban az esetben, ha az $x(t)$-t $y(t)$ üzenetként vettük. Jóllehet nem tudjuk megadni $\rho(x,y)$ explicit kifejezését ezekben az esetekben, azonban az – elvileg – elegendő kísérletezéssel meghatározható. Némely tulajdonsága a hallással kapcsolatos jól ismert kísérleti eredményekből következik, pl. a fül viszonylag nem érzékeny a fázishibára, valamint, hogy az amplitúdó ill. a frekvencia függvényében mért érzékenysége durván logaritmikus.
5. A diszkrét esetet olyan leszűkített esetként tekinthetjük, amelyben – röviden fogalmazva – a hibagyakoriság alapján végezzük az értékelést. A $\rho(x,y)$ függvényt ezután úgy definiáljuk, mint az y sorozatban szereplő azon szimbólumok számának, amelyek az x-ben lévő, megfelelő szimbólumoktól különböznek, és az x-ben lévő összes szimbólumok számának hányadosát.

---

[19] Szorosan vett értelemben azonban ez nem „metrikus”, mivel általában sem a $\rho(x,y) = \rho(y,x)$, sem a $\rho(x,y)+\rho(y,z)\geq\rho(x,z)$ kifejezést nem elégíti ki.

28. A forrás hűségértékelésére vonatkoztatott sebessége

Most már abban a helyzetben vagyunk, hogy egy folytonos forrás esetében meghatározhatjuk az információ előállítás sebességét. Adva van a forrásra *P(x)* és egy $\rho(x,y)$ – mind *x*-re mind *y*-ra folytonos – távolságfüggvény által meghatározott *v* értékelés. Egy adott $\rho(x,y)$ sajátos rendszer esetén a minőséget a $v = \int\int \rho(x,y)P(x,y)dxdy$ méri. Ezen túlmenően a *P(x,y)*-nak megfelelő bináris digitek áramlási sebessége:

$$R = \int\int P(x,y)\log\frac{P(x,y)}{P(x)P(y)}dxdy$$

Az információ előállítás egy adott v_1 minőségű helyreállításához tartozó R_1 sebességét úgy definiálhatjuk, hogy az *R* minimuma legyen, miközben *v*-t v_1-ként rögzítjük és $P_x(y)$-t változtatjuk, azaz:

$$R_1 = \underset{P_x(y)}{Min}\int\int P(x,y)\log\frac{P(x,y)}{P(x)P(y)}dxdy$$

az alábbi megszorítás mellett: $v_1 = \int\int \rho(x,y)P(x,y)dxdy$

Ez azt jelenti, hogy valójában minden, gyakorlatilag szóba jöhető és a megkövetelt hűséggel átvinni képes hírközlési rendszert figyelembe veszünk. Az átvitel bit/s-ban mért sebességét minden egyes rendszerre kiszámítjuk és ezek közül a legkisebb sebességűt választjuk. Ez utóbbi sebesség lesz az, amelyet a szóban forgó hűségnél a forráshoz hozzárendelünk. E definíció igazolását a következő eredmény adja:

21. TÉTEL:
Ha egy forrás v_1 értékeléshez tartozó sebessége R_1, lehetséges a forrás kimenetét kódolni és azt egy *C* kapacitású csatornán v_1-hez tetszés szerint közeli hűséggel átvinni, feltéve hogy $R_1 \le C$. Mindez nem lehetséges, ha $R_1 \rangle C$.

A tétel utolsó állítása közvetlenül következik R_1 definíciójából és korábbi eredményekből. Ha nem volna igaz, úgy *C* bit/s-nál nagyobb sebességgel is tudnánk információt átvinni egy C kapacitású csatornán. A tétel első részét a 11. tétel bizonyításához hasonló módon bizonyítjuk be. Először az *(x,y)* teret feloszthatjuk nagyszámú kis sejtre és a szituációt mint diszkrét esetet tekinthetjük. Ez az értékelő függvényt (ha a sejtek igen kicsinyek) nem változtathatjuk meg egy tetszés szerinti kis mértéknél jobban, mivel $\rho(x,y)$-t folytonosnak vettük. Tételezzük fel, hogy $P_1(x,y)$ az az adott rendszer, amely minimalizálja a sebességet és eredményül az R_1-et adja. A nagy valószínűségű *y*-ok közül véletlenszerűen kiválasztunk egy olyan készletet, amely $2^{(R_1+\varepsilon)T}$ tagból áll, ahol $\varepsilon \to 0$ miközben $T \to \infty$. Nagy *T* érték esetén minden kiválasztott pontot nagy valószínűségű vonalak fognak összekötni az *x*-ek egy készletével (amint az a 10. ábrán látható). A 11. Tétel bizonyításához használthoz hasonló számítás azt mutatja, hogy nagy *T* esetén csaknem az összes *x*-ek – az *y*-ok csaknem valamennyi lehetséges kiválasztása esetében – a kiválasztott *y* pontokból kiinduló görbesereg fedi. Az alkalmazandó hírközlési rendszer a következőképpen működik, a kiválasztott pontokhoz bináris számokat rendelünk hozzá. Egy *x* üzenetet előállítva az az *1*-hez közelítő valószínűséggel (miközben $T \to \infty$) rajta fog feküdni a görbesereg legalább is egyik ágán. A megfelelő bináris számot (vagy ha több van, azok közül egy, tetszés szerint kiválasztottat) megfelelő kódolással visszük át a csatornán úgy, hogy kis hibavalószínűséget érjünk el. Mivel $R_1 \leq C$, ez lehetséges. A vételi pontban a megfelelő *y*-t visszaalakítjuk és felhasználjuk, mint helyreállított üzenetet.

Az erre a rendszerre alkalmazott v_1' értékeléssel v_1-et tetszés szerinti pontossággal megközelíthetjük azáltal, hogy *T*-t elegendően nagynak választjuk. Ez annak köszönhető, hogy az *x(t)* üzenet és a helyreállított *y(t)* üzenet minden egyes hosszú mintájára az értékelés (*1* valószínűséggel) közelíti v_1-*et*.

Érdekes megjegyezni, hogy ebben a rendszerben a visszanyert üzenetekben a zajt az adó egyfajta kvantálása állítja elő és az nem a csatornában keletkezik. Ez a zaj többé-kevésbé analóg a PCM-technika kvantálási zajával.

29. A sebességek számítása

A sebesség definíciója sok tekintetben hasonlít a csatornakapacitáséhoz. Az előbbire

$$R = \underset{P_x(y)}{Min} \int\int P(x,y) \log \frac{P(x,y)}{P(x)P(y)} dxdy$$

ahol *P(x)* és $v_1 = \int\int \rho(x,y) P(x,y) dxdy$ rögzített, akkor

$$C = \underset{P(x)}{Max} \int\int P(x,y) \log \frac{P(x,y)}{P(x)P(y)} dxdy$$

ahol $P_x(y)$ rögzített és esetleg egy vagy több megszorítás (pl. az átlagteljesítmény korlátozása) $K = \int\int P(x,y) \lambda(x,y) dxdy$.

Egy forrás sebességének meghatározása során megjelenő általános maximalizálási problémára megadhatunk egy parciális megoldást. A Lagrange-módszert alkalmazva tekintsük a következő kifejezést:

$$\int\int \left[P(x,y) \log \frac{P(x,y)}{P(x)P(y)} + \mu P(x,y) \rho(x,y) + v(x) P(x,y) \right] dxdy$$

A variációs egyenlet (ha *P(x,y)*-ra az első variációt vesszük): $P_y(x) = B(x) e^{-\lambda \rho(x,y)}$

ahol λ-t úgy határoztuk meg, hogy a kívánt hűséget adja és *B(x)*-et úgy választottuk meg, hogy kielégítse: $\int B(x) e^{-\lambda \rho(x,y)} dx = 1$.

Ez azt mutatja, hogy a legjobb dekódolás mellett egy meghatározott ok, több vett *y*-hoz tartozó $P_y(x)$ feltételes valószínűsége a szóban forgó *x* és *y* közötti távolság $\rho(x,y)$ távolságfüggvényével exponenciálisan fog csökkenni. Abban a speciális esetben, amikor a $\rho(x,y)$ távolságfüggvény csupán *x* és *y* (vektor) különbségétől függ: $\rho(x,y) = \rho(x-y)$. Kapjuk, hogy $\int B(x) e^{-\lambda \rho(x,y)} dx = 1$

Ennélfogva *B(x)* konstans, (mondjuk: α) és $P_y(x) = \alpha e^{-\lambda \rho(x-y)}$.

Sajnálatos módon, egyes sajátos esetekben nehéz értékelni ezeket a formális megoldásokat és kevéssé használhatónak tűnnek. Valójában a sebességek tényleges számítását csupán néhány, igen egyszerű esetre végeztük el.
Ha a $\rho(x,y)$ távolságfüggvény az x és y közötti négyzetes középérték (effektív) bizonytalanságot jelenti és az üzenet együttes fehér-zaj, akkor a sebesség meghatározható. Ez esetben $R = Min\left|H(x) - H_y(x)\right| = H(x) - MaxH_y(x)$ és $N = \overline{(x-y)^2}$. Azonban a $H_y(x)$-nek ott van a maximuma, ahol y-x fehér-zaj, és $W_1 \log 2\pi eN$-el egyenlő, ahol W_1 az üzenetsorozat sávszélessége. Ennélfogva

$$R = W_1 \log 2\pi eQ - W_1 \log 2\pi eN = W_1 \log \frac{Q}{N}$$

ahol Q az átlagos üzenetteljesítmény. Ez a következőket bizonyítja:

22. TÉTEL:
Egy Q teljesítményű és W_1 sávszélességű fehér-zaj négyzetes-középérték hűségmértékre vonatkoztatott sebessége $R = W_1 \log \frac{Q}{N}$, ahol N az eredeti és a helyreállított üzenetek közötti megengedett effektív hiba.

Általánosabb értelemben bármely üzenetforrásra kaphatunk olyan egyenlőtlenségeket, amelyek – egy adott effektív hibakritériumra vonatkoztatva – korlátai lesznek a sebességnek.

23. TÉTEL:
Bármely W_1 sávszélességű forrás sebessége a $W_1 \log \frac{Q_1}{N} \le R \le W_1 \log \frac{Q}{N}$ szerint korlátos, ahol Q a forrás átlagos teljesítménye, Q_1 annak entrópia teljesítménye és N a megengedett effektív hiba.

Az alsó korlát következik abból, hogy $MaxH_y(x)$ egy adott $\overline{(x-y)^2} = N$ esetén akkor áll elő, amikor fehér-zajjal állunk szemben. A felső korlátot úgy kapjuk meg, hogy a (21. tétel bizonyításánál használt) pontokat nem a legkedvezőbb módon, hanem véletlenszerűen helyezzük el egy $\sqrt{Q-N}$ sugarú gömb belsejében.

Köszönetnyilvánítás

A szerző lekötelezve érzi magát a Laboratóriumnál dolgozó kollégáinak, különösképpen Dr. H.V. Bode-nak. Dr. J. R. Pierce-nek, Dr.B. McMillan-nek és Dr.B.M Oliver-nek az e-munka során nyújtott sokoldalú segítségért. Ugyancsak elismerés illeti N. Wiener professzort, akinek elegáns megoldása a stacionárius sorozatok szűrése és előrejelzése terén befolyásolta a szerző gondolkodását ezen a területen.