



**Vilniaus  
universitetas**

# **Metodinė medžiaga Įvadas į informacijos teoriją**

Mokyklos pedagogika



Kuriame  
Lietuvos ateitį  
2014–2020 metų  
Europos Sąjungos  
fondų investicijų  
veiksmų programa

## Metodinė medžiaga. Įvadas į informacijos teoriją

Informatikos ir informatinio mąstymo veiklos, metodinė medžiaga sukurta įgyvendinant projektą „Aukštųjų mokyklų tinklo optimizavimas ir studijų kokybės gerinimas Šiaulių universitetą prijungiant prie Vilniaus universiteto“, projekto Nr. 09.3.1-ESFA-V-738-03-0001, vykdomą pagal 2014–2020 metų Europos Sąjungos fondų investicijų veiksmų programos 9 prioriteto „Visuomenės švietimas ir žmogiškųjų išteklių potencialo didinimas“ 09.3.1-ESFA-V-738 įgyvendinimo priemonę „Aukštųjų mokyklų tinklo tobulinimas“, finansuojamą Europos Sąjungos fondų ir Lietuvos Respublikos valstybės biudžeto lėšomis.

„Įvadas į informacijos teoriją“, skirta Mokyklos pedagogikos studijų programos moduliui „Informatikos didaktika“. Tikslinė grupė – būsimi informatikos mokytojai. Medžiagoje glaustai paaiškinamos informacijos teorijos ištakos, pateikiamas istorinis ir dabartinis požiūris į informacijos teoriją, aiškinami pagrindiniai teorijos elementai, sąvokos: entropija, ergodinis procesas, informacijos matavimo vienetai, informacijos kiekis ir jo skaičiavimo formulė, komunikavimo kanalas, Hafmano kodas, informacijos perdavimas ir priėmimas, pateikiami pavyzdžiai, užduotys su nurodymais ir sprendimais. Kai kurioms užduotims yra galimybė jas dėstytojui koreguoti, pertvarkyti, ar dubliuoti keičiant tam tikrus parametrus. Pateikiamas pagrindinių šaltinių sąrašas.

Šios veiklos autoriai: Alvida Lozdienė, Viktoras Dagys ir prof. dr. Valentina Dagienė

Redagavo: Viktoras Dagys

Projekto vykdytojas: Vilniaus universitetas.

2022, Vilnius

## Kaip kuriama teorija?

Kai iš gausybės stebėjimų, eksperimentų išskiriama tai, ką jau galima apibendrinti, apibendrinimai atlikti ir gautos atitinkamos išvados, tik tada gimsta tam tikros srities teorija. Pavyzdžiui, Niutono dėsnų pagrindu kuriama klasikinės mechanikos teorija. Patys dėsniai dar ne teorija, jie tik sudaro teorijos pagrindą. Pavyzdžiui, geometrijos pagrindą sudaro aksiomos arba postulatai.

Teorijoje svarbu ne tik teiginiai, bet ir matematinis išvadų apdorojimas. Išvados išplaukia iš teiginių. Jei teorija teisinga, tai jos išvados turi atspindėti gamtos ar kitus reiškinius. Jei teorija neteisinga, ji niekam nereikalinga.

**Informacijos** teorija dažnai vadinama **komunikacijos** teorija, ir toliau paaiškės kodėl. Informacijos teorija yra ir labai griežta, ir matematinė. Informacijos teorija gimė tyrinėjant elektrinius signalus, kuriais gali būti perduodama informacija. Tuo pačiu informacijos teorija operuoja ir labai bendromis sąvokomis: neapibrėžtumu (angl. *uncertainty*), išranka (angl. *choice*), naudojamas universalus informacijos matavimo vienetas – **bitas**.

Labai svarbi **entropijos** sąvoka.

Informacijos teorijoje nagrinėjami informacijos šaltiniai gali sukurti bet kurį iš galimos aibės pranešimų. Informacijos kiekis, kurį šis pranešimas neša, auga, didėjant siunčiamo pranešimo neapibrėžtumui: pavyzdžiui, jei siunčiamas pranešimas yra galimas vienas iš dešimties, jis neša mažesnį informacijos kiekį, nei pranešimas, kuris tikėtinas vienas iš milijono.

Informacijos teorijoje galimų pranešimų neapibrėžtumas ir yra matuojamas **entropija**.

Kuo daugiau žinoma apie siunčiamą pranešimą, tuo mažesnis neapibrėžtumas, tuo mažesnė entropija ir tuo pačiu mažesnis informacijos kiekis.

Vienas informacijos bitas gaunamas kaip vienos iš dviejų vienodai tikėtinų galimybių realizavimo rezultatas. Šiomis galimybėmis gali būti ir pranešimai, ir skaičiai.

Informacijos teorija mums pasako, kiek bitų informacijos galima perduoti per 1 sekundę idealiu ir neidealiu ryšio kanalu, priklausomai nuo kanalo savybių. **Perdavimo aplinka** arba **kanalas** fiziškai sujungia pranešimus priimančią imtuvą su pranešimų siųstuvu (pvz., metalinis laidininkas, šviesolaidis arba laisva erdvė).

Informacijos teorija nusako, kaip išmatuoti greitį, kuriuo pranešimų šaltinis (pavyzdžiui, kalbantis ar rašantis žmogus) sukuria informaciją.

Informacijos teorija nusako, kaip geriau užkoduoti konkretaus šaltinio pranešimus duotu kanalu, pavyzdžiui, elektros grandine. Taip pat nusako, kaip išvengti klaidų perduodant informaciją.

Informacijos teorija nagrinėja labai abstrakčius ir bendrus klausimus, todėl yra gana sudėtinga ją pritaikyti sprendžiant konkrečius uždavinius.



**Klodus Šenonas** (angl. Claude Shannon; 1916–2001)

[https://en.wikipedia.org/wiki/Claude\\_Shannon#/media/File:ClaudeShannon\\_MFO3807.jpg](https://en.wikipedia.org/wiki/Claude_Shannon#/media/File:ClaudeShannon_MFO3807.jpg)

Klodus Šenonas yra informacijos teorijos pradininkas. Jis sukūrė pagrindines sąvokas, idėjas ir jų matematinės formuluotes, kurios dabar sudaro šiuolaikinių komunikacinių technologijų pagrindą. 1948 m. jis pasiūlė terminą „bitas“ mažiausiam informacijos kiekio vienetui apibūdinti (straipsnis „Matematinė komunikacijos teorija“, angl. *A Mathematical Theory of Communication*). Jis įrodė, kad entropija yra lygiavertė perduodamo pranešimo informacijos neapibrėžtumo matui. Klodus Šenonas buvo vienas pirmųjų, kuris į kriptografiją pradėjo žiūrėti moksliniu požiūriu, pirmasis suformulavo jos teorinius pagrindus ir įvedė daugelį pagrindinių sąvokų. Šenonas labai prisidėjo prie tikimybių teorijos, žaidimų teorijos, automatų teorijos ir valdymo sistemų teorijos.

## Informacijos teorijos ištakos

Matematinė informacijos (komunikacijos) teorija atsirado nagrinėjant informacijos perdavimą elektriniais signalais.

1932 m. Samuelis Morzė (Samuel Morse) pradėjo kurti pirmą elektrinį telegrafo aparatą. Pradiniame variante popieriaus juostoje buvo spausdinami taškai ir brūkšniai, kurie reiškė ne raides, o skaičius.



**Samuelis Morzė** (angl. Samuel Finley Breese Morse; 1791–1872) – amerikiečių išradėjas. 1838 m. sukūrė jo vardu pavadintą abėcėlę.

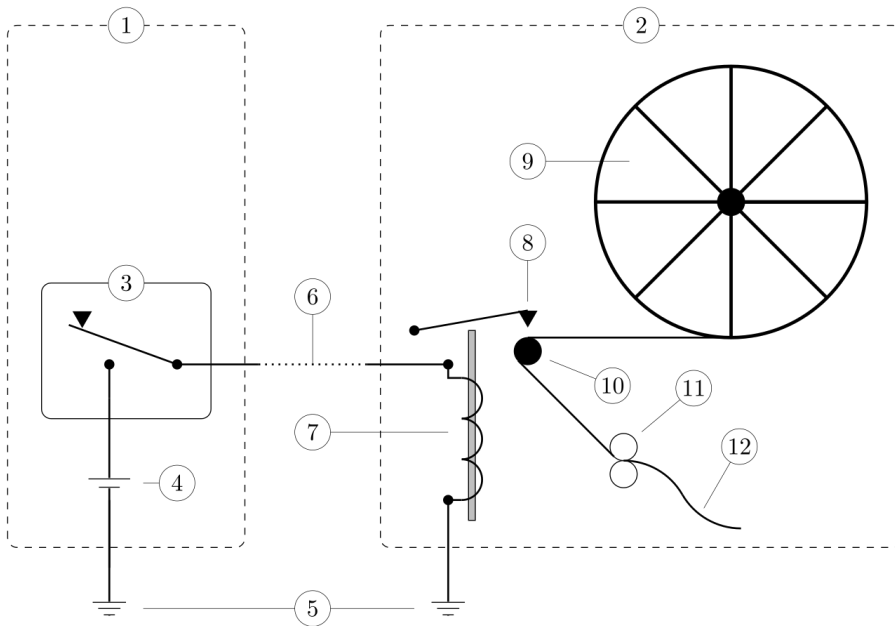
[https://upload.wikimedia.org/wikipedia/commons/8/8d/Samuel\\_Morse\\_1840.jpg](https://upload.wikimedia.org/wikipedia/commons/8/8d/Samuel_Morse_1840.jpg)

Pirmoje kodų lentelėje, kurią Morzė sukūrė 1837 m., kiekvienas skaičius reiškė vieną žodį.



### Morzės telegrafas

[https://media.izi.travel/844ea502-4b85-4410-b494-b77886b95c9d/348fd5e3-47bd-48b3-9a15-67a31febfede\\_800x600.jpg](https://media.izi.travel/844ea502-4b85-4410-b494-b77886b95c9d/348fd5e3-47bd-48b3-9a15-67a31febfede_800x600.jpg)



**Supaprastinta elektromechaninio telegrafo schema:** 1 – pranešimą siunčianti stotis; 2 – pranešimą priimanti stotis; 3 – perdavimo raktas; 4 – akumulatorius (srovės šaltinis); 5 – žeminimas; 6 – elektros linija; 7 – elektromagnetas; 8 – rašanti plunksna; 9 – popierinės juostos rulonas; 10 – velenas; 11 – prilaukiantys velenai; 12 – popierinė juosta

1838 m. sukuriamas naujas kodas – Morzės abėcėlė.

Abėcėlės raidės vaizduojamos taškais, brūkšniais ir pauzėmis. Pauzė, kai elektros srovė neteka, taškai – trumpi elektros srovės impulsai, brūkšniai – ilgesni elektros srovės impulsai. Jei taško trukmę laikysime vienetu, tai brūkšnio signalo trukmė – 3 vienetai, o pauzės tarp žodžių trukmė – 7 vienetai.

A	● ■■	U	● ● ■■
B	■■■ ● ●	V	● ● ● ■■
C	■■■ ●■■ ●	W	● ■■ ■■
D	■■■ ● ●	X	■■■ ● ● ■■
E	●	Y	■■■ ● ■■ ■■
F	● ● ■■ ●	Z	■■■ ■■ ● ●
G	■■■ ■■ ●		
H	● ● ● ●	1	● ■■ ■■ ■■ ■■
I	● ●	2	● ● ■■ ■■ ■■
J	● ■■ ■■ ■■ ■■	3	● ● ● ■■ ■■
K	■■■ ● ■■ ■■	4	● ● ● ● ■■
L	● ■■ ● ●	5	● ● ● ● ●
M	■■■ ■■	6	■■■ ● ● ● ●
N	■■■ ●	7	■■■ ■■ ● ● ●
O	■■■ ■■ ■■	8	■■■ ■■ ■■ ● ●
P	● ■■ ■■ ■■ ●	9	■■■ ■■ ■■ ■■ ●
Q	■■■ ■■ ● ■■ ■■	0	■■■ ■■ ■■ ■■ ■■
R	■■■ ● ●		
S	● ● ●		
T	■■■		

Tarptautinis Morzės kodas

Morzės telegrafo technologija labai svarbi, nes ji susijusi su pirmaisiais sėkmingais bandymais suskaitmeninti perduodamą informaciją didesniais atstumais.

Kai telegrafo linijos buvo nutiestos dideliais atstumais, buvo susidurta su keliomis problemomis. Pirmoji – informacijos perdavimo greitis. Greitį lemia telegrafo linijų fizinės savybės, tačiau informacijos perdavimo greitį galima padidinti ir kitais būdais.

Morzės abėcėlėje brūkšnį siunčiantis elektrinis signalas trunka 3 kartus ilgiau nei taško.

Kai signalas siunčiamas tik viena kryptimi, elektros srovė teka (pauzė – neteka). Siekiant didinti signalų perdavimo greitį buvo sukurta telegrafo elektros grandinė, kuria elektros srovė galėjo tekėti dvejomis kryptimis: pirmyn ir atgal. Vienos krypties telegrafijoje kodą sudarė du elementai: srovė teka ir srovės nėra. Juos galima žymėti 1 ir 0. Dviejų krypčių telegrafijoje jau yra trys būsenos, srovė pirmyn, atgal ir srovės nėra. Tokių situacijų galima žymėti +1, 0, -1.

## Užduotis

Panagrinėkite Morzės kodą anglų kalbos abėcėlei. Kodėl E raidei koduoti Morzė pasirinko tik vieną tašką, o J raidei ne tik tašką, bet ir 3 brūkšnius?

Pasiremkite šituo anglišku tekstu:

### Radiotelegraphy and aviation

The American artist Samuel Morse, the American physicist Joseph Henry, and mechanical engineer Alfred Vail developed an electrical telegraph system. It needed a method to transmit natural language using only electrical pulses and the silence between them. Around 1837, Morse, therefore developed an early forerunner to the modern International Morse code.

In the 1890s, Morse code began to be used extensively for early radio communication before it was possible to transmit voice. In the late 19th and early 20th centuries, most high-speed international communication used Morse code on telegraph lines, undersea cables, and radio circuits.

Although previous transmitters were bulky and the spark gap system of transmission was dangerous and difficult to use, there had been some early attempts: In 1910, the U.S. Navy experimented with sending Morse from an airplane. However, the first regular aviation radiotelegraphy was on airships, which had space to accommodate the large, heavy radio equipment then in use. The same year, 1910, a radio on the airship America was instrumental in coordinating the rescue of its crew.

During World War I, Zeppelin airships equipped with radio were used for bombing and naval scouting, and ground-based radio direction finders were used for airship navigation. Allied airships and military aircraft also made some use of radiotelegraphy.

However, there was little aeronautical radio in general use during World War I, and in the 1920s, there was no radio system used by such important flights as that of Charles Lindbergh from New York to Paris in 1927. Once he and the Spirit of St. Louis were off the ground, Lindbergh was truly incommunicado and alone. Morse code in aviation began regular use in the mid-1920s. By 1928, when the first airplane flight was made by the Southern Cross from California to Australia, one of its four crewmen was a radio operator who communicated with ground stations via radio telegraph.

Beginning in the 1930s, both civilian and military pilots were required to be able to use Morse code, both for use with early communications systems and for identification of navigational beacons that transmitted continuous two- or three-letter identifiers in Morse code. Aeronautical charts show the identifier of each navigational aid next to its location on the map.

[https://en.wikipedia.org/wiki/Morse\\_code](https://en.wikipedia.org/wiki/Morse_code)

Nubraižykite stulpelinę diagramą ir dar kartą paanalizuokite Morzės kodo raidžių ilgio pasirinkimą. Išrikiuokite diagramą raidžių pasikartojimo mažėjimo tvarka.

Tyrimai rodo, kad pakeitus Morzės kodus į racionaliau panaudojamus, informacijos perdavimo greitis padidėtų apie 10 procentų. Pakomentuokite šį teiginį.

A, a	B, b	C, c	D, d	E, e	F, f	G, g	H, h	I, i	J, j	K, k	L, l	M, m	N, n	O, o	P, p	Q, q	R, r	S, s	T, t	U, u	V, v	W, w	X, x	Y, y	Z, z
178	23	58	84	227	35	38	79	165	1	3	73	58	141	140	37	3	145	120	151	56	19	33	3	27	1

Apskaičiuokite tarpų ir raidžių tikimybes. Palyginkite su „Britannica.com“ skelbiama angliško teksto raidžių ir tarpų tikimybių lentele.

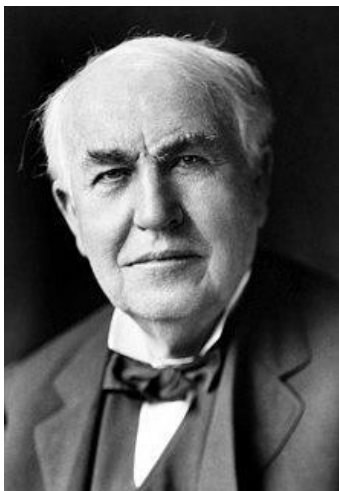
<https://www.britannica.com/science/information-theory/Applications-of-information-theory>

Simbolis	tarpas	E	T	A	O	I	N	S	R	H	L	D	U	C
Tikimybė	0.1859	0.1031	0.0796	0.0642	0.0632	0.0575	0.0574	0.0514	0.0484	0.0467	0.0321	0.0317	0.0228	0.0218

Simbolis	F	M	W	Y	P	G	B	V	K	X	Q	J	Z
Tikimybė	0.0208	0.0198	0.0175	0.0164	0.0152	0.0152	0.0127	0.0083	0.0049	0.0013	0.0008	0.0008	0.0005



1874 m. Tomas Edisonas sukūrė kvadrupolinę telegrafiją, kai naudojamos ne tik dvi srovės kryptys, bet ir dvi srovės vertės. Edisonas du pranešimus sugebėjo perduoti vienu metu. Vienas pranešimas buvo perduodamas varijuojant vienos krypties srove, o kitas – priešingos krypties srove.



**Tomas Alva Edisonas** (angl. Thomas Alva Edison; 1847–1931) – amerikiečių išradėjas ir verslininkas. Vienas produktyviausių išradėjų, savo vardu užpatentavęs net 1093 išradimus. Kai kurie jų buvo tik senų patentų patobulinimai, dalį atrado jo darbuotojai. Svarbiausi išradimai – elektros lemputė (išradėjas yra Heinrichas Gėbelis, tačiau Edisonas jį komercializavo), fonografas, diktofonas, kinetoskopas, elektros kėdė.

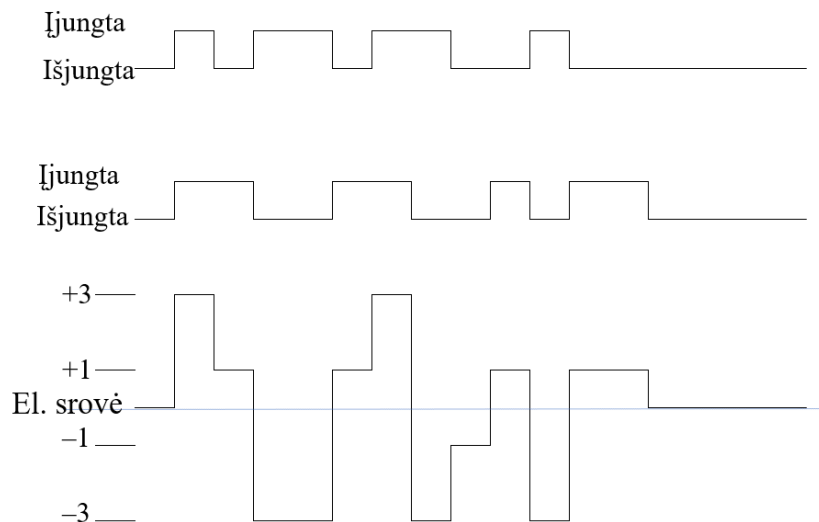
[https://lt.wikipedia.org/wiki/Thomas\\_Edison#/media/Vaizdas:Thomas\\_Edison2-crop.jpg](https://lt.wikipedia.org/wiki/Thomas_Edison#/media/Vaizdas:Thomas_Edison2-crop.jpg)

Tarkime, kad srovių stipriai vienas nuo kito skiriasi dviem vienetais, tuomet visas keturias sroves galima žymėti: +3, +1, –1, –3.

Abu pranešimai perskaitomi naudojant lentelę.

Signalą perduodančios elektros srovės stiprio atitikmuo	Pirmas pranešimas	Antras pranešimas
+3	Įjungta	Įjungta
+1	Išjungta	Įjungta
–1	Išjungta	Išjungta
–3	Įjungta	Išjungta

Dviejų nepriklausomų pranešimų kodavimo pavyzdys, kai naudojama kvadrupolinė telegrafija:



Matome, kad taip perduodant signalus, vienu metu perduodami du nepriklausomi pranešimai: 0 arba 1 pirmajame pranešime, ir 0 arba 1 antrajame pranešime. Taigi per tą patį laiką perduodama du kartus daugiau informacijos. Pavyzdžiui, per 1 minutę perduodama 2 kartus daugiau raidžių, nei perduodant vienos krypties ir vienos vertės elektros srovė.

Tačiau didesnis skaičius perduodamų signalų skirtingų verčių sukelia ir sunkumų. Perdavimo linijose atsiranda pašaliniai (triukšmo) signalai ir srovės.

Dar viena techninė galimybė yra susijusi su signalo stiprinimu. Tačiau tuo pačiu metu didėjant srovės stipriui, didėja ir įtampa, o tai gali sukelti trumpuosius jungimus, nes laidai stipriai įkaista, lydosi ir linija yra pažeidžiama.

Morzės telegrafija labai svarbi aiškinantis informacijos perdavimą ir sprendžiant svarbų informacijos teorijos uždavinį – informacijos kodavimą.

## Kodavimas

Kodavimas atsirado labai seniai. Tada pagrindinis jo tikslas buvo užšifruoti – apsaugoti siunčiamo pranešimo informaciją nuo nepageidaujamų asmenų.

Paprasčiausias šifravimo pavyzdys – vieni žodžiai pakeičiami kitais. Keitimas nurodomas šifro žinyne.

Gali būti naudojamas ir specialus šifravimo kodas, kai raidės ar skaitmenys keičiami pagal iš anksto sutartą sistemą.

Kodavimas – tai bet kokio objekto atvaizdavimas įvairiomis priemonėmis (fiziniais signalais). Pavyzdžiui, Morzės kodas leidžia užkoduoti bet kokį tekstą.

## Užduotis

Koks anglišką žodis čia užkoduotas? Pasinaudokite Morzės abėcėle.

• — — • • — • — — — — — • — • • • • • — — — — — • — •

Atsakymas

PROCESSOR

P • — — • R • — • O — — — — — C — • — • E • S • • • S • • • O — — — — — R • — •

Informacijos teorija nagrinėja ne tik diskrečius signalus (tekstas), bet ir tolydžius (kalba, muzika, paveikslas).

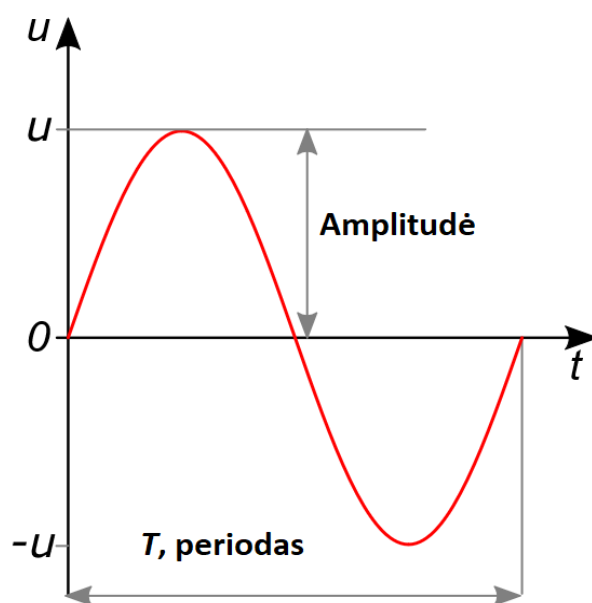
Viena iš svarbiausių informacijos teorijos teoremų (*Sampling theorem*) sako, kad bet kokį tolydų signalą galima pavaizduoti ir atkurti šio signalo matavimų seka, kai matuojama per vienodus laiko tarpus. Matuoti reikia tokiais laiko intervalais, kad jie būtų mažesni ar lygūs didžiausio dažnio, esančio signale, periodo pusei.

Dažnai signalai vaizduojami kaip sinuso arba kosinuso funkcijos. Tai susiję su tuo, kad XIX amžiuje prancūzų matematikas Furjė (*Fourier*) įrodė, kad bet kokią kintančią laiką ir signalą vaizduojančią matematinę funkciją galima pavaizduoti, kaip sumą trigonometrinių funkcijų, kurios viena nuo kitos skiriasi faze ir dažniu.

Taigi kinta ir yra svarbiausi šie signalo parametrai: **amplitudė, dažnis arba periodas, sklidimo greitis.**

Jei signalą grafiškai vaizduotume tokia viena funkcija, tai horizontalioje ašyje vaizduojant laiką būtų stebima, kaip bėgant laikui keičiasi terpės taško padėtis pusiausviros padėties atžvilgiu. Jei signalas fizinis ir sklinda terpėje, pavyzdžiui garso signalas (kamertonu išgaunamas pastovus *la* garsas), sklinda ore, tai didžiausias oro sutankėjimas arba praretėjimas (fiksuoja kaip slėgis) ir bus amplitudė. Jei sklinda elektromagnetinė banga (pavyzdžiui, radijo, optiniai ar infraraudonieji spinduliai), tai ji gali sklisti ne tik terpe, bet ir vakuume. Elektromagnetinių bangų amplitudė siejama su intensyvumu. Labai svarbus laiko intervalas, per kurį terpės taškai ar elektromagnetinis laukas svyruoja sinchroniškai. Tokį intervalą vadiname svyravimo periodu ir žymime  $T$ , matuojame sekundėmis arba jos dalimis. Atstumą, kurį banga nusklinda per vieną periodą, vadiname bangos ilgiu ir žymime  $\lambda$  (lambda) raide. Bangos ilgį padalinę iš periodo gauname bangos sklidimo greitį. Vienalytėje terpėje tos

pačios prigimties bangos sklinda vienodu greičiu. Greitis nepriklauso nuo bangos ilgio ar periodo. Kai tik tos pačios prigimties bangos pereina į kitą terpę, iš karto keičiasi sklidimo greitis ir, tai yra labai svarbu, **keičiasi ir bangos ilgis, bet ne periodas**. Dažniausiai yra naudojama ne periodo sąvoka, bet bangos periodui atvirkščias dydis dažnis  $f$ . Jei bangos periodas yra 0,1 s, tai reiškia, kad per 1 sekundę įvyksta 10 svyravimų, tai yra dažnis lygus 10 hercų. Hercas žymimas Hz ir  $1 \text{ Hz} = 1 \text{ s}^{-1}$ . ( $T = 1 / f$ , o  $f = 1 / T$ ). Jei informacijos šaltinis artėja arba tolsta nuo stebėtojo, pasireiškia Doplerio efektas ir dažnis gali pasikeisti. Tai visi patiria, kai girdimos greitosios pagalbos ar ugniagesių automobilio sirenos. Jų garsas aštresnis, kai automobilis artėja link stebėtojo (dažnis arba tonas aukštesnis), o sirenai tostant garsas tampa duslesnis, storesnis (dažnis retesnis, tonas žemesnis).



Sinuso dėsnio kintanti įtampa.

## Entropija

Ekonometrijoje ir signalų apdorojime stochastinis procesas laikomas **ergodiniu**, jei jo statistines savybes galima nustatyti iš vienos pakankamai ilgos atsitiktinės proceso imties. Tai grindžiama tuo, kad bet kokia atsitiktinių proceso imčių aibė turi atspindėti viso proceso vidutines statistines savybes. Kitaip tariant, nepriklausomai nuo to, kokios yra atskiros imtys, imčių rinkinys turi atspindėti visą procesą. Ir atvirkščiai, procesas, kuris nėra ergodinis, yra procesas, kuris kinta nepastoviu greičiu.

Tikimybių teorijoje ir susijusiose srityse stochastinis arba atsitiktinis procesas yra matematinis objektas, paprastai apibrėžiamas kaip atsitiktinių kintamųjų šeima. Stochastiniai procesai plačiai naudojami kaip sistemų ir reiškinių, kurie kinta atsitiktinai, matematiniai modeliai.

## Paprastas entropijos paaiškinimas

Informacijos teorijoje pranešimų šaltinis traktuojamas kaip ergodinis procesas, kuriam vykstant sukuriama simbolių (pranešimų) seka, kuri iki tam tikro lygio yra neprognozuojama.

Įsivaizduokime, kad pranešimų šaltinis išrenka pranešimą atsitiktiniu būdu ir yra ergodinis. Tegul iš pradžių šaltinis gali pasirinkti bet kurį vieną iš dviejų simbolių – arba  $X$ , arba  $Y$  – ir naujam pasirinkimui jokios įtakos neturi anksčiau buvęs pasirinkimas. Tokiu atveju yra žinoma, kad  $X$  pasirinkimo tikimybė yra  $p_0$ , o  $Y$  –  $p_1$ . Šias tikimybes galima apibrėžti išnagrinėjus pakankamai ilgą simbolių, kuriuos sukuria šis šaltinis, seką. Svarbu tai, kad jei šaltinis ergodinis, tai šios tikimybės nepriklauso nuo laiko (kintant laikui, tikimybės nesikeičia).

Tokiu paprasčiausiu atveju šaltinio pranešimų entropija  $H$  apibrėžiama taip:

$$H = -(p_0 \log_2 p_0 + p_1 \log_2 p_1) \text{ bitų vienam simboliui.}$$

Taigi tik du galimus pranešimus kuriančio šaltinio **entropija** – tai suma dviejų sandaugų su minuso ženklu. Viena – tai tikimybės  $p_0$ , kad bus pasirinktas simbolis  $X$ , ir jos logaritmo sandauga, o kita – tikimybės  $p_1$ , kad bus pasirinktas simbolis  $Y$ , ir jos logaritmo sandauga.

Informacijos teorijoje naudojamas logaritmas, kurio pagrindas 2.

Panagrinėkime monetos mėtymą. Tikimybės, kad moneta nukris herbu ar skaičiumi yra vienodai tikėtinos ir lygios  $1/2$ . Tai entropija tokiu atveju bus:

$$H = -(1/2 \log_2 1/2 + 1/2 \log_2 1/2)$$

$$H = -(1/2 (-1) + 1/2 (-1))$$

$$H = 1 \text{ bitas vienam metimui}$$

Jei pranešimų šaltinis yra monetos metimo rezultatų – herbu ar skaičių – seka, tai pranešimas, kuria puse atvirto moneta, atitinka vieną bitą informacijos.

Jei herbą pažymėsime vienetu, o skaičių – nuliu, tai monetos metimų seką galima pavaizduoti nulių ir vienetų seka. Šiuo atveju entropija – 1 bitas vienam monetos metimui ir dvejetainių skaitmenų sekoje skaičius yra vienodi. Šiuo atveju dvejetainių skaitmenų, reikalingų perduoti šį seką atspindintį pranešimą skaičius yra lygus informacijos šaltinio entropijai.

Pasinaudokime ypatinga moneta. Ji pagaminta taip, kad tikimybė, jog moneta nukris herbu į viršų yra lygi  $3/4$ , o kad skaičiumi –  $1/4$ . Tada entropija bus skaičiuojama taip:

$$p_1 = 3/4$$

$$p_0 = 1/4$$

$$H = -(1/4 \log_2 1/4 + 3/4 \log_2 3/4)$$

$$H = -(1/4 (-2) + 3/4 (-0,415)), \text{ nes } \log_2 3/4 = -0,415$$

$H = 0,811$  bito vienam metimui.

**Užduotis**

Yra ypatinga moneta, kurią metant tikimybė, kad moneta atvirs herbu lygi 0,9, o skaičiumi – 0.1. Kam lygi informacijos entropija sužinojus vieno metimo rezultatą?

$$H = -(p_0 \log_2 p_0 + p_1 \log_2 p_1)$$

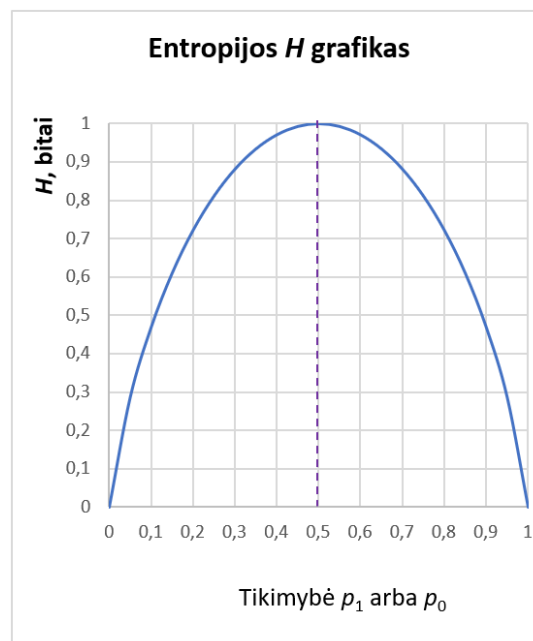
$$H = -(0,1 \log_2 0,1 + 9/10 \log_2 0,9)$$

$$H = 0,469$$

Intuityviai aišku, kad jei herbas iškrenta dažniau, tai apie baigtį yra žinoma daugiau, nei kai tikimybė buvo lygi 0,5. Jei tikimybė iškristi herbei būtų lygi vienetui, o skaičiui – nuliui, tai nebūtų jokio pasirinkimo.

Jei herbo pasirinkimo tikimybę bet koku atveju pažymėsime  $p_1$ , tai skaičiaus pasirinkimas lygus  $1 - p_1$ . Taigi, jei žinoma  $p_1$ , tai žinoma ir  $p_0$ . Galima apskaičiuoti entropiją  $H$  kiekvienu atveju ir nubraižyti grafiką  $H = H(p_1)$

$H$  reikšmė yra maksimali (lygi 1), kai  $p_1$  lygu  $1/2$  ir virsta nuliui, kai  $p_1$  lygi 1 arba 0. Ta pati kreivė gautųsi, jei  $p_1$  pakeistume  $p_0$ .



Jei šaltinis sukuria vieną iš galimų  $n$  simbolių arba  $n$  žodžių, tačiau nauji pranešimai nepriklauso nuo prieš tai buvusių, tokiu atveju entropija apibrėžiama kaip suma sandaugų:

$$H = - \sum_{i=1}^n p_i \log_2 p_i \text{ bitų vienam simboliui (arba pasirinkimui)}$$

Arba

$$H = \sum_{i=1}^n p_i \log_2 \frac{1}{p_i}$$

$p_i$  – tai tikimybė, kad bus pasirinktas  $i$ -asis simbolis. Kai  $n = 2$ , turime paprasčiausią jau nagrinėtą atvejį.

### Pavyzdys

Tarkime turime dvi monetas. Abi monetas išmetamos vienu metu. Yra keturios galimybės:

SS arba 1

SH arba 2

HS arba 3

HH arba 4

Kiekvienos baigties tikimybė vienoda ir lygi  $\frac{1}{4}$ , entropija:

$$H = -(1/4 \log_2 1/4 + 1/4 \log_2 1/4 + 1/4 \log_2 1/4 + 1/4 \log_2 1/4)$$

$$H = -(-1/2 -1/2 -1/2 -1/2)$$

$H = 2$  bitai išmetus 2 monetas vienu metu.

Visas baigtis galime aprašyti 2 dvejetainiais skaičiais (00, 01, 10, 11), todėl ir šiuo atveju entropija sutampa su dvejetainių skaitmenų skaičiumi.

Jei yra  $n$  simbolių ir kiekvieno tikimybė yra vienoda ir lygi  $1/n$ , skaičiuojant entropiją po sumos ženklų  $\Sigma$  bus  $n$  narių, kurių kiekvienas skaičiuojamas  $1/n \log_2 1/n$  ir visa suma bus lygi  $\log_2 1/n$ , o entropija

$$H = - \log_2 1/n \text{ arba } H = \log_2 n \text{ bitų simboliui}$$

Metant vieną kauliuką visų akių nuo 1 iki 6 iškritimo tikimybė vienoda ir lygi  $1/6$ . Reiškia entropija lygi  $\log_2 6 = 2,58$  bito vienam kauliuko metimui.

Bendresniu atveju imamas pasirinkimas vieno iš  $N$ -ženklio dvejetainio skaičiaus. Tokių dvejetainių skaičių yra  $2^N$ , todėl

$$n = 2^N$$

Todėl  $\log_2 2^N = N$  ir entropija šaltinio, kuris su vienoda tikimybe kuria vieną pasirinkimą iš  $N$ -ženklų dvejetainių skaičių yra lygi  $N$  bitų vienam skaičiui.

## Užduotis



Turime kauliuką su 8 sienomis . Yra 8 vienodų tikimybių galimybių, kad metant kauliuką iškris 1, 2, ..., 8.

Kokia entropija, kai sužinoma vieno metimo baigtis?

$$H = - \sum_{i=1}^8 \frac{1}{8} \log_2 1/8 = \log_2 8 = 3 \text{ bitai vienam metimui.}$$

Jei kauliuko sienų skaičių padidinsime iki 16, tai ir galimybių iškristi nuo 1 iki 16 padaugės 2 kartus.

Entropija bus  $H = - \sum_{i=1}^{16} \frac{1}{16} \log_2 1/16 = \log_2 16 = 4$  bitai vienam metimui. Entropija padidėja 1 bitu.

**Galimybės padidėjus lygiai 2 kartus, entropija padidėja 1 bitu.**

Jei sienų skaičius lieka toks pat (8 sienos), tai ant sienos gali būti užrašyti 2 kartus didesni skaičiai:  $A_x = \{2, 4, 6, 8, 10, 12, 14, 16\}$ . Kiekvienos vertės (sienos) iškritimo galimybė lieka tokia pati (1/8), todėl ir entropija nesikeičia: 3 bitai vienam metimui.

Tolesni kauliuko metimai nepriklauso nuo buvusių metimų sienų verčių, todėl tokios pranešimų vertės yra nepriklausomos (angl. *independent and identically distributed, iid*).

Jeigu paimsime rišlų tekstą, tai pranešimų sekos elementai nėra nepriklausomi, todėl kiekvienos raidės atsiradimas pranešime nebėra visiškai atsitiktinis – tokių pranešimų, sudarytų iš priklausomų elementų (raidžių) entropija yra mažesnė, nes sumažėja atsitiktinių galimybių arba neapibrėžtumų.

Jei bėgant laikui pranešimų šaltinis siunčia vienodus pranešimus, tai toks šaltinis vadinamas *stacionariu*.

### Bitai, šenonai ir bansai

Jei turime diskrečius kintamuosius, tai informacijos maksimumas, susijęs su šiais kintamaisiais yra  $m$  vienodų galimybių įgyti diskrečias ir nepriklausomas vertes logaritmas. Kadangi dvejetainis kintamasis gali įgyti  $m = 2$  būsenas, tai jis neša  $n = 1$  bitą informacijos. Naudojant bitus kaip informacijos matavimo vienetą, naudojamas logaritmas, kurio pagrindas yra 2.



Jei informacija matuojama ir skaičiuojama naudojant natūrinį logaritmą, kurio pagrindas yra  $e = 2,72$ , tai informacijos matavimo vienetai vadinami natsais (angl. *nats*). Jei logaritmo pagrindas yra 10, tada matavimo vienetas yra bansas (angl. *bans*).

Norint atskirti bito naudojimą kaip informacijos matavimo vienetą ir dvejetainį skaitmenį, vietoje bito kaip informacijos kiekio (entropijos) matavimo vieneto naudojamas šenonas (angl. *Shannon*), trumpinys *Sh*.

## Komunikavimo kanalas

Turime informacijos šaltinį, kuris kuria (generuoja) pranešimus. Pranešimas – tai tam tikra tvarka išrikiuotų  $k$  simbolių seka:  $\mathbf{s} = (s_1, \dots, s_k)$ , čia kiekvienas simbolis atitinka atsitiktine tvarka gautą skaičių ar raidę (žymima kaip atsitiktinis kintamasis  $S$ ). Visa seka pažymėta pajuodinta  $\mathbf{s}$  raide, o sekos simboliai turi apatinius indeksus ir apibrėžiami skliausteliais.

Kiekvienas simbolis yra išvestis atsitiktinio kintamojo  $S$ , kuris gali įgyti bet kurią vertę iš  $\alpha$  skirtingų simbolių:  $A_s = \{s_1, \dots, s_\alpha\}$ ,

Pranešimų šaltinio kiekvieno simbolio generavimo tikimybė apibrėžiama taip:

$p(S) = \{p(s_1), \dots, p(s_\alpha)\}$ ,  $\alpha$  – tai galimų generuoti simbolių skaičius

Visų tikimybių suma turi būti lygi vienetui:

$$\sum_{i=1}^{\alpha} p(s_i) = 1$$

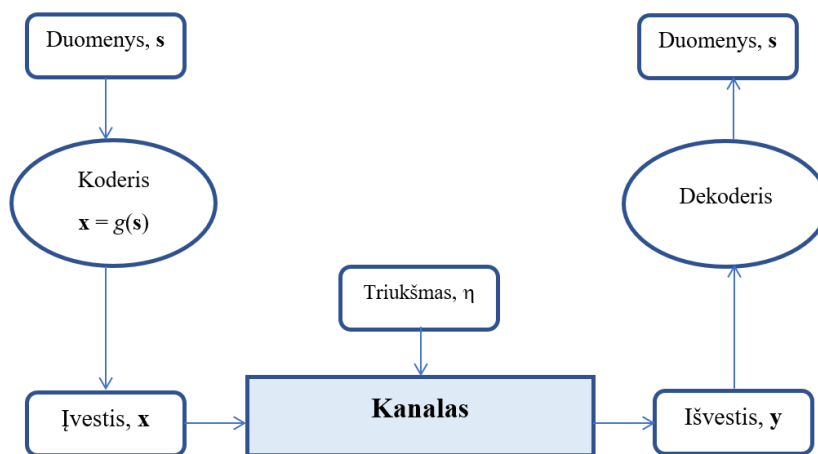
Komunikavimo kanalas naudojamas perduoti duomenims nuo jų įvesties iki išvesties. Jei duomenys kanalu perduodami be klaidų, tai komunikacija įvyksta sėkmingai.

Prieš perduodant kiekvienas pranešimas  $\mathbf{s}$  yra koderio užkoduojamas ir tai galima atvaizduoti kaip funkciją  $g$ , kuri pranešimą užkoduoja į kodinių žodžių seką  $\mathbf{x} = g(\mathbf{s})$ , čia  $\mathbf{x} = (x_1, \dots, x_n)$ , ir kiekvienas kodinis žodis yra reikšmė atsitiktinio kintamojo  $X$ , kuris gali įgyti bet kurią vieną iš  $m$  skirtingų kodų lentelės (angl. *code book*) reikšmių.

$A_x = \{x_1, \dots, x_m\}$ ,

Kiekvieno kodinio žodžio tikimybė yra pasiskirstymo (angl. *distribution*) tikimybė:

$p(S) = \{p(x_1), \dots, p(x_m)\}$ ,



Komunikacijos kanalas. Pranešimas  $s$  yra užkoduojamas kodiniais žodžiais  $x$  prieš patekdamas į perdavimo kanalą. Kanale koduotas pranešimas gali būti sugadinamas triukšmo  $\eta$  ir gaunama išvestis bus  $y = x + \eta$ . Pranešimo gavėjas dekoduoja išvestį  $y$ , kad atgautų įvestį  $x$ , kuri interpretuojama kaip pranešimas  $s$ .

Galime taip užkoduoti pranešimą, kad  $x$  būtų identiškas  $s$  ( $x = s$ ).

Tačiau keli simboliai gali būti apjungti į vieną kodinį žodį arba pranešimas gali būti suglaudintas (suspaustas, angl. *compressed*) atsisakant pasikartojimų. Tai reiškia, kad kodinių žodžių skaičius gali nesutapti su užkoduoto pranešimo ( $x = g(s)$ )  $s$  simbolių skaičiumi.

Jei iš suglaudintų duomenų galima visiškai tiksliai atkurti pirminį jų pavidalą, tada turime **nuostolingąjį glaudinimą** (angl. *lossless compression*). Taip glaudinami tekstai, programos ir kiti duomenys, kuriuose bent vieno bito praradimas gali sukelti nepageidaujamų padarinių.

**Nuostolingasis glaudinimas** (angl. *lossy compression*) – tai duomenų glaudinimo būdas, kai juos suspaudžiant prarandama dalis informacijos. Iš taip suglaudintų duomenų nebegalima visiškai tiksliai atkurti pirminio jų pavidalo. Tačiau galima pasiekti didesnę glaudinimo laipsnį, negu glaudinant nuostolingai. Taip dažniausiai glaudinami garso ir vaizdo failai ir kiti duomenys, kuriuose nedidelio informacijos kiekio praradimas nesukelia rimtesnių nepageidaujamų padarinių. Dar vartojama sąvoka *glaudinimas su praradimais*.

Kodas – tai simbolių ir juos atitinkančių kodinių žodžių sąrašas. Pavyzdys – seka dešimtainių skaičių, kurie didėja kas 3 vienetai:

Simbolis	Kodinis žodis	Simbolis	Kodinis žodis
$s_1 = 3$	$x_1 = 000$	$s_5 = 15$	$x_5 = 100$
$s_2 = 6$	$x_2 = 001$	$s_6 = 18$	$x_6 = 101$
$s_3 = 9$	$x_3 = 010$	$s_7 = 21$	$x_7 = 110$
$s_4 = 12$	$x_4 = 011$	$s_8 = 24$	$x_8 = 111$

Sekos simboliai (dešimtainiai skaičiai), užkoduoti kaip kodiniai žodžiai (dvejetainiai skaičiai).

Pranešimo, atkeliavusio iš kanalo išvestis yra  $\mathbf{y} = (y_1, \dots, y_n)$ . Kiekvienas išvesties elementas yra reikšmė kintamojo  $Y$ , kuris įgyja bet kurią vieną reikšmę iš  $m$  galimų:  $A_y = \{y_1, \dots, y_m\}$ .

Kiekvienos išvesties tikimybė apibrėžiama kaip pasiskirstymas tikimybių:

$$p(Y) = \{p(y_1), \dots, p(y_m)\}$$

Kiekviena sekos  $\mathbf{y}$  išvestis interpretuojama kaip sekos  $\mathbf{x}$  konkreti įvestis. Jei kanale pasireiškia triukšmai, išvesties  $y_j$  gali skirtis nuo perduoto kodinio žodžio  $x_j$ .

Jei ši išvestis dėl kanalo triukšmo yra su klaidomis, tai rezultatas laikomas klaida.

Dekoderis konvertuoja kiekvieną išvesties  $y$  seką į pranešimą, kuris su tam tikra tikimybe gali būti klaidingas. Kodavimo klaidų dažnis (angl. *error rate*) yra klaidingų išvesčių skaičius, padalintas iš visų galimų išvesčių skaičiaus.

Kanalo pralaidumas (angl. *channel capacity*) – tai ryšių kanalo pajėgumas praleisti tam tikrą duomenų kiekį per laiko vienetą. Matuojamas bitais (kilobitais, megabitais) per sekundę (b/s, Kb/s, Mb/s). Priklauso nuo kanalui išskirtos dažnių juostos pločio ir moduliavimo metodo.

Kanalo pralaidumas nebūtinai sutampa su informacijos perdavimo sparta kanalu. Sparta gali būti mažesnė, bet niekada nebus didesnė už kanalo pralaidumą. Kanalo pralaidumas – tai maksimali informacijos perdavimo sparta.

Pranešimas, sudarytas iš simbolių sekos  $\mathbf{s} = (s_1, \dots, s_k)$  yra užkoduojamas pagal funkciją  $\mathbf{x} = g(\mathbf{s})$  į seką kodinių žodžių, kurių skaičius nebūtinai lygus simbolių skaičiui. Kodinių žodžių seka keliauja komunikacijos kanalu, o gauta išvestis  $\mathbf{y} = (y_1, \dots, y_n)$  yra dekoduojama į pranešimą  $\mathbf{s}$ .

Tarkime, kad ryšio kanalu per 1 sekundę galima perduoti  $C$  impulsų (pavyzdžiui, aukšta įtampa, žema įtampa). Vadinasi per 1 sekundę tokiu kanalu galima perduoti  $C$  dvejetainių ženklų ir kiekvienas dvejetainis ženklas neša 1 bitą informacijos. Galima sakyti, kad tokio kanalo informacijos perdavimo sparta yra  $C$  bitų per 1 sekundę. Jei pranešimų šaltinio entropija  $H$ , išmatuota bitais per sekundę, yra mažesnė už  $C$ , tai tokiu kanalu galima perduoti informaciją koduojant vadinamuoju Hafmeno metodu (angl. *Huffman Coding*).

Pagrindinė informacijos teorijos teorema buvo Šenono suformuluota taip: Jei informacijos šaltinio entropija yra  $H$ , o kanalo informacijos perdavimo sparta  $C$ , tai galima koduoti šaltinio pranešimus taip, kad simbolių perdavimo vidutinė sparta būtų  $C / H$ . Perduoti didesne sparta nei  $C / H$  neįmanoma.

Tai reiškia, kad bet koks diskretus kanalas, kuriuo perduodama informacija, turi jam būdingą pralaidumo gebą  $C$  (informacijos perdavimo spartą). Ergodinis šaltinis turi apibrėžtą entropiją  $H$ . Jei  $H$  mažesnė už  $C$  arba jai lygi, tai šiuo kanalu įmanoma perduoti šaltinio pranešimus. Jei  $H$  didesnė už  $C$ , tai perdavimas neįmanomas.

Galime padaryti išvadą, kad išmatavus šaltinio entropiją bitais, aiškiai nurodoma, kiek dvejetainių ženklų ar būsenų (išjungta-įjungta) reikia ženklui per sekundę, kad būtų įmanoma perduoti pranešimus, kuriuos kuria šaltinis. Tuo pačiu žodis bitas – tai tiesiog dvejetainio skaitmens trumpesnis pavadinimas.

Informacijos teorijoje bitu matuojamas informacijos kiekis ir bitu vadinamas dvejetainis skaitmuo.

**Praktikoje entropija, išreikšta bitais, skiriasi nuo naudojamų dvejetainių skaitmenų kiekio.**

Pavyzdžiui, turime ryšio kanalą, kuriuo per 1 s galima perduoti 10 000 atsitiktinai pasirinktų dviejų būsenų impulsų. Žinoma, tokio kanalo pralaidumas bus 10 000 bitų per sekundę. Bet jei per 1 sekundę šiuo kanalu bus perduota 10 000 vienodų (nesikeičiančių) impulsų, tai perduotas informacijos kiekis (entropija) bus 0 bitų per 1 sekundę, nors kanalu perduodama 10 000 impulsų.

**Diskretaus ir idealaus (be triukšmo) kanalo pralaidumo formulė**

Šenonas diskretaus ir idealaus kanalo pralaidumą apibrėžė kaip dvejetainį logaritmą  $N(t)$  koduotų pranešimų, kurie **gali būti** perduodami kanalu per laiko  $t$  intervalą:

$$C = \lim_{t \rightarrow \infty} \frac{N(t)}{t} \text{ bitų per sekundę (b/s)}$$

Bendruoju atveju kodiniai žodžiai gali būti pakankamai ilgi, todėl laikas  $t$  artėja į begalybę, paprastai  $t$  prilyginamas vienai sekundei.

Pavyzdžiui, žmogus per sekundę gali pasakyti vieną žodį, tačiau šis žodis – vienas iš žodyno, kuriame yra 1024 žodžiai ( $1024 = 2^{10}$ ). Tokio žmogaus (pranešimų šaltinio) kanalo vidutinis pralaidumas (sparta) per sekundę bus  $\log_2 1024 = \log_2 2^{10} = 10$  bitų per sekundę.

Tokia sparta pasiekama, jei žmogus kuria žodžius, kurie visi yra vienodai tikėtini, nepriklauso vienas nuo kito. Jei tam tikri žodžiai pasikartoja dažniau arba yra tiesiog prognozuojami, pavyzdžiui „labas vakaras“, tuomet pranešimai perduoda mažiau informacijos. Bet koks perduodamos informacijos kiekio sumažėjimas reiškia, kad kanalu perduodama mažiau informacijos, nei yra tokio kanalo apibrėžta sparta.

Tarkime, turime kanalą, kuris per vieną sekundę perduota 9 dvejetainius skaitmenis. Kiekviena devynių dvejetainių skaitmenų seka sukuria vieną iš  $N = 2^9$  dvejetainių skaičių. Pagal Šenono formulę gauname, kad

$$C = \frac{2^{9t}}{t} = \frac{9t}{t} = 9 \text{ b/s}$$

Jei kanalas gali perduoti 9 dvejetainius skaitmenis per sekundę, tai maksimumas informacijos, kurį kanalas gali perduoti ir yra devynių bitų per sekundę vidurkis.

Idealaus kanalo **pralaidumas C (sparta)** yra maksimumas bitų, kuriuos jis gali perduoti per sekundę ir C yra lygus skaičiui dvejetainių skaitmenų, kuriuos per sekundę kanalas gali perduoti.

## Duomenų glaudinimas



Turime du šešiasienius kauliukus. Jei abu kauliukus išmesime vienu metu, tai 36 iškritusių akių sumos galimybės pasiskirstys taip:

Simbolis	Suma	Iškritusios akys	Dažnis	$p(s_i)$	$\log_2 1/p(s_i)$	Huffman kodas x
$s_1$	2	1:1	1	0,028	5,170	10000
$s_2$	3	1:2; 2:1	2	0,056	4,170	0110
$s_3$	4	1:3; 3:1; 2:2	3	0,083	3,585	1001
$s_4$	5	2:3; 3:2; 1:4; 4:1	4	0,111	3,170	001
$s_5$	6	2:4; 4:2; 1:5; 5:1; 3:3	5	0,139	2,848	101
$s_6$	7	3:4; 4:3; 2:5; 5:2; 1:6; 6:1	6	0,167	2,585	111
$s_7$	8	3:5; 5:3; 2:6; 6:2; 4:4	5	0,139	2,848	110
$s_8$	9	3:6; 6:3; 4:5; 5:4	4	0,111	3,170	010
$s_9$	10	4:6; 6:4; 5:5	3	0,083	3,585	000
$s_{10}$	11	5:6; 6:5	2	0,056	4,170	0111
$s_{11}$	12	6:6	1	0,028	5,170	10001

Entropija bus:  $H(S) = \sum_{i=1}^{11} p(s_i) \log_2 \frac{1}{p(s_i)} = 3,27$  bitų / simboliui. Tai reiškia, kad vidutinis informacijos kiekis išmetus du kauliukus yra 3,27 bito. Todėl kanalo sparta turi užtikrinti galimybę per 1 sekundę perduoti 3,27 dvejetainius skaičius, tenkančius vienam simboliui.

Jei kiekvieną simbolį koduosime 4 dvejetainiais skaitmenimis, tai kodinio žodžio ilgis  $L(X) = 4$ , o kodavimo efektyvumas

$$\frac{H(S)}{L(X)} = \frac{3,27 \text{ bito/simboliui}}{4 \cdot \text{dvejetainiai skaitmenys/simboliui}} = 0,818 \frac{\text{bito}}{\text{dvejetainiui skaitmeniui}}$$

### Hafmano (Huffman) kodas

1952 metais sukurtas Hafmano kodas remiasi tuo, kad dažnai pasikartojantys simboliai (pranešimai) turi trumpus kodinius žodžius, o retai kuriami pranešimai – ilgus kodinius žodžius. Tokiu metodu koduojant siunčiami pranešimai labai artimi jų entropijai (informacijos kiekiui).

Bet kuriam diskrečiam kintamajam  $X$  su entropija yra  $H(X)$  egzistuoja simbolio kodas, kurio ilgis  $L(X)$  tenkina nelygybę:

$$H(X) \leq L(X) < H(X) + 1$$

Naudojant Hafmano kodą informacija neprarandama.

Hafmano kodu yra užkoduoti 11 simbolių, kurie gaunami kaip pranešimai išmetus du kauliukus vienu metu. Dviejų šešiasienių kauliukų metimus aprašančioje lentelėje tai paskutinis stulpelis. Jei paskaičiuosime vidutinį kodinio žodžio ilgį  $L(X)$ :

$$L(X) = \sum_{i=1}^{11} p(x_i) L(x_i) = 3,31 \text{ dvejetainis skaitmuo / simboliui}$$

Hafmano kodo efektyvumas:

$$\frac{H(S)}{L(x)} = \frac{3,27 \text{ bito/simboliui}}{3,31 \cdot \text{dvejetainiai skaitmenys/simboliui}} = 0,99 \frac{\text{bito}}{\text{dvejetainiui skaitmeniui}}$$

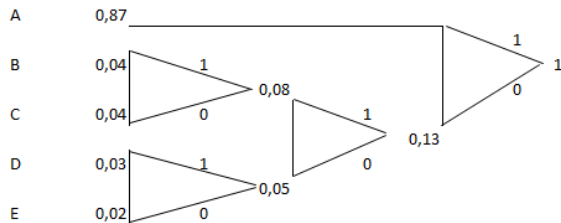
Taikant Hafmano kodą pradedama nuo dviejų mažiausiai tikėtinų simbolių ir kuriant tų simbolių įsivaizduojamą kompozitą. Tada šie du mažiausiai tikėtini simboliai pakeičiami kompozitu, kurio tikimybė yra šių pradinių simbolių tikimybių suma.

Šis procesas kartojamas, kol lieka tik vienas simbolis (tikimybė lygi vienetui). Kai du simboliai komponuojami, visada viršutinė šaka žymima 1, o apatinė 0. Taip gaunamas medis iš kairės į dešinę. Kai medis užpildytas, kodai skaitomi iš dešinės į kairę (žr. 1-ą pavyzdį, mažiausia tikimybė čia žymima 0). Hafmano medis vaizduojamas ir vertikaliai, tuomet kodai skaitomi nuo viršaus žemyn (žr. 2-ą pavyzdį, mažiausia tikimybė čia žymima 1).

## 1 pavyzdys

Turime 5 pranešimų šaltinio kuriamus simbolius, kurių tikimybės yra tokios:

A	B	C	D	E
0,87	0,04	0,04	0,03	0,02



Simbolių kodai

A	B	C	D	E
1	011	010	001	000

## 2 pavyzdys

Turime 5 pranešimų šaltinio kuriamus simbolius, kurių tikimybės yra tokios:

A	B	C	D	E
0,1	0,35	0,16	0,2	0,19

Išrikiuojame didėjimo tvarka (galima ir mažėjimo):

A	C	E	D	B
0,1	0,16	0,19	0,2	0,35

Jungiamo mažiausias tikimybės A ir C (AC – 0,26) ir vėl išrikiuojame:

E	D	AC	B
0,19	0,2	0,26	0,35

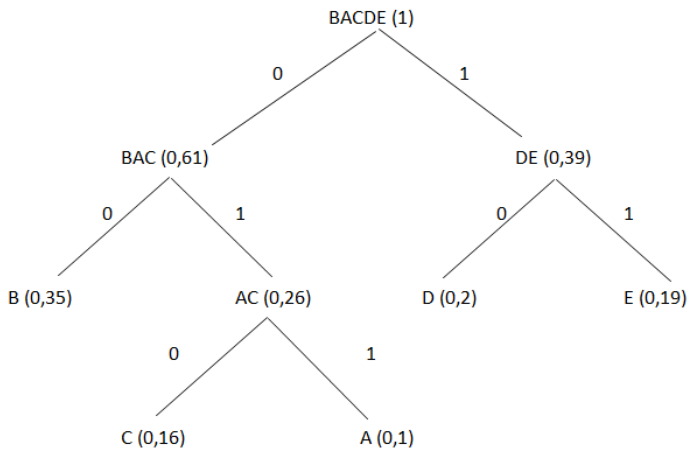
Vėl jungiamo mažiausias tikimybės E ir D (DE – 0,39) ir vėl išrikiuojame:

AC	B	DE
0,26	0,35	0,39

Vėl jungiamo mažiausias tikimybės AC ir B (BAC – 0,61) ir vėl išrikiuojame:

BAC	DE
0,61	0,39

Sujungus paskutinį kartą gauname vienetą (BACDE – 1)



Ženklas	Tikimybė	Kodinis žodis
A	0,1	011
B	0,35	00
C	0,16	010
D	0,2	10
E	0,19	11

### Kaip perduoti garso signalo informaciją dideliais atstumais?

Buvo aptarta, kad viena iš svarbiausių informacijos teorijos teoremų (*Sampling theorem*) sako, kad bet kokį tolydų signalą galima visiškai atvaizduoti ir atkurti šio signalo matavimų seka, kai matuojama vienodais laiko tarpais. Matuoti reikia tokiais laiko intervalais, kad jie būtų mažesni ar lygūs didžiausio dažnio, esančio signale, periodo pusei.

Tarkime, kad garso signalo dažniai kinta nuo 0 iki 4000 Hz, tai per 1 sekundę reikia matuoti 8000 kartų.

Televizijos signalo dažniai yra nuo 0 iki 4 MHz, todėl reikia imti 8 000 000 matavimų per 1 s.

Jei signalo dažnio diapazonas  $f$  Hz, tai norint tiksliai atvaizduoti tolydžiai kintantį signalą, reikia imti mažiausiai  $2f$  matavimų per sekundę.

Nebūtina signalą atvaizduoti absoliučiai tiksliai. Pavyzdžiui, atvaizduojant kalbą, pakanka 1 proc. tikslumo.



## Informacijos perdavimas palydovine televizija

Televizijos stotis perduota koduotą ir suglaudintą signalą į palydovą. Televizoriuje yra speciali įranga, kuri gautą signalą dekoduoja ir ekrane matome vaizdus ir girdime garsus.

Skaitmeninė kamera įrašo originalius vaizdo duomenis maždaug 1,5 gigabitų per sekundę sparta. Tai atitinka 1500 milijonų dvejetainių skaitmenų. Laikome, kad ekrane yra 1920 horizontalių eilučių ir 1080 vertikalų. Per sekundę signalas apibėga kiekvieną ekrano elementą 30 kartų. Kiekvienas elementas (taškas) ekrane yra vaizduojamas trimis pikseliais: raudonu, žaliu ir mėlynu. Kiekvienos spalvos intensyvumas kinta nuo 0 iki 255. Reiškia, kiekvienos spalvos pikselis neša 8 bitus informacijos, nes 256 skirtingi dvejetainiai skaičiai yra aštuonženkliai ( $2^8 = 256$ ). Iš viso ekrane yra apie 6 milijonus pikselių ( $3 \times 1920 \times 1080 = 6\,220\,800$ ). Šį skaičių padauginę iš 8 gauname beveik 50 milijonų dvejetainių skaitmenų vienam ekrano vaizdui, o per sekundę tokių vaizdų įrašoma 30. Gauname, kad kamera užfiksuoja 1500 milijonų dvejetainių skaitmenų per sekundę arba 1500 milijonų bitų per sekundę. Tačiau signalo perdavimo kanalo sparta yra tik 19,2 dvejetainių skaitmenų per sekundę.

Signalai yra glaudinami ir glaudinami nuostolingai (angl. *lossy compression*). Paprastai glaudinama naudojant MPEG (angl. *Moving Pictures Experts Group*) kodavimo schemas.

Naudojantis keliais algoritmais garso įrašas suspaudžiamas visiškai pašalinant žmogaus ausiai negirdimus garsus. Šitaip gaunama garso įrašo kokybė beveik nesiskiria nuo nesuglaudinto garso įrašo kokybės, o suglaudinto failo dydis sumažėja iki 10–12 kartų. Analogiškai glaudinamas ir vaizdas.

Nuostolingi glaudinimo būdai pakankamai sudėtingi, tačiau žmogus žiūrėdamas ir klausydamas nejaučia informacijos praradimo.

## Radio ryšys

Garsas – labai patogi ir populiari žmonių bendravimo priemonė. Gamta žmonijai suteikė puikius prietaisus – balsą ir ausis, kurie perduoda ir priima informaciją. Šie prietaisai nėra tobuli, o garsas aplinka sklinda labai trumpais atstumais.

Radio bangos gali perduoti garsą ar bet kokią kitą siunčiamą informaciją dideliais atstumais ir labai greitai (šviesos  $c = 3 \cdot 10^8$  m/s greičiu). Radio bangų dažnis užima platų diapazoną elektromagnetinių bangų spektre: nuo dešimčių kilohercų iki tūkstančių megahercų, o žmogus girdi dažnius nuo 20 Hz iki 20 000 Hz.

Radijo ryšio sistema gali perduoti skirtingų rūšių informaciją, kuri gali būti garsas, vaizdas ar įvairiai pateikti duomenys. Siunčiamas signalas sudarytas iš dviejų – vienas jų yra pastovaus dažnio ir pastovios amplitudės radijo signalas, kuris dažniausiai vadinamas nešančiąja banga (nešančiuoju dažniu), o kitas – informacijos signalas. Priešingai nei nešančioji banga, informacijos signalas nuolat kinta, jis gali visiškai nutrūkti ir po to vėl atsirasti – priklausomai nuo to, kaip kinta informacija.

Informacijos signalo pavyzdys – garsinis signalas, kuris formuojamas, kai kalbama naudojantis mikrofonu, kuris keičia grandinės elektrinę įtampą (tuo pačiu ir grandinė tekančios elektros srovės stiprį). Oru sklindanti garso banga savo informaciją perduoda elektriniam signalui.

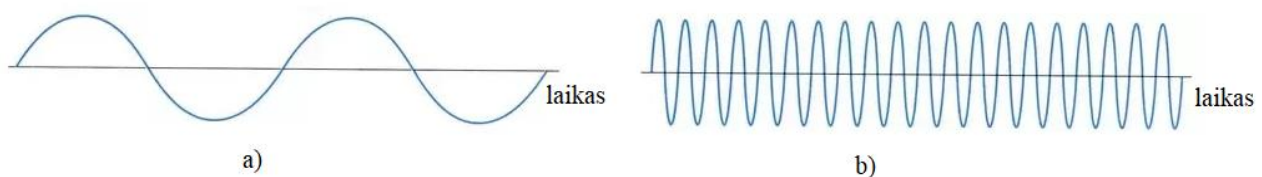
Siunčiamas signalas – tai informacijos signalas drauge su nešančiąja banga. Šių signalų susiejimas vadinamas moduliacija.

## Moduliacija

Siunčiamą signalą galima moduluoti įvairiais būdais. Paprasčiausias būdas, kuriuo galima perduoti užkoduotą informaciją nesudėtingai, pavyzdžiui Morzės abėcėle, yra tiesiog įjungti ir išjungti radijo signalą atitinkamais laiko momentais.

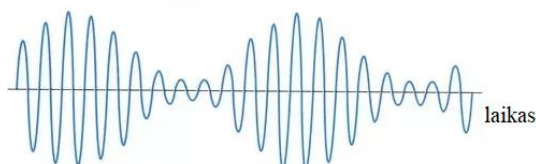
Norint perduoti sudėtingesnę garsinę informaciją, pavyzdžiui, kalbą arba muziką, tenka naudoti sudėtingesnius moduliacijos metodus. Radijo transliacijai naudojami du pagrindiniai moduliacijos būdai: amplitudinė moduliacija (AM) ir dažninė moduliacija (DM).

**Amplitudinė moduliacija** gaunama, kai informacijos signalas (a) sujungiamas su nešančiąja banga (b). Nešančiosios bangos amplitudė pradeda kisti.



Harmoniškai kintantys žemo dažnio (a) ir aukšto dažnio (b) signalai

Modulatoriaus – prietaiso, kuriame signalai sujungiami – išėjime gaunamas gana sudėtingas signalas (c):



c)

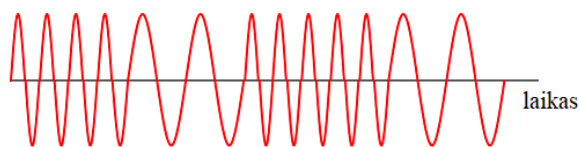
Moduliuotas signalas (amplitudinė moduliacija)

Tarkime, kad garso signalas – tik paprastas vieno dažnio tonas. Tarkime, jo dažnis  $f_m$ , o nešančiojo signalo radijo dažnį pažymėkime  $f_c$ . Nešančiojo signalo dažnis yra žymiai didesnis. Modulatorius taip sumaišo signalus, kad išėjime gaunami dviejų skirtingų dažnių signalai, atitinkantys šių signalų sumą ir skirtumą. Tad iš viso turime 4 dažnių skirtingus signalus:

$$f_c, f_m, (f_m - f_c), (f_m + f_c),$$

Trys signalai yra panašaus dažnio, tačiau garso signalo dažnis  $f_0$  yra žymiai mažesnis. Paskutinė siųstuvo dalis yra stiprintuvas, kuris garsinio dažnio signalą nufiltruoja, o minėti trys signalai keliauja į anteną.

**Dažninė moduliacija** gaunama, kai moduluojamas nešančiojo signalo dažnis (d). Tai sudėtingesnė moduliacija, tačiau turinti privalumų. Dažnine moduliacija perduodamas signalas yra geresnės kokybės, jo beveik neiškraipo trikdžiai. Aparatūroje dažninės moduliacijos diapazonai žymimi angliškais raidėmis FM (angl. *frequency modulation*).



d)

Moduliuotas signalas (dažninė moduliacija)

Trikdžiai – tai visokie atsitiktiniai signalai, kurie atsiranda elektrinėse grandinėse arba juos registruoja radijo imtuvo antena. Trikdžius taip pat generuoja elektros įrenginiai, pavyzdžiui, elektriniai grąžtai ar dulkių siurbiai.

Bet koks trikdžių šaltinis yra signalas, kuris sąveikauja su siunčiamo (gaunamo) signalo amplitude. Todėl dažnai klausantis radijo laidų, kurios perduodamos amplitudinės moduliacijos metodu, girdime traškesius ir švilpesius.

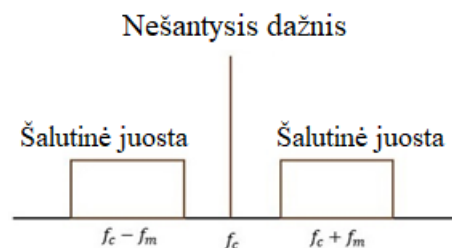
Dažninės moduliacijos signalas yra kintamojo dažnio signalas, todėl juo pernešamos informacijos kokybei neturi įtakos trikdžių sukelti tokio signalo amplitudės iškreipimai.

Dažninės moduliacijos metodas naudojamas perduoti stereo signalams, kur reikia aukštos garso atkūrimo kokybės.

### Juostos plotis

Realiai perduodamas garso signalas yra daug sudėtingesnis nei ką tik nagrinėtas vieno tono (dažnio) signalas. Realų garso signalą, ar tai būtų žmogaus kalba, ar muzikantų grupės atliekamas kūrinys, sudaro tam tikrų dažnių sritis.

Amplitudinėje moduliacijoje garsas perduodamas nuo 300 Hz iki 4 kHz. Šis intervalas vadinamas **garso dažnių juosta**. Kai visa juosta naudojama amplitudinei nešančiojo signalo moduliacijai, išėjime gauto signalo spektrą sudaro nešantysis dažnis ir greta jo dvi šalutinės juostos.



Juostos plotis =  $(f_c + f_m) - (f_c - f_m) = 2f_m$ , čia  $f_m$  – maksimalus garso dažnis.

Paveikslo kairėje esanti juosta vadinama žemesnio dažnio šalutine juosta, o dešinėje – aukštesnio dažnio šalutine juosta. Centre – nešantysis signalas.

Kiekvienai radijo stočiai skiriama ne 8 kHz, o 10 kHz dažnių juosta, kad radijo stočių siunčiami signalų dažniai nepersiklotų.

Dažninės moduliacijos signalo spektras yra gerokai sudėtingesnis nei amplitudinės moduliacijos signalo spektras. Jame yra daug daugiau šoninių juostų, o pati juosta platesnė.

### Amplitudinės moduliacijos ir dažninės moduliacijos apribojimai

Neįmanoma pasiekti techninio lankstumo ir duomenų perdavimo patikimumo. Amplitudinės moduliacijos ir dažninės moduliacijos signalai yra **atviri**: išspinduliuotus siųstuvo signalus gali išgirsti bet kas, turintis suderintą imtuvą. Abi šios duomenų perdavimo sistemos veikia realiu laiku – signalas gali būti priimamas tik jo perdavimo metu. Radijo bangų sritis beveik perpildyta, todėl signalo perdavimo aplinkybės vis labiau komplikuojasi. Todėl vis sparčiau diegiami skaitmeniniai komunikacijos metodai.

## Skaitmeniniai signalai

Skaitmeniniai signalai yra ne tokie jautrūs elektriniams trikdžiams, gali skliti didžiuliais atstumais, taip pat užima siauresnę dažnių juostą.

Vaizdai ir garsai yra paverčiami į skaitmeninius signalus (kodus), kurie perduodami kaip impulsai. Tada atitinkami imtuvo įrenginiai impulsus vėl paverčia į pradinis vaizdus ir garsus. Skaitmeninius signalus lengva priimti. Vienintelis dalykas, ką turi galėti imtuvas, tai užregistruoti, ar impulsas yra įjungtas, ar išjungtas (ar signalo lygis yra aukštas, ar žemas).

Radijo bangomis perduodami skaitmeniniai signalai kinta.

Pradinis taisyklingos formos impulsas, nusklidęs tam tikrą atstumą, pasikeičia. Tačiau net ir šiuo atveju, jei tik detektorius (imtuvas) įstengia atskirti žemą signalo lygį nuo aukšto, tai pradinis signalas gali būti tiksliai atkurtas.

Analoginiams (tolydžiai kintantiems) signalams nebūdingi staigūs signalo šuoliai nuo aukšto iki žemo lygio. Jie kinta tolygiai, todėl lengvai sutrikdomi ir jais sklindanti informacija lengvai iškraipoma esant daug mažesniam atstumui.

## Signalų keitimas

Daugelis perduodamų signalų yra analoginiai. Kalba ir muzika kinta laikui bėgant tolydžiai (nenutrūkstamai). Perduodant informaciją analoginiai signalai keičiami skaitmeniniais.

## Kodinė impulsinė moduliacija

Analoginius signalus pakeisti skaitmeniniais galima keliais būdais. Vienas jų – kodinė impulsinė moduliacija.

Pirmiausia reikalingas įrenginys, kuris analoginį signalą (garsą, muziką, vaizdą) pakeistų į elektrinį analoginį signalą, kurio įtampa ir amplitudė kinta laike.

Kodinė impulsinė moduliacija yra kelių pakopų. Pirmiausia analoginio signalo amplitudė (įtampa) daug kartų išmatuojama labai mažais laiko tarpais. Šis etapas (procesas) vadinamas **diskretizacija** arba imčių parinkimu. Tipinis garsinio dažnio analoginis signalas išmatuojamas 8000 kartų per sekundę. Tai reiškia, kad kas 125  $\mu$ s ( $1 \text{ s} / 8000 = 125 \mu\text{s}$ ) atliekamas vienas matavimas. Matavimo rezultatas – labai trumpi impulsai, kurių visuma atkartoja pradinio analoginio signalo formą.

Impulsų skaičius per sekundę vadinamas **išrankos dažniu**. Šioje pakopoje gautas signalas vadinamas impulso amplitudė moduliotu signalu.

Per kitą pakopą kintamos amplitudės impulsai paverčiami dvejetainiu kodu.

Didžiausia analoginio signalo įtampos vertė sudalijama į tam tikrą skaičių lygių. Šis procesas vadinamas **kvantavimu**, o skirtingos amplitudės vertės – kvantavimo lygiais. Kiekvienam kvantavimo lygiui priskiriamas dvejetainis skaičius. Pavyzdžiui jei maksimali analoginio signalo įtampa 10 voltų, tai naudojant 16 kvantavimo lygių, po nulio pirmasis lygis bus 0,625 V ( $10 / 16 = 0,625$ ).

Tada kiekvienas išmatuotas impulsas palyginamas su šiais lygiais ir kiekvienam impulsui priskiriamas artimiausias jo amplitudę atitinkantis kvantavimo lygis. Kiekvienam amplitudė moduluotam impulsui suteikiamas skaitmeninis kodas – dvejetainis skaičius. Dvejetainio skaičiaus dydis (jį sudarančių bitų skaičius) priklauso nuo to, kiek yra kvantavimo lygių. Jei yra 8 lygiai, tai užtenka 3 bitų, jei 16 lygių – 4 bitų. Kiekvienas impulsas atvaizduojamas vienodu bitų skaičiumi. Dvejetainis kodas vadinamas žodžiu.

Bet kuris dvejetainis kodas perduodamas nuoseklia seka, pradedant nuo vieneto.

### **Užduotis**

Apskaičiuokite signalo, kurio išrankos dažnis 8000 Hz, perdavimo spartą, jei naudojama 16 kvantavimo lygių.

Norint 16 kvantavimo lygių užkoduoti dvejetaine sistema, reikia 4 bitų ilgio skaičių.

perdavimo sparta = išrankos dažnis × bitų skaičius vienam matavimui = 32 kb per sekundę

### **Skaitmeninis analoginis keitimas**

Įtūve informacija išskiriama priešinga tvarka. Gauta kodinės impulsinės moduliacijos signalo kiekvienas dvejetainis žodis dekoduojamas – paverčiamas atitinkamos amplitudės trumpu impulsu. Visus impulsus nuosekliai sujungus gaunamas analoginis signalas, kurio forma idealiu atveju turi būti tokia pati, kaip buvo siųstuvo įėjime.

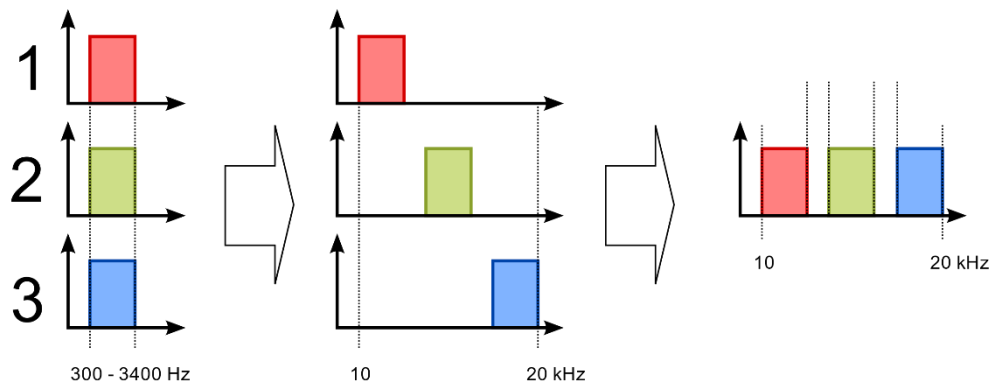
Pradinio signalo atkūrimo kokybė priklauso nuo analoginio signalo kvantavimo lygių skaičiaus. Kuo daugiau lygių, kuo jie yra arčiau vienas kito, tuo mažesnė paklaida daroma priskiriant impulsų amplitudei dvejetainį kodą. Tačiau kai yra daug lygių, dvejetainiams kodams reikia vis daugiau bitų. Norint perduoti ilgesnius dvejetainius kodus per tą patį laiką, tenka mažinti dvejetainio skaitmens (0 arba 1) impulso trukmę.

### **Multipleksavimas**

Jau aptarėme skirtingus informacijos perdavimo būdus. Dabar išsiaiškinkime: ar galima, pavyzdžiui, šviesolaidžiu tuo pačiu metu pasiųsti daugiau nei vieną signalą? Arba, kaip

palydovinė ryšių technika vienu metu gali patikimai sujungti tūkstančius telefonų? Tai įgyvendinama naudojant procesą, vadinamą multipleksavimu. Multipleksavimas – tai kelių skirtingų signalų sujungimas ir siuntimas kartu.

### Dažninis multipleksavimas



Dažninis multipleksavimas iš esmės yra analoginė technologija. Multiplekseryje keli signalai sujungiami į vieną multipleksinį signalą. Signalų nešančiųjų dažnių diapazonai yra skirtingi. Paveiksle parodyti trys signalai, kurių nešančieji signalai užima intervalą nuo 10 kHz iki 20 kHz. Vienas iš labiausiai paplitusių taikymų yra tradicinis radijo ir televizijos transliavimas iš antžeminių, mobiliųjų ar palydovinių stočių arba kabelinės televizijos. Klientą pasiekia tik vienas kabelis, tačiau paslaugų teikėjas tuo kabeliu gali vienu metu siųsti kelis televizijos kanalus ar signalus visiems abonementams be trukdžių. Imtuvai turi nusistatyti atitinkamą dažnį, kad pasiektų norimą signalą.

### Laikinis multipleksavimas

Skaitmeniniai signalai dar multipleksuojami ir laikiniu multipleksavimu.

Pavyzdyje, kuriame nagrinėjome kodinę impulso moduliaciją, analoginio signalo skleidimas 8 kHz dažnis buvo nuskaitomas. Tai reiškia, kad signalas matuojamas kas 125  $\mu$ s. Kadangi kiekviena imtis atvaizduojama 2–3  $\mu$ s impulsu, tai tarp impulsų yra didesni nei 120  $\mu$ s tarpai, kurių metu sistema neperduoda jokios informacijos – sistema tuščiai eikvoja laiką.

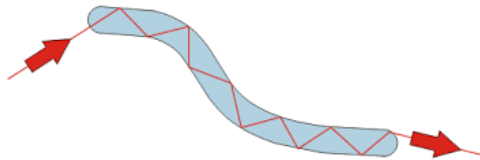
Laikinio multipleksavimo metu į laisvą vietą įterpiami kiti signalai.

### Šviesolaidžiai

Šviesolaidžiuose elektriniai signalų impulsai paverčiami trumpesnėmis nei radijo elektromagnetinėmis bangomis. Dažniausiai naudojami infraraudonieji spinduliai.

Šviesolaidis (optinė skaidula) gaminamas iš plono stiklinio cilindro, kuriuo skersmuo apie 5 mm. Tai šviesolaidžio šerdis, įvilкта į mažesnio lūžio rodiklio stiklo sluoksnį. Jei šviesa patenka į šerdį pakankamai dideliu kampu, tai dėl visiško vidaus atspindžio nuo šerdies iki apvalkalo ribos šviesa tarsi uždara šerdyje.

Šviesos sklindimas šviesolaidžiu pavaizduotas paveikslėlyje:

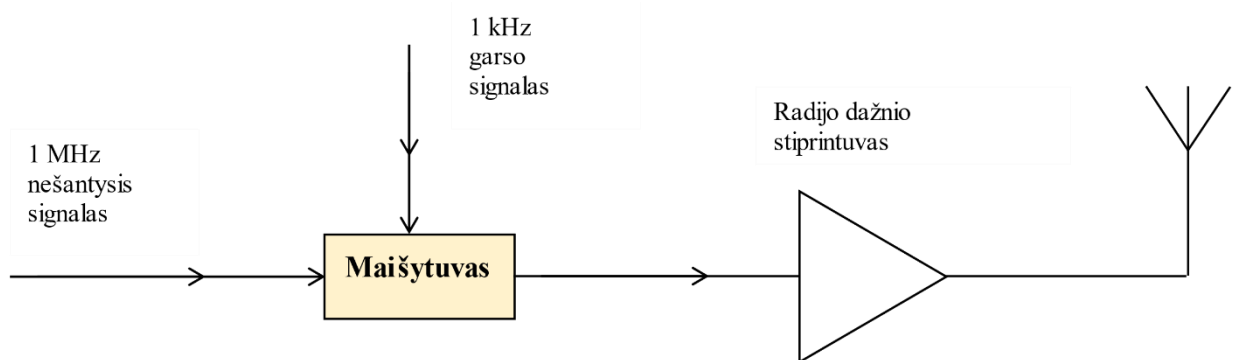


Šviesolaidžiu sklindantys signalai neiškraipomi, nesukuria magnetinio lauko, kaip laidininku tekanti elektros srovė.

Antžeminėje palydovinio ryšio stotyje optinis signalas paverčiamas elektriniu signalu, tačiau pastarasis išlieka skaitmeninis. Tuomet šis signalas mikrobangų pavidalu perduodamas į palydovą, kuris skrieja stacionaria Žemės orbita (jo padėtis virš Žemės paviršiaus nekinta). Palydovas retransliuoja signalą žemyn į kitas antžemines palydovinio ryšio stotis.

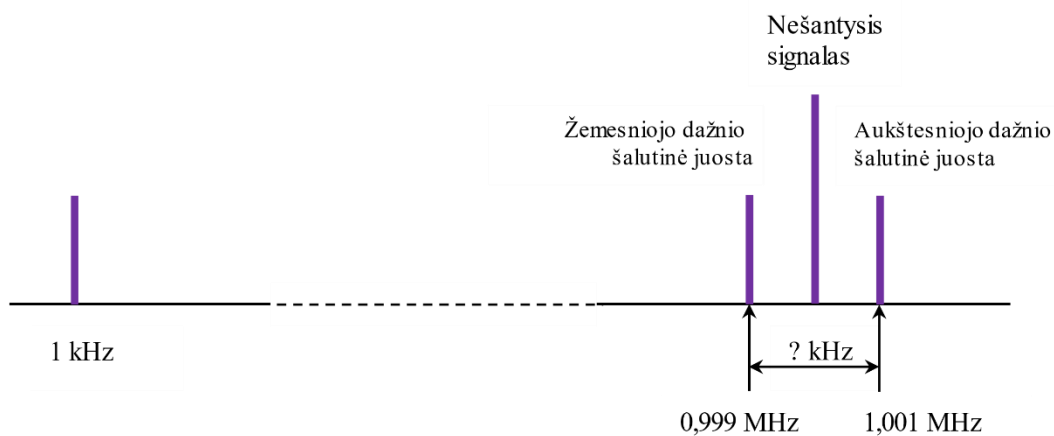
Šiuo metu tarpžemyniniai telefono pokalbiai perduodami ne per ryšių palydovą, o vandenyno dugne nutiestu šviesolaidiniu kabeliu. Taip yra dėl didžiulio šviesolaidinių sistemų pralaidumo.

## Užduotys



Amplitudinės moduliacijos sistemos blokinė schema





Dažnių spektras, kuris atsiranda 1 MHz nešantįjį signalą moduluojant 1 kHz garso signalu

Radijo dažnio stiprintuvas nestiprina žemiausiojo dažnio signalo. Koks yra perduodamo signalo juostos plotis?

*Atsakymas*

Juostos plotis 2 kartus didesnis už garso signalo dažnį, todėl lygus 2 kHz.

Garsas pakeičiamas garsiniu signalu, kurį sudaro įvairūs dažniai nuo 300 Hz iki 3,4 kHz.

Kokius dažnius generuoja maišytuvas, moduluodamas 1) 300 Hz dažniu ir 2) 3,4 kHz dažniu?

Koks juostos plotis?

Kiek radijo stočių, transliuojančių tokius signalus, gali tilpti į dažnių diapazoną nuo 800 kHz iki 1200 kHz?

*Atsakymas*

1) 300 Hz, 1 MHz, 0,9997 MHz, 1,0003 MHz

2) 3,4 kHz, 1 MHz, 0,9966 MHz, 1,0034 MHz

Juostos plotis 6,8 kHz.

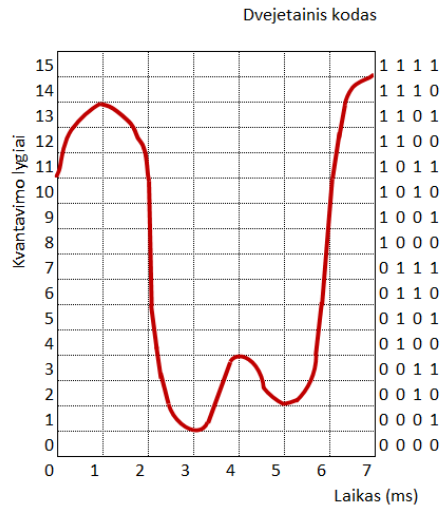
Maždaug 58 stotys.

Aukštos kokybės garsas turi būti perduodamas garsiniu signalu, kurio dažnių spektras yra nuo 20 Hz iki 20 kHz. Kiek radijo stočių, transliuojančių aukštos kokybės signalus, galėtų tilpti dažnių srityje nuo 800 kHz iki 1200 kHz?

*Atsakymas*

Ne daugiau kaip 10 stočių.

Paveiksle pateiktas analoginis signalas yra išmatuotas kas sekundę naudojant 16 kvantavimo lygių:



Kiekvieno matavimo rezultatą paverskite dvejetainiu kodu ir pavaizduokite grafiškai jo skaitmeninį atitikmenį. Kokio didžiausio dažnio analoginį signalą galima tiksliai paversti skaitmeniniu esant tokiam išrankos dažniui?

Laikas (ms)	0	1	2	3	4	5	6	7
Dvejetainis kodas	1011							
Skaitmeninis išėjimo signalas								

Atsakymas

Laikas (ms)	0	1	2	3	4	5	6	7
Dvejetainis kodas	1 0 1 1	1 1 1 0	1 0 0 1	0 0 0 1	0 1 0 0	0 0 1 0	1 0 1 0	1 1 1 1
Skaitmeninis išėjimo signalas								

D. Britanijos telekomas vartoja 8 kHz išrankos dažnį garso signalams nuo 300 Hz iki 3,4 kHz pakeisti į skaitmeninį pavidalą. Signalų lygiams nuskaityti vartojama 8 bitų kodavimo sistema. Keli signalai siunčiami tuo pačiu kanalu naudojant laikinį multipleksavimą.

Kiekvienas signalas užima 25  $\mu\text{s}$  intervalą, kai nešančiojo signalo dažnis 1 MHz. Synchronizavimo impulsai siunčiami tuo pačiu 25  $\mu\text{s}$  ruožu.

Kiek gali būti išrankos dažnių, kai naudojama 8 bitų kodavimo sistema?

*Atsakymas*

8 bitai – tai  $2^8 = 256$ , todėl galima užkoduoti 256 skirtingus išrankos lygius.

Apskaičiuokite intervalą tarp signalo imčių.

$$1 / 8000 = 125 \mu\text{s}.$$

Paaiškinkite, kaip laikinis multipleksavimas įgalina 5 signalus perduoti vienu kanalu.

*Atsakymas*

Kadangi vieno signalo trukmė 25  $\mu\text{s}$ , o intervalas tarp išrankos dažnio 5 kartus ilgesnis, todėl kaip tik telpa 5 signalai:  $5 \times 25 \mu\text{s} = 125 \mu\text{s}$ .

Skaitmeninės sistemos aprašomos perdavimo sparta. Tai – per sekundę perduodamų dvejetainių skaitmenų skaičius. Perdavimo spartos vienetai bitai per sekundę (b/s), kilobitai per sekundę, megabitai per sekundę, terabitai per sekundę.

Vieno kanalo perdavimo sparta apskaičiuojama remiantis tokia lygtimi:

$$\text{perdavimo sparta} = \text{bitų skaičius} \times \text{išrankos dažnis}$$

Telefono sistemos išrankos dažnis 8 kHz, o kiekviena imtis koduojama 8 bitų dvejetainiu skaičiumi.

Kiek imčių užregistruojama kiekvieną sekundę?

$$8000$$

Apskaičiuokite perdavimo spartą

$$8 \text{ bitai} \times 8000 \text{ s}^{-1} = 64\,000 \text{ b} \cdot \text{s}^{-1} = 64 \text{ kb} \cdot \text{s}^{-1}$$

Laikinio multipleksavimo būdu galima vienu kabeliu ar šviesolaidžiu tuo pat metu siųsti kelių kanalų duomenis. Kokia būtų perdavimo sparta, jei tokiu būdu sujungtume 32 telefono kanalus?

$$32 \times 64 \text{ kb} \cdot \text{s}^{-1} = 2048 \text{ kb} \cdot \text{s}^{-1}$$

$$2048 \text{ kb} \cdot \text{s}^{-1} = 2 \text{ Mb} \cdot \text{s}^{-1}$$

Tarkime, kad konkrečios optinės telefono sistemos tarpžemyninis kabelis veikia  $5 \text{ Gb s}^{-1}$  greičiu. Kiek telefono pokalbių šis kabelis gali perduoti vienu metu?

$$5\,000\,000\,000 / 64\,000 = 78125$$

Amplitudinė moduliacija atliekama susiejant du labai skirtingų dažnių signalus. Nagrinėkime paprastą atvejį, kai tam tikro dažnio  $f_0$  garso nešančiojo signalo dažnis daug didesnis ir lygus  $f_n$ .

Nešančiojo signalo amplitudė laiko momentu  $t$  gali būti apskaičiuojama pagal tokią formulę:

$$a = a_0 \sin(2\pi f_n t), \text{ čia } a_0 - \text{maksimali amplitudė.}$$

Dviejų šalutinių juostų signalai yra vienodi, tačiau jų dažniai šiek tiek skiriasi:

$$b = b_0 \sin(2\pi(f_n - f_0)t), \quad c = c_0 \sin(2\pi(f_n + f_0)t), \text{ čia } b_0 \text{ ir } c_0 \text{ yra šoninių juostų maksimalios amplitudės, jos paprastai yra lygios.}$$

Visos trijų signalų amplitudės kiekvienu laiko momentu reikia apskaičiuoti ir sudėti, po to nubraižyti grafiką.

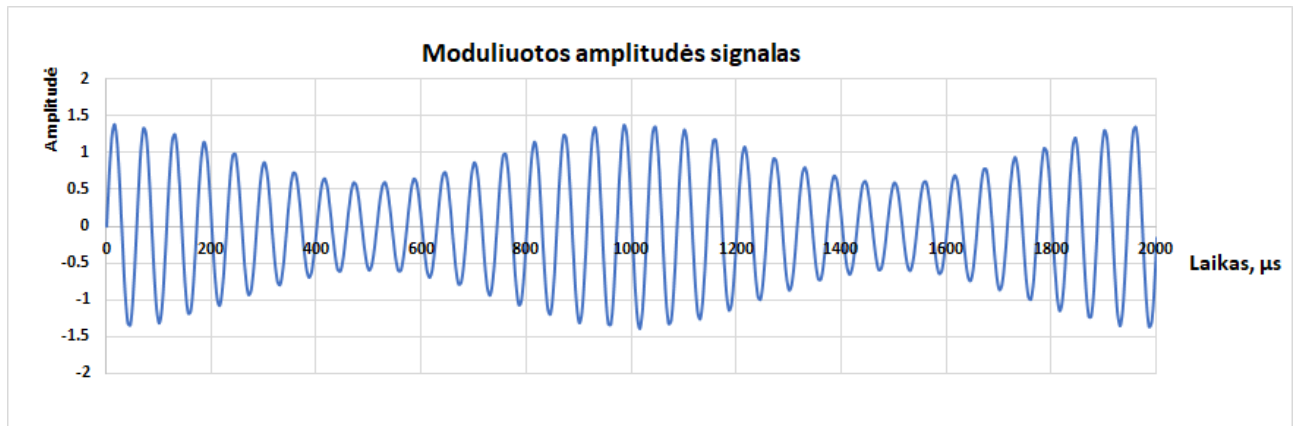
Užduotis atliekama „Excel“ (ar analogiška) skaičiuokle.

Paveiksle laikas kinta kas  $5 \mu\text{s}$ , atitinkamai parenkamas nešančiojo signalo dažnis, amplitudė prilyginama vienetui.

Laikas ( $\mu\text{s}$ )	Funkcijų suma	Nešantysis signalas	1-a šalutinė juosta	2-a šalutinė juosta
0	0	0	0	0
5	0,73106	0,5222609	0,10975531	0,09904608
10	1,24635	0,89075332	0,18350409	0,17209487
15	1,39401	0,99698259	0,19705218	0,19997279
20	1,13099	0,80967179	0,14595498	0,17536252
25	0,53567	0,38397055	0,04697552	0,10472369
30	-0,2156	-0,1547825	-0,0674148	0,00659733
35	-0,90091	-0,6479633	-0,1596889	-0,09326066
40	-1,31858	-0,9503651	-0,1995752	-0,16863992
45	-1,3467	-0,9729524	-0,1739885	-0,19975513
50	-0,97884	-0,7090748	-0,0913228	-0,1784393
55	-0,32541	-0,2364254	0,02130245	-0,11028731
60	0,41959	0,30583439	0,12693914	-0,01318749
65	1,03635	0,75804784	0,19093195	0,08737375
70	1,34436	0,98707042	0,19228727	0,16500141

Tada šalutinių juostų amplitudės bus lygios 0,2.

Nubraižytas grafikas atrodo taip:



Naudodamiesi šiuo grafiku apskaičiuokite nešantįjį dažnį ( $\approx 17,5$  kHz) ir moduliuojantį dažnį ( $\approx 500$  Hz).

Parenkite lentelę, nubraižykite grafiką ir pakeiskite kintamuosius: šalutinių juostų amplitudes, nešančiojo ir moduliuojančio (garso) signalo dažnius. Pakeiskite laiko matavimo intervalą (jei jį sumažinsite, gausite glotnesnę kreivę).

### Skaitmeninės moduliacijos būdai

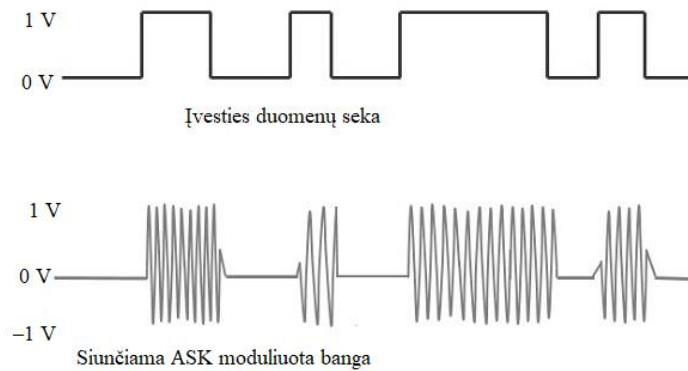
Skaitmeninė moduliacija užtikrina didesnę informacijos pralaidumą, aukštą duomenų saugumą, greitesnę sistemos pasiekiamumą ir puikią ryšio kokybę. Skaitmeninės moduliacijos būdai dažniau naudojami, nes jais galima perduoti didesnius duomenų kiekius nei analoginiais būdais.

Skaitmeninės moduliacijos metodų yra daugybė rūšių, galime naudoti net ir šių metodų derinius. Šiame skyriuje aptarsime svarbiausius skaitmeninės moduliacijos būdus.

### Amplitudės poslinkio kodavimas

Rezultato išėjimo amplitudė priklauso nuo įvesties duomenų – ar tai turėtų būti nulinis lygis, ar teigiamo ir neigiamo lygio pokytis – priklausomai nuo nešančiojo dažnio.

Amplitudės poslinkio kodavimas (angl. *Amplitude Shift Keying*) yra amplitudės moduliacijos rūšis, kuri dvejetainius duomenis pateikia signalo amplitudės kitimo pavidalu.

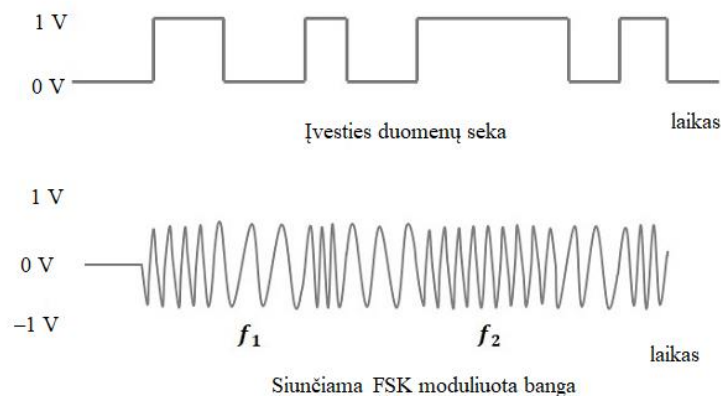


Amplitudės poslinkio kodavimu moduluojamos bangos formos ir jos įvesties schema

### Dažnio poslinkio kodavimas

Priklausomai nuo įvesties duomenų, išėjimo signalo dažnis bus aukštas arba žemas.

**Dažnio poslinkio kodavimas** (angl. *Frequency Shift Keying*) – tai skaitmeninės moduliacijos metodas, kai nešančiojo signalo dažnis kinta priklausomai nuo diskrečių skaitmeninių pokyčių.

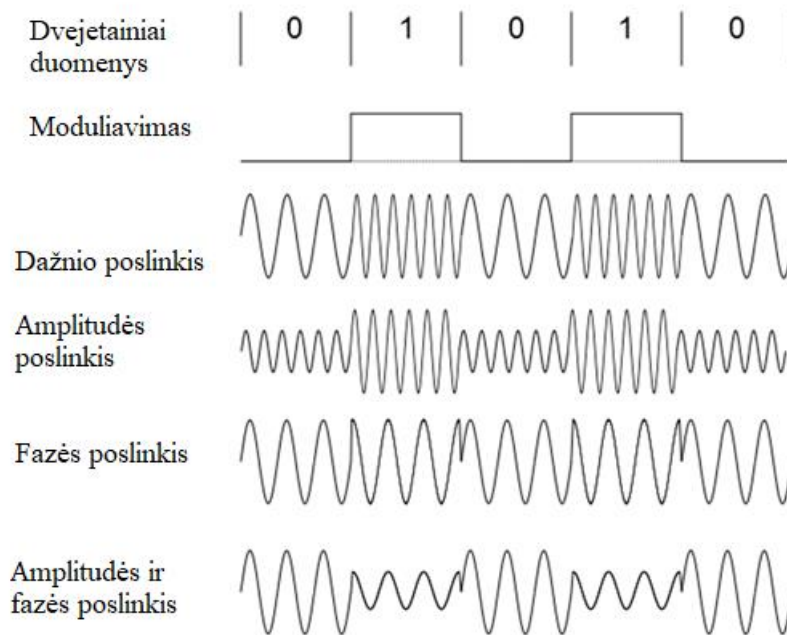


Dažnio poslinkio kodavimu moduluotos bangos formos schema ir jos įvestis

### Fazės poslinkio kodavimas

Išėjimo signalo fazė pasislenka priklausomai nuo įvesties.

Fazės poslinkio kodavimas (angl. *Phase Shift Keying*) – skaitmeninis moduliavimo metodas, kai nešančiojo signalo fazė pasikeičia tam tikru metu keičiant sinusinį ir kosinusinį jėjimus. Fazės poslinkio kodavimo technika naudojama belaidžiuose vietinių tinklų (angl. *local area network*, LAN), biometrinių, bekontakčių operacijų, mikrobanginio (angl. *Bluetooth*) ryšių srityje.



Pranešimų moduliavimo pavyzdžiai

Dažniausiai informacinis signalas netinka tiesioginiam perdavimui. Todėl siųstuvas jį transformuoja į tinkamesnį pavidalą. Ši signalų transformacija vadinama aukšto dažnio signalo moduliavimu informaciniu signalu. Pvz., gali būti, kad duotajame kanale sklindančių bangų dažnis turi būti žymiai didesnis už mikrofono generuojamos įtampos dažnį. Tokiu atveju perduodamas ne pats informacinis signalas, o aukšto dažnio signalas, kurio momentinė amplitudė, dažnis arba fazė proporcinga informaciniam signalui (paprasčiausiu atveju).

Skaitmeninė ryšių sistema gali būti naudojama analoginės informacijos perdavimui, o analoginė ryšių sistema gali būti naudojama skaitmeninės informacijos perdavimui. Pirmuoju atveju siųstuvas prieš moduliavimą paverčia analoginį informacinį signalą skaitmeniniu signalu, o antruoju atveju skaitmeninis informacinis signalas paverčiamas analoginiu. Įmtuvai po demoduliavimo atlieka priešingą transformaciją.

## Apibendrinimas

Informacijos teorija nagrinėja informacijos kūrimo spartą arba ergodinio pranešimų šaltinio entropiją; informaciją perduodančio kanalo savybes (triukšmas yra ar nėra) ir pralaidumą (informacijos perdavimo spartą); taip pat efektyvų pranešimų kodavimą ir glaudinimą, kuris leistų pranešimus perduoti be klaidų ir greičiau, kuris būtų artimas kanalo pralaidumo spartai. Klasikinė informacijos teorija dažnai vadinama komunikacijos teorija. Informacijos teorija taikoma lingvistikoje, psichologijoje, informatikoje, telekomunikacijų inžinerijoje.

## Literatūra

1. James, V Stone. Information theory: A tutorial introduction. – First edition, Sebtel Press, 2015. – 243 p.
2. Pierce, John R. An introduction to information theory. Symbols, signals & noise. – Second, revised edition. – Dover Publications, New York, 1980. – 305 p.
3. Dagienė, Valentina; Jevsikova, Tatjana. Aiškinamasis kompiuterijos terminų žodynas. – Vilnius, Vilniaus universitetas, 2016. – 484 p.
4. Encyclopædia Britannica, Inc., <https://www.britannica.com/science/information-theory>, 2022-03-18
5. Signalai telekomunikacijų sistemose, parengė Poškus A., Vilniaus universitetas, 2004, [https://www.ff.vu.lt/external/ff/files/institutai/ChFI/Studentams/Signalai\\_telekomunikacij%C5%B3\\_sistemose.pdf](https://www.ff.vu.lt/external/ff/files/institutai/ChFI/Studentams/Signalai_telekomunikacij%C5%B3_sistemose.pdf), 2022-03-18.