



Tiedonsiirron perusteet ja fyysinen kerros

Tietoliikenne kohtaa todellisuuden
OSI-mallin alimmainen kerros
Kirja sivut 43 - 93



- Data: koneiden tai ihmisten käsiteltävissä oleva tiedon esitysmuoto
- Informaatio: datan merkityssisältö
- Tieto: ihmisen ymmärrystä
- Näistä määritelmistä on variaatioita ja vaihtelua eri tieteenaloilla
 - Aihe on lähellä filosofiaa ja siitä on kirjoitettu kirjoja
- Esimerkkejä:
 - Pakkaamattomalla TIFF- ja pakatulla GIF-formaatin kuvalla saattaa olla sama informaationsisältö, mutta GIF-kuvan käyttämä datamäärä on pienempi
 - JPEG-pakkaus hävittää kuvasta silmälle näkymätöntä informaatiota
 - TCP ja IP-protokollalle sovellustason siirrettävät tavut ovat dataa; HTTP kohtelee otsaketietoja informaationa mutta siirrettäviä HTML-tiedostoja datana; selaimelle HTML on informaatiota



- Informaatio *koodataan symboleina* dataan, jota siirretään tietyn *kaistanleveyden* ja muita ominaisuuksia omaavan *median* ylitse käyttäen sopivaa *linjakoodausta*



- Kuparijohto
- Kierretty parikaapeli
- Koaksiaalikaapeli
- Valokuitu
- Myös massamuistivälineet



- Sähkömagneettinen säteily
 - Radioaallot
 - Satelliittiyhteydet
 - Mikroaallot
 - Laser
 - Infrapuna
- Ääni



Digitaalisen datan siirtäminen analogisessa maailmassa

- Bitit on koodattava jotenkin
- Kantataajuusmenetelmässä bitit muutetaan suoraan esim. jännitteeksi, toimii yhteyksillä, joissa on suhteellisen vähän häiriöitä
- Häiriöisillä yhteyksillä kantoaalto auttaa erottamaan signaalin taustakohinasta, signaali muokkaa kantoaaltoa
 - Amplitudimodulaatio (AM)
 - Taajuusmodulaatio (FM)
 - Vaihemodulaatio (PM)
- Nykyään on myös monimutkaisempia koodauksia
 - Esim. UMTS:n käyttämä CDMA

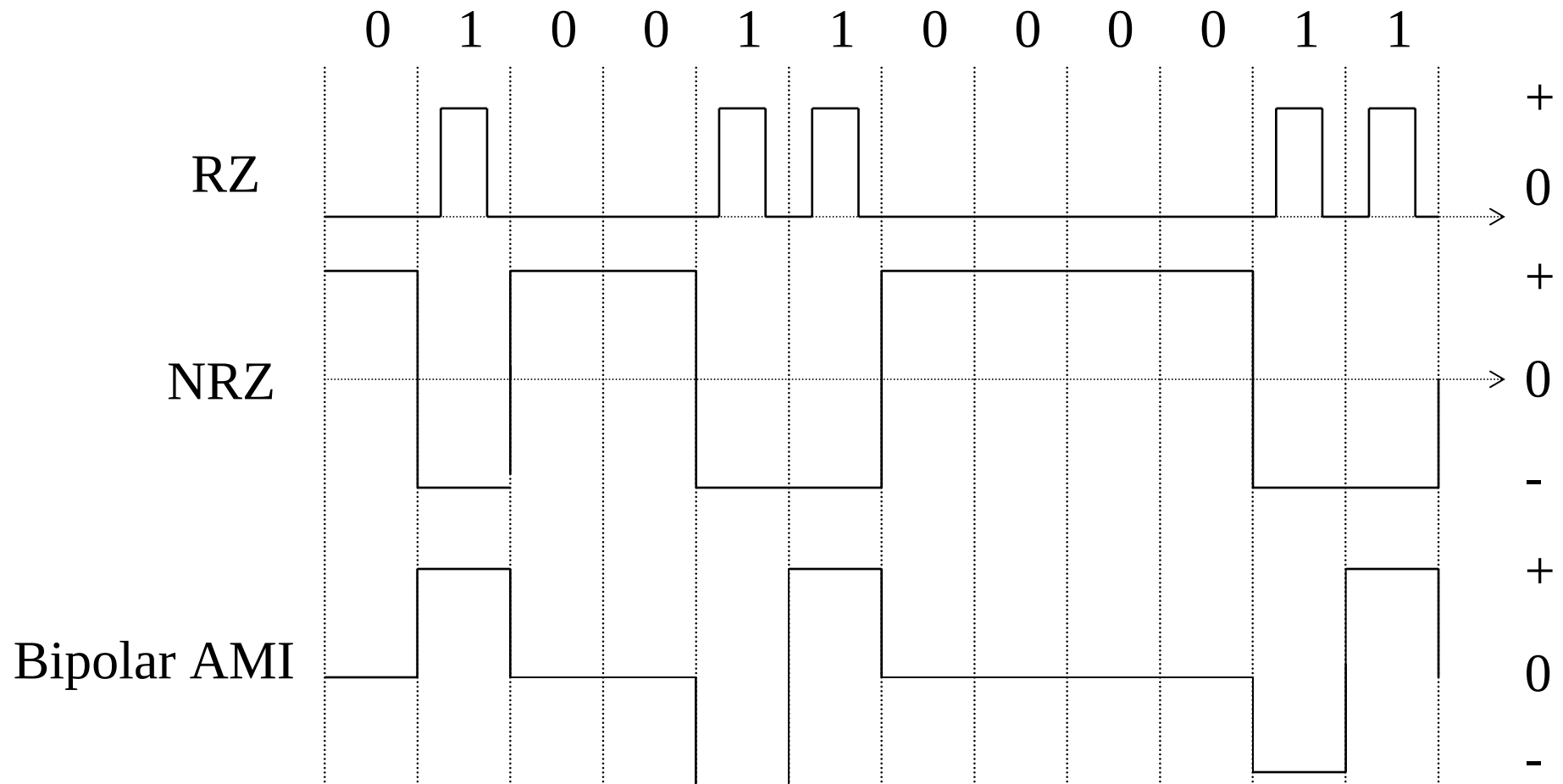


Synkronoitu ja synkronoimaton tiedonsiirto

- Vastaanottajan pitää tunnistaa data
- Synkronoimattomassa tiedonsiirrossa vastaanottaja tunnistaa viestin (yksi tavu dataa) alun ja ottaa näytteitä riittävän usein saadakseen viestin
 - Esim. start-bitti, 8 databittiä, pariteetti ja stop-bitti
- Synkronoidussa tiedonsiirrossa lähettäjällä ja vastaanottajalla on kellot samassa ajassa, vastaanottaja tietää milloin bitin pitäisi tulla
 - Tehokkaampaa

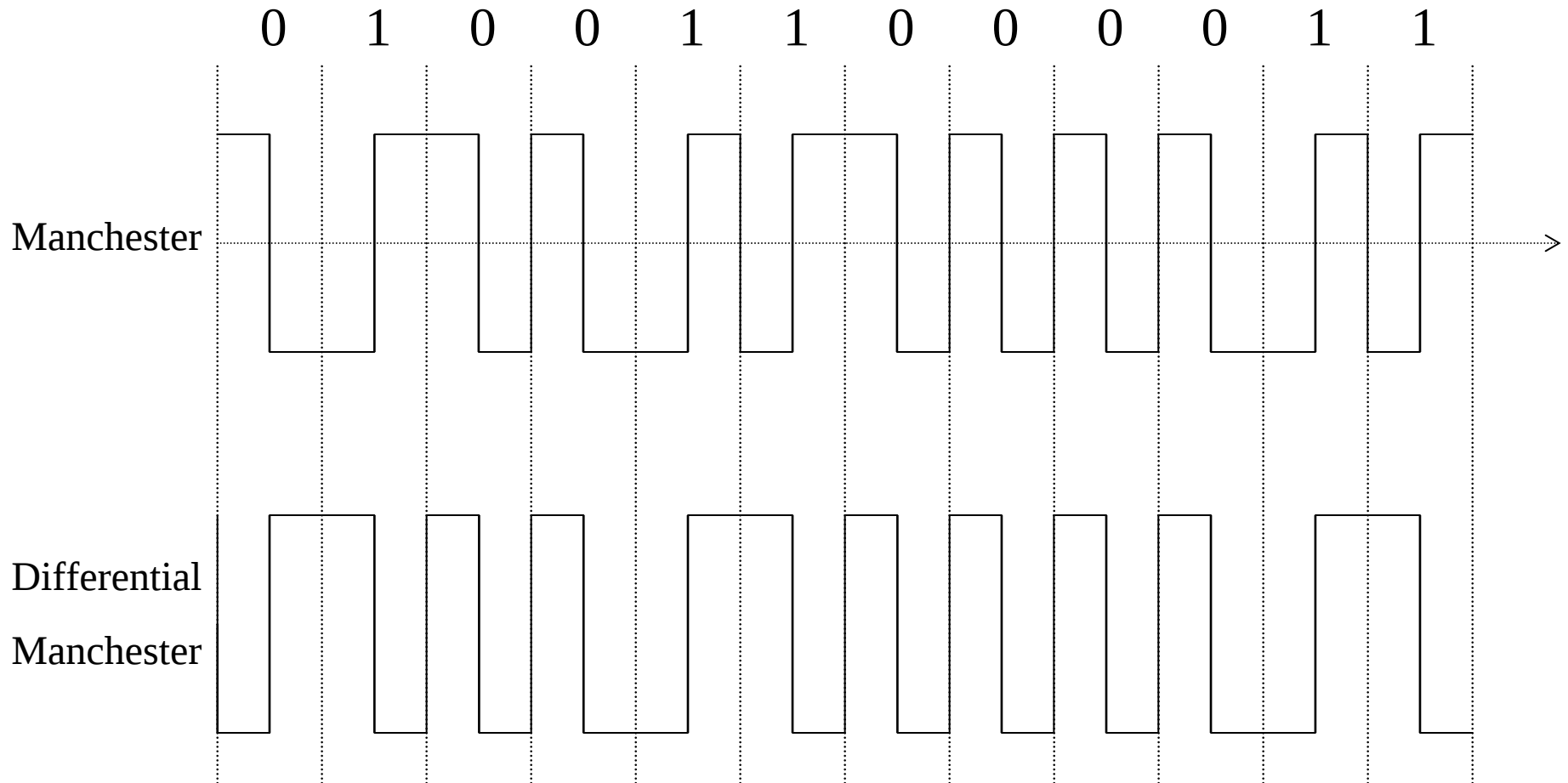


Synkronoitu kantataajuuskoodaus



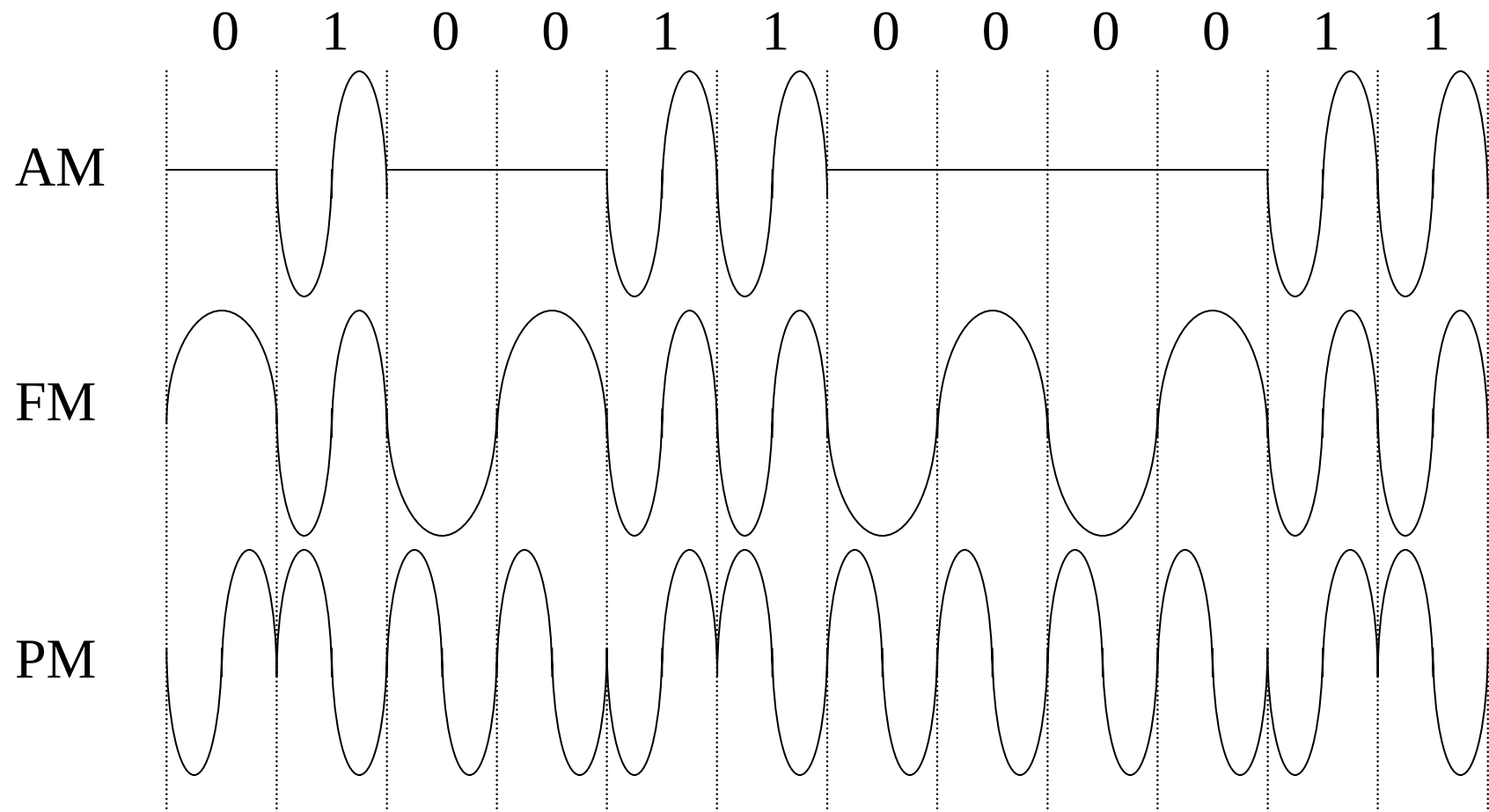


Kellosignaali kantataajuus- koodauksen mukana





Kantoaallon modulointi



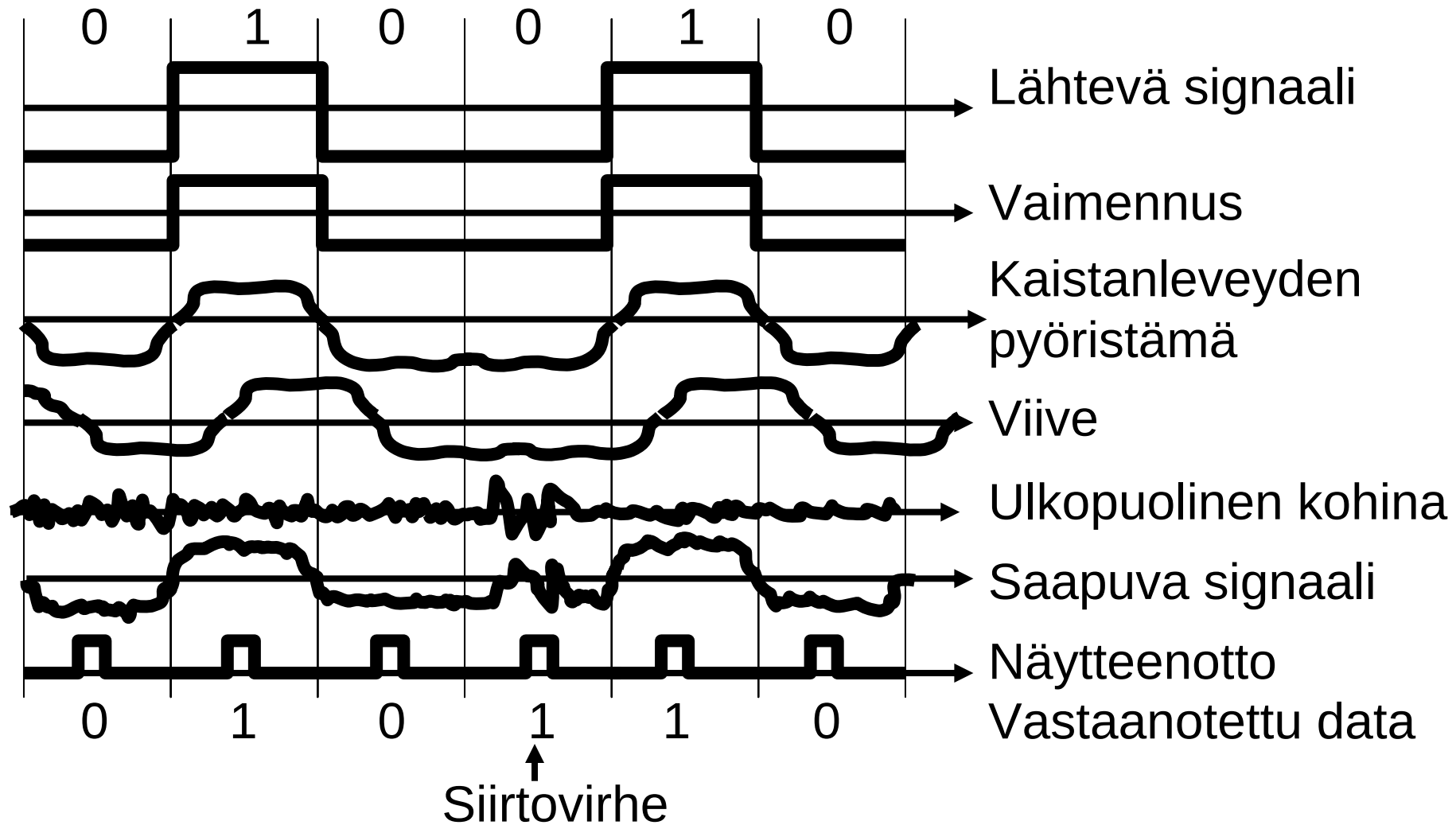


Siirtotien laatua kuvaavia tekijöitä

- Kaistanleveys (bandwidth)
 - Käytettävissä oleva taajuusalue
- Kohina (noise)
 - Taustalla olevat häiriösignaalit, joista viesti on erotettava
- Siirtokapasiteetti (capacity)
 - Montako bittiä sekunnissa saadaan siirrettyä (lasketaan edellisistä)
- Virhetaajuus (error rate)
 - Montako databittiä sekunnissa muuttuu
- Viive (latency)
 - Kauanko bitiltä kestää matkata päästä päähän
- Huojunta (jitter)
 - Perättäisten viestien aikavaihtelut
- Vaimennus (attenuation)
 - Kuinka usein signaalia on vahvistettava (riippuu mediasta)



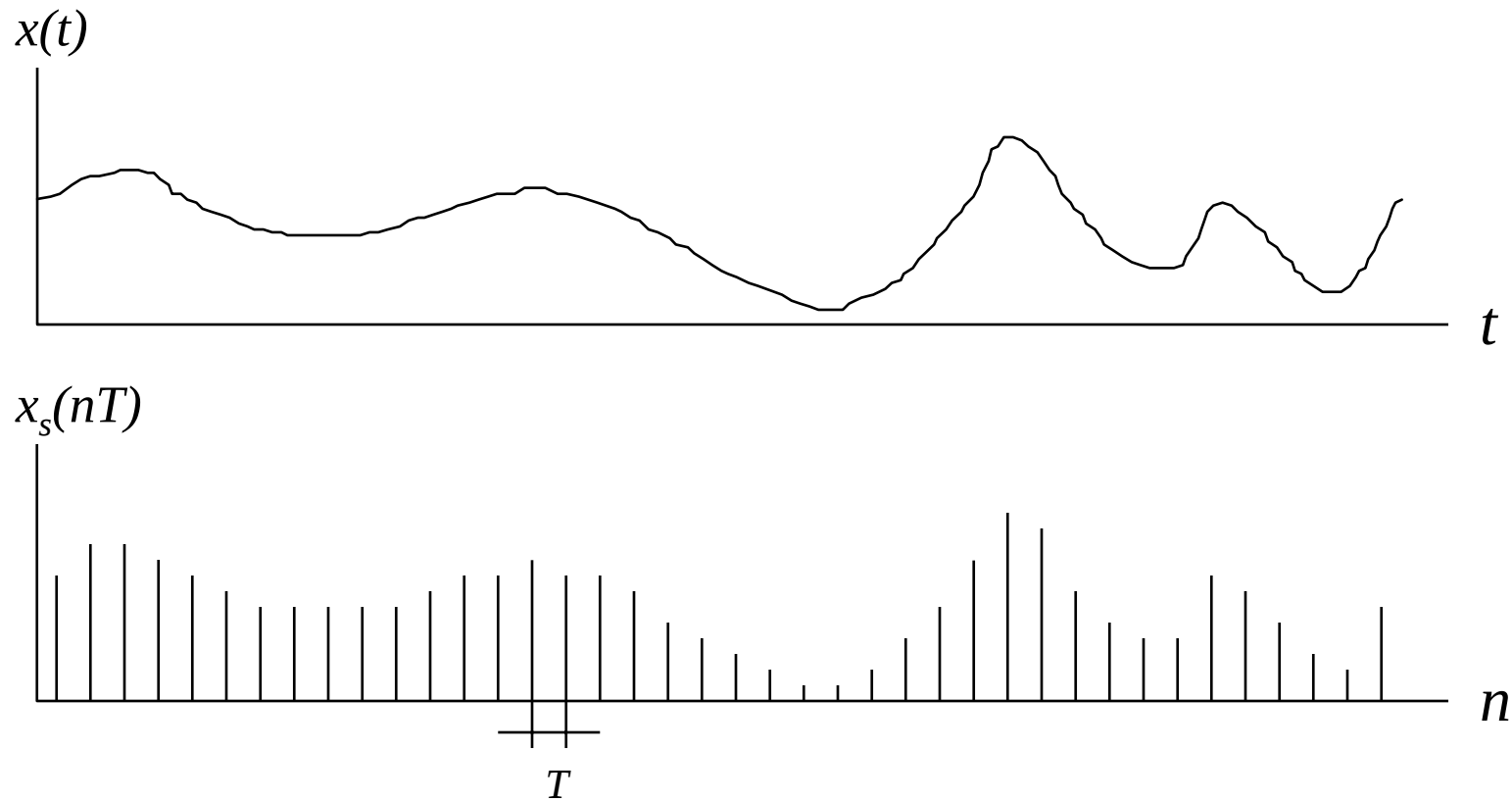
Signaali luonnossa

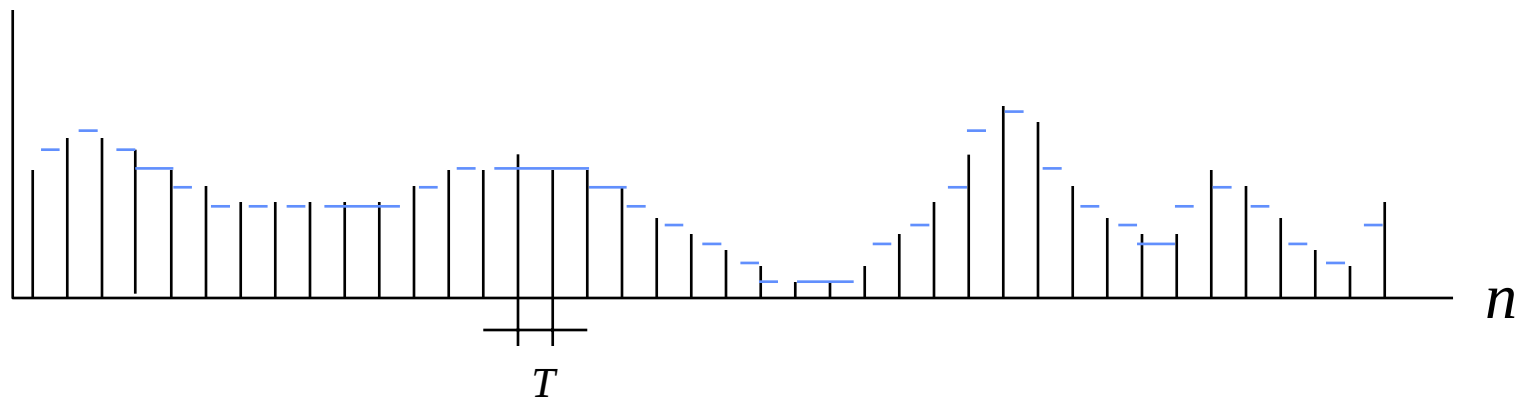
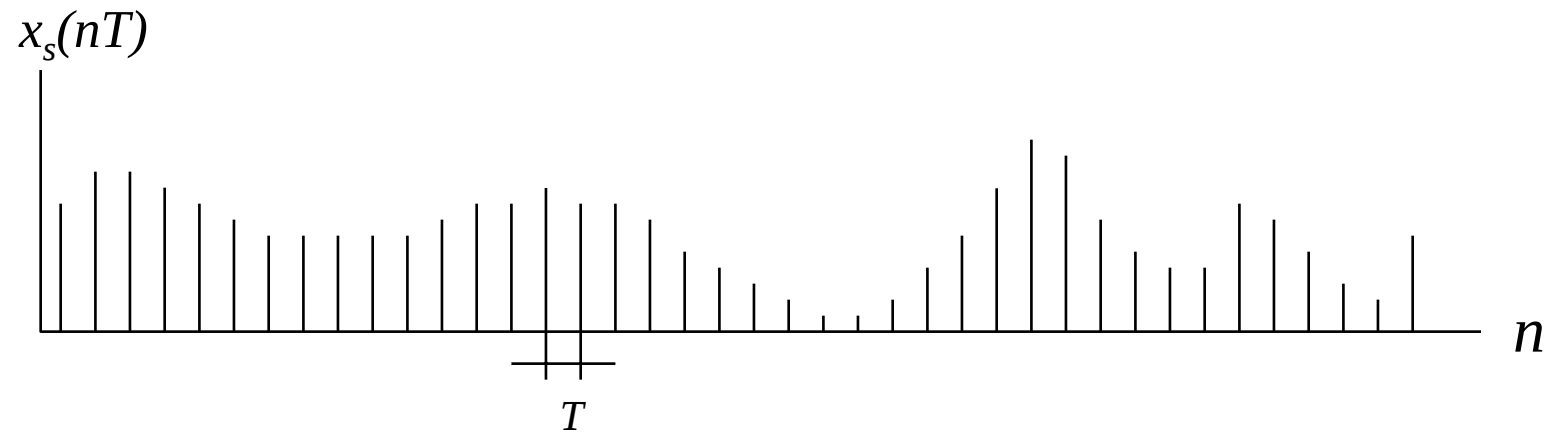




Analogisen datan koodaus digitaaliseksi

- Ihmisen puhe on analogista dataa
 - Liki rajaton määrä äänensävyjä
- Nykyään ääni digitoidaan
 - Puhelinverkot, musiikki CD-levylle
- Äänestä otetaan näytteitä (sampling)
- Näytteet koodataan datavirraksi
- Näytteenottoteorian mukaan näytteidenottotaajuuden on oltava vähintään kaksi kertaa signaalin taajuus
 - Digitaalinen puhelinverkko on rakennettu näytteenottotaajuuden 8000 näytettä/s perusteella, jotta n. 4000 Hz puheäänestä toistuisi
 - Ihmisen kuuloalue on noin 20 - 20 000 Hz
- Kvantisointi kuvaa näytteet diskreetiksi informaatioksi esim. 8-bittiseksi luvuksi (puhelinverkko)







- Analogisissa siirtoteissä on aina erilaisia häiriöitä
- Signaalinvoimakkuuden suhde kohinan voimakkuuteen on hyvin tärkeä mitta siirtotien laadulle
 - SNR, Signal to Noise Ratio
- Ilmaistaan yleensä desibeleissä
- Esim. jossain analogisessa puhelinverkossa voidaan mitata signaalin olevan 10 000 kertaa taustakohinaa voimakkaampi
 - Tällöin SNR on $10 \log_{10} 10000 = 40 \text{ db}$



Kaistanleveys ja kapasiteetti

- Kaistanleveys tarkoittaa tietyn kanavan läpäisevää taajuusaluetta
 - Tyypillisesti siirtotien sähköiset ja optiset ominaisuudet määrittävät lähetettävälle pulssille tietyn minimipituuden, jotta se voitaisi havaita luotettavasti vastaanottopäässä
 - Pulssin minimipituudesta saadaan vastaava korkein taajuus, joka voidaan siirtää
 - Ilmaistaan Hertzeinä, Hz
- Kapasiteetti on tietyn kanavan läpi siirrettävän datan määrä
 - Ilmaistaan bitteinä sekunnissa, b/s
- Nyquistin kaava tiedonsiirrolle on
$$r_{\max} = 2 W \log_2 L$$

$r_{\max}$ on suurin saavutettava bittien siirtokapasiteetti
 W on ideaalikanavan kaistanleveys
 L on signaalitasojen määrä



Realistinen kaistanleveys ja kapasiteetti

- Käytännön kanavilla on kohinaa, joka rajoittaa signaalitasojen määrää
- Shannonin kaava kanavakapasiteetille on
$$C = W \log_2 (1 + \text{SNR}) \text{ b/s}$$

C on kanavan siirtokapasiteetti b/s
W on kaistanleveys
SNR on kanavan signaali/kohinasuhde

 - jos ilmaistu desibeleinä: $10^{(\text{SNR}/10)}$
- Signaalitasojen määrä ei vaikuta, kaava määrittelee suurimman mahdollisen bitinsiirtokapasiteetin



Diskreetin informaation koodaus

- Kirjoitettu kieli muodostuu aakkosista eli symboleista, jotka voidaan numeroida
 - ASCII kirjaimille a-z, numeroille ja välimerkeille
 - ISO-8859 -perhe länsimaisille kielille
 - Unicode (ja UTF) kaikille kielille
- Kuvat voidaan digitoida kuten ääni ja muu analoginen data
 - Myös liikkuva kuva



- Idea perustuu Shannonin Informaatioteorialle
 - N erilaisten vaihtoehtojen esittämiseen tarvitaan $\log_2 N$ bittiä
- Mikäli datassa on toistuvia informaationsymboleja, niiden esittämiseen voidaan käyttää sovittuja bittikuvioita
- Pakkaamiseen ja purkamiseen tarvitaan algoritmi, jonka ohjelmistototeutus on asennettuna valmiiksi
- Algoritmeja ja pakkausformaatteja on paljon ja ne soveltuvat eri datatyypeille
 - Lempel-Ziv, Huffman, JPEG, MPEG-2, MPEG-1 layer 3



- Kohinaa
 - Vaikeuttaa signaalin havaitsemista
 - Signaalin tulovoimakkuus ei ole niin tärkeä kuin signaalin tulovoimakkuus suhteessa kohinaan, josta se erottuu
 - Lämpöenergia komponenteissa
 - Radiolinkeillä muut radiolähetykset ja universumin taustakohina
- Häiriöpurskeita
 - Esim. matkapuhelimet
- Signaalien ylikuuluminen viereisistä johtimista
- Katkoksia



- Virheet voidaan korjata
 - Etukäteen lisäämällä ylimääristä dataa, redundanssia
 - Esim. Hamming-koodaus, vaikkapa 4 bitin viestiin lisätään 3 bittiä dataa, joilla voidaan korjata kaikki yhden bitin siirtovirheet
 - Vrt. tietokoneiden RAID-levyt
 - Jälkikäteen tunnistamalla virheet ja lähettämällä viesti uudestaan
 - Jättää korjaamatta
 - Esim. puhe ja video sisältävät usein itsessään niin paljon redundanssia, *että pinet vireet eivät estä ymmärtästä viestiä*



Virheiden tunnistaminen

- Cyclic Redundancy Check -tarkistussummat
 - Perustuvat polynomeihin
 - Esim. CRC-CCITT tunnistaa 16 bitin tai lyhyemmät purskevirheet
- Pariteettibitti
 - Yleinen pääteyhteyksien ja tietokoneiden muistin suojaaja
 - Esim. data-elementissä on aina pariton määrä 1-bittejä:
1110 0010 -> 1110 0010 1
- Kryptografiset tarkistussummat (hashit eli tiivistet) suojaavat tahallisia muutosyrityksiä vastaan
- Käytettävä virheentunnistus valitaan yleensä odotetun virheen mukaisesti



Termejä: simplex ja duplex

- Simplex: yksisuuntainen tiedonsiirto
 - Esim. yleisradio
- Half-duplex: vuorotteleva kaksisuuntainen tiedonsiirto
 - Esim. radiopuhelin
- Duplex (full duplex): samanaikainen molemmansuuntainen tiedonsiirto
 - Esim. puhelin
- Internet ja muut tietoverkot rakennetaan usein yhdistelemällä simplex-yhteyksiä, mm. koska liikenne on epäsymmetristä (kuten WWW-liikenne)
 - Käyttäjälle näkyvä palvelu on kuitenkin duplex



Termejä: montako vastaanottajaa

- Täsmälähetys, unicast: lähetetään yhdelle vastaanottajalle
 - Normaali Internet-liikenne
- Yleislähetys, broadcast: lähetetään kaikille vastaanottajille
 - Yleisradioliikenne
 - Tiettyjen palveluiden löytäminen lähiverkossa
- Ryhmälähetys, multicast: kohdistetaan ryhmälle (palvelun tilanneita) vastaanottajia
 - Käytetään Internetissä esim. videokuvan jakamiseen, vastaanottaja tilaa liikenteen liittymällä multicast-ryhmään
- Jokulähetys, anycast: lähetetään viesti yhdelle vastaanottajalle joukosta, vastaanottaja valitaan ad-hoc



- Fyysinen kerros edustaa sovellettuja luonnontieteitä, lähinnä fysiikkaa ja fysiikan ilmiöitä
- Tälle kerrokselle on määritelty suuri määrä erilaisia tekniikkoja, joiden avulla tietoa yritetään siirtää
- Mikään tekniikka ei ole ylivoimainen, vaan valinnat tehdään tilanteen mukaan
 - Tarpeet ja resurssit vaikuttavat