IGOR LETURIA AZKARATE
Informatikaria eta ikertzailea

CLAUDE SHANNON
mundu digitalaren aita

Aurtengo apirilean bete dira 100 urte Claude Shannon matematikaria jaio zela. Publiko zabalarentzat ezezaguna bada ere, gizartearen aurrerapenari ekarpenik handienak egin dizkioten inoizko zientzialaririk garrantzitsuenen artean dago; batzuen ustez, ia Newton, Einstein eta Darwinen parekoa da. Izan ere, guk guztiok egunero hainbat alditan erabiltzen ditugu hari esker garatu diren gailu eta teknologia ugariak: ordenagailuak, smartphoneak eta gailu elektroniko guztiak; Internet, telefonia mugikorra eta komunikatzeko edozein modu digital; CDak, MP3 eta ZIP formatuak, eta informazioa digitalki gordetzeko metodo denak. Ezagut dezagun Claude Shannon, gure mundu digitalaren aita.

Informazioaren eta Komunikazioaren Teknologien munduan izan diren pertsona garrantzitsuenak zeintzuk izan diren galdetuta, jende gehiena ziurrenik ez da gai izango Bill Gates, Steve Jobs eta halako enpresa-gizonenez haragoko izenik emateko. Informatikaren mundua hobetoxeago ezagutzen dugunok beste izen batzuk esango genituzke ziur aski, hala nola Alan Turing edo Tim Berners-Lee. Gure artean ere, beharbada, ez lukete hainbestek aipatuko Claude Shannon. Arrazoia zein den ez dakigu, baina Shannonek ez dauka, inondik inora, bere ekarpen ikaragarriek merezi duten ospea; Shannon izango da, apika, IKTen munduan izan den izenik garrantzitsuena.

Claude Elwood Shannon 1916ko apirilaren 30ean jaio zen, AEBko Michigan estatuko herrixka batean. Txikitatik erakutsi zuen eskolan zientzia eta matematika gaietarako gaitasuna, eta baita gailu mekaniko eta elektrikoak asmatu eta eraikitzeko zaletasuna ere. Hala, 1932an (16 urte besterik ez zuela), Michigango Unibertsitatean hasi zituen unibertsitate-ikasketak, eta 1936rako, matematikako eta ingeniaritza elektrikoko karrerak amaitu zituen. Ondoren, 1936an, ingeniaritza elektrikoko masterra egiten hasi zen Massachusetts Institute of Technology edo MIT ezagunean. Eta 1940an, matematikako doktorego-tesia defendatu zuen, erakunde horretan bertan.

ELEKTRONIKA DIGITALAREN OINARRIAK

Hala ere, zientziari egin zizkion bi ekarpen handienetarik lehena tesiaren aurretik egin zuen, masterreko proiektuan hain zuzen, 21 urte besterik ez zuela. Hainbat adituren iritziz, inoizko master-tesirik garrantzitsuena izan zen bere hura.

Masterra egiten ari zela, orduko ordenagailu ahaltsuenetako batekin lan egiteko aukera izan zuen Shannonek, Vannevar Bush-en analizatzaile diferentzialarekin. Orduko ordenagailu haiek analogikoak ziren; polea, engranaje eta halako mekanismoen bidez zebiltzan, eta toki handia hartzen zuten. Era berean, matematikako ikasketetan, Booleren aljebra ezagutzeko aukera izan zuen. Booleren aljebran aldagaiek ez dute zenbakizko baliorik, *Egia* eta *Gezurra* (edo 0 eta 1) balioak baizik, eta oinarrizko eragiketak ez dira batuketa eta biderketa, baizik eta konjuntzioa, disjuntzioa eta ukapena (And, Or eta Not). Eragiketa horiek egiaren taula hauen bidez definitzen dira:

Konjuntzioa (And, Eta, $\wedge$)

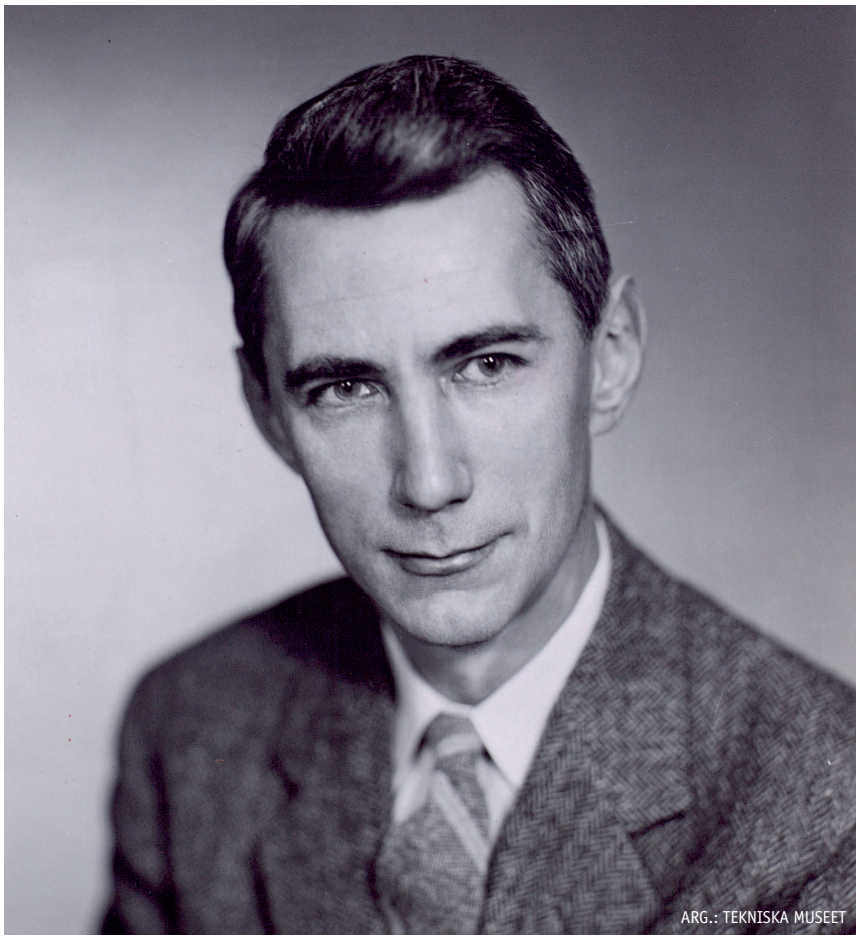
x	y	x eta y
0	0	0
0	1	0
1	0	0
1	1	1

Disjuntzioa (Or, Edo, $\vee$)

x	y	x edo y
0	0	0
0	1	1
1	0	1
1	1	1

Ukapena (Not, Ez, ¬)

x	ez x
0	1
1	0


ARG.: TEKNISKA MUSEET

Booleren aljebra inplementatzen zuten etengailu-zirkuitu elektrikoak diseinatu zituen Shannonek, masterreko proiektuan. Gainera, frogatu zuen zirkuitu horien konbinazioen bidez edozein operazio matematiko edo logiko egin zitekeela, eta zirkuitu horietako batzuk diseinatu ere egin zituen. Lan horrek zirkuitu digitalak diseinatzeko oinarriak jarri zituen, eta hortik etorri zen elektronika digitala. Ordenagailuak eta beste gailu elektroniko eta digital guztiak haren masterreko proiektu horretan izan zuen ikuspenari esker etorri dira.

INFORMAZIOAREN TEORIA, INFORMAZIOA BILTEGIRATZEAREN ETA KOMUNIKAZIO DIGITALAREN OINARRIA

Ekarpen hori berez ikaragarria izanik ere, eta harrigarria bada ere, haren beste ekarpen bat jotzen da garrantzitsuentzat. Shannon batez ere ezaguna da informazioaren teoriaren ezagutza-arloa sortzeagatik eta ia osorik garatzeagatik, 1948an, Bell laborategi entzutetsuetan lanean ari zela, idatzi zuen "A mathematical theory of communication" artikuluan.

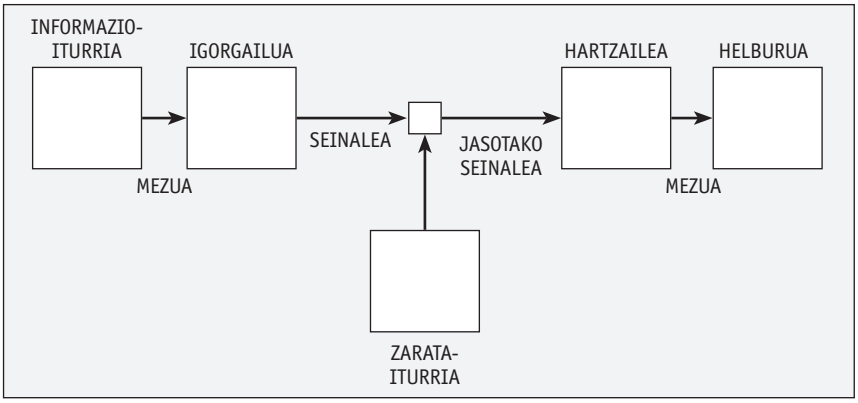
Artikulu horretan, Shannonek beheko irudiaren bidez definitu zuen komunikazio-sistema bat.

Hala, moduak aztertu eta proposatzen ditu mezua modu egokienean kodetzeko. Irudia berdin aplika dakioke biltegiratze-sistema bati, non informazio bat idazten den medio batean, hor gordetako seinaleak degradazio bat jasan dezakeen, eta irakurtzean informazioak idatzitako bera izan behar duen.

Artikuluan, Shannonek proposatu zuen informazioaren biltegiratze eta garraiorako, informazioa oinarri bitarrera bihurtzea; hau da, 0 eta 1 zenbakien bidez egitea. Informazio-unitate horiei artikulu hartan esan zitzaien lehenbizikoz *binary digit* edo *bit*, nahiz eta izena J. W. Tukeyk, Shannonen lankideak, proposatutakoa zen.

Informazioa digital bihurtzea izugarrizko aurrerapena zen. Izan ere, ordura arte erabiltzen zen informazio analogikoak egoera posible gehiago ditu (infinitu egoera, egia esan), eta egoera batetik besterako desberdintasuna txikia denez, biltegiratzean edo transmisioan dagoen zarata edo aldaketa txiki orok seinale bat beste bat bihurtzea dakar, eta ezinezkoa da jatorrizko informazio zehatza berreskuratzea. Horrek arazo handiak eragiten zituen seinaleen transmisioan, adibidez, aldiro-aldiro seinalearen potentzia errepikagailuekin berriz altxatuta ere, konpondu ezin zitekeen informazioaren degradazioa eta galera gertatzen baitzen. Aldiz, informazioa oinarri bitarrera eramanda, bi egoera posible baino ez daude, 0 eta 1, eta bien arteko diferentzia nahi bezain handia izan daiteke seinale fisikoan (esaterako, 0 volt eta 5 volt). Hala, nahiz eta seinalea apur bat degradatu edo aldatu, zailagoa da 0 bat 1 bat bihurtzea edo alderantziz. Informazioa errazago eta era fidagarriago ba-

Shannonek irudi honen bidez definitu zuen komunikazio-sistema bat.



tean biltegiratzen da horrela, eta transmisioa nahi bezain urrutira egin daiteke errepikagailuen bidez, informaziorik galdu gabe.

Hori, bere aldetik, aurrerapauso itzela bazen ere, Shannon harago joan zen: biltegiratze eta komunikazioen konpresiorako eta fidagarritasunerako gaur egun erabiltzen diren metodoentzako teknika eta formula matematiko ia guztiak definitu zituen.

"Komunikazioen fidagarritasunerako erabiltzen diren metodoentzako teknika eta formula matematiko ia guztiak definitu zituen"

Horretarako, hainbat kontzeptu definitu zituen. Horietako bat da mezu batek ematen duen informazioa. Izan ere, Shannonek esan zuen mezu guztiek ez dutela informazio bera ematen: mezu bat zenbat eta probableagoa izan mezu posible guztien artean, orduan eta informazio gutxiago ematen du mezu horrek. Halaber, mezu bateko ikur edo karaktereen artean, maizago agertzen den ikur batek informazio gutxiago ematen du. Ikur baten informazioa neurtzeko $I = -log_2\ p$ formula eman zuen, non p den ikurraren probabilitatea edo maiztasuna. Informazioaren neurri honek, gainera, adierazten du ikur hori transmititzeko zenbat bit behar diren. Hala, txanpon bat botatzean atera daitezkeen bi emaitza (leon edo kastillo) ekiprobableetako bakoitza adierazteko bit bat beharko genuke, formularen arabera. Edo 27 letratako alfabeto bateko letretako bakoitza adierazteko, denek probabilitate bera balute, 4,76 bit beharko lirateke.

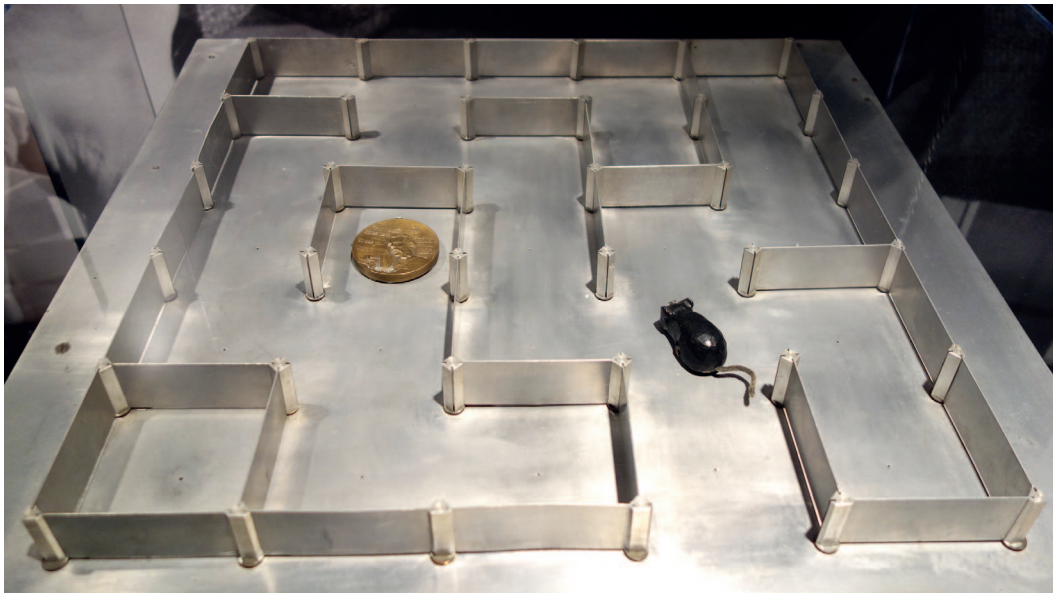
Mezu baten entropiaren definizioa ere eman zuen. Mezu batek batez beste duen informazioa da mezuaren entropia, beraz, honela definitzen da: $H = -\sum_i p_i log_2 p_i$. Horrez gain, ikur batzuk probableagoak izaten direnez testuinguru batzuetan beste batzuetan baino (letren kasuan, adibidez, ohikoena izaten da q letraren ondoren u letra joatea, edo kontsonanteen ondoren normalagoa da bokal bat joatea beste kontsonante bat baino), entropia baldintzatua ere definitu zuen Shannonek, non probabilitateak kalkulatzeko aurreko ikurra ere kontuan hartzen den. Bada, entropiaren neurri horrek esaten du mezu bat zenbateraino konprima daitekeen informaziorik galdu gabe! Hizkuntza naturaleko mezuetan, adibidez, letren maiztasunak eta maiztasun baldintzatuak kontuan izanda, bat baino txikiagoa da entropia; beraz, letra bakoitza bit bat baino gutxiagoko seinaletan kodetu daiteke. Eta hori nola egin ere adierazi zuen Shannonek. Gaur egun hainbeste erabiltzen diren galerarik gabeko konpresio-sistemak (ZIP, RAR edota komunikazioetan erabiltzen diren beste asko…) Shannonen formula eta metodoetan oinarrituta daude.

Horrekin nahikoa ez, eta galeradun konpresioaren inguruan ere teorizatu zuen Shannonek, eta konpresio-mota hori erabiltzen duten sistema guztiek (JPG, MP3, DivX…) dute hori oinarrian.

Eta hori guztia gutxi balitz, akatsak detektatu edo/eta konpontzen dituzten kodeen inguruan ere idatzi zuen artikulu hartan. Jasotako edo irakurritako informazio batean akatsak dauden detektatzeko edo akatsak konpontzeko, funtsean, informazio gehigarri erredundantea bidaltzen da. Horren adibide dira Espainiako nortasun-agiriaren zenbakiari gehitu dioten letra edo kontu korronteen zenbakietako kontrol-zenbakiak. Horiek, finean, *checksum* edo kontroleko baturak dira. Zenbakiren batean huts egin badugu, kontroleko batura ez da bat etorriko, eta jakingo dugu zenbakian okerren bat dagoela. Akatsa konpontzeko, eskuz begiratu beharko dugu; izan ere, adibide horietan akatsak detektatzeko soilik erabiltzen da kodea.

Baina badago modua jasotako mezuan akatsak egonik ere jatorrizko mezua osatzeko. Horretarako, informazio gehigarri erredundante gehiago sartu behar da, eta, batez ere, kodean bidal daitezkeen hitz edo zenbaki posible guztiek gutxieneko distantzia bat izan behar dute elkarrekiko. Hala, demagun karaktere posible guztiek beren artean duten distantzia hiru bitekoa dela, bit batean akatsa gertatzen bada, nahikoa ziur jakin dezakegu zein zen bidali nahi zen karakterea, baldin eta seinale baten bi bit oker jasotzeko probabilitatea nahikoa txikia bada. Hala, kanalaren fidagarritasunaren arabera, kodeko seinaleen arteko distantzia (bit-kopurua) handiagoa edo txikiagoa egingo dugu eta informazio erredundante gehiago edo gutxiago sartuko, baina jasotzaileak mezu akastunak ere konpondu ahal izango ditu.

Akatsak konpontzen dituzten kodeak ez zituen Shannonek asmatu. Haren lankide Richard Hamming zen horretan aitzindari. Baina Shannonek formalizatu zuen matematikoki akatsak konpontzeko kodeen teoria eta definitu zuen zenbat bit gehigarri behar diren gehienez kanalaren hainbat akats-probabilitaterentzat. Eta formulazio hori baliatzen dute gaur egun akatsak konpontzeko gai diren kode guztiek, milaka erabilera dituztenak: espazio-ontziekiko ko-



Shannonek garatutako Theseus sagu automatikoa gai da labirinto bateko irteera aurkitzeko eta bidea ikasteko.
ARG.: DADEROT/CC 1.0.

munikazioetan eta bestelakoetan, CD eta pendrive-tan informazioa gordetzean, eta abar.

Ikusi duzuenez, "A mathematical theory of communication" artikuluan Shannonek egindako ekarpenak itzelak izan ziren. Bere garaiari asko aurreratu zitzaion, etorkizuneko arazo praktiko askorentzat irtenbideak proposatu zituen, eta dena oso modu dotorean, formula eta teorema matematikoekin, eta teoremak frogatuz.

...ETA ARE GEHIAGO

Ekarpen nagusi horiez gain, beste ekarpen "txiki" asko ere egin zituen Shannonek. "Txiki", nolabait esatearren; izan ere, beste ekarpenekin alderatuta txikiak dira, baina horietako batzuk nahikoa izango lirateke beste edozein famatu egiteko.

II. Mundu Gerraren garaian, bere artikulu ezagguna idatzi aurretik, kriptografian aritu zen Bell laborategietan, eta kriptografiaren arloko zenbait aurkikuntza egin zuen. Horri esker, Alan Turing ezagutzeko aukera izan zuen. Turingek bere makina unibertsalaren ideiaz hitz egin zion, egungo ordenagailuen kontzeptuaz, azken finean, zeina oso ongi osagarritzen zen Shannonen ideiekin. Garai hartan, bestalde, seinale-fluxu diagramak asmatu zituen Shannonek.

Shannonek xakea atsegin zuen, eta horren inguruan ere aritu zen lanean. 1950ean, ordenagailuek xakean jokatzeko programen problematikaren inguruko lehenengoetariko artikulu bat idatzi zuen, alorra abiarazi zuena. Artikulu horretan zenbatetsi zuen xake-jokoaren konbinazio posibleak 10^{120} direla, gutxienez; gaur egun, Shannonen zenbaki esaten zaio horri.

1950ean, Theseus sagu automatikoa egin zuen; sagua gai zen labirinto bateko irteera aurkitzeko, eta hurrengo baterako, bidea ikasten zuen. Mota horretako adimen artifizialeko lehen gailua omen da. Bestalde, malabarismoak ere gogoko zituen, eta 70eko hamarkadan, lehen robot malabarista eraiki zuen. Halaber, Rubik-en kuboa ebazten zuen makina bat ere eraiki zuen.

Kide batzuekin batera Las Vegasera egiten zituen bidaietan dirutza irabazi zuen; izan ere, joko-teoria aplikatuz, kartak kontatzen zituzten, eta ordenagailu txiki eta ezkutagarri bat ere egin zuten (batzuk lehen *wearable*tzat jotzen dutena), jokoan probabilitateak kalkulatzeko. Eta antzeko metodoak erabilita, burtsan are gehiago irabaztea lortu zuten.

Shannon 2001ean hil zen, azken urteak alzheimer gaitzak jota eman ondoren. Inguruaren kontzientzia galdu zuenerako, haren ekarpenek aplikazio ugari izan zituzten jada; baina ez zuen ezagutzeko aukerarik izan azken urteetan Internetek, ordenagailuek eta bestelako gailuek izan duten aurrerapena eta, batez ere, zabalkundea. Harritu egingo litzateke ziurrenik. Guk harritu beharko genukeen bezala, aldaketa ikaragarri hori gauzatzea ahalbidetu zuen pertsona ia ezezaguna delako. ●

"etorkizuneko arazo praktiko askorentzako irtenbideak proposatu zituen, oso modu dotorean"