Informazioa neurritzen

Egan, 34 zk., 1974

«Informazio» hitzak «mass media» direlakoak bururatzen dizkigu, alegia, egunkariak, aldizkariak, telebista, erratia, etabar. Hau ez da, ordea, artikulu honetan «*informazio*» hitzari ematen diogun esanahia.

Informazioa deitzen duguna, ez da bakarrik giza-munduan agertzen, fisikako munduan non nahi ere aurkitzen dugu. Non nahi ere, Natura guztian, ibiltzen dira, batetik bestera, *informazio eramale* diren ezin kontako eta izkutuzko *gaztiguak*.

Gaurko Zibernetikak frogatzen duanez, Naturaren mugimendu eta harreman guztiak, holako mandatuen bitartez egiztatzen dira.

Alde honetik, eguneko zienzian, garrantzirik handienetako konzeptu bat da informazioa, Unibersoaren sail guztietan sartzen dalako.

Gaztigu batek eramandako informazioa hamaika motakoa izan daiteke. Batzutan *agindu* bat izaten da, izan batek beste bati bialdutakoa[1]. Bigarren honek, agindua hartzen duanean, bera betetzen abiatzen da, hortarako programaturik bait dago. Beste batzutan, galdera bat izan ohi da, erantzun bat eskatzen duena, edo *ohar* bat, hartzaileak kontuan izan behar duena, bere jokaera bera zuzentzeko. *Berri* bat izan daiteke ere, baten batek beste edonorri «hemen, hau gertatzen den» jakin arazi diona. *Deiezko* eta deadarrezko gaztiguak ba daude ere, zenbaitek espaziora zabaltzen dituanak laguntza eskatzeko, SOS bezalatsuak. Zenbait hori ez da beti pertsona bat, bizi gabeko bat izan daiteke ere (izar bat edo molekula bat, esate baterako) conplementoko elementu bat eskatzen duena.

Ikusten dugunez, informazioa konzeptu guztiz orokarra da.

Gainera, *gaztigu* eta *informazio* konzeptuak guztiz loturik daude. Baina gaztigu batean dagoen informazioa neurri ahal daitekean zerbait da. Informazio *kantitate* bat da, magnitude neurrigarri bat, alegia.

Hartley injinadorearen hastapenen ondoan, Claude Elwood Shannon izan da informazioaren neurri eta kalkuloaren egiazko sortzailea. 1948 urtean, Zienzia Historiara igaroko diren bi artikulu argitaratu zuen, gai honi buruz. Biak

Iparramerikako «Bell Systema Journal» aldizkarian agertu ziren, eta hoietan, zienzia berri baten oinarriak jarri zituen Shannon jakintsuak.

Matematikalari fina zen —eta da, oraindik bizi bait da—. Nahiz eta konzeptu sinple eta elementalak erabili, Matematika zienziari bide berri bat iriki dio.

Shannon jakintsuak asmatutako matematika berria «informazioaren teoria» deitzen zaio. Haren helbururik nagusiena *informazio kantitatea* mugatu eta neurritzea da, eta kalkulatzen honetik hamaika ondore garrantzizkoak sortzea dira, teknika eta zienziaren sailean.

Informazioaren neurria

«Baina nola izan daiteke informazioa neurritzea?» —galdetuko du bat edo batek.

Informazio baten garrantzia eta saria, gauza guztiz subjetiboa da. Alde honetik informazio neurritzerik ez dago. Askotan gertatzen da, balio handienetako informazio batek, informazio teoriaren aldetik oso gutxi baliotzea.

«Zer izan da, mutiko ala neskato?» —galdetzen dio emaginari erditu berri dan emakumeak.

«Analisaren emaitza, zer izan da, baiezkoa ala ezetzkoa?» galdetzen dio midikuari mingaizto batez mehatxaturik dagoen gaisoak.

«Zigorra, zein izanen da, herioa ala presondegia?» —prozesatuak abokatari.

«Nixon lendakariak bere kargua utziko du, ala ez, 1974. urtean?». Hau jakitea kauka-mauka ederra izanen litzateke bat noski batzuentzat.

Holako galderak, galderatzailentzat ezin esanezko balioa izan arren, Informazio teoriaren aldetik baliorik txikiena dute, «*bit*» unitatea dalakoa, alegia.

Aipatutako kasu guztietan, *erantzun bakar bat* eskatzen du galderak: *Bai* ala *ez*. Berehala ikusiko dugunez, holako kasuetan behar den informazioaren neurria «bit» bat da, besterik ez. Hori baino txikiagoa ez dago praktikan.

Beraz, Informazioaren Teoriak, informazioaren mamia alde batera uzten du, eta kantitatea besterik ez du begiratzen. Zenbat «bai ala ez» behar da galdera honi erantzuteko? Zenbat baialaezko entsaio edo proba egin behar da, halako fenomenoa garbitzeko? Zenbat sinbolo erabili behar da gertaera hau adiarazteko? Mintzaera jakin batean, zenbat da edozein letra batek eskuratutako

informazioa? Nola jokatu behar da, informazio bat ematekotan, energeiaren gastorik txikiena egiteko? Etabar...

Hona hemen Teoria honen helburuen exenplo sinple batzuk. Zienzia honen aplikazio-alorra nolako zabala denez, irakurlea errez konturatuko da, duda gabe.

Baina mintza gaitezen, lehenengoa informazioko *entropia* delakoaz.

Zalantzazko egoera baten entropia

Eman dezagun zalantzazko egoera batean gaudela. Adibidez, bi bide dugula aurrean. Gure mugarat heltzeko nondik jo ez dakigu, ordea. Seguro gaude, bidea, bata ala bestea izan behar dela, baina ez dakigu zein den. Bestaldetik, aukera egiteko, arrazoin berdinak ditugu, bi aldetatik, guretzat bi bideek probabilidade berdina dutelako.

Kasu honetan probabilidadeen banaketa (1/2; 1/2) dela esango dugu.

Zalantzazko egoera honetik ateratzeko, igarotzen den lehenengoari galdetzen diogu. «Mezedez, jauna, halako herrira nondik joaten da?».

Baina galdera hau, Informazio Teoriaren aldetik, ez da ondo egina. Beste gisaz egin behar dugu. Adibidez: «Halako herrira joateko, ezkerretikako bidea izango al da, ote?». Besteak «bai» ala «ez» erantzungo digu, eta kito.

Baialaezko erantzun batekin garbiturik dugu, beraz, kasu honetan, zalantzazko egoera.

Egoera honen entropia, *bat* dala esango dugu. Baina zer motako *bat*? Hemen, neurritze guztian bezala, unitate bat behar da. Esan dugunez, *bit* delakoa da, unitate hori. Aipatutako egoeraren entropia, «bit» bat dela esan dezakegu. (BINIT edo BIT, «bynary digit» ingeles esanbidetik dator.)

Entropia, zalantzazko egoera baten *zalantzaren neurria* da, eta zalantza zuritzeko *behar diren erantzun zenbakiaren arauaz neurritzen da*.

Hau pixka bat hobe erakusteko, ipuitxo bat kontatuko dizuegu.

Erregina eta zortzi damatxoak

Erregina batek zortzi damatxo omen zituen, eta hauen artean bere lehenengo dama aukeratzea erabaki zuen. Haietako bat autatu zuen, beraz, baina

honen izena «in petto» gorde zuen, Aita Santuak, kardenalak izendatzean, egiten omen duen bezala.

Damatxoek gogokoaren izena jakin nahi zuten, jakina, eta galdezka hasi zintzaizkion.

Erreginak, galdera batzuk egin zitzatela esan zien, eta berak erantzungo ziela, bi baldintzakin, ordea, alegia, «bai» ala «ez» baizik erantzungo ziela, eta galderaen zenbakia ahalik eta txikiena izan behar zela. Behar ziren galderak baino gehiago egiten baldin ba zuten, tutik ere ez zuela esango jakin arazi zien.

Damatxoak, orduan, nola jokatu asmatzen asi ziren.

Alfabetako ordenez Ainize, Dido, Eder, Erlea, Gabon, Krysis, Psykis eta Safo ziren, damatxoen izenak.

Erlea izena zuenak, nahiko sinplea zena, galderak bana banaka egitea proposatu zuen. Hau da: «Majestadea, Ainize izanen da, ote, zure gogokoa?»... «Majestadea, Dido izanen ote da, zure gogokoa?» etabar, izena agerrarazi arte.

Ainizek proposatutako jokabidea ez zen oso ona. Zori ona baldin ba zen, bai, izena berehala agertuko zen. Baina damatxo berezia azkenetako bat izanez gero, Psykis adibidez, galdera asko beharko zen, eta galdera gehiegi izanaz, erregina galdeketaren erdian ixilik geldituko zen.

Orduan, Krysis zelakoa, oso argia zena, beste proposamendu bat egin zuen, benetan ona.

Lehenengo galdera, honoko hau izango litzateke: «Majestadea, zure autatua, lehenengo lauen artean izango al da, ote?» (Lehenengo lauak, alfabetako ordenez, Ainize, Dido, Eder eta Erlea ziren, dakigunez).

Baina ala ez erantzungo zuen erreginak. Eman dezagun «ezetz» izan zala haren jardespena.

Orduan, bigarren galdera honoko hau izanen litzateke: «Majestadea, Gabon eta Krysis-en artean izanen da, ote, «in petto» daukazuna?». Bai ala ez ihardetsiko zuen erreginak. Onar dezagun *bai* izan zela jardespena.

Azkeneko galdera honoko hau izango litzateke, erreza da hau jakitea: «Gabon da, ote, zure gogokoa?».

Erreginak, nahiz «bai», nahiz «ez», erantzun, galdeketa bukaturik zegoen, esan beharrik ez, hirugarren galdera honekin. Hiru baialaezko galdera nahikoa zen egia jakiteko.

Aipatutako egoeraren *entropia* hiru «bit» dela esan dezakegu, beraz.

Gauza garrantzi handiko bat gertatzen da hemen. Frogatzen ahal danez, kasu honetan egia zuritzea, *hiru galdera baino gutxiagoren bidez ezin daiteke*.

Egoera jakin batean, *galdera zenbakirik txikiena, entropiak erakusten digu* hain zuzen ere.

Kalkulutxo batzuk, entropiari buruz

Ikusi dugunez, hiru zatiketa egiten dira «Krysis-en metodoaz».

Lehenengoa, bi erdietan, bakoitza lau damatxoko.

Bigarrena, bi laurdenetan, bakoitza bi damatxoko.

Hirugarrena, eta azkena, bi sortzirenetan, bakoitza damatxo bakarrekoa.

Beraz 8/2 = 4 4/2 = 2 2/2 = 1.

Eta horretatik 8 = 2 x 2 x 2 egiztatzen da, $8 = 2^3$ alegia.

Zortzi aterabide ditu egoera honek, eta entropia *3* da. Erlazioa da, beraz,

$$2^{\text{entropia}} = \text{aterabide zenbakia.}$$

Edozein kasoan, entropia ematen duan erlazioa honoko hau da

$$2^x = n$$

(*x* entropia, *n* aterabide zenbakia)

Hemendik erreza da beste honoko formula hau ateratzea

$$x = \frac{\log n}{\log 2} \text{ «bit»}$$

Adibidez, egoera batek hamar atera baldin badu, haren entropia

$$\frac{\log 10}{\log 2} = 3 \text{ bit gutxi gora bera da.}$$

Beraz, egoera jakin batean, *aterabide guztiek probabilidade berdina baldin ba dute* (baldintza hau nahi ta nahiezkoa da hontaz) entropia, eman berri dugun formulaz kalkulatzen da.

Aterabide guztiek probabilidade berdina baldin ez ba dute, entropiaren kalkuloa pixka bat zailagoa da.

Hau da, hain zuzen ere, Hartley eta Shannon-en teorien arteko desberdintasuna. Esan dugunez, Hartley izan zen zalantzaren neurria aztertzen hasi zen lehenengoa.

Baina Hartley, edozein egoeraz, aterabide guztiei probabilidade berdina ezartzen zien beti. Shannon, berriz, aterabide bakoitzaren probabilidade desberdina kontuan izaten hasi zen. probabilidade banaketa (p1, p2...) baldin ba da, entropia neurritzeko honoko formula hau asmatu eta frogatu zuen Shannon jaunak:

$$\frac{p_1 \log 1/p_1 + p_2 \log 1/p_2 + ...}{\log 2}$$

(besteak beste). Baina ez gera hemen sartuko holako kalkuluetan, irakurle matematikalariak beste tokian aurki ditzakelako.

Informazioaren kantitatea. Kriptografiarekiko exenplo bat

Hau ere, exenplo guztiz errezaren bitartez adiaraziko dugu.

Kriptolari bat gaztigu bat deskorapilatzen saiatzen da.

Gaztigu hortan, hitz bat agertzen da, kontestoaren guztiz berezia. Aginduzko hitz zifratu bat da, edo, garrantzi handia duana.

Zoritxarrez lehengo letra galdu da. Besteak, bai, heldu dira kriptolarien eskuetara.

Beste letra hauek *a*, *r* eta *a* dira. Hitza beraz, honela idazten da

A R A

eta toki utsan zenbait letra ezarri behar da.

Hitza, zer hizkuntzazkoa da? Occitanoa, portugesa, grekotarra, alemana, esperantoa...? Ez da jakiten. Gauza bat ba daki kriptolariak, alegia, falta den letra, latin alfabetako hogeitazortzi letraen artean aukeratua dela (eman dezagun 28 letra direla).

Lehen esan dugun arauaz, egoera honen entropia, 4'8 bit da, hau da, log 28/log 2.

Baina, zenbaitek *informazio* baliotsu bat ematen dio kriptolariari, azterketan zegoelarik.

Informazio honek, misteriozko hitza *euskal hitz bat dela dio.*

Agiri da, informazio hau hartzean, *zalantza gutxitu dela.* Zalantza edo entropia, guretzat gauza bera delako.

Lehenengo entropia, ondoko entropia baino handiagoa zen, duda gabe.

Azter dezagun, beraz, entropiaren balore berria, hau da, hartutako informazioaren ondoan zenbat balio du egoeraren entropiak.

Oker ez ba gaude, euskeraz, lau letrako *ara* bukaera duten hitzak, honoko hauek dira: *bara*, *gara*, *hara*, *jara*, *kara*, *mara*, *para*, *zara*, *xara.*

Bederatzi aterabide daude, eta hauen artean egin behar du kriptolariak bere aukera.

Egoera berriaren entropia, edo ondoko entropia, log 9/log 2 izanen da, alegia, 3'2 bit.

Ikusten dugunez entropia gutxitu da, informazioari esker.

Gutxitze hau 4'8 – 3'2 = 1'6 bit da.

Eta *hori da*, hain zuzen ere, kriptolariak hartutako *informazioaren balioa.*

Egoera jakin batean, gaztigu batek emandako informazioaren balioa, *lehenengo entropia eta ondoko entropiaren arteko liferentzia da.*

Kinieletan

Igandero, espainol gehienak kiniela egitea du bere eginkizun nagusia.

Kiniela osoak 314 aterabide desberdinak ditu. Kiniela jokalariaren zalantzak 22 bit du, entropiatzat (gutxi gora bera): log 3^{14} / log 2

Eman dezagun orain, partidu batean ondorea ematen zaigula.

Zenbat da informazio honen balioa?

Partidu batek, 3 aterabide ditu. Beraz, partidu baten ondorea jakiteak log 3 / log 2, alegia, 1'6 bit informazio ematen digu. Partidu baten informazioa, 1'6 bit inguru da.

Informazio hori hartu ondoan entropia 22 – 1'6 = 20'4 bit izango da.

Bi informazio izanez gero, 18'8 bit izango zen ondoko entropia.

Etabar.

Jakiña, 14 informazio hartzen ba ditugu, ondoko entropia ZERO izanen da, kiniela guztia ezagutzen dugulako.

Xuhurra eta diru faltsua

Xuhur batek mila urrezko diru gordeturik zituen, eltze handi baten ondoan. diruak onak ziran eta ederrak. Haien dirdirek, zikoitzaren atsegin eta zoriontasuna egiten zuten.

Baina itzalaren batek goibelarazten zuen gizonaren bizia.

Diruetako bat, faltsua zen, besteak baino pisu txikiagoa. Gainera, liferentzia oso txikia zen eta diru faltsua eskuz kausitzea ezina zen.

Hori zela eta, egunen batean, bera xuhur bezain lagun bati, balantzatxo bat eskatu zion.

Besteak, prestatu ez, alogeratuko ziola esan zion ordea. Alogera zenbat izango zen, galdetu zion, orduan, gure gizonak.

«Haztatze edo hazta bakoitza gaitik txanpon bat eman behar didazu —erantzun zion lagunak—. Bestaldetik, pisuak edo pisu neurriak ez dauzkat eta, nola konpondu, zuronek ikusiko duzu».

Gure zikoitza haztakinakin nola jokatu pentsatzen hasi zan, ahalik eta haztatzerik gutxiena egiteko.

Egoerak mila aterabide zituen. Beraz haren entropia 10 bit zan, zenbaki borobilatuez (Log 1.000/Log 3).

Bestaldetik hazta bakoitzak hiru aterabide izanik, dagokion informazioa 1'6 bit da.

Argia da 10/1'6 hazta egin behar zala, gutxienez informazio osoa lortzeko. Sei hazta ez zan nahikoa. Zazpiak egin behar ziren, gutxienez, eta zazpi txanpon ordaindu.

Hona hemen xuhurrak nola jokatu zuen.

333 diru jarri zituen haztakinaren platera batean, eta beste 333, bestean. Haztakina plomuan gelditzen baldin ba zan, gainerako 334 diruen artean aurkitzen zen, duda gabe, diru faltsua. Platera bat gora egiten baldin ba zen, ordea, diru faltsua platera hortan bertan aurkitzen zen.

Ematen baldin ba dugu bi platerek pisu berdina zeukatela, zalantza oraindik 334 diruen artean dago (Beste kasuan 333, izanen litzateke besterik ez, eta arazoa errezagoa izan).

Zer egin orduan? 111 diru platera batean jarri eta beste 111, bestean. Lehen bezala, kasorik txarrenaz, 112 kasoen artean gelditzen zan zalantza. Etabar.

Ez gara jarraituko. Irakurleak erreza du kalkulo hau bukatzea. Eta, kasurik txarrenean ere, zazpi hazta egin eta, diru faltsua agertzen dela ikusiko du.

Bukaera

Exenplo sinple ta errez batzuk erakutsi ditugu, besterik ez. Umekeria ba dirudi ere, teoria hau oso ederra da, eta garrantzi handikoa. Aldizkari honetan zerbait gehiago esatea ez da posible, ordea.

Irakasleek hortaz pixka bat ikasiko balute, ikastoletan oso ariketa politak egin litezke.

---

1. Nahasietatik itzuritzeko, *mandatu* eta *manu* alde batera utzita, *gaztigu* eta *agindu* hitzak erabiliko ditugu gehienetan: «*gaztigu*» («mensaje, message»); «*agindu*» («orden, ordre»).