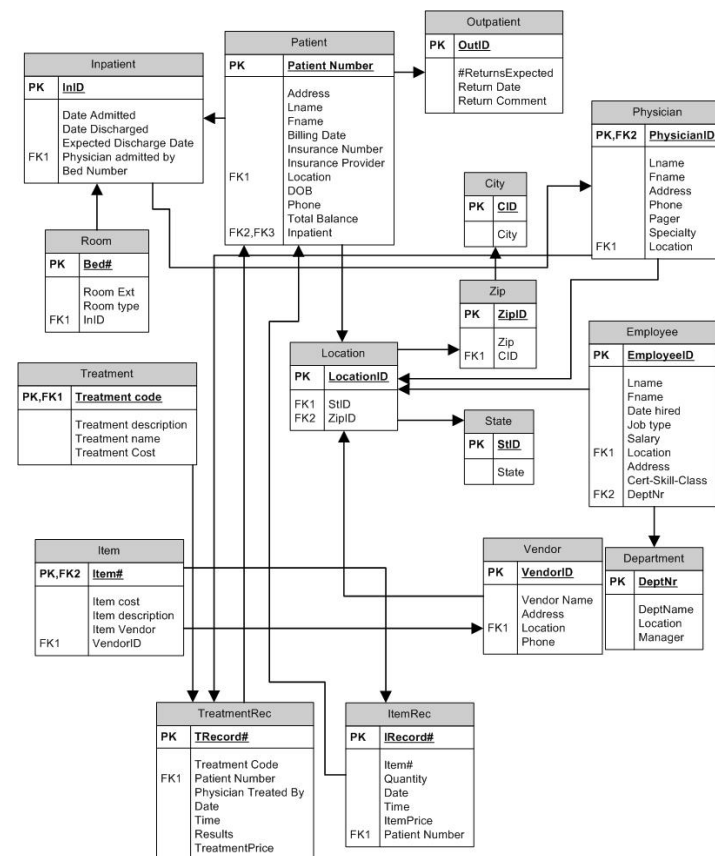
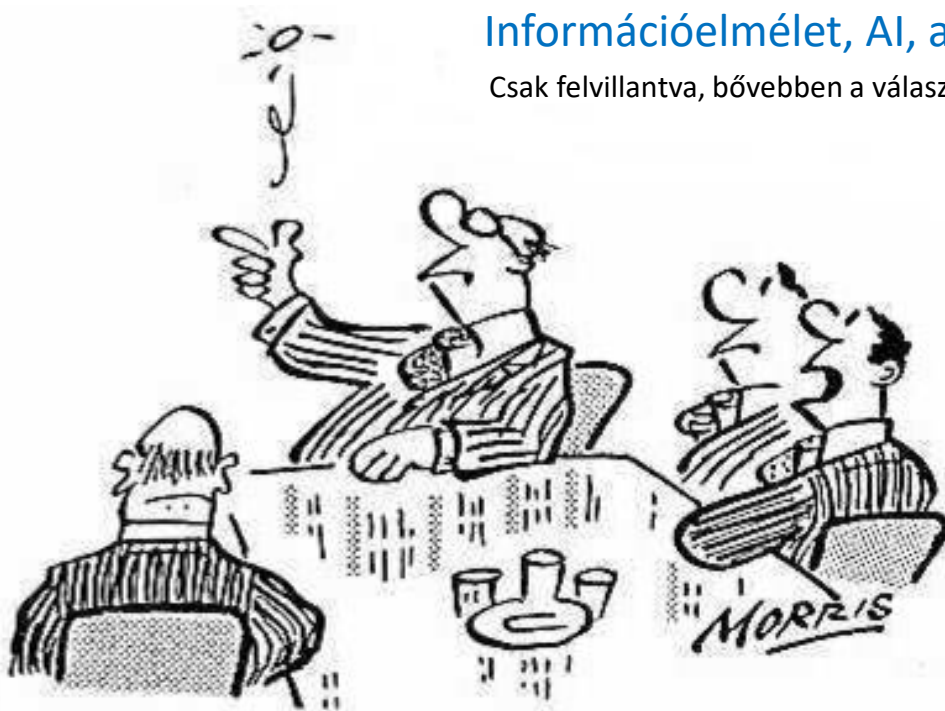
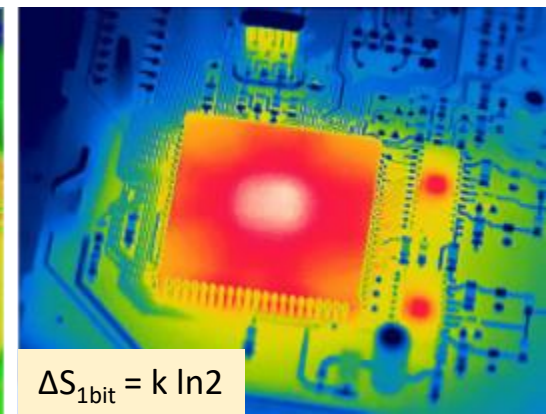
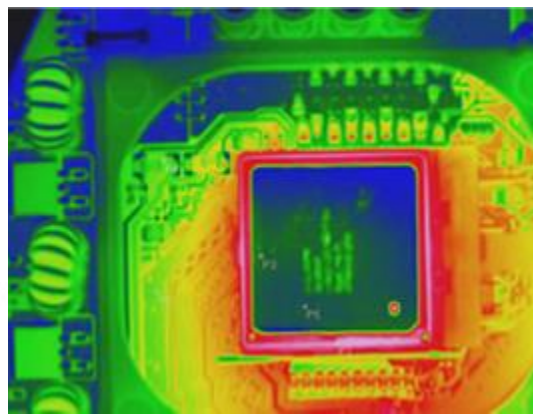
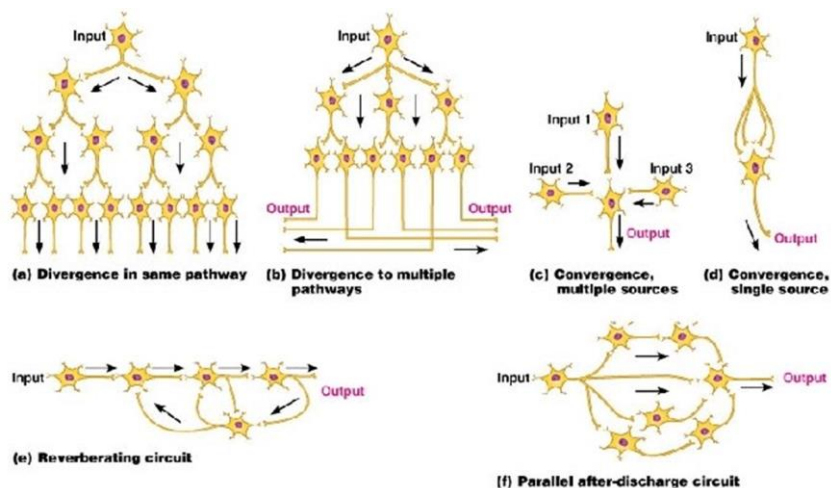


Információelmélet, AI, adatbázisok

Csak felvillantva, bővebben a választható kurzusban



Bárcsak olyan nyugodt lehetnék mint J.B. amikor fontos döntésekről van szó!



$$\Delta S_{1\text{bit}} = k \ln 2$$

Az információ fogalma (példával)

Intuitívan

"informare" (Lat.) : „**az elmét formálni**”, tanítani, utasítani valakit

Azaz: akkor tudunk tanulni, vélekedésünket megváltoztatni, ha *információhoz* jutunk

vagy:

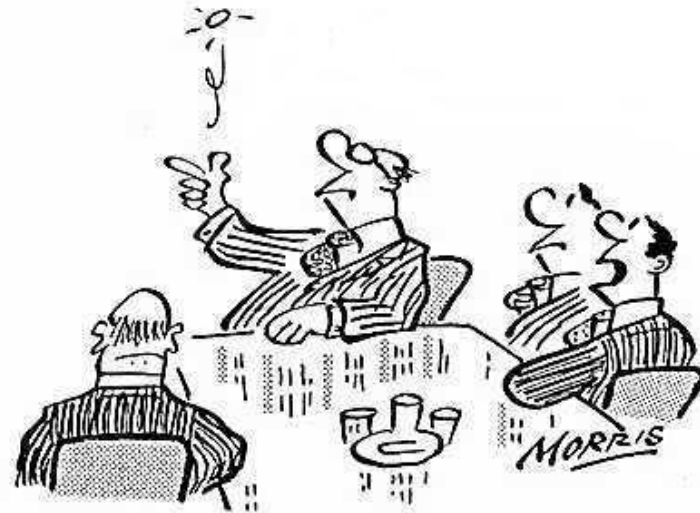
„egy eszközbe vagy élőlénybe bevitt jel, mely választ vált ki”

(Pl. Pavlovi reflex: táplálék illata → nyáleválasztás, mozdulatok)

vagy:

„ az információ olyan mintázat amely más mintázathoz

(Pl. DNS szekvencia → fehérje szerkezet)



Bár csak olyan nyugodt lehetnék mint J.B. amikor fontos döntésekről van szó!

Információátvitel - kódolás

általánsságban

Információ forrás

kódolás



Átviteli csatorna

dekódolá
s



**(át)vevő
cél(pont)**

A lehetséges események közül
melyik következett be?

eseményeket SZÁMOKKAL reprezentálunk

SZÁMOKBÓL visszaállítjuk az
eseményeket

(hír)

Információátvitel - kódolás

általánosságban

Információ forrás

kódolás

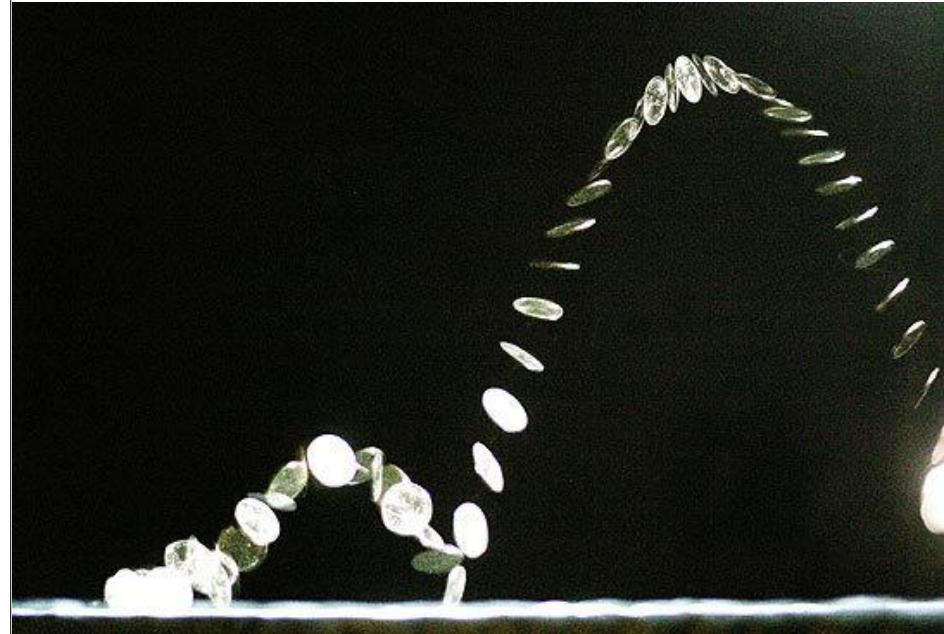
Átviteli csatorna

dekódolá
s

**(át)vevő
cél(pont)**

példa

Fej vagy írás?



Információátvitel - kódolás

SZÁMOKKÁ
alakítjuk az információt.

általánosságban

Információ **forrás**

kódolás

Átviteli **csatorna**

dekódolás

(át)vevő
cél(pont)



A példában

Fej vagy írás?

kódolás

Fej/Írás → **1,0**

Beszéd, SMS, rádió, stb.

dekódolás

1,0 → Fej/Írás

Ki nyert?



Információátvitel – digitális kódolás

Hány **bitre** van szükségünk a kódoláshoz?

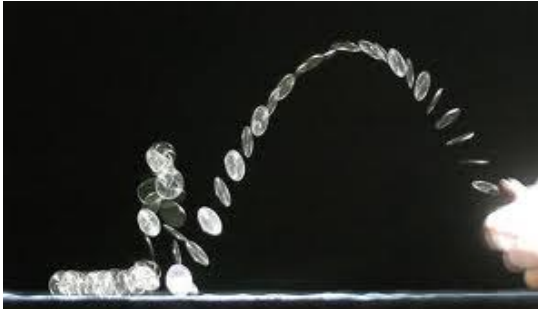
Bit: **b**inary
digit



0 vagy
1

Számjegyek : csak *két* jegy van: 0 és 1.
(10-es rendszerben 9 jegy van: 0,1,2,...,9)
A számjegyeknek helyiértéke van, amit az alap hatványai adnak meg:
 2^2 , 2^1 , 2^0
(10-esben $10^2 = 100$, $10^1 = 10$, $10^0 = 1$)







$$\text{Pl: } 101 = 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 4 + 0 + 1 = 5$$

Információátvitel – digitális kódolás


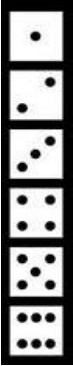


Esemény		szám	digitális kód
	:	1	1
	:	0	0









	:	1	001
	:	2	010
	:	3	011
	:	4	100
	:	5	101
	:	6	110

Információátvitel – kódolási hatások

Esemény szám		digitális kód		Bitek száma		Maximális eseményszám
	:	1	1	1		2
	:	0	0			
	:	1	001	3		8
	:	2	010			
	:	3	011			
	:	4	100			
	:	5	101			
	:	6	110			
		7	111			
		0	000			

Itt 6 eseményt kódoltunk 3 biten, csak hogy $2^3 = 8$, azaz 8 eseményt is lehetne kódolni ennyi bit segítségével

Információátvitel – kódolási hatások

Esemény	szám	digitális kód	Bitek száma	Maximális eseményszám
 :	1	001	3	8
 :	2	010		
 :	3	011		
 :	4	100		
 :	5	101		
 :	6	110		
	7	111		
	0	000		

Itt 6 eseményt kódoltunk 3 biten, csak hogy $2^3 = 8$, azaz 8 eseményt is lehetne kódolni ennyi bit segítségével

Egy jobb hatásfokú kódolás:

$\{X_1 X_2 X_3\}$ csoportosítsuk az eseményeket 3-asával

Ez az előbbiek szerint
3x3 bit = 9 bit igényt jelent

Összesen $6^3 = 216$ lehetőség van
de $2^8 = 256$, így 8 bit is elég
(sőt sok is).

1 bittel kevesebb is bőven elég!

azaz $8/3 = 2.66$ bit is elegendő

Információátvitel – kódolási hatások

Az információtartalom definiálható a legjobb hatásfokú kódolással

azaz:

Egy jel, esemény információtartalma megadható azzal, hogy minimálisan hány bitre van szükség az átviteléhez.

Ez egyben a kódolási hatásfok elméleti határa is.

Többször nem egyetlen átvitelre vagyunk kíváncsiak, hanem arra, hogy egy több lehetséges eseményből álló *eseménytér* hogyan kódolható.

Információátvitel – kódolási hatások

Az információtartalom definiálható a legjobb hatásfokú kódolással

azaz:

Egy jel, esemény információtartalma megadható azzal,
hogy

minimálisan hány bitre van szükség az átvitléhez.

Ez egyben a kódolási hatásfok elméleti határa is.

Hogyan kapcsolódik ez az intuitív
információtartalomhoz?

	p	q
-Fej vagy írás?	$\frac{1}{2}$	$\frac{1}{2}$
-Ma reggel nincs dugó.	$\frac{1}{4}$	$\frac{3}{4}$
-Holnap esni fog.	1%	99%
-nyertem a lottón!	1/13,983,816	0.999....

Információátvitel – kódolási hatások

Az információtartalom definiálható a legjobb hatásfokú kódolással

azaz:

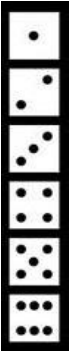
Egy jel, esemény információtartalma megadható azzal, hogy minimálisan hány bitre van szükség az átvitléhez. Ez egyben a kódolási hatások elméleti határa is.

Hogyan kapcsolódik ez az intuitív információtartalomhoz?

	p	q	
-Fej vagy írás?	$\frac{1}{2}$	$\frac{1}{2}$	nem tudjuk előre
-Ma reggel nincs dugó.	$\frac{1}{4}$	$\frac{3}{4}$	
-Holnap esni fog.	1%	99%	
-nyertem a lottón!	$\frac{1}{13,983,816}$	0.999	valószínű nem nyert...

A nyert információ az esemény valószínűségével fordítottan arányos!

információátvitel – az információ mértéke

„rendes”	p	valószínűség	kódolási példa	bit igény	p^* (bitek száma)
	1/6	0.17	000	3	0.5
	1/6	0.17	001	3	0.5
	1/6	0.17	010	3	0.5
	1/6	0.17	011	3	0.5
	1/6	0.17	100	3	0.5
	1/6	0.17	101	3	0.5

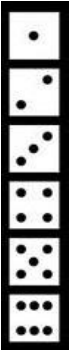
itt nem tudunk előre
semmit, azaz
a bizonytalanság
maximális

bitek számának várható értéke:

3
(2.66)

cinkelt P

A cinkelt kocka eredményeinek kisebb a „hírértéke”, kevesebbet tanulunk belőle

	1/2	0.5	0	1	0.5
	1/4	0.25	10	2	0.5
	1/8	0.13	110	3	0.38
	1/16	0.06	1110	4	0.25
	1/32	0.03	11110	5	0.16
	1/32	0.03	11111	5	0.16

itt van előismeretünk
az 1-est esetleg a 2-
est
várjuk.

Átlagosan
kevesebbet
„tanulunk” a cinkelt
kockával

bitek számának várható értéke:

1.94

információátvitel – az információ mértéke

Hogyan lehet **matematikailag** definiálni az információtartalmat (H)? (Shannon 1948)

1.: **H legyen p folytonos függvénye** (ha p egy kicsit változik \rightarrow H sem „ugorhat” nagyot)

2.: **Ritka események információtartalma nagyobb mint a gyakoriaké:**

H fordítottan legyen arányos p -vel

Ha minden p egyforma, ($p = 1/n$)

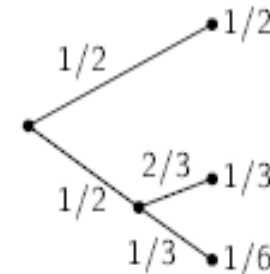
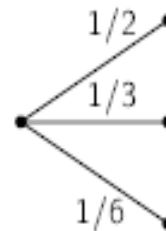
Akkor H az események számának (n) monoton növekvő függvénye legyen.

Ha több esemény lehetséges, akkor több választásra van lehetőség, nagyobb a bizonytalanság
Ekkor egy esemény bekövetkezése nagyobb a hírértéke.

3.: **Elágazó választási lehetőségek:**

H értéke nem függhet attól, hogy milyen úton jutottunk az adott választáshoz
(esemény bekövetkezéséhez)

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



információátvitel – az információ mértéke

Shannon :

$$H = p \cdot \log_2 \left(\frac{1}{p} \right)$$

H az átlagos, várható információtartalmat adja meg, ez hozható kapcsolatba a kódolással.

Szokás megadni egyetlen esemény információtartalmát is (I):

$$I = \log_2 \left(\frac{1}{p} \right) \quad \text{Ezzel } H = p \cdot I, \text{ azaz a bekövetkezési valószínűséggel súlyozott információtartalom.}$$

Ha a teljes eseményteret I_a akarjuk fedni, akkor összegezni kell minden eseményre

$$H = \sum_i p_i \cdot \log_2 \left(\frac{1}{p_i} \right) = \sum_i -p_i \cdot \log_2 p_i$$

az információ mértéke - entrópia

Fair érme: $p = \frac{1}{2}$

Fej vagy írás

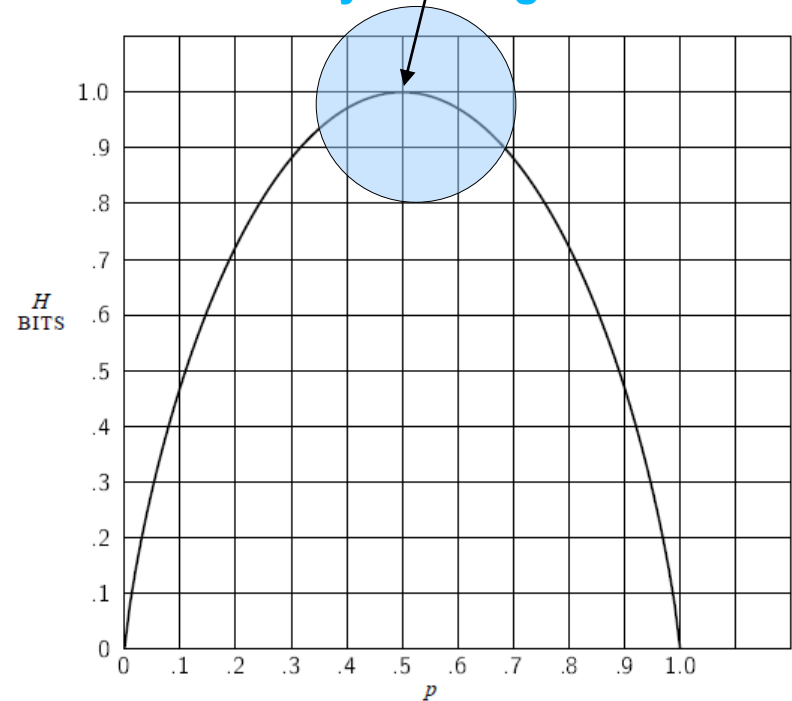


p



$q = 1-p$

Nincs előfeltevés,
Maximális
bizonytalanság



$$\begin{aligned} H &= \sum_i -p_i \cdot \log_2 p_i = -p \cdot \log_2 p - q \cdot \log_2 q \\ &= -p \cdot \log_2 p - (1-p) \cdot \log_2 (1-p) \end{aligned}$$

az információ mértéke - entrópia

Fair érme: $p = \frac{1}{2}$

Fej vagy írás



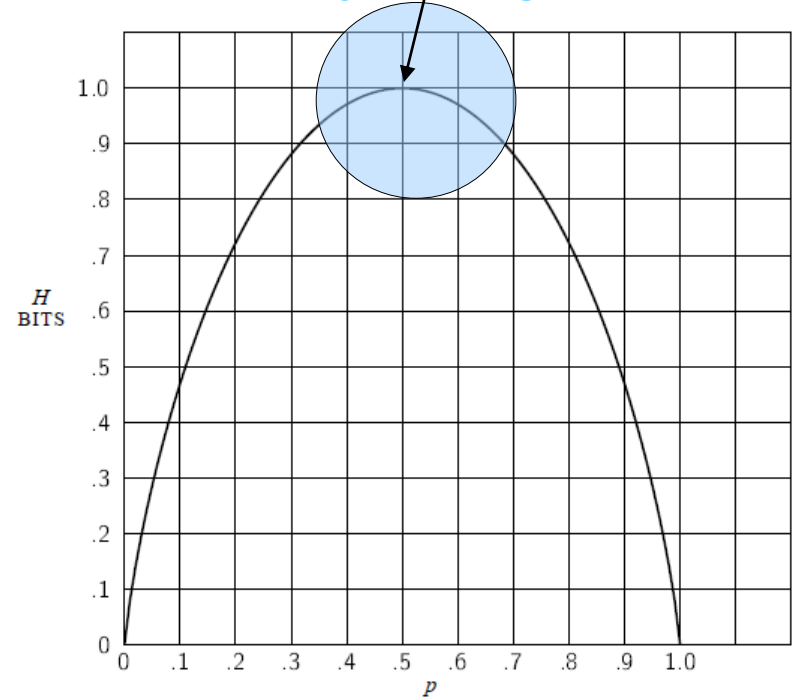
p



$q = 1 - p$

H másik neve: **Shannon-entrópia**

Nincs előfeltevés,
Maximális
bizonytalanság



H akkor **maximális** ha semmit sem tudunk előre azaz minden eseményhez egyforma p ($= 1/n$)

Minden lehetőség egyformán valószínű, legtöbbféle előfordulást látjuk

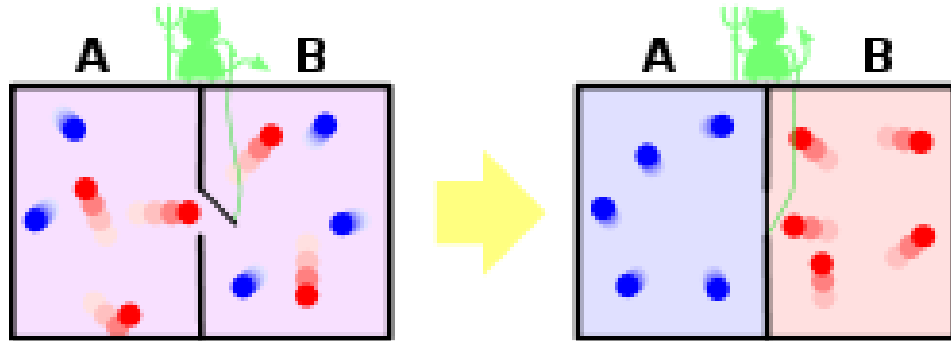


A fizikai entrópia (S) maximuma is ott van ahol a mikroállapotok száma maximális

Információs és fizikai entrópia kapcsolata

„ izolált rendszerben az entrópia nem csökkenhet” Termodinamika II. főtétele

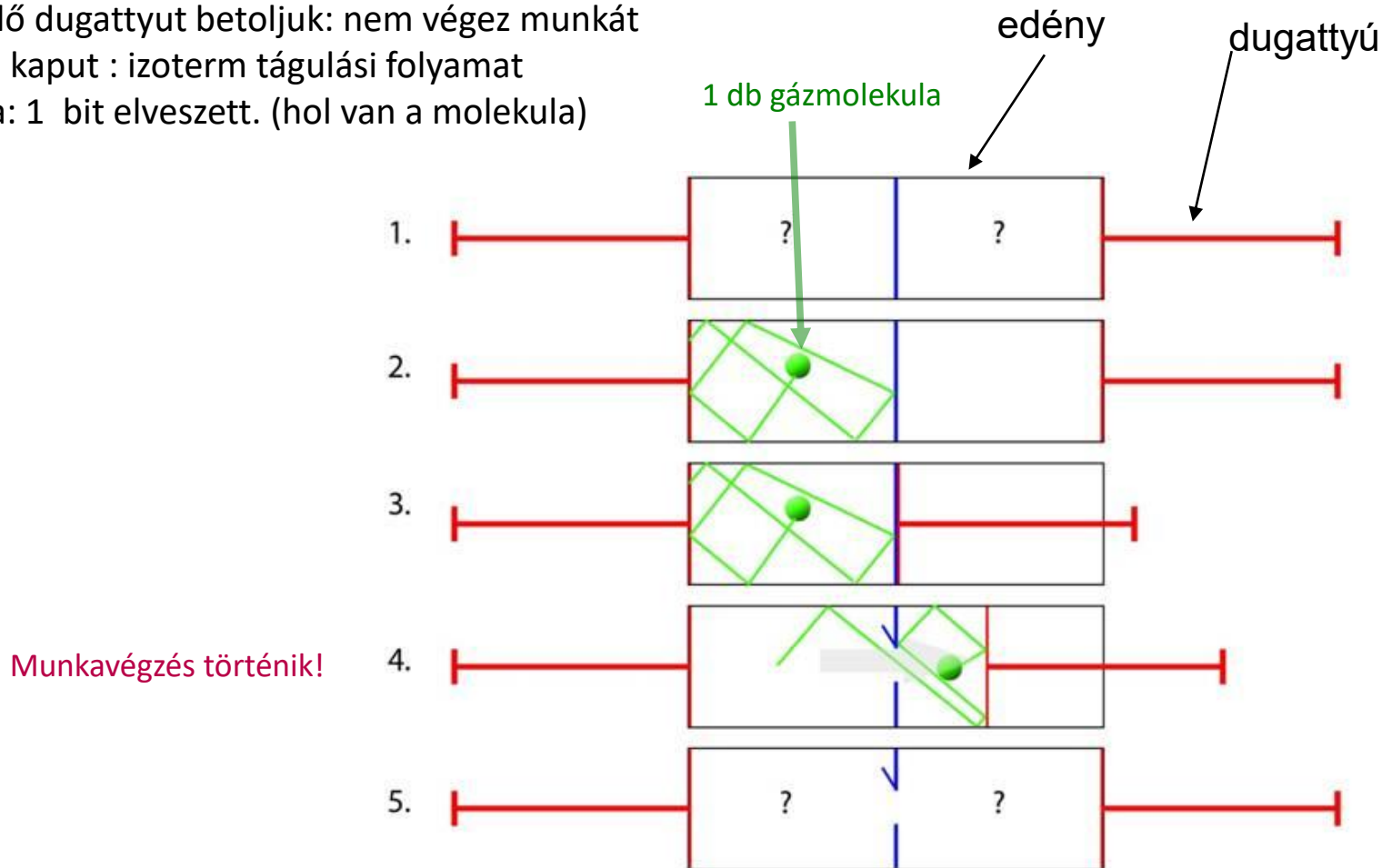
A Maxwell démon



A hőmérséklet a két oldalon: A **csökken**, B **növekszik** → II. főtétele ? (A zárt rendszerünk = A+B)

megoldás: NEM, mert a Maxwell-démon is része a rendszernek, vagy ha nem akkor nem izolált a rendszer. Ha része, akkor viszont a démon **információt** szerez vagy veszít!

1. : molekula helyzete ismeretlen
2. : megmérjük melyik oldalon van: információ = 1 bit
3. : a megfelelő dugattyút betoljuk: nem végez munkát
4. : kinyitjuk a kaput : izoterm tágulási folyamat
5. : ajtó nyitva: 1 bit elveszett. (hol van a molekula)



Ingyen munka??

1. : molekula helyzete ismeretlen
2. : megmérjük melyik oldalon van: információ = 1 bit
3. : a megfelelő dugattyút betoljuk: nem végez munkát
4. : **kinyitjuk a kaput : izoterm tágulási folyamat**
5. : ajtó nyitva: 1 bit elveszett. (hol van a molekula)

Izoterm tágulási folyamat:

$$W_{A \rightarrow B} = NkT \ln \left(\frac{V_A}{V_B} \right)$$

Itt most:

$N=1$

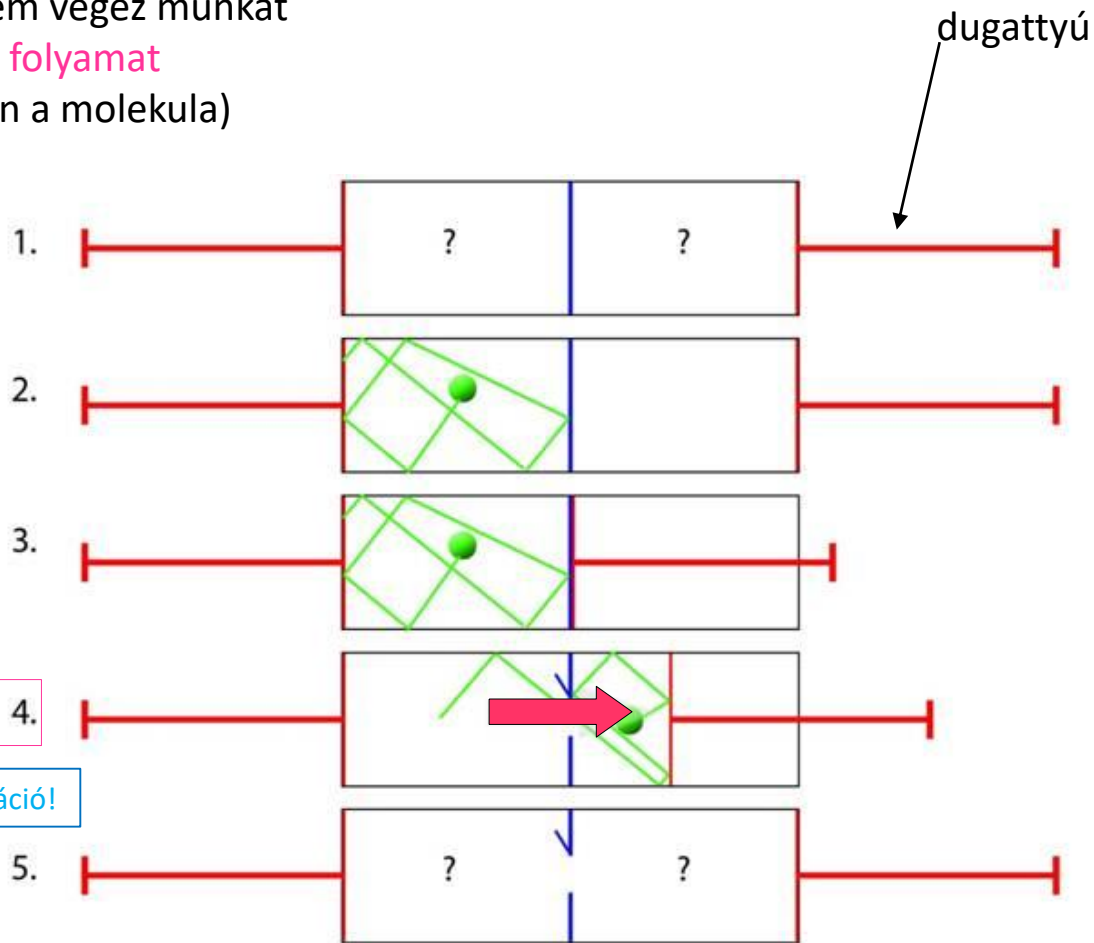
$V_A/V_B = 2$

azaz

$$W = kT \ln(2)$$

Munkavégzés történik!

Itt pedig elvész 1 bit információ!



Információs és fizikai entrópia

Szilárd Leó:

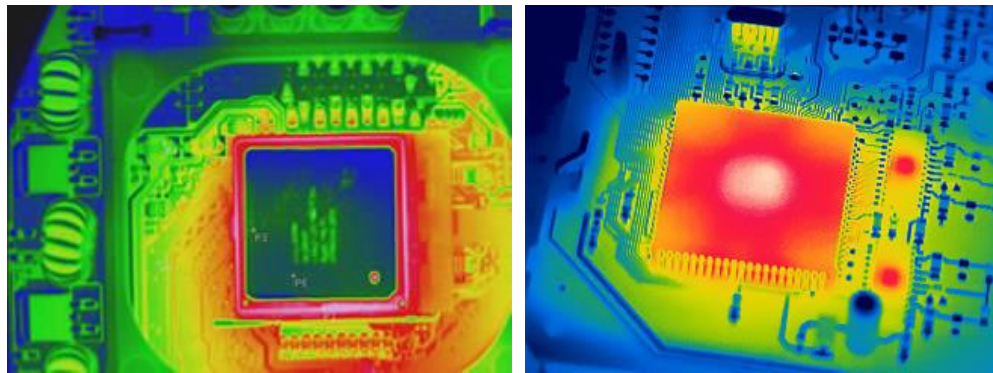
A II. főtételből, és tudva hogy $W = T\Delta S$

$$T\Delta S = kT \ln 2$$

$$\Delta S = k \ln 2$$

1 bit információ elvesztése egyenértékű a fizikai entrópia $k \cdot \ln 2$ mértékű növekedésével.

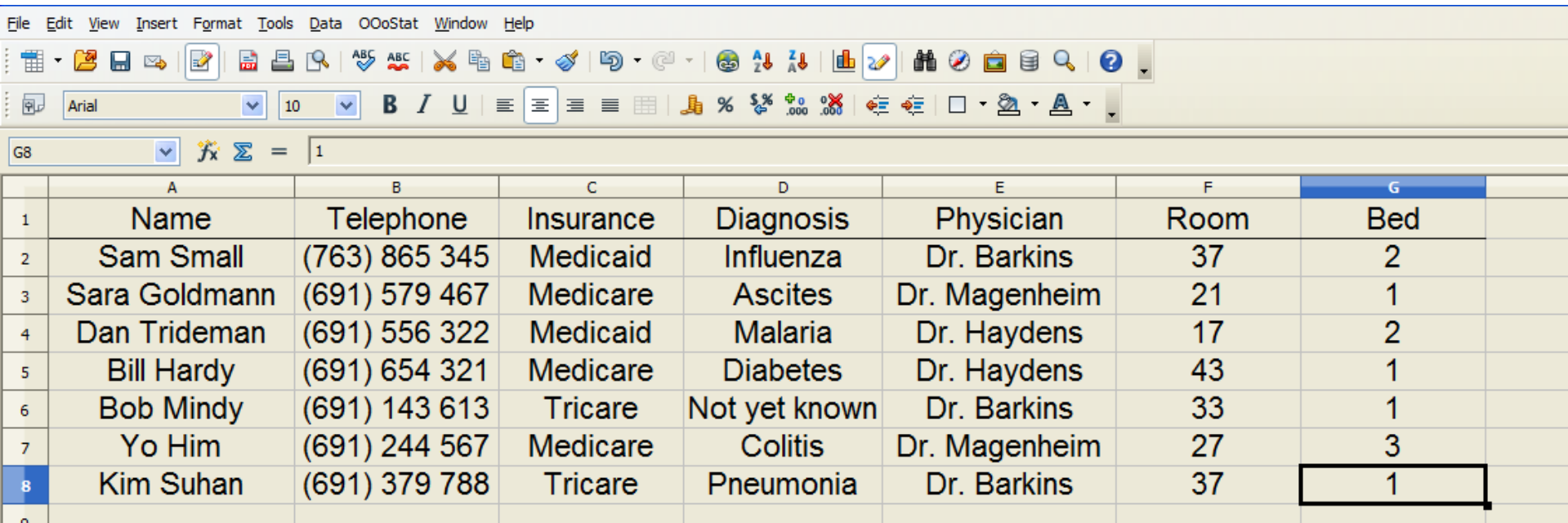
(Landauer 1971, logikailag nem visszafordítható folyamatok, pl. ÉS, VAGY)



Adatbázisok

Adatbázisokban ***információt*** tárolnak, rendszereznek és olvasnak ki.

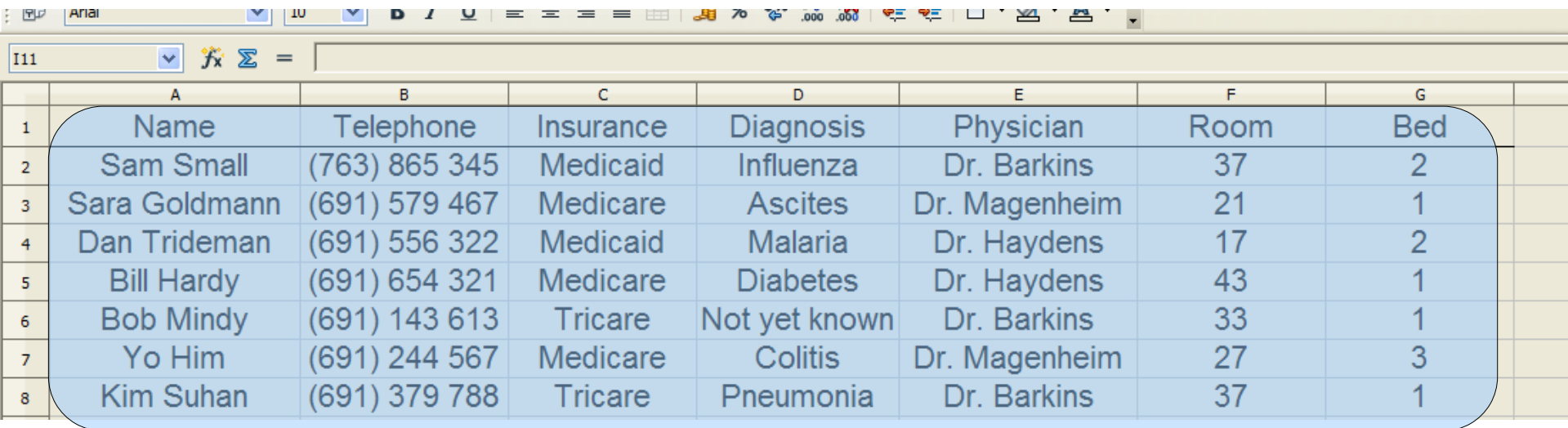
Adatbázisok - információtárolás



The screenshot shows a spreadsheet application interface. At the top is a menu bar with options: File, Edit, View, Insert, Format, Tools, Data, OOoStat, Window, and Help. Below the menu is a toolbar with various icons for file operations, editing, and formatting. A status bar at the bottom of the toolbar shows the font 'Arial', size '10', and bold, italic, and underline formatting options. The spreadsheet itself has a grid with columns labeled A through G. The data is organized into a table with 8 rows. The first row (row 1) contains headers: Name, Telephone, Insurance, Diagnosis, Physician, Room, and Bed. The subsequent rows (rows 2-8) contain patient data. The cell at row 8, column G (containing the value '1') is highlighted with a black border. The formula bar at the top of the spreadsheet shows 'G8' and the value '1'.

	A	B	C	D	E	F	G
1	Name	Telephone	Insurance	Diagnosis	Physician	Room	Bed
2	Sam Small	(763) 865 345	Medicaid	Influenza	Dr. Barkins	37	2
3	Sara Goldmann	(691) 579 467	Medicare	Ascites	Dr. Magenheim	21	1
4	Dan Trideman	(691) 556 322	Medicaid	Malaria	Dr. Haydens	17	2
5	Bill Hardy	(691) 654 321	Medicare	Diabetes	Dr. Haydens	43	1
6	Bob Mindy	(691) 143 613	Tricare	Not yet known	Dr. Barkins	33	1
7	Yo Him	(691) 244 567	Medicare	Colitis	Dr. Magenheim	27	3
8	Kim Suhan	(691) 379 788	Tricare	Pneumonia	Dr. Barkins	37	1

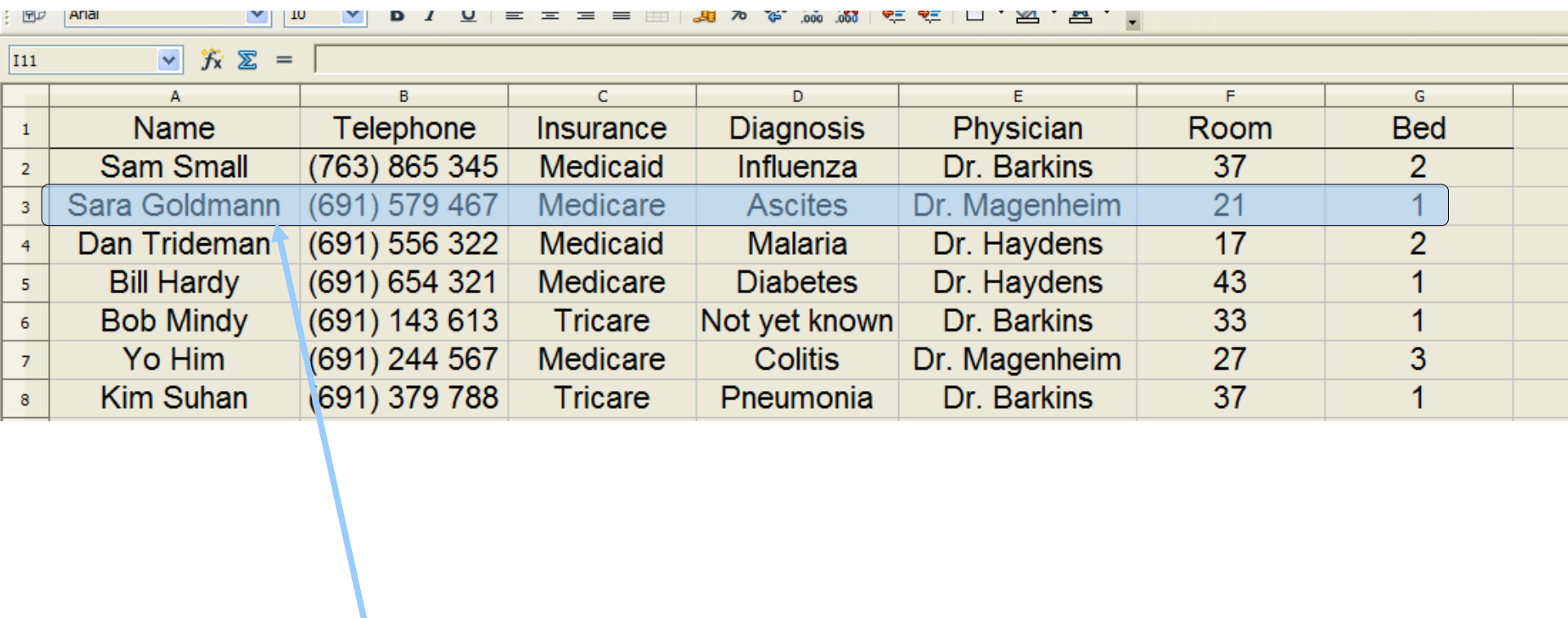
Adatbázisok - információtárolás



	A	B	C	D	E	F	G
1	Name	Telephone	Insurance	Diagnosis	Physician	Room	Bed
2	Sam Small	(763) 865 345	Medicaid	Influenza	Dr. Barkins	37	2
3	Sara Goldmann	(691) 579 467	Medicare	Ascites	Dr. Magenheim	21	1
4	Dan Trideman	(691) 556 322	Medicaid	Malaria	Dr. Haydens	17	2
5	Bill Hardy	(691) 654 321	Medicare	Diabetes	Dr. Haydens	43	1
6	Bob Mindy	(691) 143 613	Tricare	Not yet known	Dr. Barkins	33	1
7	Yo Him	(691) 244 567	Medicare	Colitis	Dr. Magenheim	27	3
8	Kim Suhan	(691) 379 788	Tricare	Pneumonia	Dr. Barkins	37	1

Tábla : rendezett adathalmaz

Adatbázisok - információtárolás



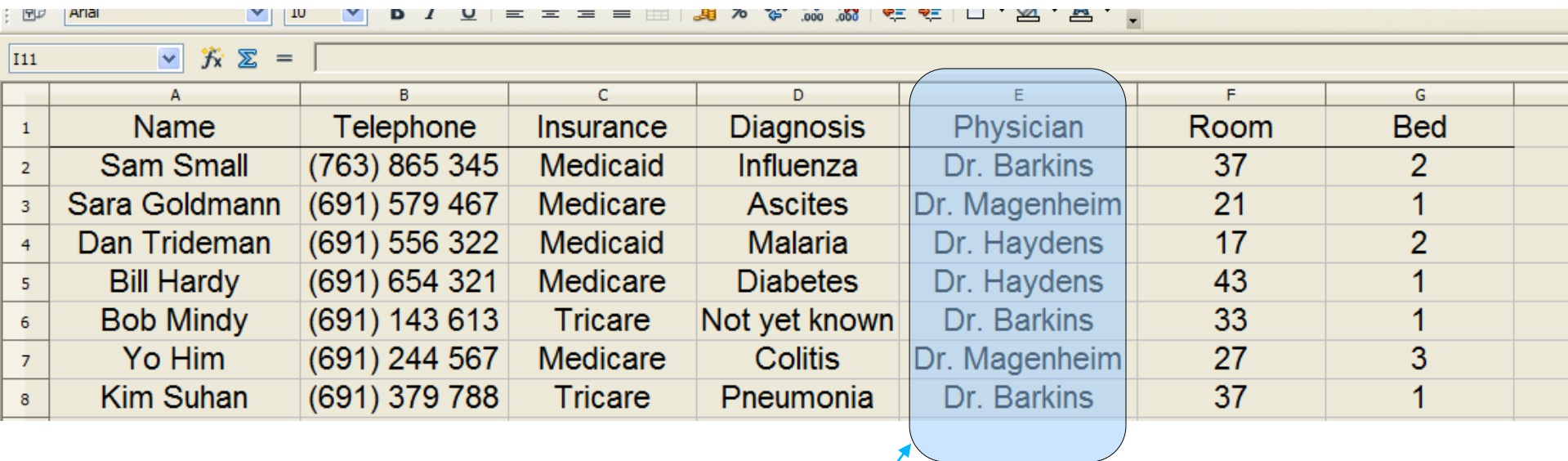
	A	B	C	D	E	F	G
1	Name	Telephone	Insurance	Diagnosis	Physician	Room	Bed
2	Sam Small	(763) 865 345	Medicaid	Influenza	Dr. Barkins	37	2
3	Sara Goldmann	(691) 579 467	Medicare	Ascites	Dr. Magenheim	21	1
4	Dan Trideman	(691) 556 322	Medicaid	Malaria	Dr. Haydens	17	2
5	Bill Hardy	(691) 654 321	Medicare	Diabetes	Dr. Haydens	43	1
6	Bob Mindy	(691) 143 613	Tricare	Not yet known	Dr. Barkins	33	1
7	Yo Him	(691) 244 567	Medicare	Colitis	Dr. Magenheim	27	3
8	Kim Suhan	(691) 379 788	Tricare	Pneumonia	Dr. Barkins	37	1

Rekord : csoportosított információcsomag
(egy sor a táblában)

Minden sor egy adat-csoport

Minden sor szerkezete azonos

Adatbázisok - információtárolás

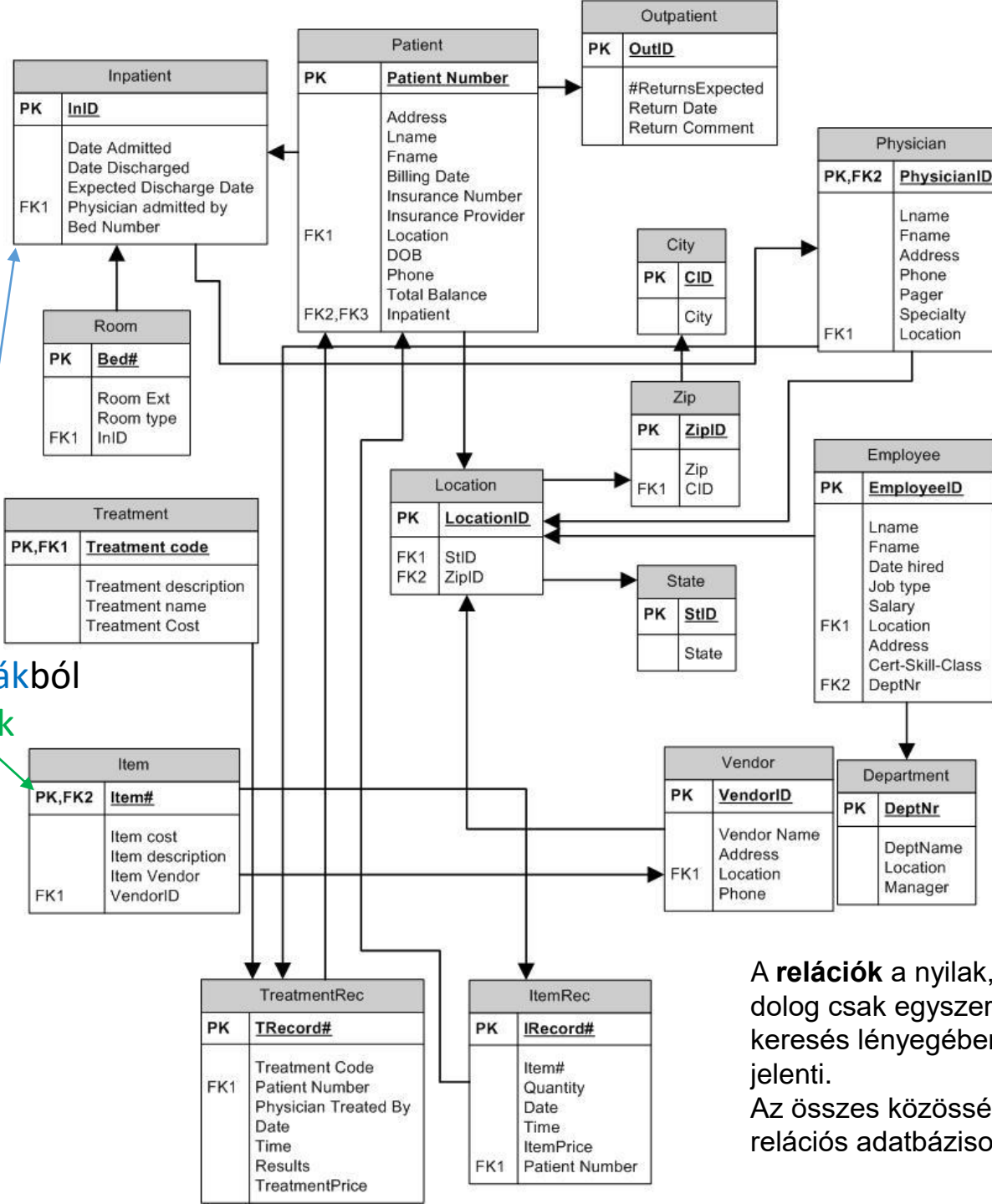


	A	B	C	D	E	F	G	
1	Name	Telephone	Insurance	Diagnosis	Physician	Room	Bed	
2	Sam Small	(763) 865 345	Medicaid	Influenza	Dr. Barkins	37	2	
3	Sara Goldmann	(691) 579 467	Medicare	Ascites	Dr. Magenheim	21	1	
4	Dan Trideman	(691) 556 322	Medicaid	Malaria	Dr. Haydens	17	2	
5	Bill Hardy	(691) 654 321	Medicare	Diabetes	Dr. Haydens	43	1	
6	Bob Mindy	(691) 143 613	Tricare	Not yet known	Dr. Barkins	33	1	
7	Yo Him	(691) 244 567	Medicare	Colitis	Dr. Magenheim	27	3	
8	Kim Suhan	(691) 379 788	Tricare	Pneumonia	Dr. Barkins	37	1	

oszlop: adat **típus**

A relációs adatbázisban minden csak egyszer tárolódik, az egyes adat-típusok közötti kapcsolatokat tároljuk.

Ehhez egy matematikai leírás és programozási nyelv is tartozik, tetszőleges feladatra összeállítható egy adatbázis-szerkezet. Az adatbázisban a kapcsolatok az információ-áramlás irányát is jelzik.



Az adatbázis táblákból áll, amiket kulcsok kötnék össze

A relációk a nyilak, az adatbázisban egy dolog csak egyszer tárolódik, és egy keresés lényegében a nyilak követését jelenti.
Az összes közösségi oldal mögött pl relációs adatbázisok dolgoznak...

Orvosi alkalmazások

Diagnosztikus tesztekől nyert információ

Szakértői rendszerek

Genetikai adatbázisok

Proteomika

...

a : preteszt valósz.

b : poszt teszt valósz.

$$D(b||a) = \sum_{i=1}^n b_i \log_2(b_i/a_i)$$

Testing Situation	Pretest Probability of Disease	Test Operating Characteristics: Sensitivity/Specificity	Test Result	Posttest Probability of Disease	Information Gained
Breast cancer screening with mammography	0.01	0.75/0.94	Positive Negative	0.11 0.003	0.25 bits 0.006 bits
Mammography given palpable breast mass	0.2	0.80/0.90	Positive Negative	0.67 0.05	0.74 bits 0.13 bits
Screening for HIV with antibody test	0.001	0.99/0.998	Positive Negative	0.33 0.00001	2.4 bits 0.001 bits
Presence of tonsillar exudate in diagnosing infection with group A streptococci	0.1	0.45/0.84	Positive Negative	0.24 0.07	0.11 bits 0.01 bits
Colon cancer screening by fecal occult blood testing	0.005	0.40/0.90	Positive Negative	0.02 0.003	0.02 bits 0.0005 bits

Kullback–Leibler Divergencia

$$D_{KL}(P|Q) = \sum_i p_i \cdot \log \left(\frac{p_i}{q_i} \right)$$

$$D_{KL}(P|Q) = \int_{-\infty}^{+\infty} p(x) \cdot \log \left(\frac{p(x)}{q(x)} \right) dx$$

Shannon-Entropiával kapcsolatos:

$$H = \log(N) - D_{KL}(p(x)|PE_N)$$

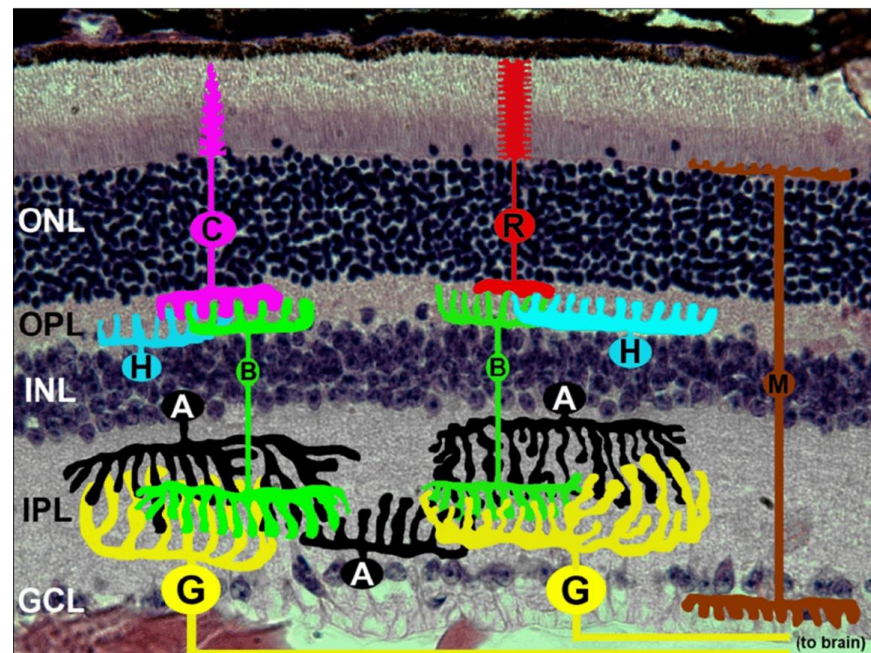
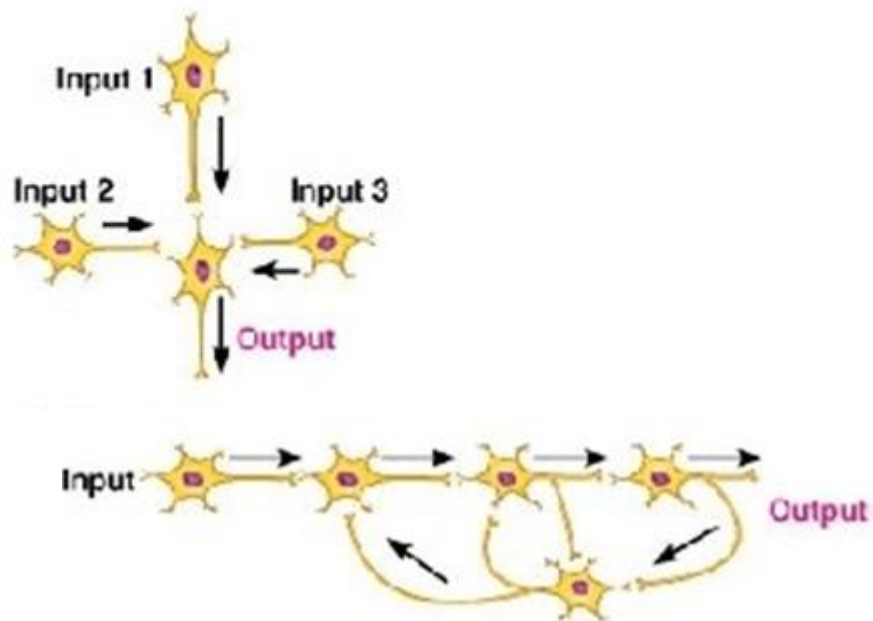
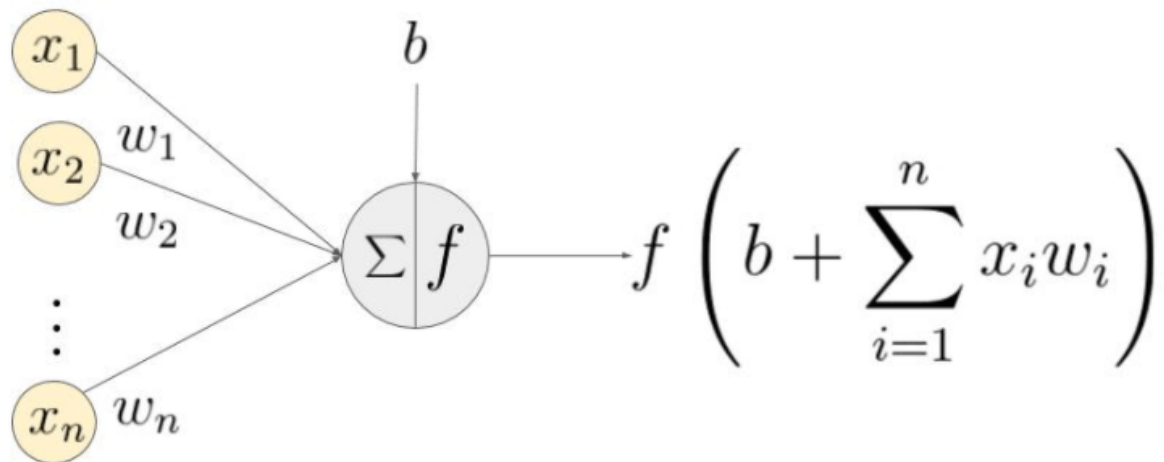
A kapott,
kérdéses eloszlás

PE_N : N elemű egyeletes eloszlás
(legkevésbé informált prior)

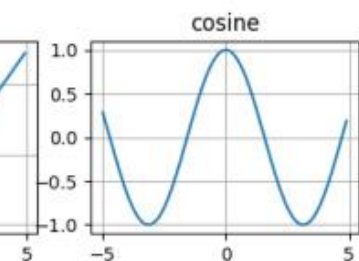
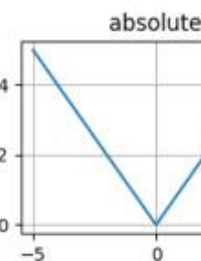
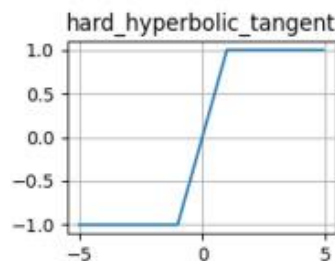
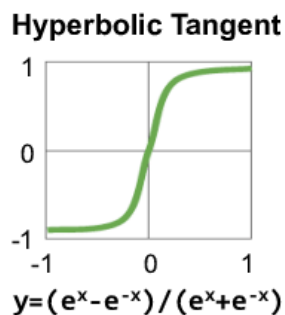
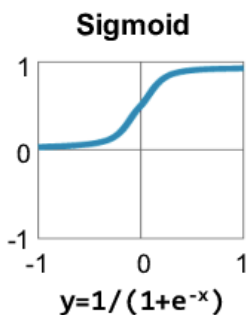
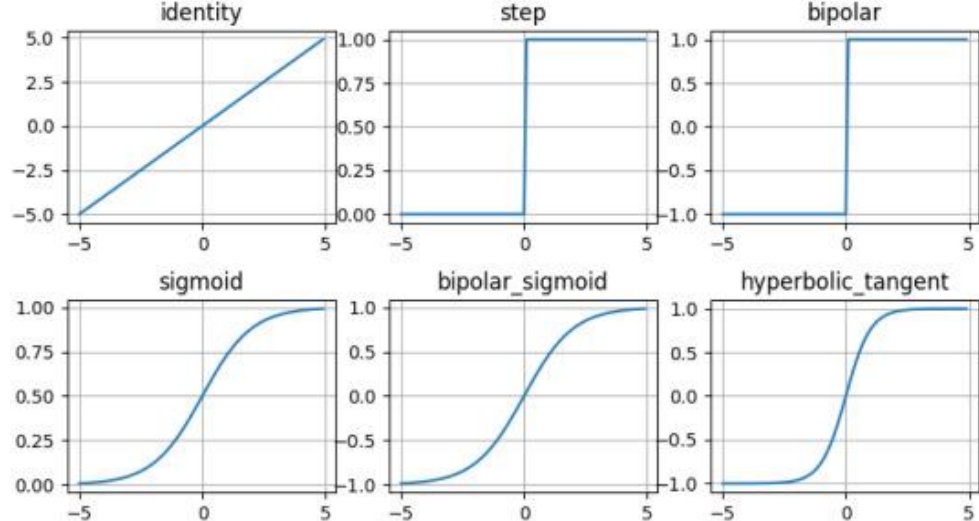
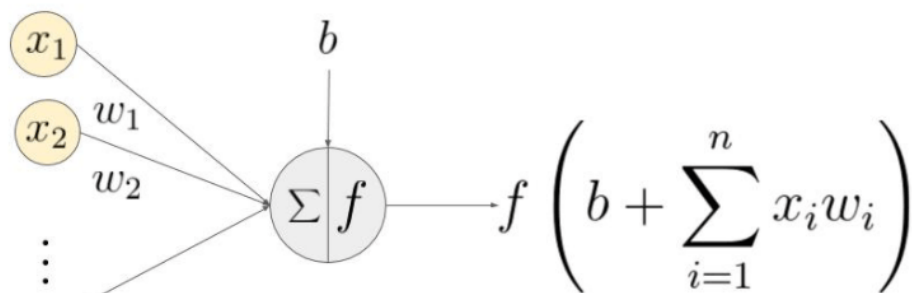
Tehát megmondja hogy mennyivel *kevesebb* bitre van szükség mint az egyenletes eloszláshoz.
Ennyit tanultunk, több előismeret kevesebb bizonytalanságot jelent!

Mesterséges neuron -> a biológiát másoljuk amennyire sikerül.

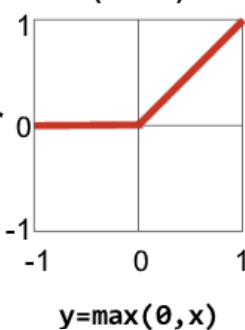
(a jó ötletet mindig érdemes felhasználni)



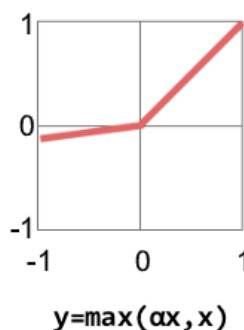
Aktivációs függvények



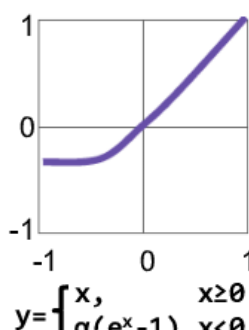
Rectified Linear Unit (ReLU)



Leaky ReLU

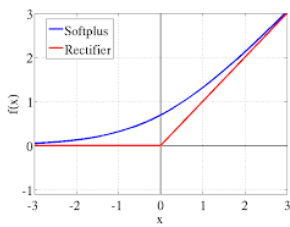


Exponential LU

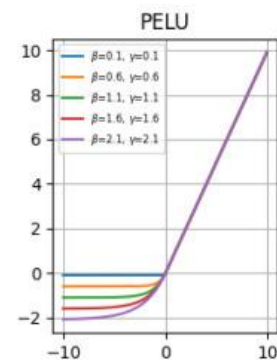
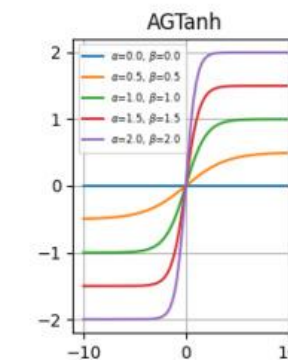
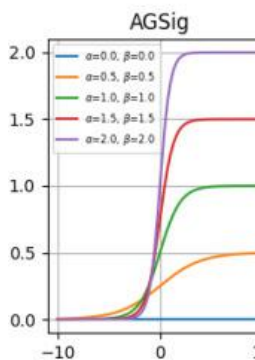
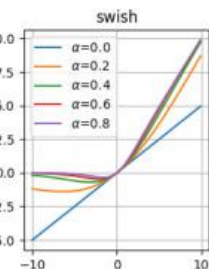
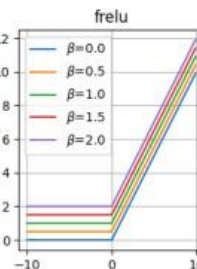
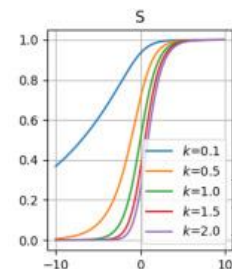


$\alpha = \text{small const. (e.g. 0.1)}$

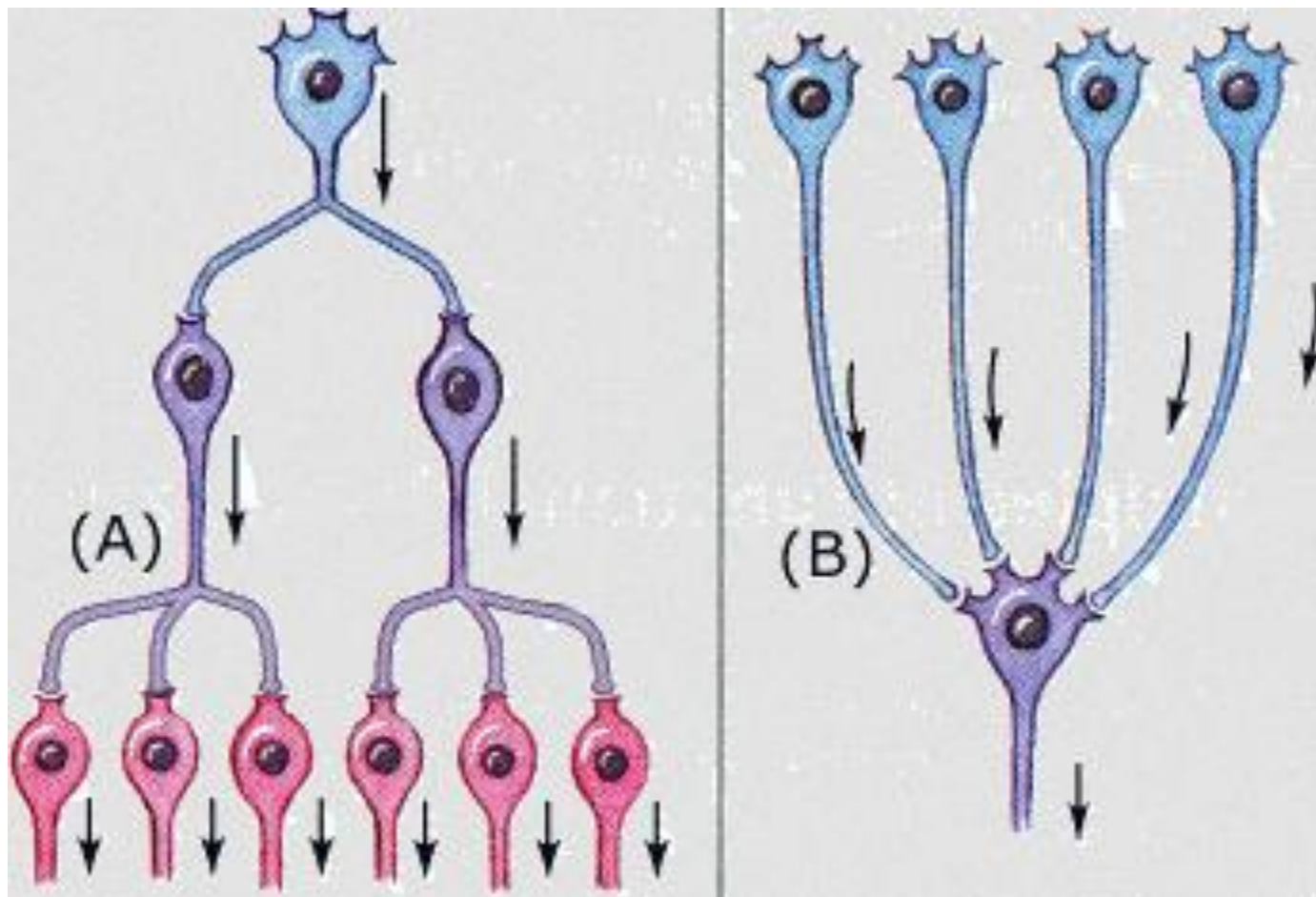
Modern Non-Linear Activation Functions



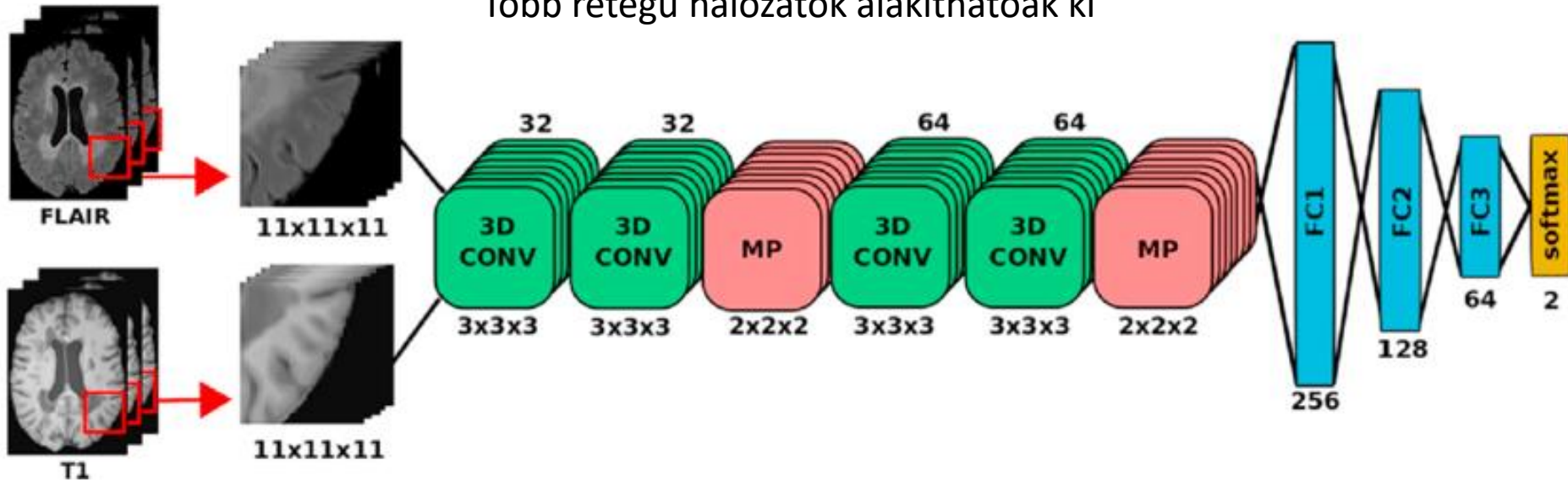
Softplus: $y = \ln(1 + e^x)$



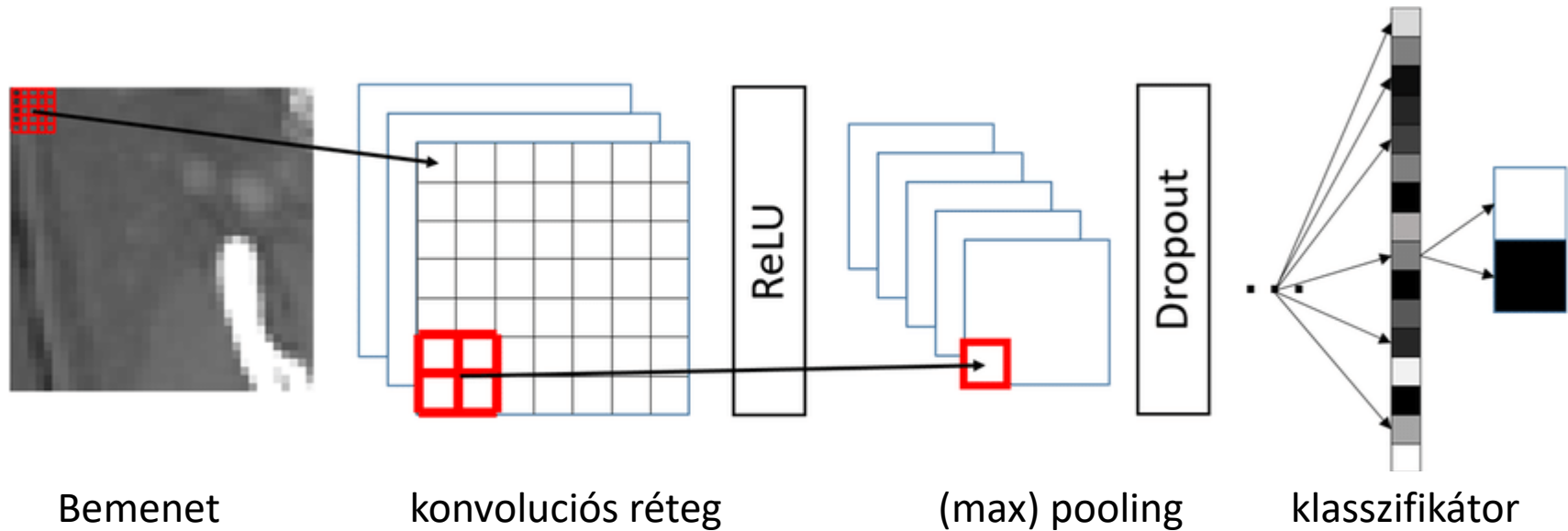
Konvergenca és divergencia



Több rétegű hálózatok alakíthatóak ki



Valverde, Sergi & Salem, Mostafa & Cabezas, Mariano & Pareto, Deborah & Vilanova, Joan C & Ramió-Torrentà, Lluís & Rovira, Alex & Salvi, Joaquim & Oliver, Arnau & Llado, Xavier. (2018). One-shot domain adaptation in multiple sclerosis lesion segmentation using convolutional neural networks. *NeuroImage: Clinical*. 21. 101638. 10.1016/j.nicl.2018.101638.



Ibragimov, Bulat & Xing, Lei. (2016). Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. *Medical Physics*. 44. 10.1002/mp.12045.