Rekenen met neuronen 2b: Introductie Decoding

Fleur Zeldenrust

Van Perceptie tot Bewustzijn, 2017

Overzicht

- Introductie neurale netwerken en neural coding
- Encoding modellen
    - college 1a: binair neuron & feed-forward perceptron
    - college 1b: rate neuron & recurrente ‘attractor' netwerken
    - college 2a: integrate-and fire neuron & recurrente ‘balanced' netwerken
- Decoding
    - college 2b: wat is informatie?

College 2a/b

Iets complexer / biologisch realistischer:

2a: Waar komt onregelmatige hersenactiviteit vandaan?

2b: Decoding: hoe interpreteer ik gemeten data?

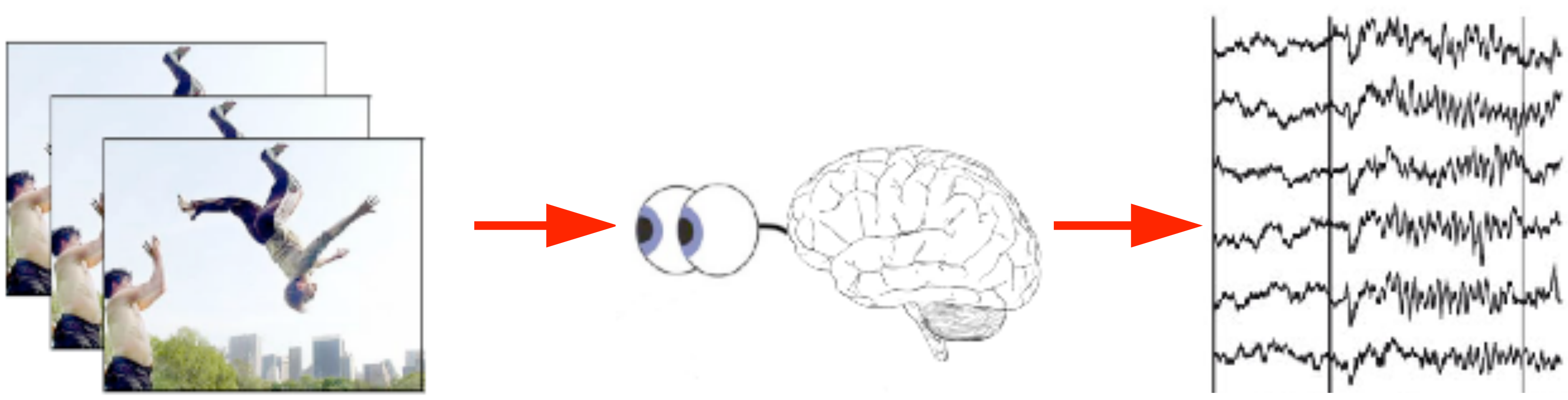
- Entropie: mogelijke informatie
- Informatie: wat is de relatie stimulus - hersenactiviteit?
- Bias bij het meten van informatie

Neural coding

Er is een relatie tussen de wereld, de activiteit van onze hersenen, ons gedrag en wat we waarneming noemen.

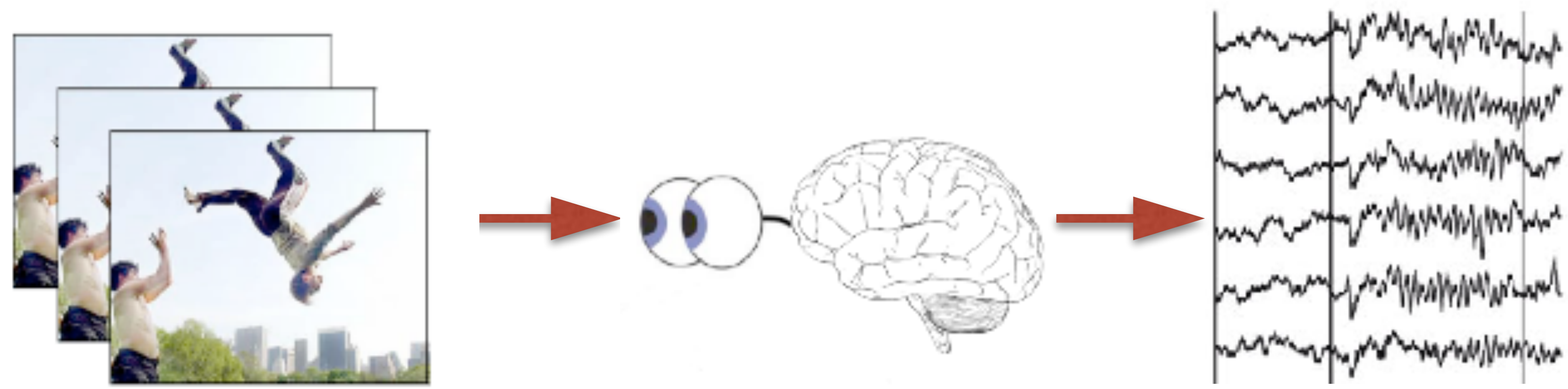
Wat is deze relatie? Neural coding

Representaties: wordt wat we waarnemen weergegeven in hersenactiviteit?



Encoding / Decoding

encoding model: hoe reageert een neuron / netwerk op stimuli?



decoding: wat vertelt neurale activiteit me over de stimulus?

Neural Decoding

Tot nu toe:

- **coding**: Hoe kan ik een probleem oplossen dmv een neuraal netwerk?
- **encoding**: Waar hangt de activiteit van een neuraal netwerk vanaf?

Nu: **decoding**

- Als ik hersenactiviteit meet, hoe weet ik wat die betekent?
- Hoeveel informatie geeft hersenactiviteit me over input?
- In andere wiskundige taal: kansrekening

Wat is informatie?

- Hoeveel informatie geeft hersenactiviteit me over input?
- Hoe meet ik informatie?
    - Wat is 1 bit (en 1 byte)?

- Experiment: hoeveel informatie geeft hersenactiviteit me over input?

- Hoeveel informatie zat er in de input?

- Hoe kwantificeer ik dat?

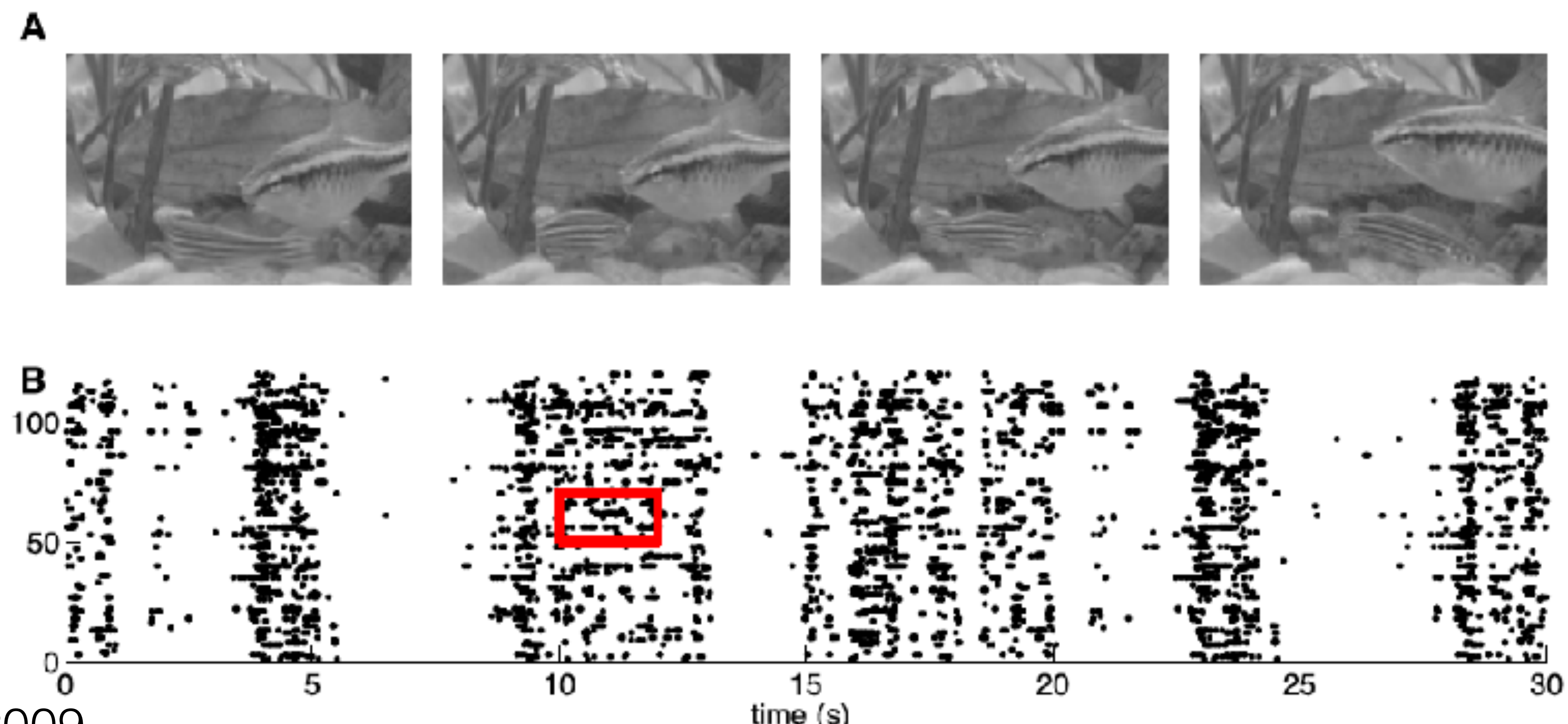


Image: Tkačik et al., 2009

- Experiment: hoeveel informatie geeft hersenactiviteit me over input?
- Hoeveel informatie zat er in de input?
- Hoe kwantificeer ik dat?

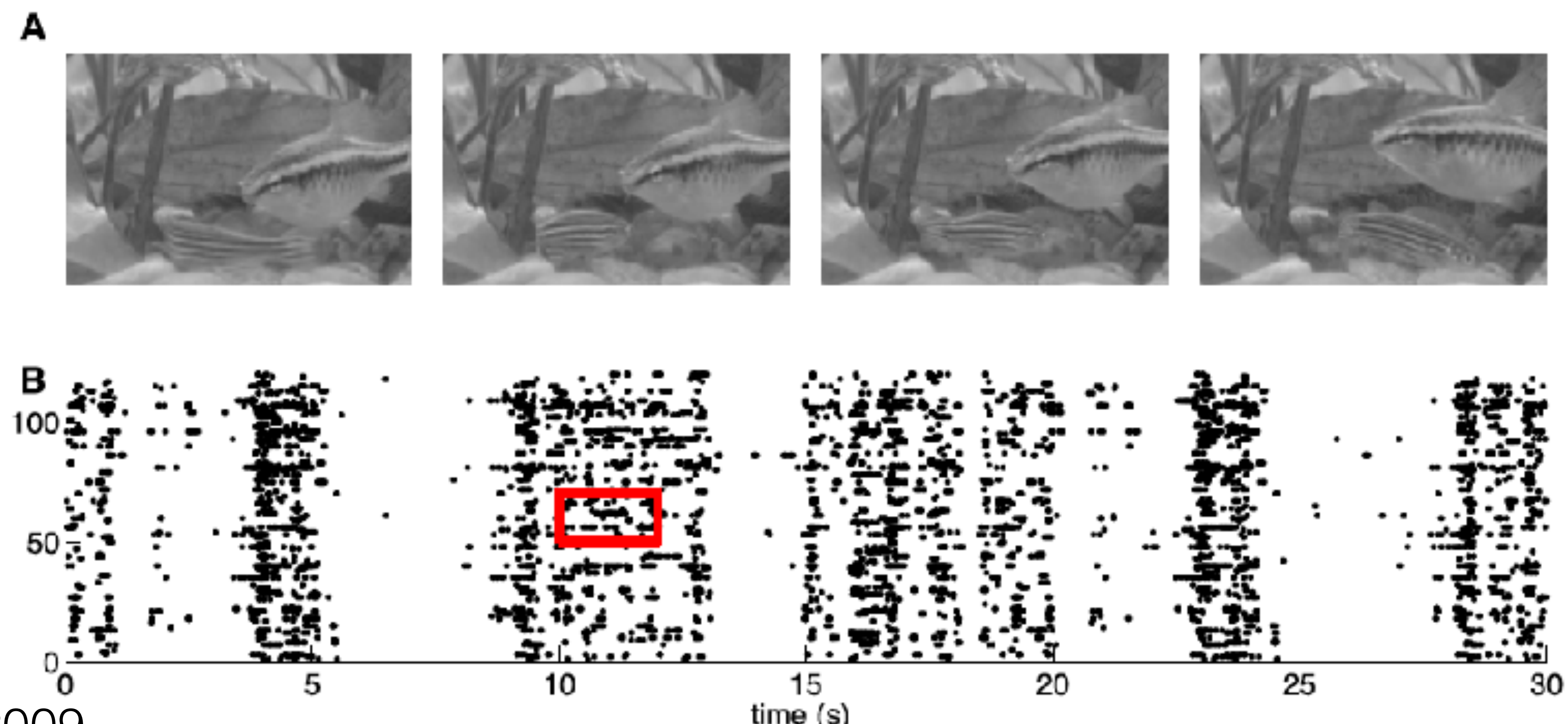


Image: Tkačik et al., 2009

College 2a/b

Iets complexer / biologisch realistischer:

2a: Waar komt onregelmatige hersenactiviteit vandaan?

2b: Decoding: hoe interpreteer ik gemeten data?

- Entropie: mogelijke informatie
- Informatie: wat is de relatie stimulus - hersenactiviteit?
- Bias bij het meten van informatie

Entropie: mogelijke informatie

Als ik met een dobbelsteen gooi, hoe verbazingwekkend is een 6?

Hangt af van

- aantal mogelijkheden
- kans op een 6

Entropie = ‘Surprise’, link aantal mogelijkheden, onzekerheid

Hoe kwantificeren?

Entropie: mogelijke informatie

kwantificeer ‘surprise’ S:

1. S is **positief**

2. S **neemt af** met kans op uitkomst: S(P(x)) is dalende functie (en S(1) = 0)

3. **optellen**: voor twee ***onafhankelijke*** gebeurtenissen is S de som van de twee surprises: S(P(x)*P(y)) = S(P(x))+S(P(y))

Alleen de log functie voldoet hieraan!

$$S(P[x]) = S(X) = -\sum_x P(x) \log_2 P(x)$$

Entropie: mogelijke informatie

voorbeeld: gooi een munt

Entropie S = -(P(kop) $\log_2$ P(kop) + P(munt) $\log_2$ P(munt))

Entropie S = -(P(kop) $\log_2$ P(kop) + (1-P(kop)) $\log_2$ (1-P(kop)))
= 1 bit

Wat als ik de kans P(kop) verander?

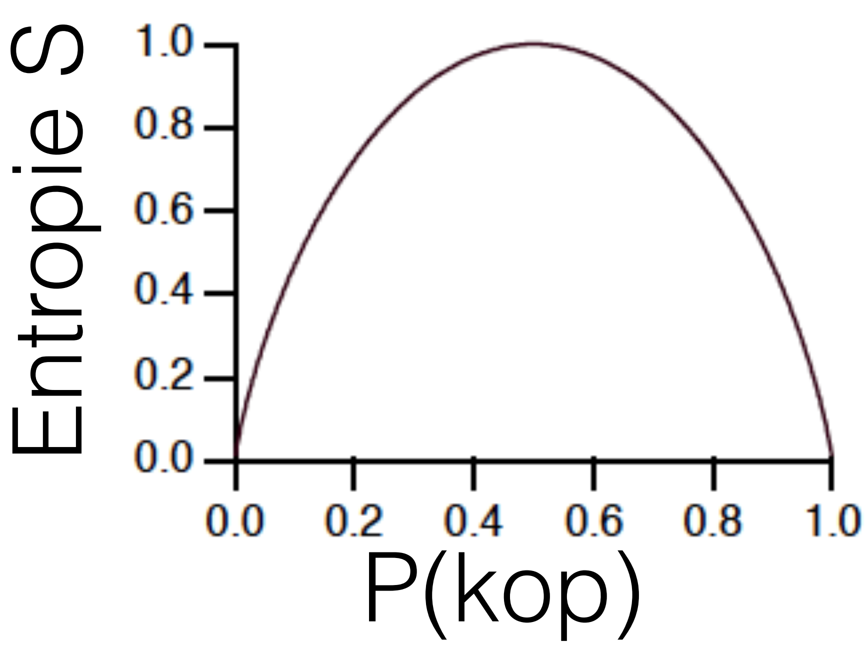


Image: Dayan&Abbott, 2001

Dus entropie het grootst als
kans kop = kans munt!

$$S(P[x]) = S(X) = -\sum_x P(x) \log_2 P(x)$$

Rekenvoorbeelden

1 dobbelsteen:

- P(elke uitkomst) = 1/6
- S = -6*(1/6 * $\log_2$ 1/6) = -$\log_2$ 1/6 = 2,6 bit

2 dobbelstenen, onafhankelijk

- P(elke uitkomst) = 1/36
- S = -36*(1/36 * $\log_2$ 1/36) = -$\log_2$ 1/36 = 5,2 bit

- Dus: twee dobbelstenen kunnen meer informatie opslaan dan 1!

$$S(P[x]) = S(X) = -\sum_x P(x)\log_2 P(x)$$

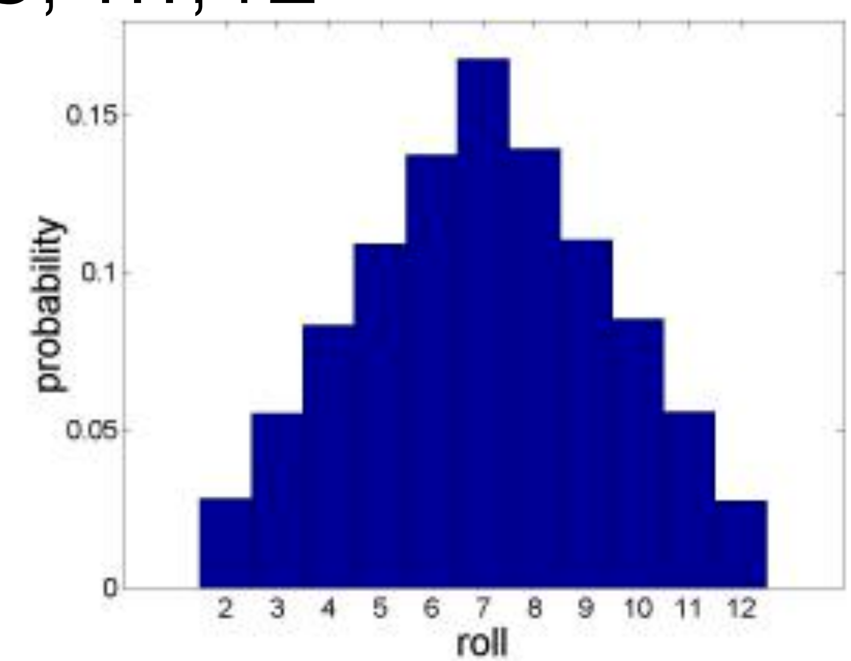
Rekenvoorbeelden

paardenrace: Storm heeft 1/4 kans om te winnen, Thunderbolt ook, Speedy 1/2. Wat is de entropie?

- $S = -1/4 * \log_2 1/4 - 1/4 * \log_2 1/4 - 1/2 * \log_2 1/2$
  $= -1/2 (\log_2 1/4 + \log_2 1/2) = 1{,}5$ bit

2 dobbelstenen, wat is de entropie van de som?

- 36 mogelijke uitkomsten, som = 2,3, …,12
- kans op 2 is 1/36, op 3 is 2/36, etc
- dus $S = -6/36 * \log_2 6/36$
  $-2(1/36 * \log_2 1/36 + \ldots)$
  $= 3{,}3$ bit



Entropie: intuïtie

- ‘Verbazing’, ‘surprise’, 'variabiliteit'
- gerelateerd aan het aantal mogelijkheden → onzekerheid
- Gemiddeld aantal bits nodig om een variabele te omschrijven
- Mogelijkheid tot informatieopslag, mogelijke informatie

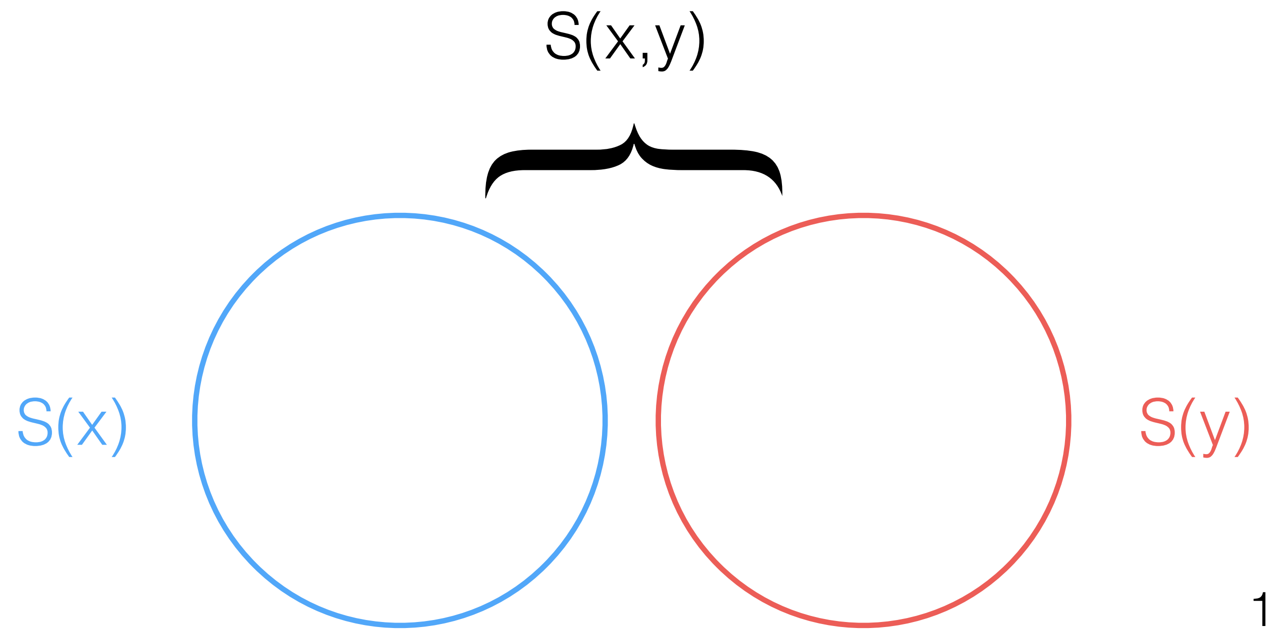
Welke heeft de grootste entropie?

Images: Natasha Devroye, 2009

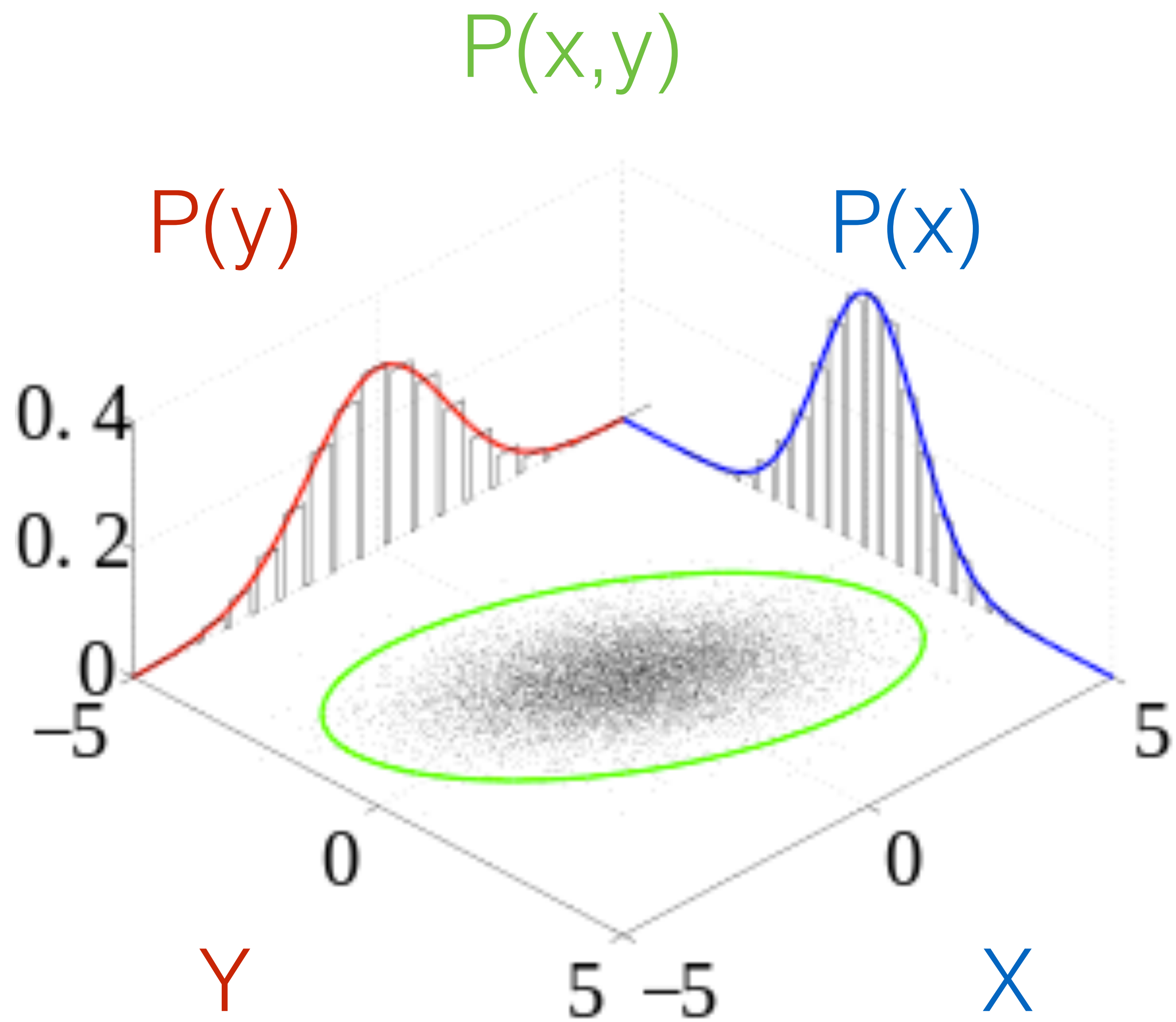
Joint Entropy

Joint Entropy: Entropie van twee variabelen: S(x,y)

- Onafhankelijke variabelen: som S(x,y) = S(x) + S(y)



Joint Probability Distribution



Joint Entropy

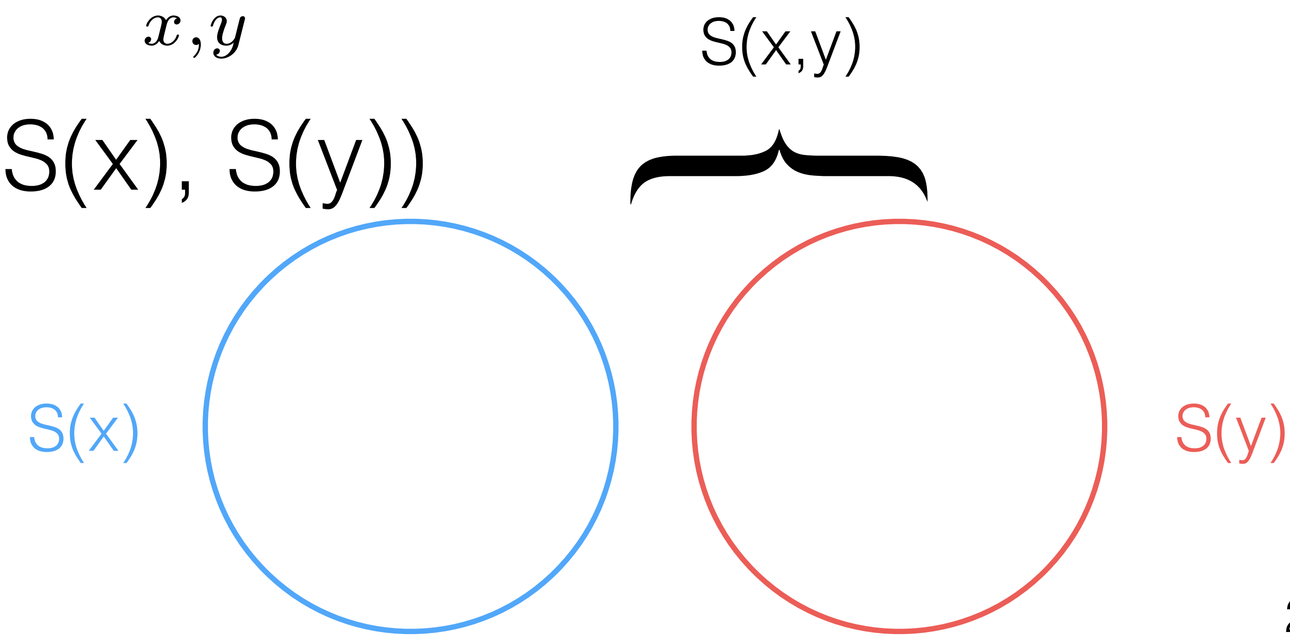
Joint Entropy: Entropie van twee variabelen: S(x,y)

- Onafhankelijke variabelen: som S(x,y) = S(x) + S(y)

- Niet onafhankelijke variabelen: let op ‘overlap’!

$$S(P[x], P[y]) = S(X, Y) = -\sum_{x,y} P(x, y) \log_2 P(x, y)$$

- S(x)+S(y) ≧ S(x,y) ≧ max(S(x), S(y))



Voorbeeld: het weer

In de zomer is de kans dat het warm weer is groter op een dag dat het niet regent dan als het wel regent (dus niet onafhankelijk).

Stel: P(zon, warm) = 1/2; P(zon, koel) = 1/12, P(regen, warm) =1/6, P(regen, koel) = 1/4

P(zon) = 1/2+1/12 = 7/12; P(regen) = 1/6+1/4 = 5/12
S(weer) = $-7/12 * \log_2(7/12) - 5/12 * \log_2(5/12) = 0{,}98$ bit

P(warm) = 1/2+1/6 = 2/3; P(koel) = 1/12+1/4 = 1/3
S(temperatuur) = $-2/3 * \log_2(2/3) - 1/3 * \log_2(1/3) = 0{,}92$ bit

De joint entropy is gelijk aan
S(weer, temperatuur) = $-1/2 * \log_2(1/2) - 1/12 * \log_2(1/12) - 1/6 * \log_2(1/6) - 1/4 * \log_2(1/4) = 1{,}7$ bit

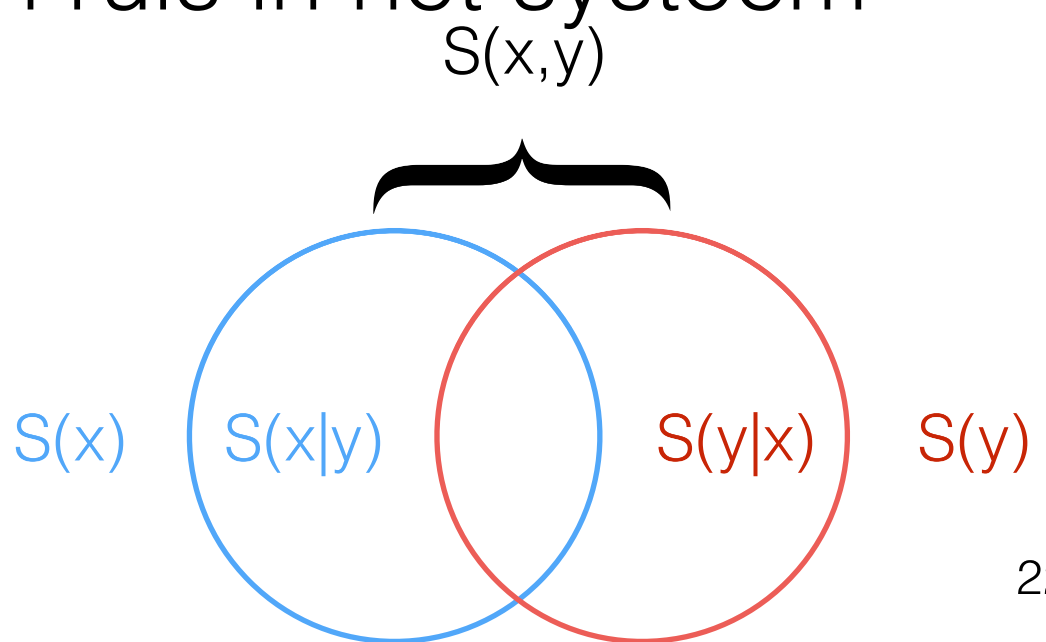
Dus S(weer)+S(temperatuur) ≧ S(weer, temperatuur) ≧ max(S(weer), S(temperatuur))

Conditional Entropy

Conditionele entropie: wat als ik een van de twee variabelen al weet?
(entropie van respons als ik stimulus weet)

$$S(X|Y) = -\sum_{x,y} P(y)P(x|y)\log_2 P(x|y) = S(X,Y) - S(Y)$$

- Ook wel ‘noise entropy’ (als x=respons neuron, y=input): entropie afkomstig van ruis in het systeem
- altijd kleiner of gelijk aan S(X)!

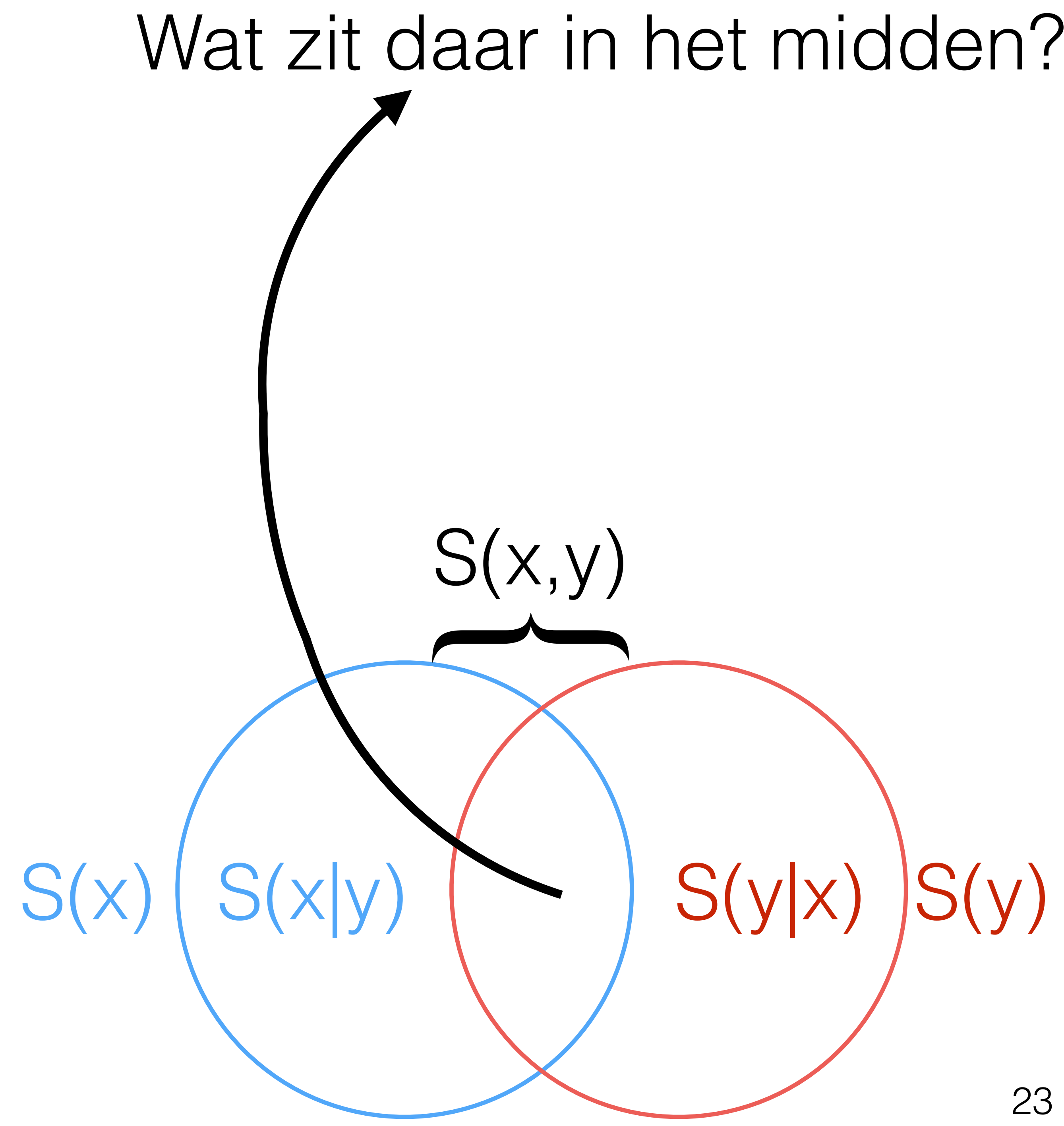


Conclusie Entropie

- Entropie = 'Verbazing', 'surprise', 'variabiliteit', gerelateerd aan het aantal mogelijkheden → onzekerheid, mogelijkheid tot informatie

- Joint Entropy = Entropie van twee variabelen

- Conditionele entropie: wat als ik een van de twee variabelen al weet? = entropie van respons als ik stimulus weet



College 2a/b

Iets complexer / biologisch realistischer:

2a: Waar komt onregelmatige hersenactiviteit vandaan?

2b: Decoding: hoe interpreteer ik gemeten data?

- Entropie: mogelijke informatie
- Informatie: wat is de relatie stimulus - hersenactiviteit?
- Bias bij het meten van informatie

- Experiment: hoeveel informatie geeft hersenactiviteit me over input?

- Hoeveel informatie zat er in de input?

- Hoe kwantificeer ik dat?

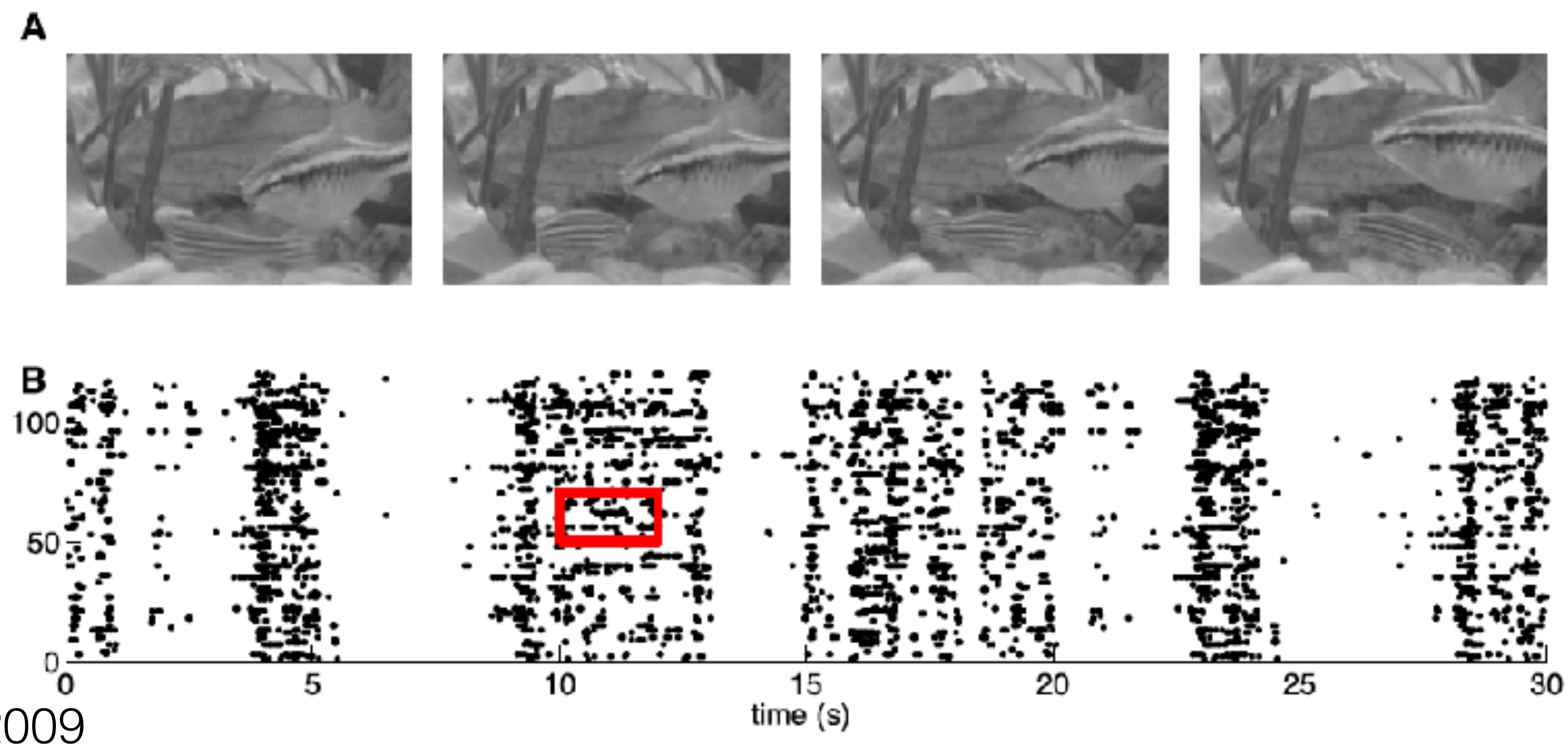
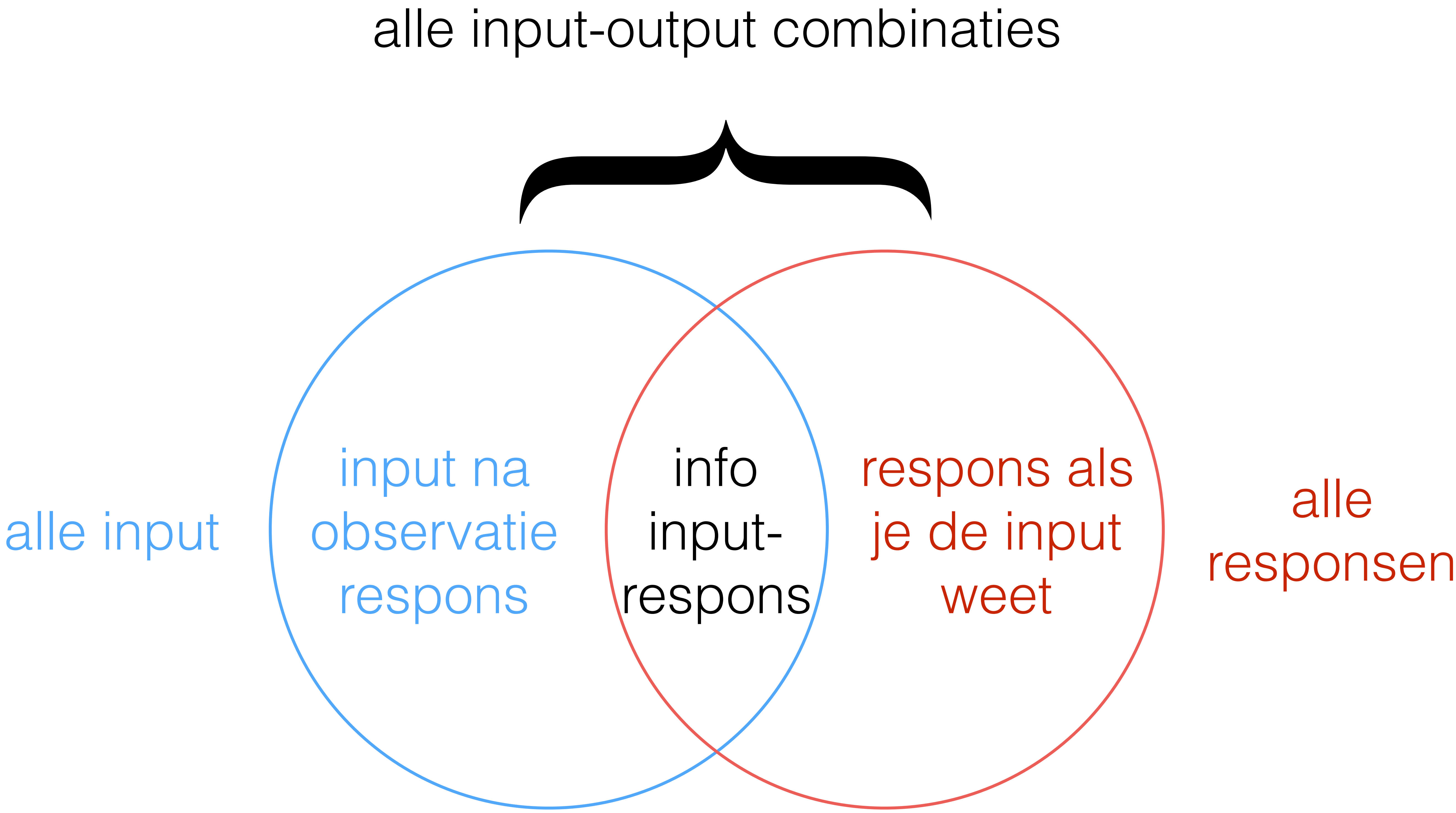
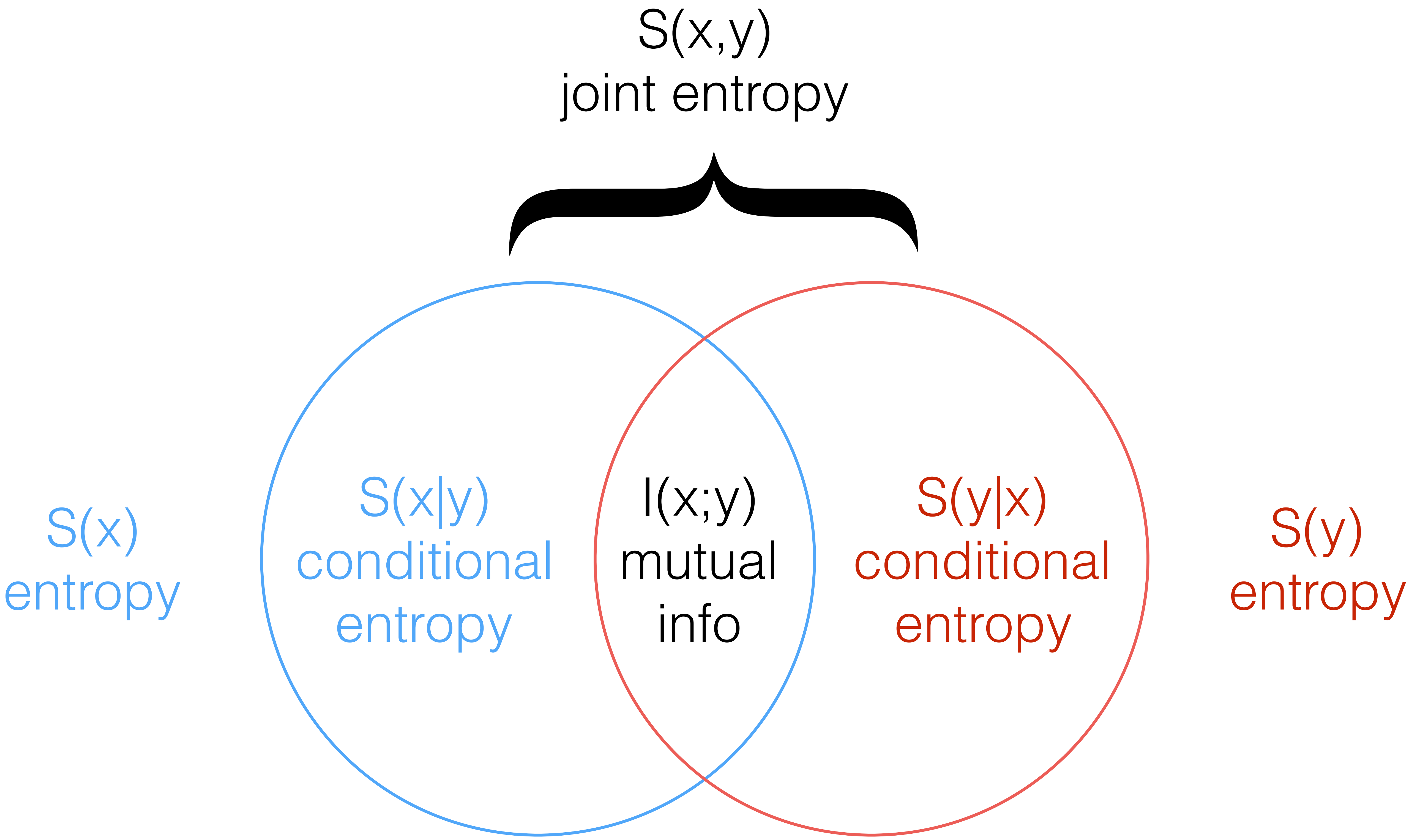


Image: Tkačik et al., 2009

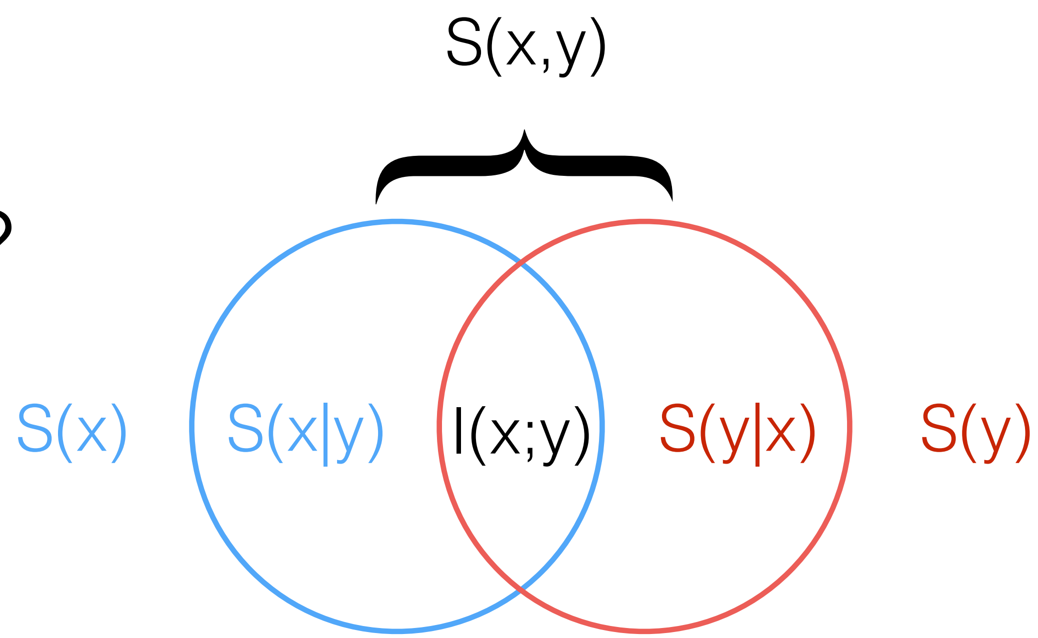

alle input-output combinaties
alle input
input na
observatie
respons
info
input-
respons
respons als
je de input
weet
alle
responsen


S(x,y)
joint entropy
S(x)
entropy
S(x|y)
conditional
entropy
I(x;y)
mutual
info
S(y|x)
conditional
entropy
S(y)
entropy

Mutual Information (I)
Claude Shannon (1948)

- Afname van onzekerheid: hoeveel meer weet ik van y door x te observeren?

- Hoeveel meer weet ik van de input door een spike train te observeren?
  - Dus de **afname in entropie**



$$I(X;Y) = \sum_{x,y} P(x,y) \log_2 \frac{P(x,y)}{P(x)P(y)} = S(X) - S(X|Y)$$

Voorbeeld Mutual Information

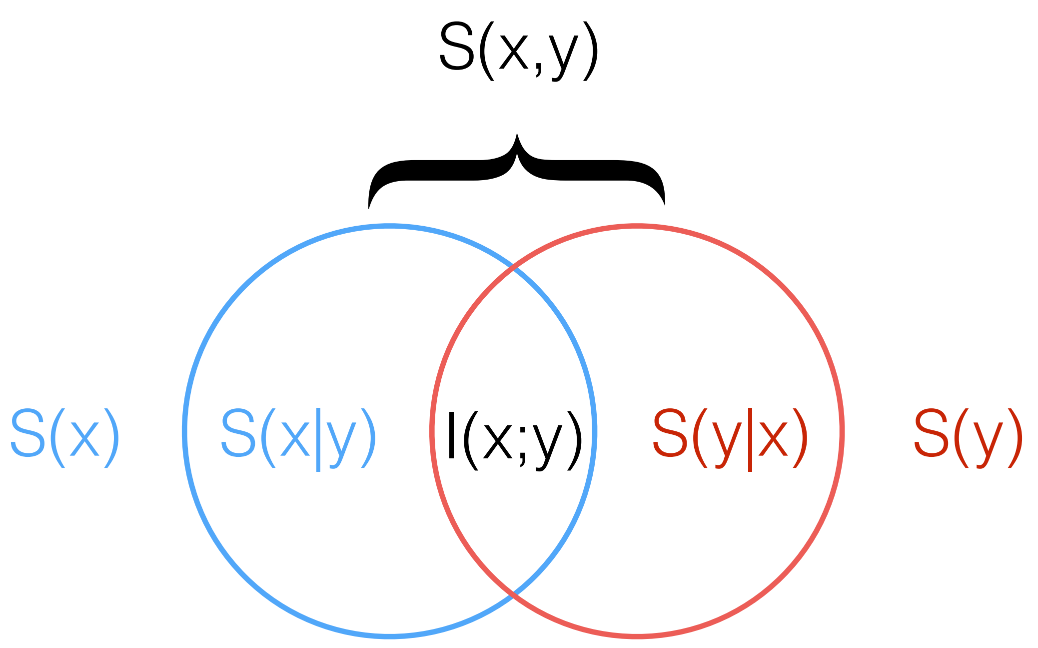
Voorbeeld 1: stel een neuron wordt niet beïnvloedt door de stimulus.

- Onafhankelijk, dus P(r,s) = P(r)P(s)

$$I(R;S) = \sum_{r,s} P(r,s) \log_2 \frac{P(r,s)}{P(r)P(s)}$$

$$= \sum_{r,s} P(r)P(s) \log_2 \frac{P(r)P(s)}{P(r)P(s)}$$

$$= \sum_{r,s} P(r)P(s) \log_2 1 = 0$$



Voorbeeld Mutual Information

Voorbeeld 2: Experiment visuele cortex V1

- 2 stimuli: rode stip en blauwe stip;
- we meten de vuurfrequentie van een V1 neuron in respons.
- Ik doe 20 experimenten.

Het neuron vuurt meestal met een hogere frequentie als we de rode stip laten zien.

1. Wat is de entropie van de input?
2. Wat is de kans op een hoge vuurfrequentie van het neuron?
3. Wat is de mutual information tussen de respons van het neuron en de input?

stimulus → respons ↓	rode stip	blauwe stip
hoog $r>15$ Hz	7/20	5/20
laag $r<15$ Hz	3/20	5/20

Antwoord

1. S(Input)?
   P(rood) = 7/20 + 3/20 = 1/2
   P(blauw) = 1/2
   S(input) = 1 bit

2. P(hoog)=
   7/70+5/20 = 12/20 = 3/5;
   P(laag) = 2/5

stimulus → respons ↓	rode stip	blauwe stip
hoog r>15 Hz	7/20	5/20
laag r<15 Hz	3/20	5/20

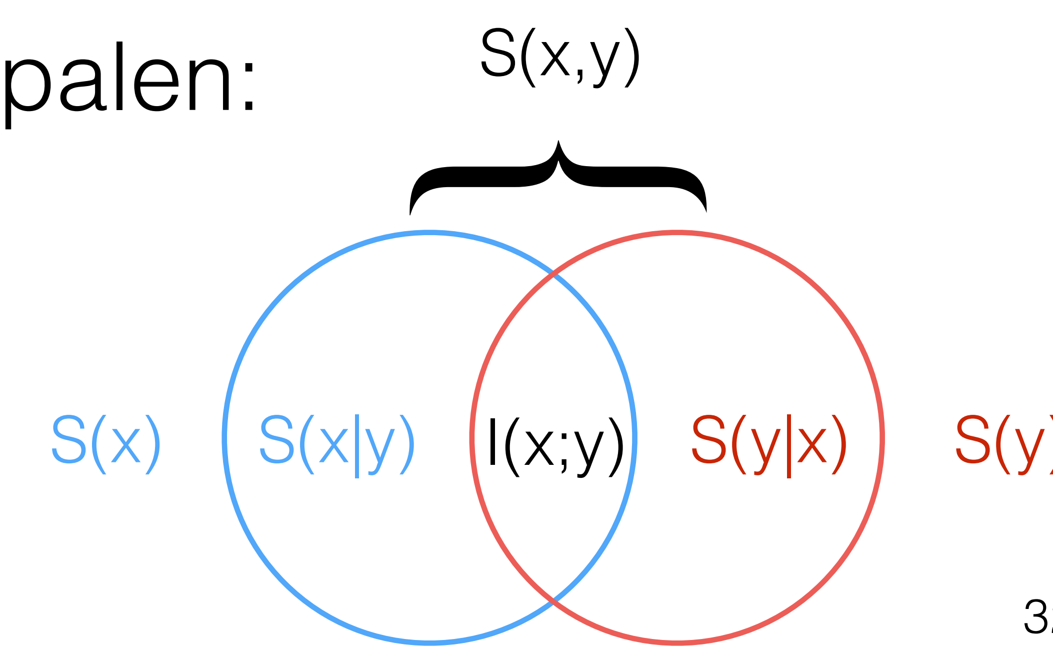
$$
\begin{aligned}
I(\text{stimulus; frequentie}) =& P(\text{rood, hoog}) \log_2 \frac{P(\text{rood, hoog})}{P(\text{rood})P(\text{hoog})} + \text{...rood, laag...} + \text{...} \\
=& 7/20 \log_2 \frac{7/20}{1/2 * 3/5} + 3/20 \log_2 \frac{3/20}{1/2 * 2/5} \\
& + 5/20 \log_2 \frac{5/20}{1/2 * 3/5} + 5/20 \log_2 \frac{5/20}{1/2 * 2/5} \\
=& 0.03 \text{ bit}
\end{aligned}
$$

Eigenschappen info

- Symmetrisch: I(x;y) = I(y;x)

- Je kunt geen informatie maken uit het niets:
    - MI positief (of 0)
    - S(X) ≧S(X|Y)
    - informatie is nooit groter dan entropie in de input: 'entropie neemt altijd af' (of blijft het gelijk)

- Probleem: P(x,y) vaak moeilijk te bepalen: dan zou je alle mogelijke stimulus-respons paren uitgebreid moeten onderzoeken!



College 2a/b

Iets complexer / biologisch realistischer:

2a: Waar komt onregelmatige hersenactiviteit vandaan?

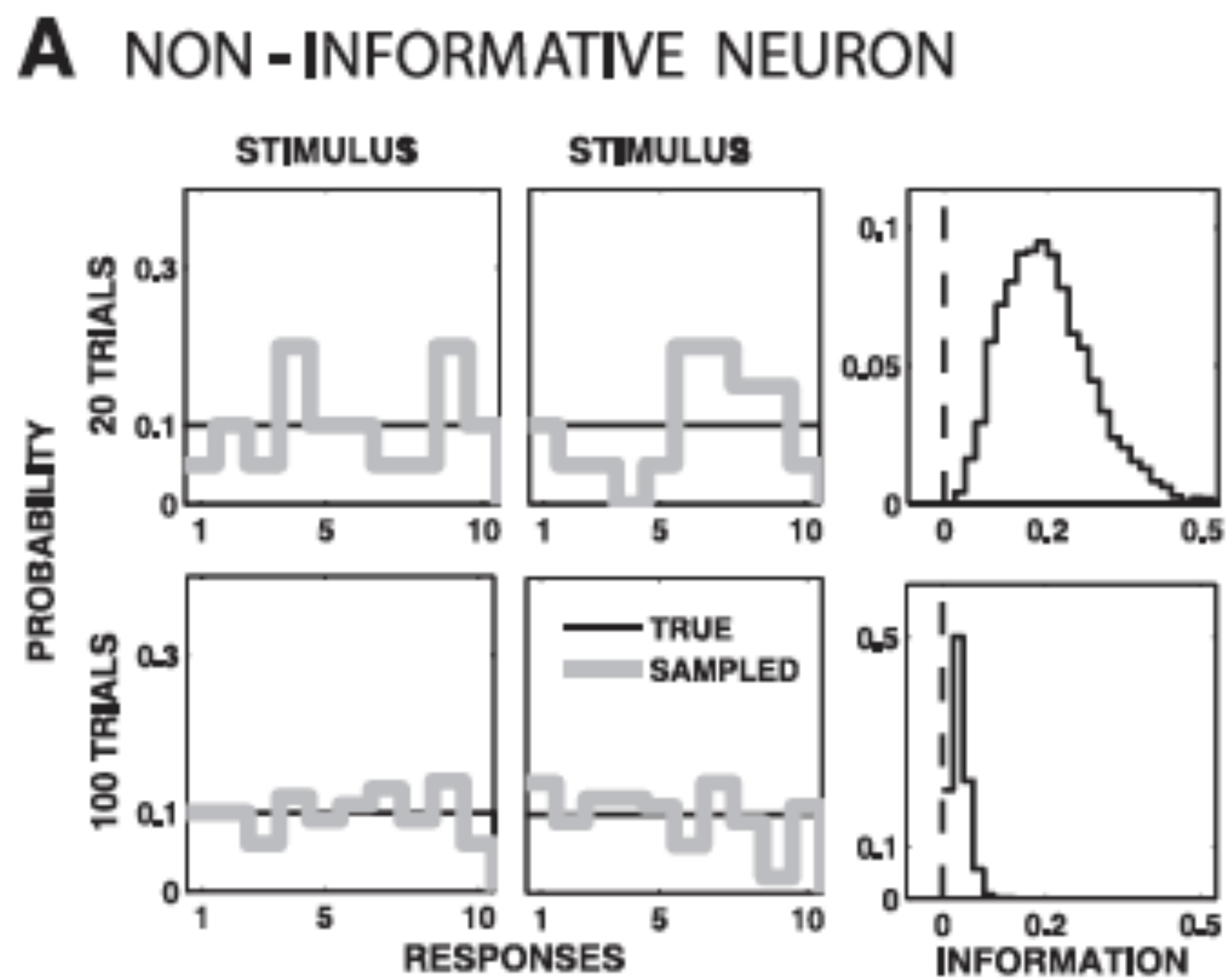
2b: Decoding: hoe interpreteer ik gemeten data?

- Entropie: mogelijke informatie
- Informatie: wat is de relatie stimulus - hersenactiviteit?
- Bias bij het meten van informatie

Bias door undersampling

In een experiment met een **beperkt aantal metingen**, kunnen de echte, onderliggende kansen afwijken van de 'gemeten kansen'

Hierdoor overschat je de hoeveelheid informatie: **bias**

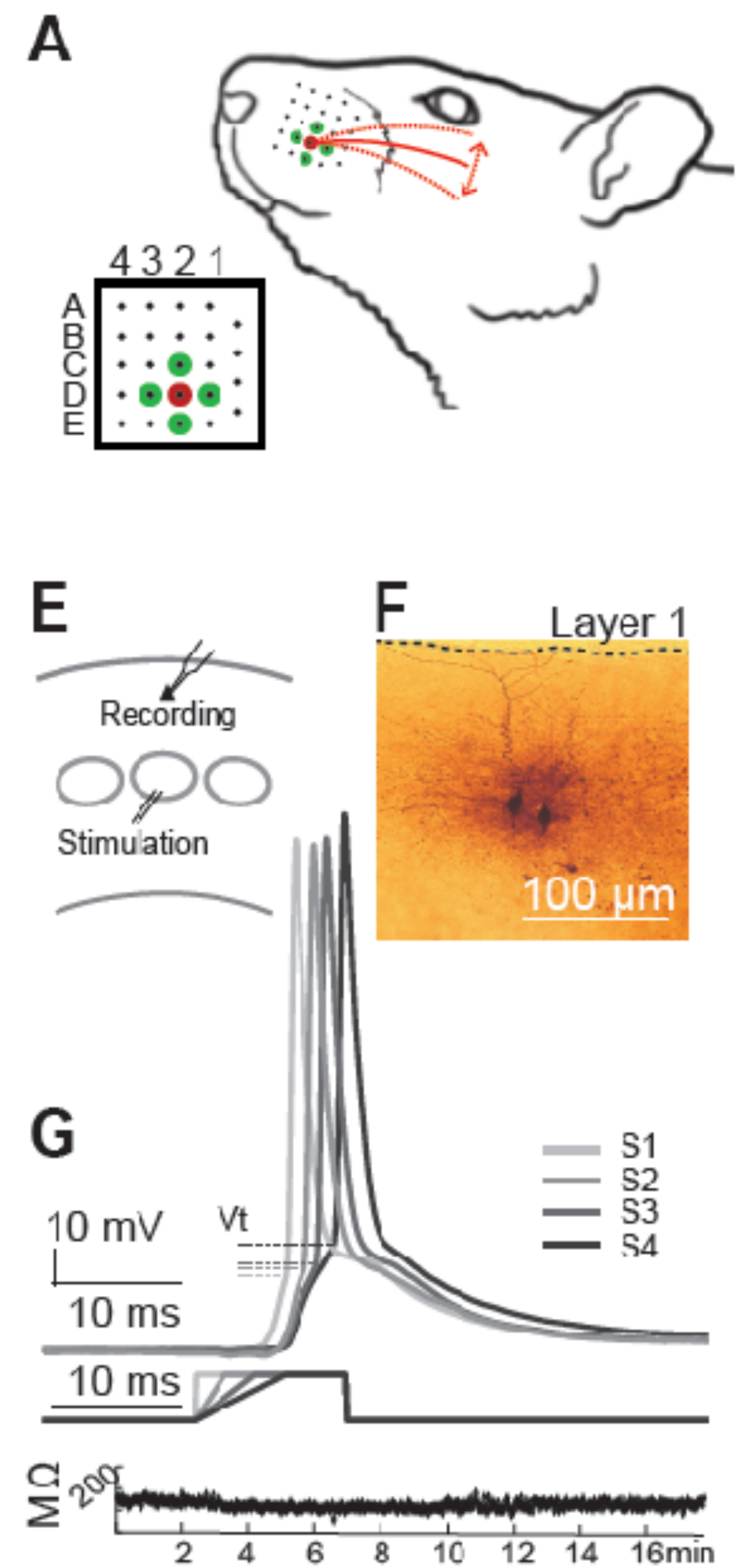


Panzeri et al. 2007

Informatie in neuroscience

Voorbeeld: Barrel cortex (snorharen)
Chao et al. (under review)

Beweeg snorhaar rat, meet de respons neuronen L2/3 barrel cortex.



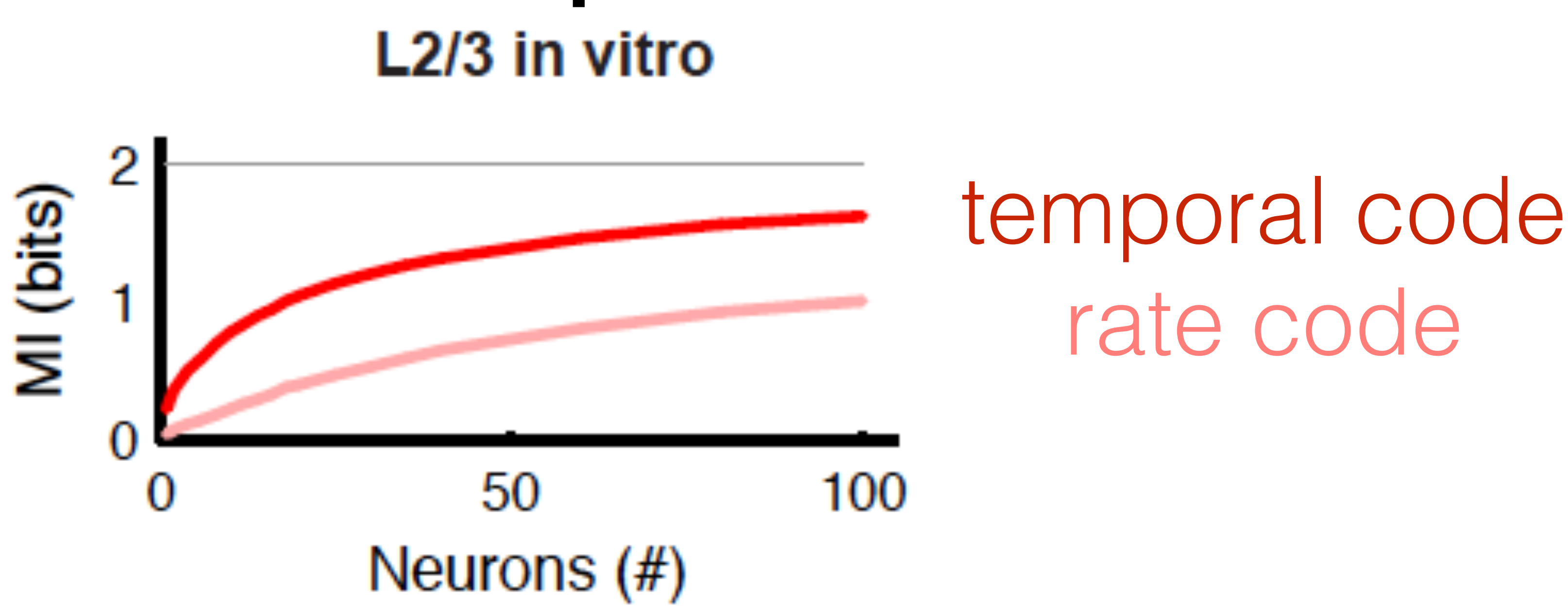
Informatie in neuroscience

Voorbeeld: Barrel cortex (snorharen) Chao et al. (in progress)

Beweeg snorhaar rat, meet de respons neuronen L2/3 barrel cortex.

Kan ik afleiden hoeveel snorhaar bewogen werd uit activiteit neuronen in barrel cortex?

Gaat dat beter met een **temporal code** of met een **rate code**?



Signaal en ruis

- Uitgangspunt: tussen de wereld (input) en de representatie (neurale activiteit) zit ruis

- **Decoding**: probeer de wereld te reconstrueren aan de hand van neurale data

- ruis = ‘echte' input - gereconstrueerde input

- Signal-to-Noise Ratio (**SNR**) = var(signal) / var(noise)

- Informatie is gerelateerd aan SNR: hoe groter de ruis, hoe kleiner SNR en hoe minder informatie
  (MI= 0.5 log2 (1+SNR) voor een Gaussian channel)

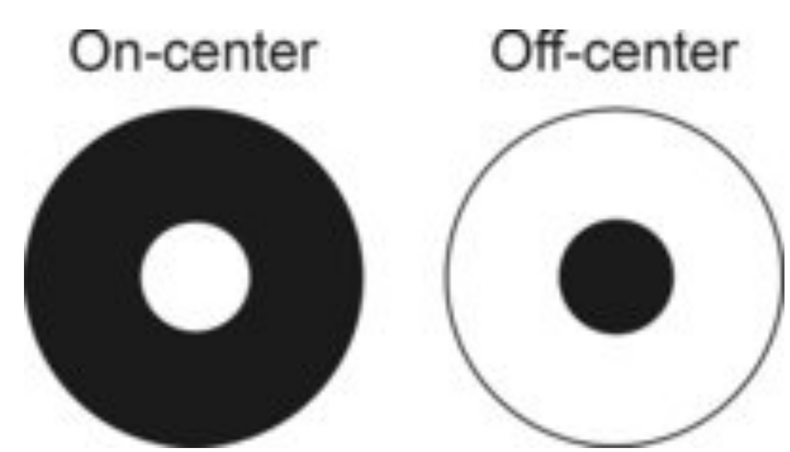
Receptive field

- Tot nu toe: receptive field = ‘waar een neuron op reageert’

- Kun je het ook omdraaien: dat wat een neuron representeert?

- Kun je de input reconstrueren dmv receptive fields?

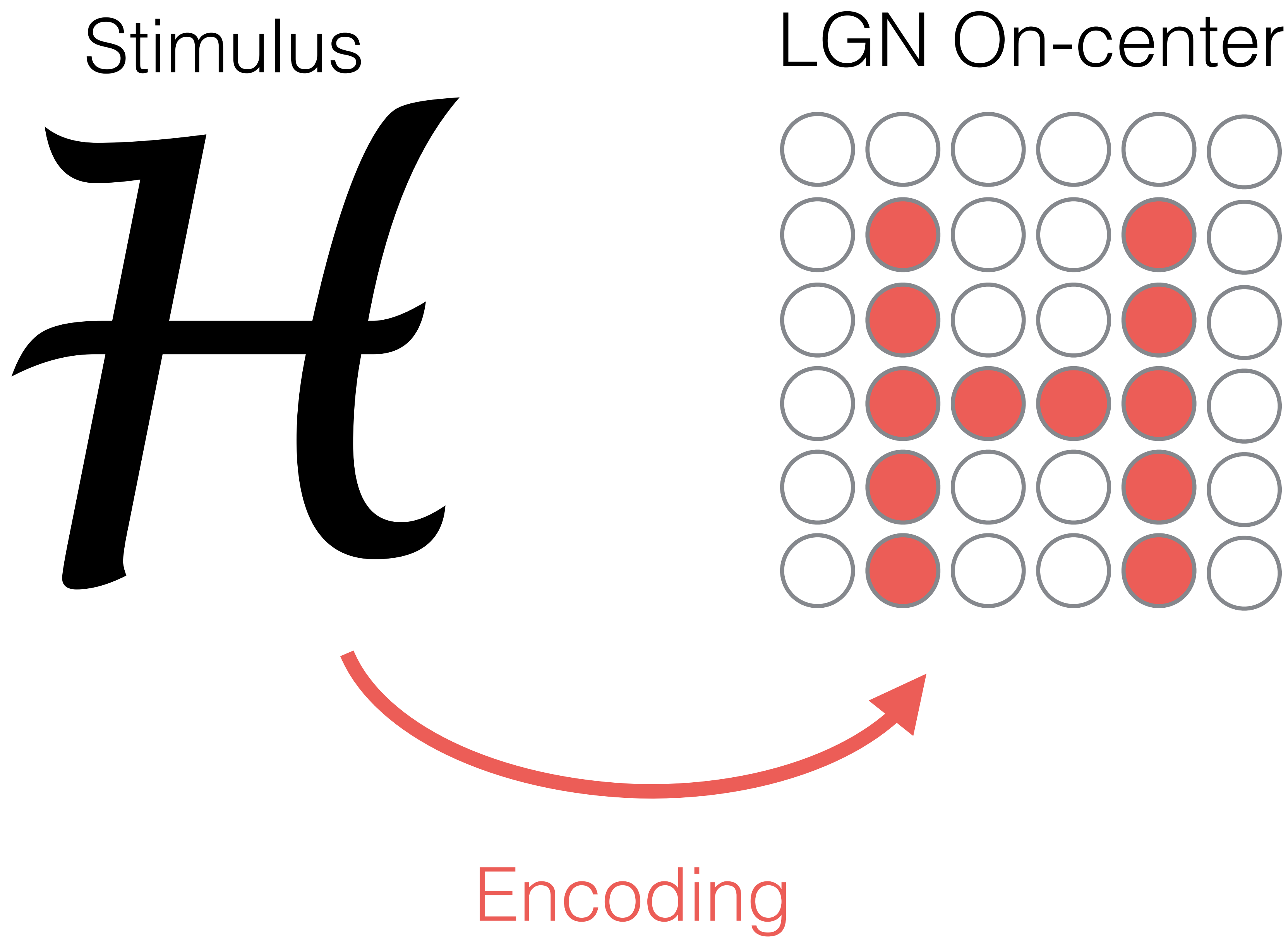
Dit noemen we een **decoder**



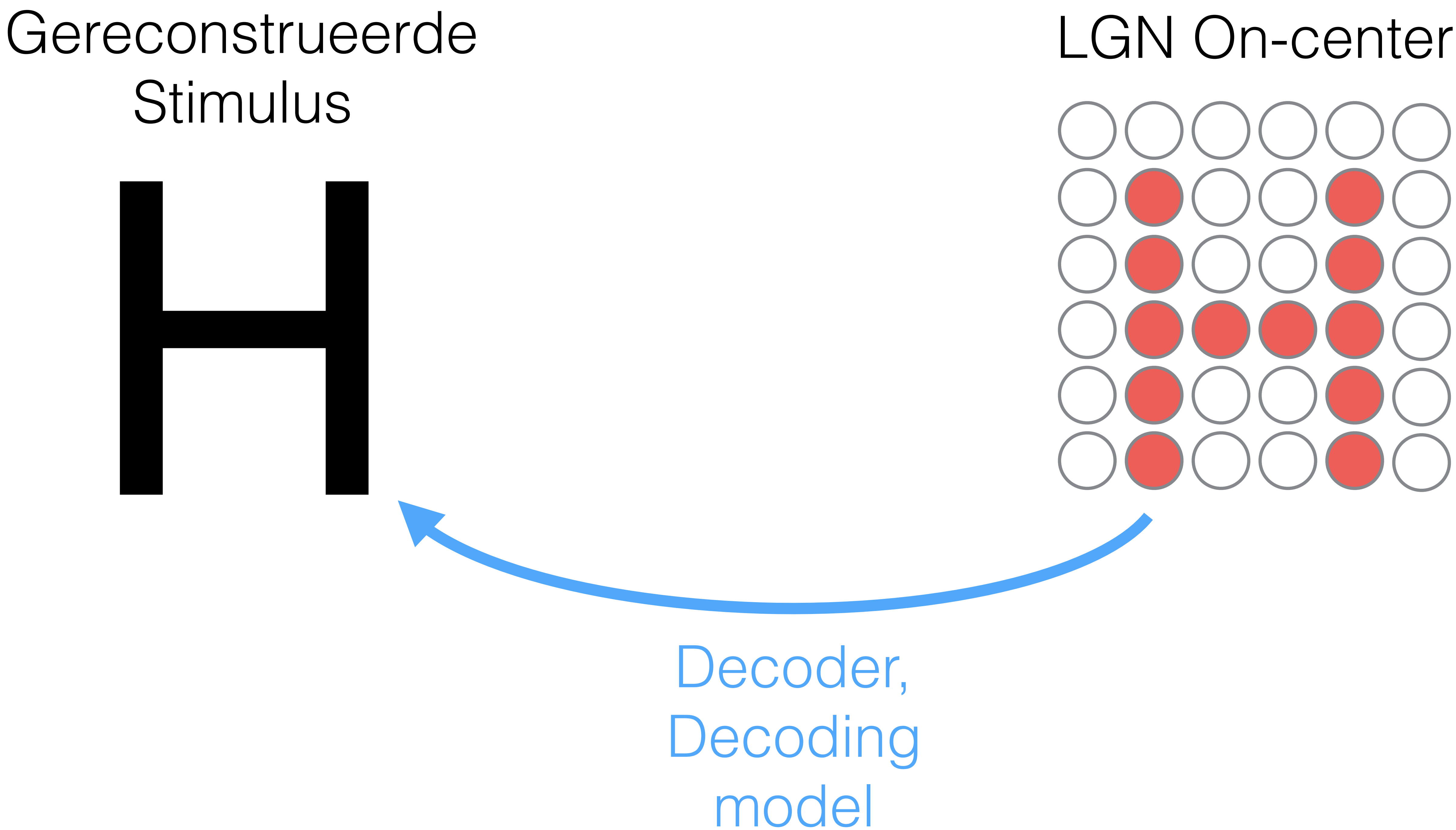
Decoder = model lezen hersenactiviteit

- Rekenvoorbeeld: ‘hoge vuurfrequentie = rode stip, lage vuurfrequentie = blauwe stip’
- Barrel cortex voorbeeld: ‘vuren L2/3 cel = beweging snorhaar’
- LGN On-center: ‘pixel'

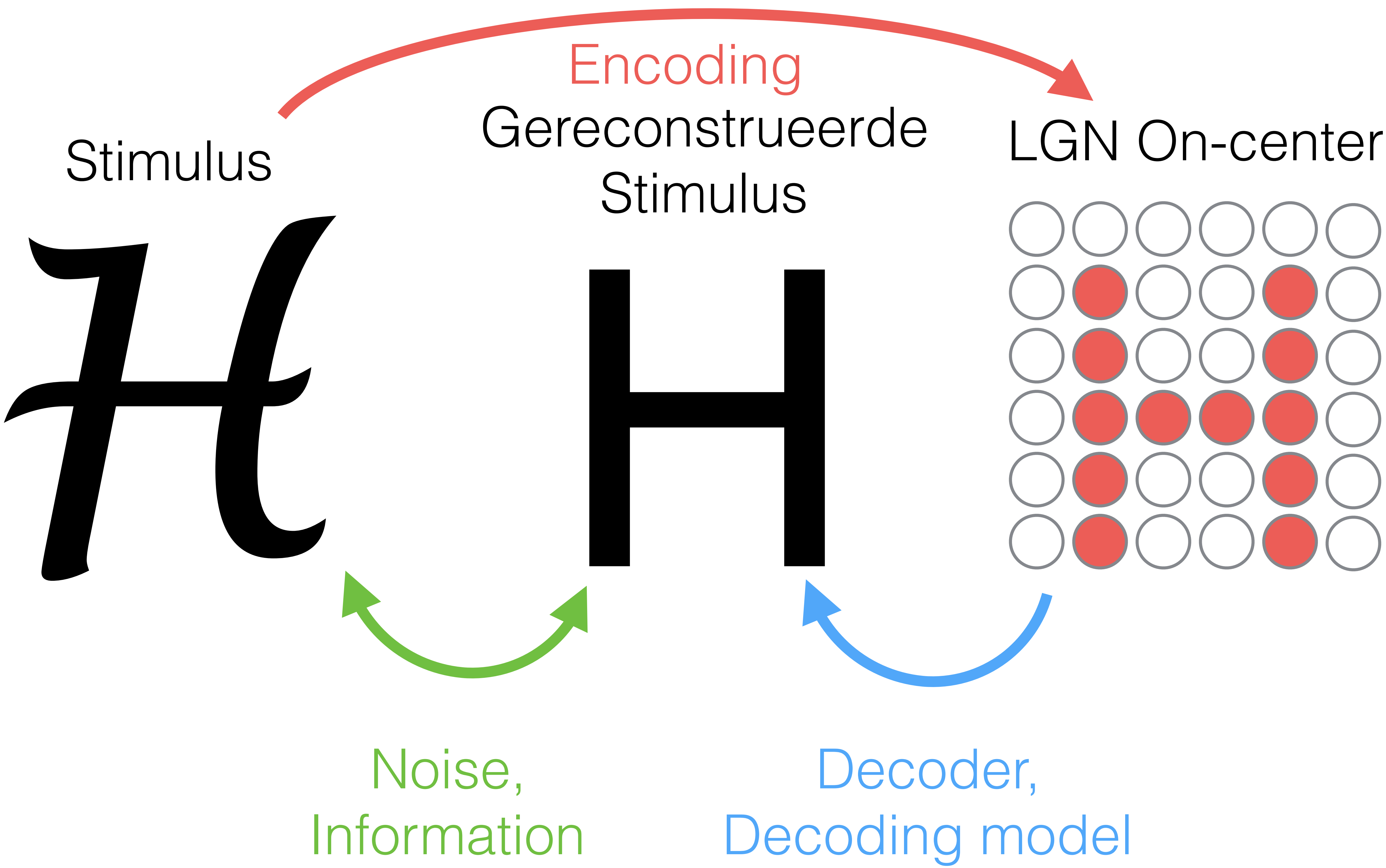
Encoding = model maken hersenactiviteit



Decoding = model wat hersenactiviteit betekent



Mutual information & ruis: hoeveel lijkt gereconstrueerde stimulus op stimulus?



Samenvatting decoding

Hoeveel informatie geeft hersenactiviteit me over input?

- Hoeveel informatie zat er in de input (**entropie**)?
- Hoeveel weet ik van de input met behulp van hersenactiviteit (**mutual information**)?
- Probleem: door undersampling wordt MI vaak overschat (bias)

Voor reconstrueren van de input en berekenen informatie moeten we een aanname maken over wat hersenactiviteit ‘betekent': een **decoder** → decoding modellen

Voor wie meer wil

- Dayan & Abbott: ‘Theoretical Neuroscience’
- Izhikevich: 'Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting’
- Rieke et al.: 'Spikes: exploring the neural code’

Dit was het!

- Introductie neurale netwerken en neural coding

- Encoding modellen
    - college 1a: binair neuron & feed-forward perceptron
    - college 1b: rate neuron & recurrente ‘attractor' netwerken
    - college 2a: integrate-and fire neuron & recurrente ‘balanced' netwerken

- Decoding
    - college 2b: wat is informatie?

NB Verder met neurale netwerken
- Leren in modelnetwerken: Leren en Geheugen
- Data-analyse en modellen programmeren: Cognitive Computational Neuroscience en Signal Analysis