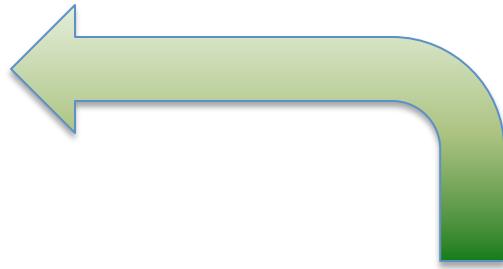
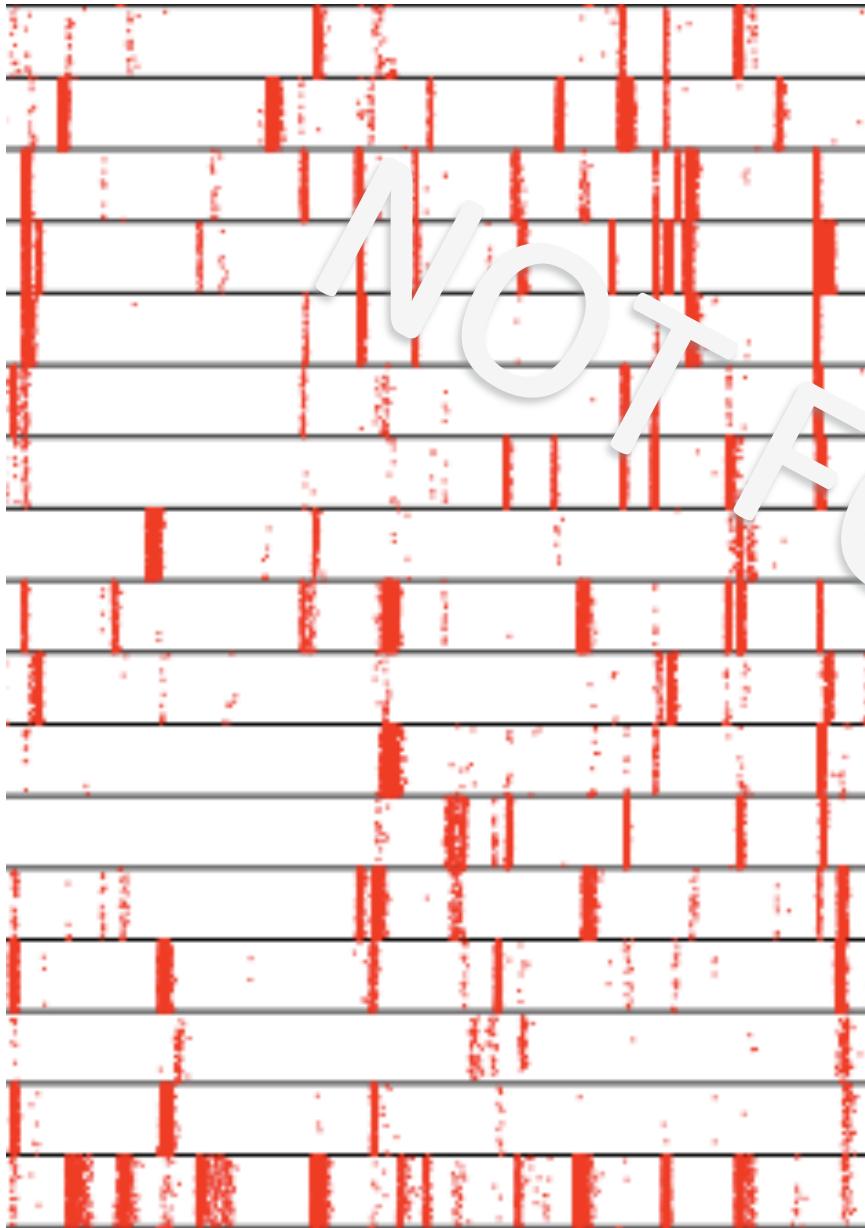


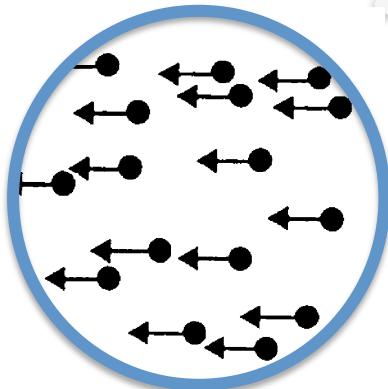
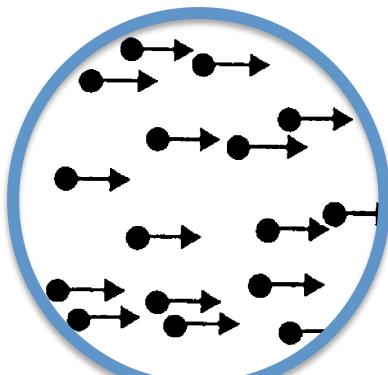
Information and coding principles



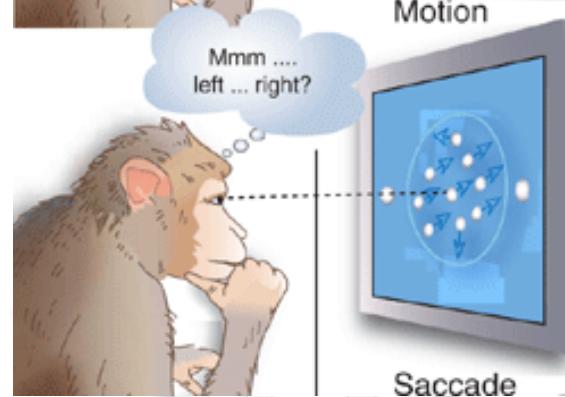
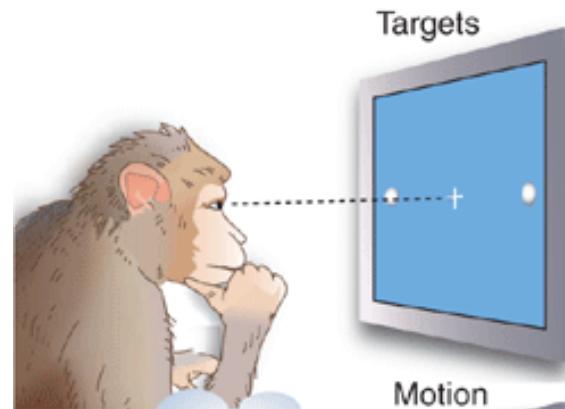
Information and coding principles

- Defining entropy and information
- Computing information for neural spike trains
- What can information tell us about coding?

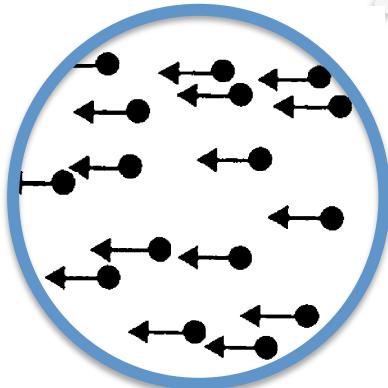
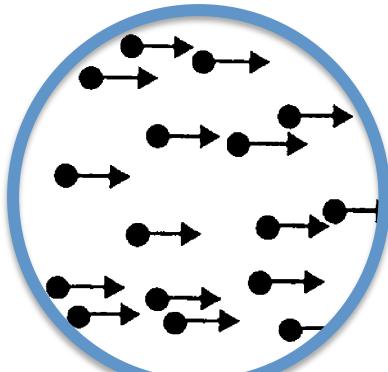
How good is my code?



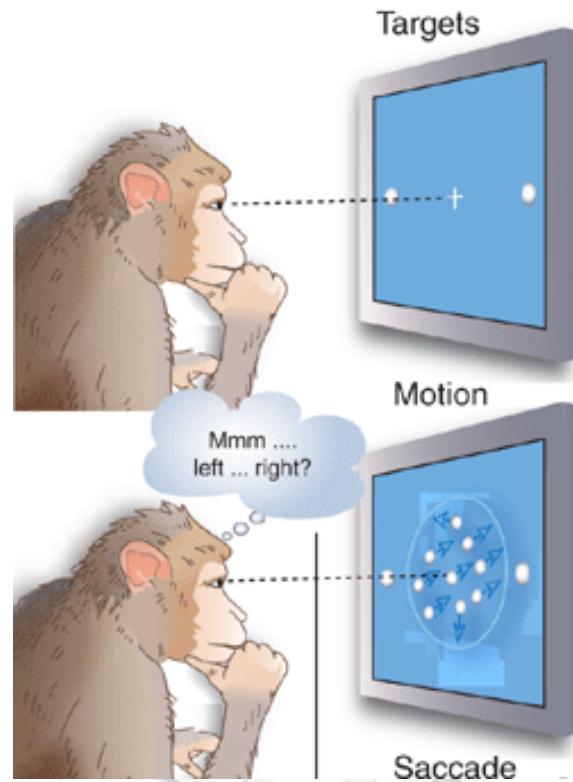
T FOR USE



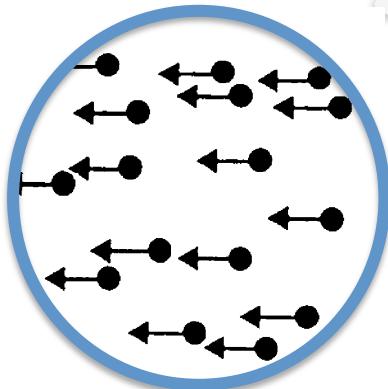
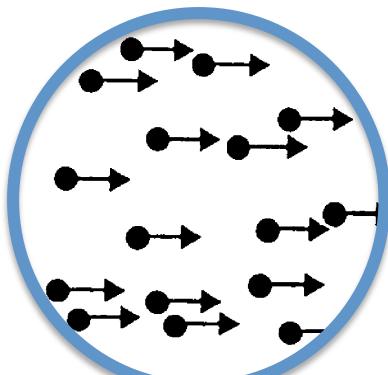
How good is my code?



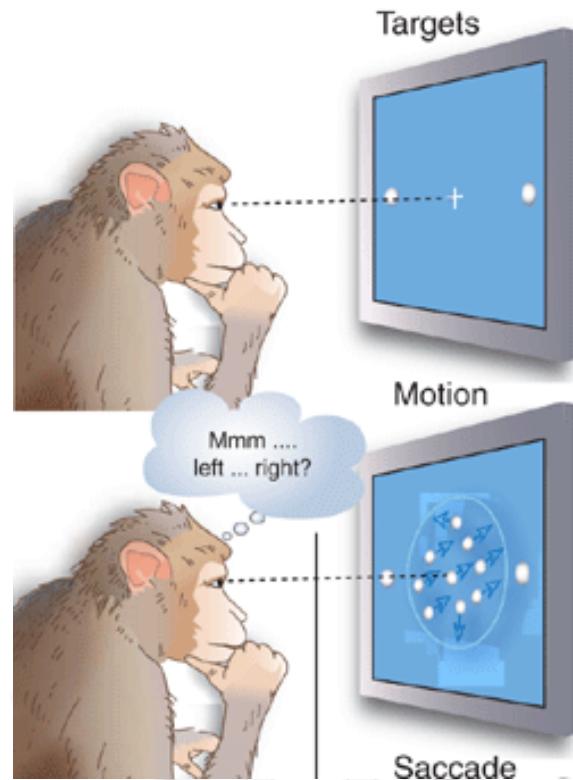
T FOR USE



How good is my code?

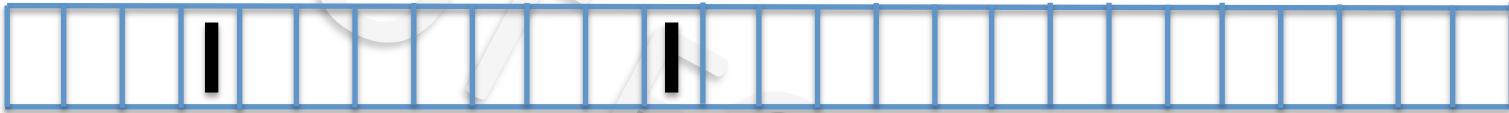


T FOR USE



Surprise!

Information



$$P(1) = p$$

$$P(0) = 1 - p$$

$$\text{information}(1) = - \log_2 p$$

$$\text{information}(0) = - \log_2 (1-p)$$

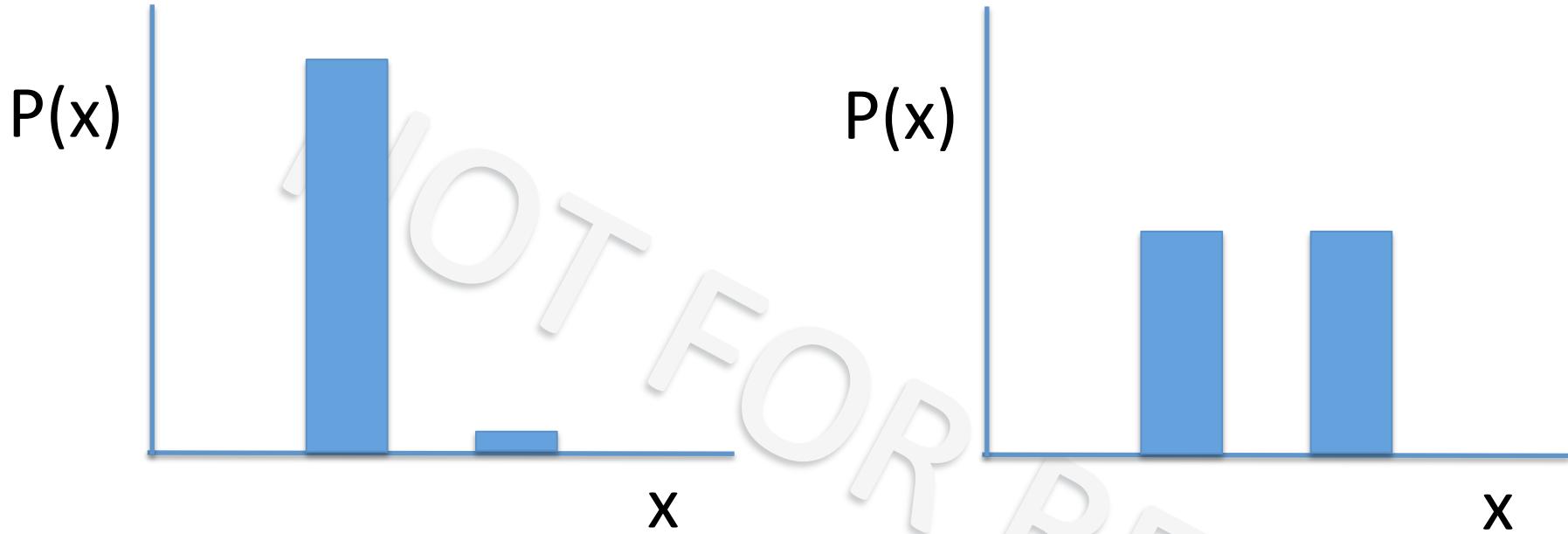
Information



Each *bit* of information specifies location by an additional factor of 2.



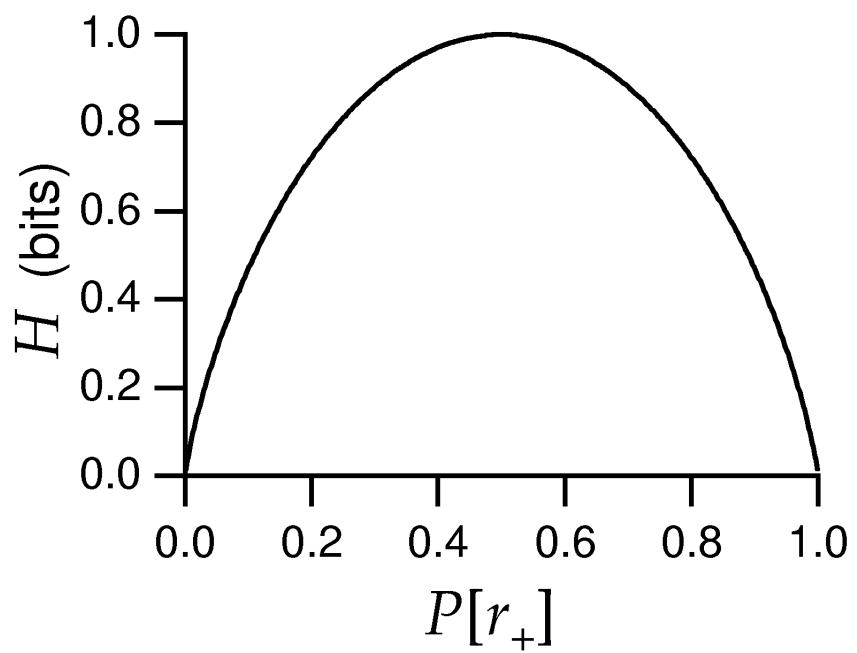
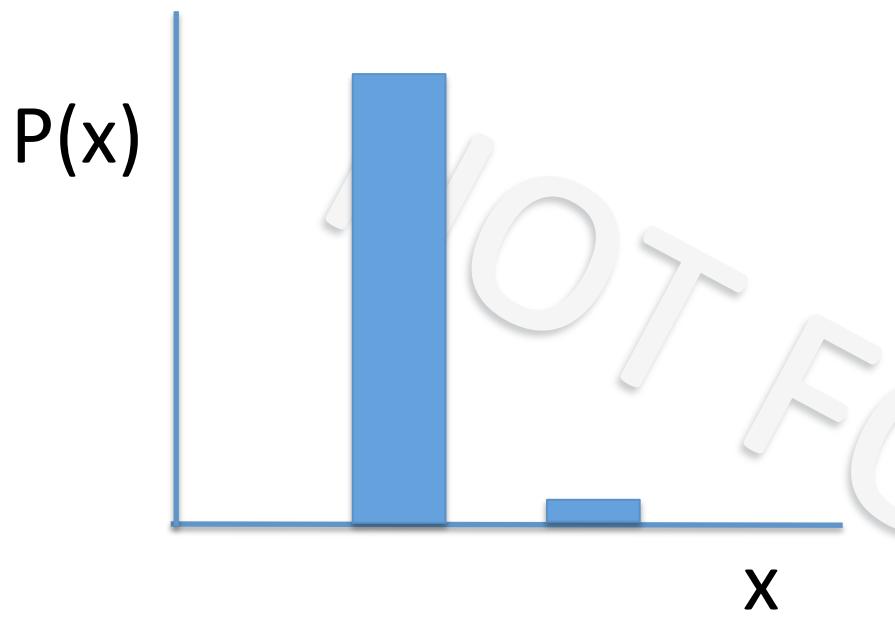
What is entropy?



Entropy = average information
 = $-\sum p_i \log_2 p_i$
 = $-\int dx p(x) \log_2 p(x)$

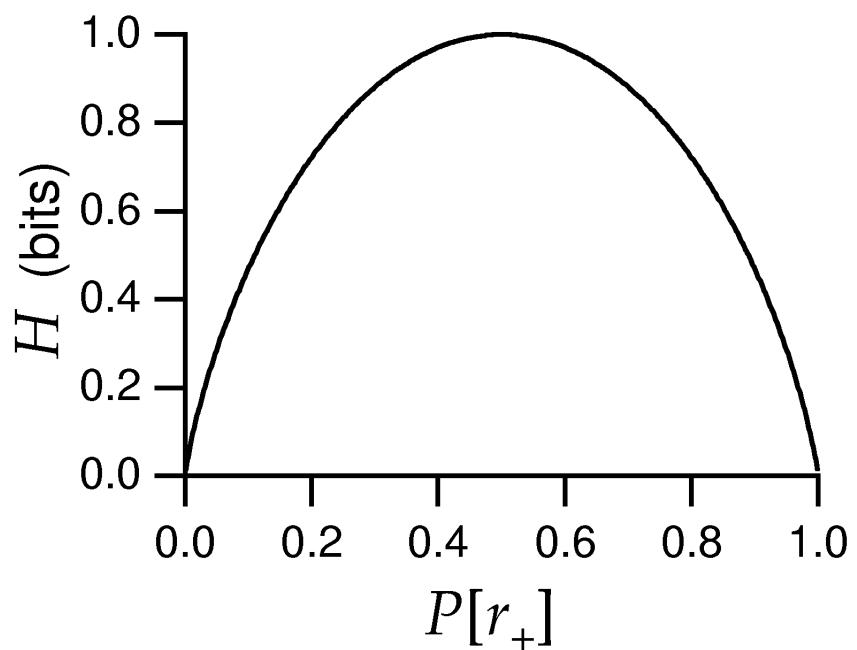
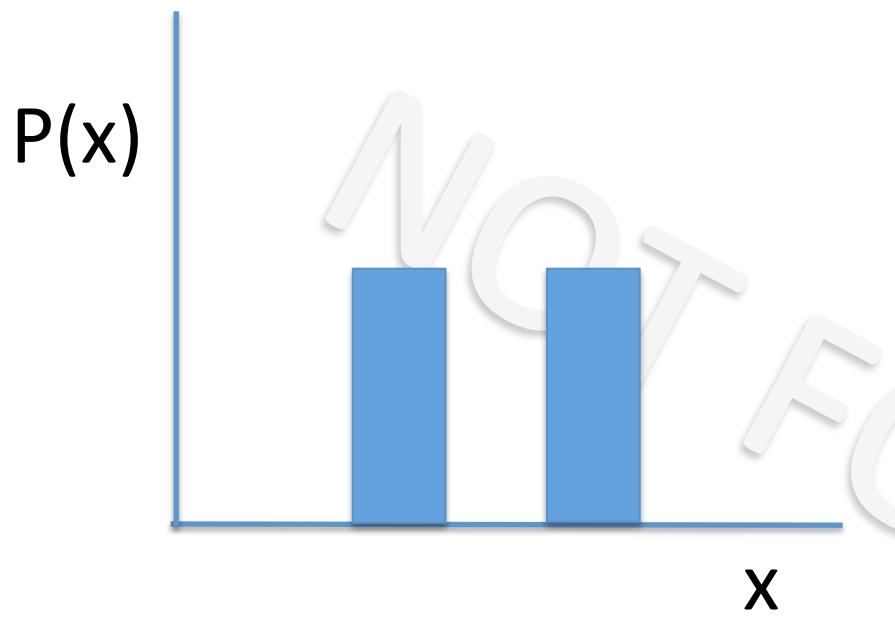
Units are *bits*

Maximizing the entropy



$$\text{Entropy} = - \sum p_i \log_2 p_i$$

Maximizing the entropy



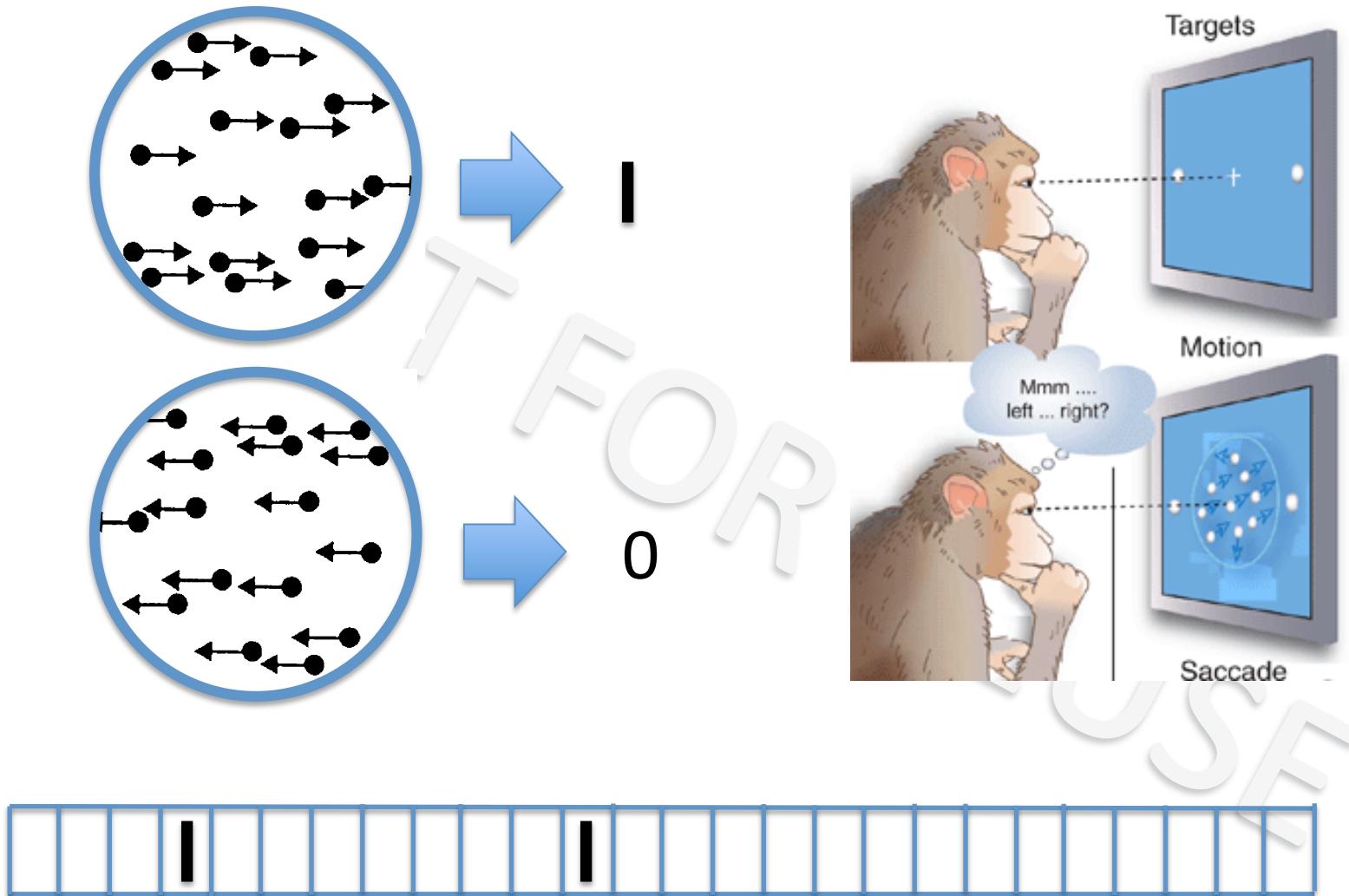
$$\text{Entropy} = - \sum p_i \log_2 p_i$$

Maximum entropy squash

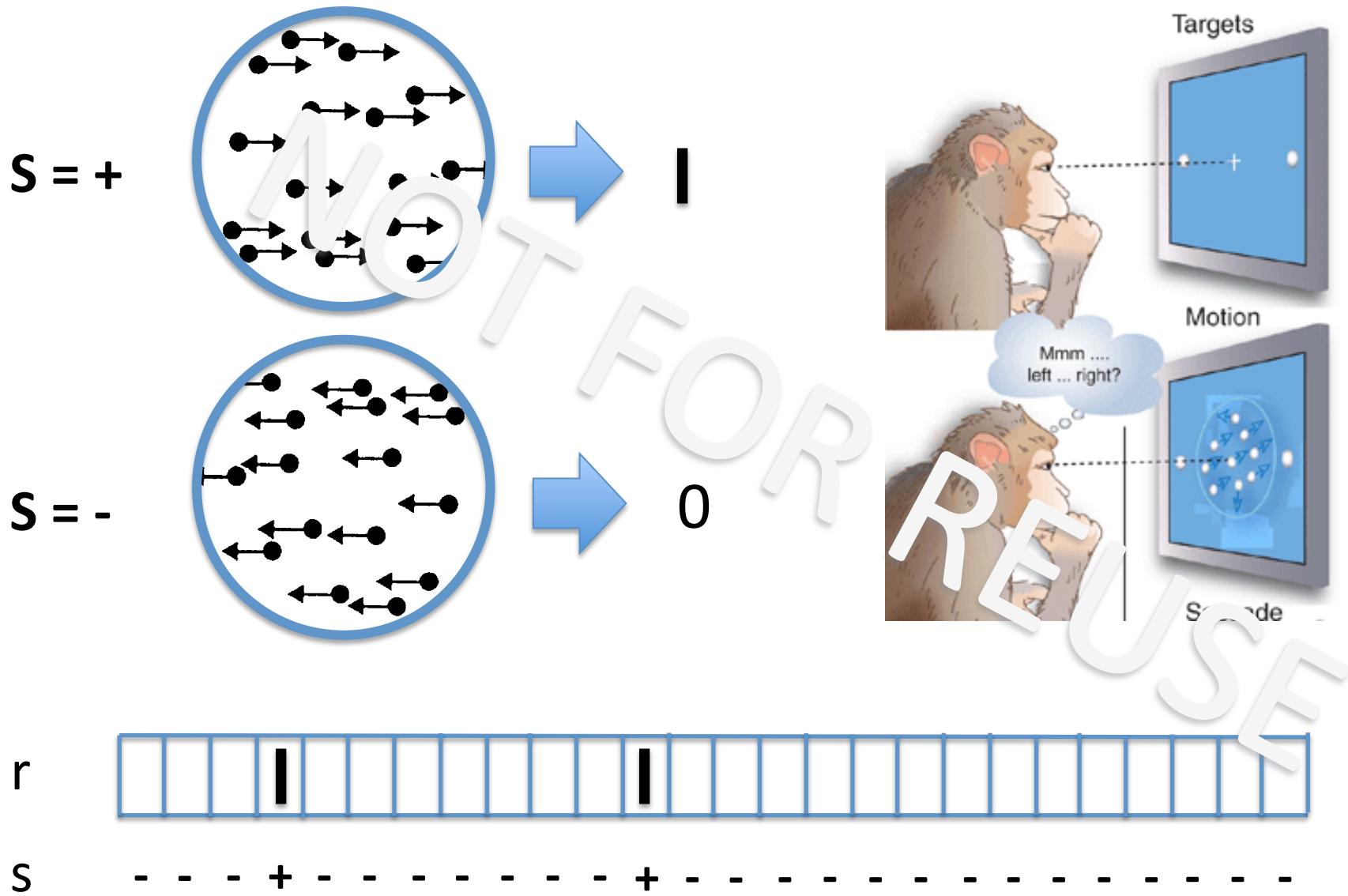


Generally $P(x)$ is not uniform... but it would be best for you if it were!

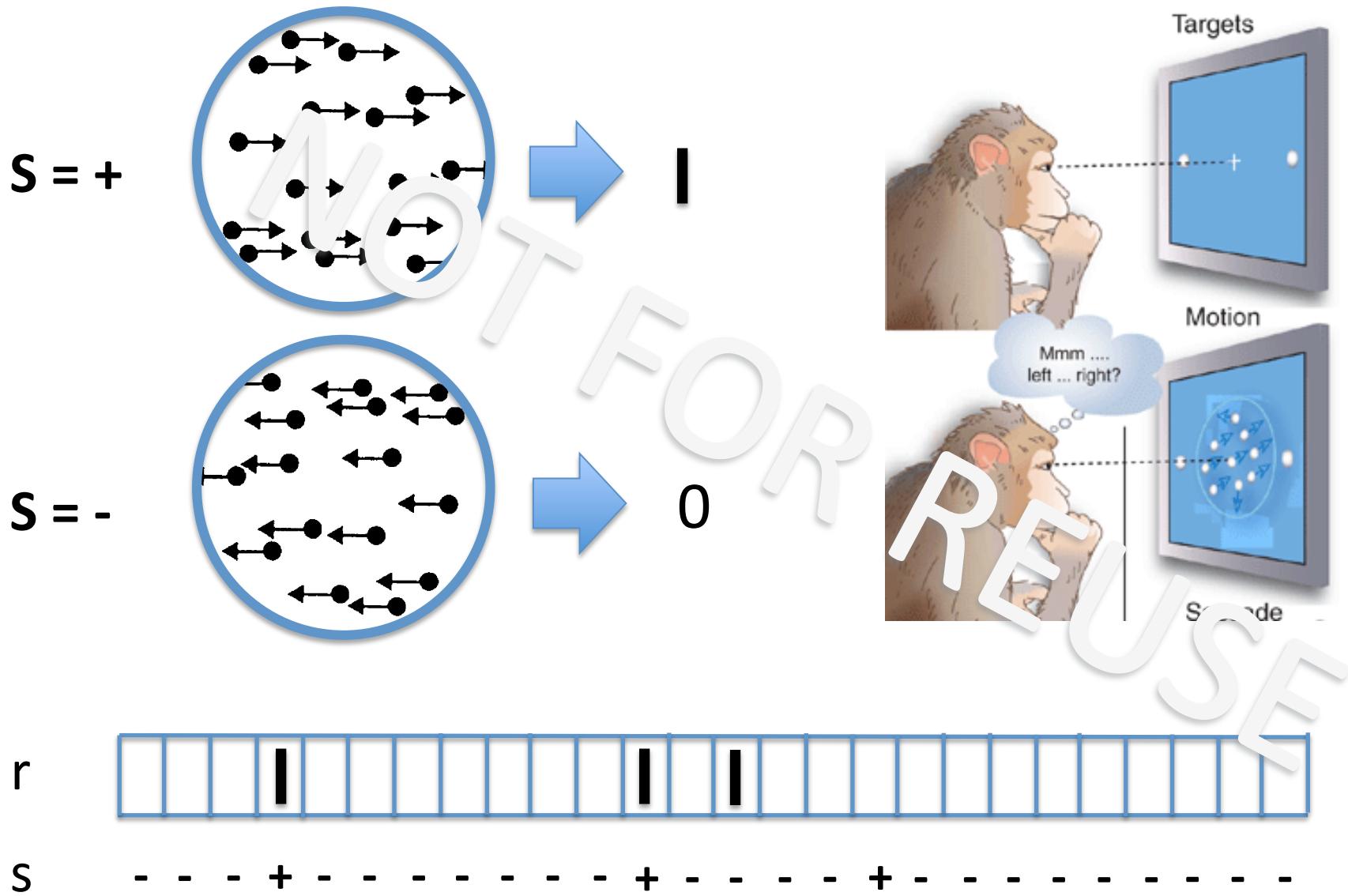
Back to our spike code: how about the stimulus?



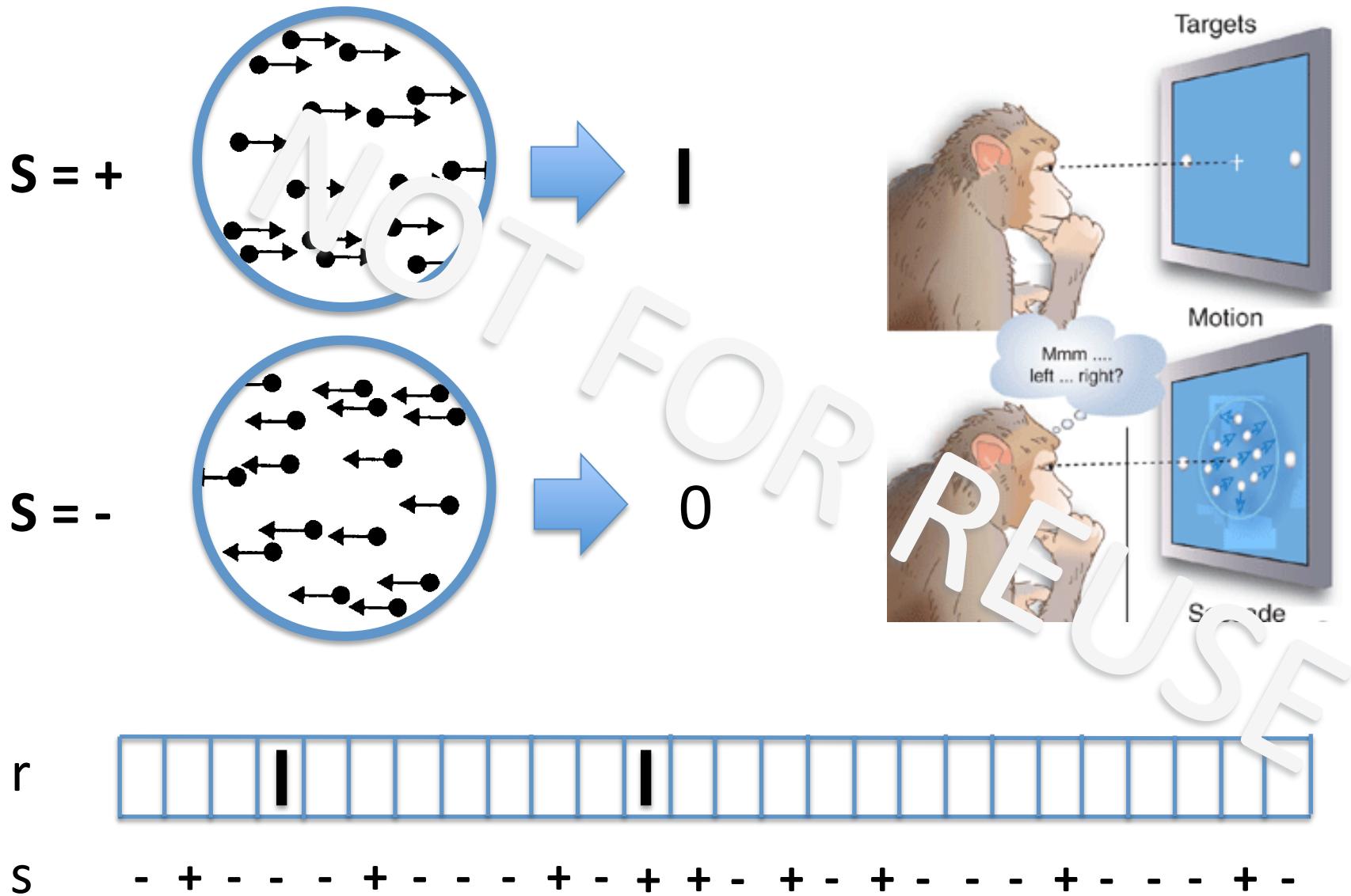
How about the stimulus?



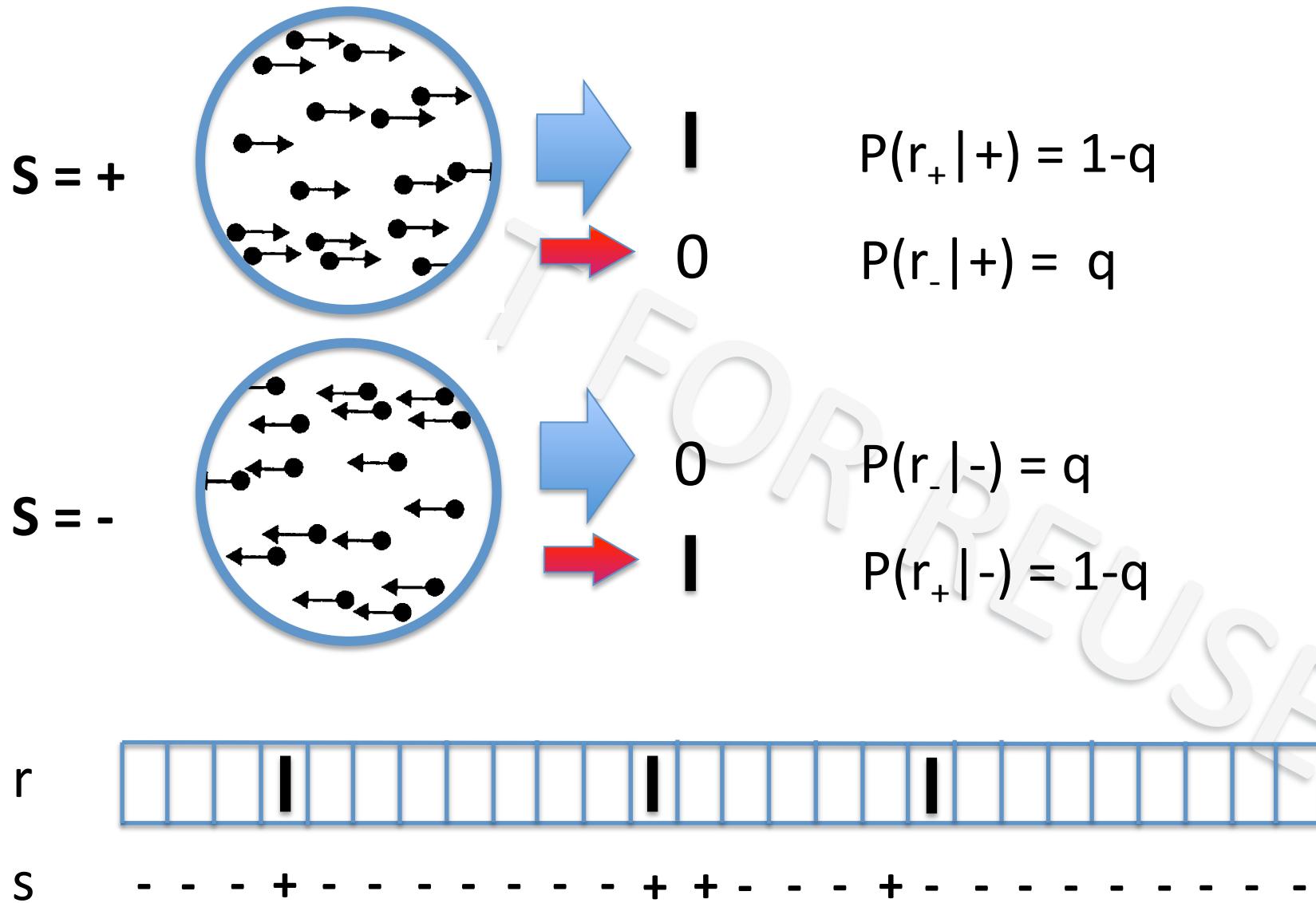
How about the stimulus?



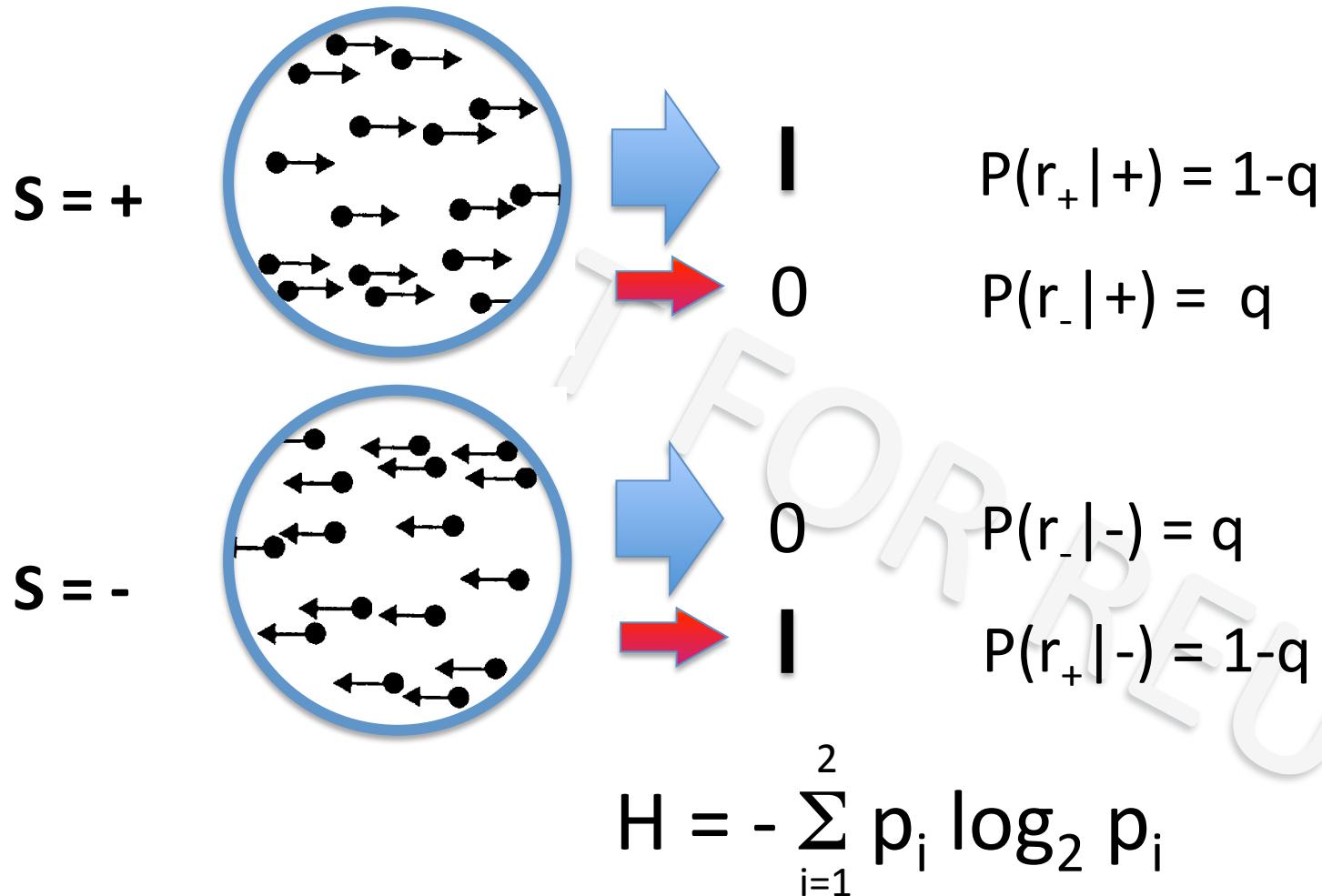
How about the stimulus?



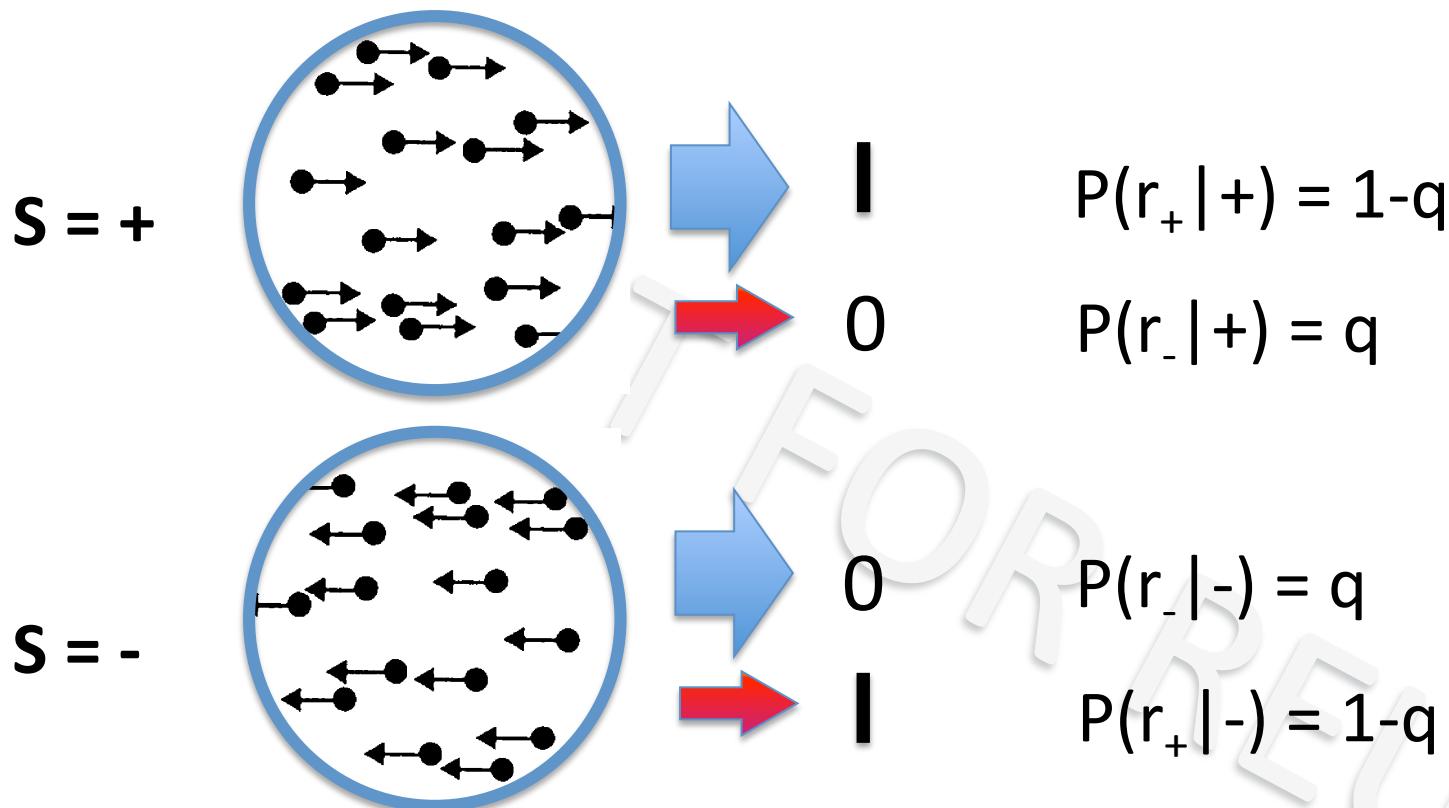
How much of the variability in r is encoding s ?



How much of the variability in r is encoding s ?



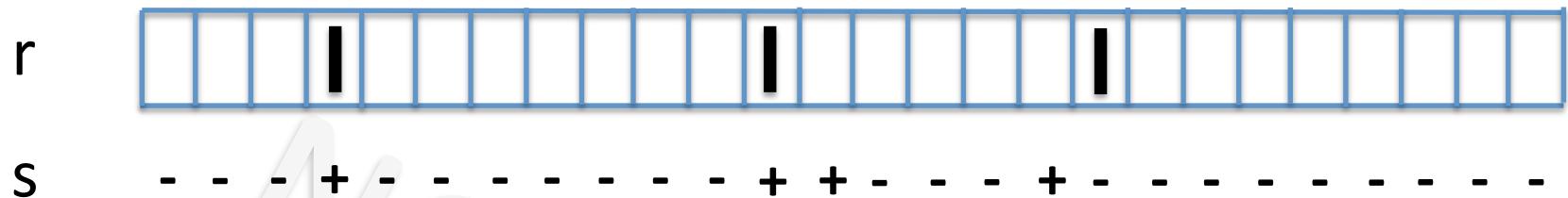
How much of the variability in r is encoding s ?



Total entropy: $H[R] = -P(r_+) \log P(r_+) - P(r_-) \log P(r_-)$

Noise entropy: $H[R|+] = -q \log q - (1-q) \log (1-q)$

Mutual information

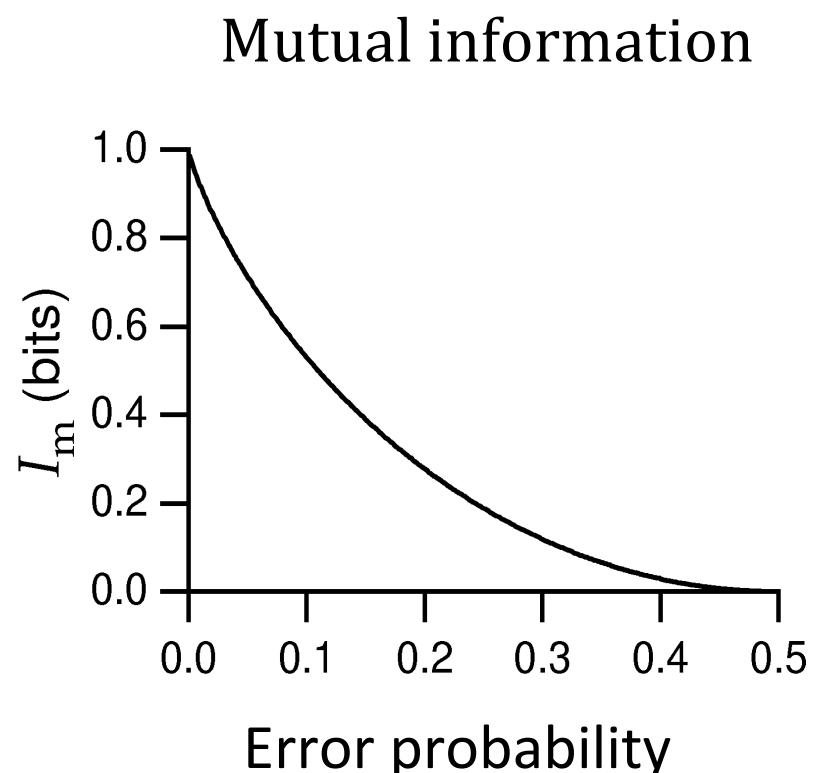
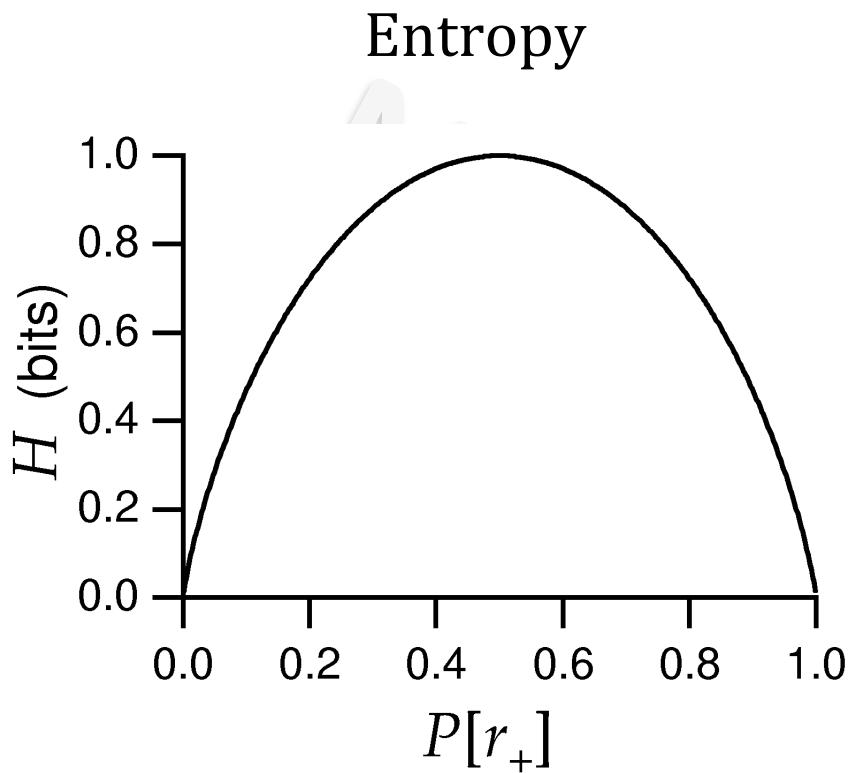


The amount of entropy that is used in coding the stimulus

$$MI(S,R) = \text{Total entropy} - \text{average noise entropy}$$

$$MI = - \sum_r p(r) \log_2 p(r) - \sum_s p(s) \left[- \sum_r p(r|s) \log_2 p(r|s) \right]$$

Entropy and information



Mutual information you can calculate in your head

1.

r



s

- + - - - + - - - + - + + - + - + - - - + - - - + -

Response is unrelated to stimulus

- What is $p(r|s)$?
- What is the MI ?

2.

r

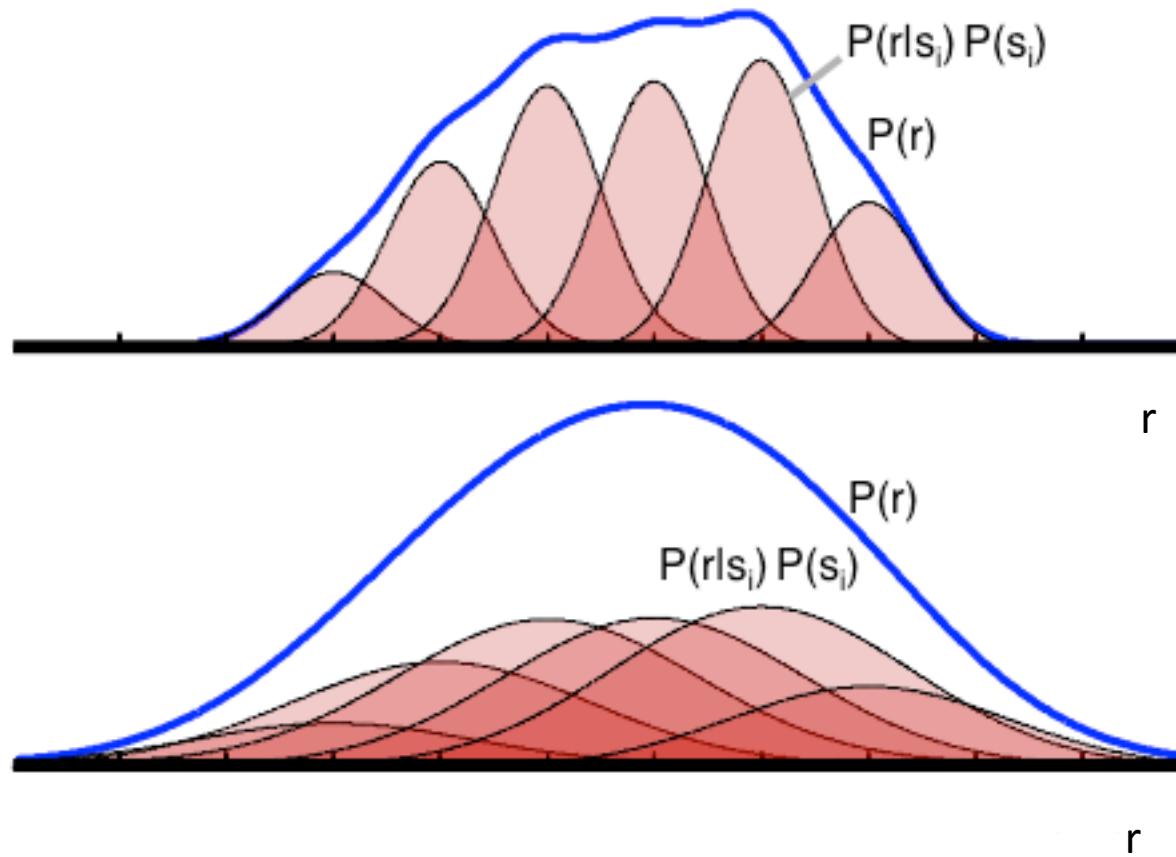


s

- - - - + - - - - - - - + - - - - - - - - - - - - - -

Response is perfectly predicted by stimulus

Entropy and information in continuous variables



Mutual information measures relationships

$$\begin{aligned} I(R,S) &= - \sum_r p(r) \log_2 p(r) - \sum_s p(s) [\sum_r p(r|s) \log_2 p(r|s)] \\ &= H[R] - \sum_s p(s) H[R|s] \end{aligned}$$

Information quantifies how *independent* R and S are:

$$I(S;R) = D_{KL} [P(R,S), P(R)P(S)]$$

Mutual information

Information quantifies how *independent* R and S are:

$$I(S,R) = D_{KL} [P(R,S), P(R)P(S)]$$

$$I(S,R) = H[R] - \sum_s P(s) H[R|s].$$
$$I(S,R) = H[S] - \sum_r P(r) H[S|r].$$

Calculating mutual information

Mutual information is the difference between the total response entropy and the mean noise entropy:

$$I(S;R) = H[R] - \sum_s P(s) H[R|s] .$$

Grandma's famous mutual information recipe

Take one stimulus s and repeat many times to obtain $P(R|s)$.

Compute variability due to noise: *noise entropy* $H[R|s]$

Repeat for all s and average: $\sum_s P(s) H[R|s]$.

Compute $P(R) = \sum_s P(s) P(R|s)$ and the total entropy $H[R]$