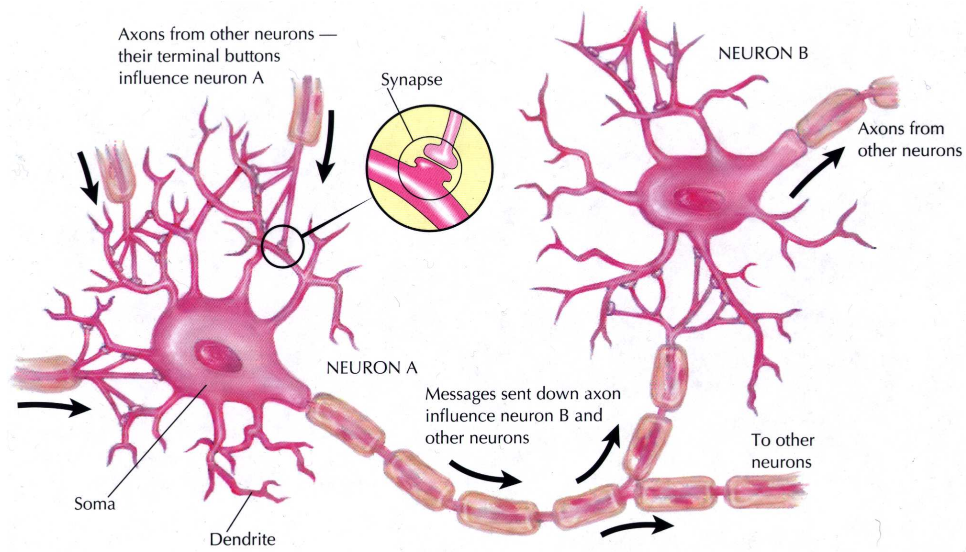
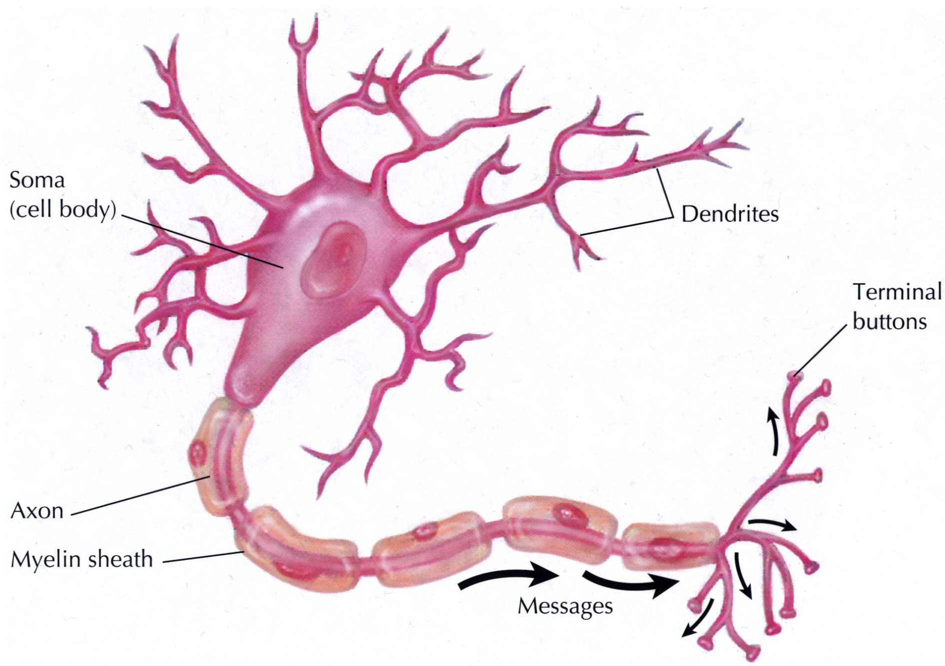
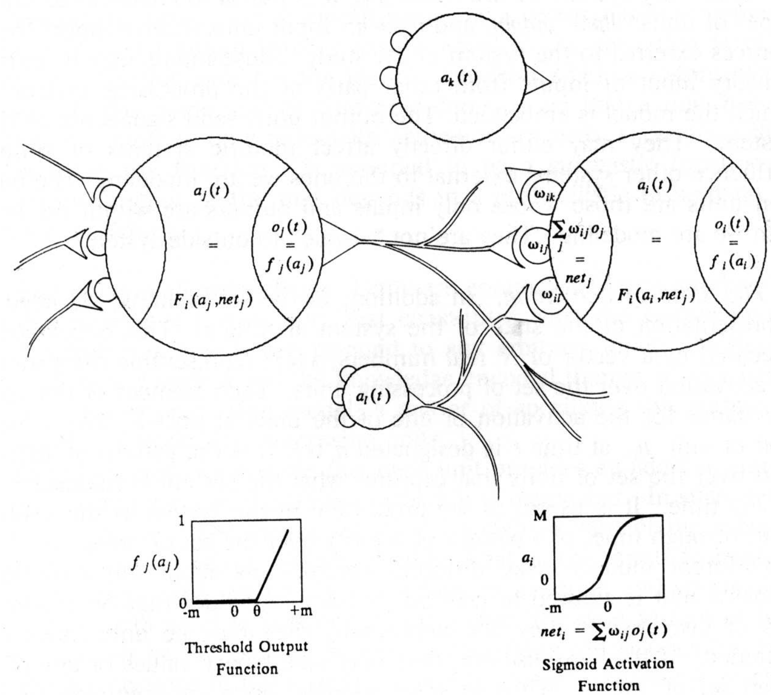
Understanding complex information-processing systems

Marr (1982)

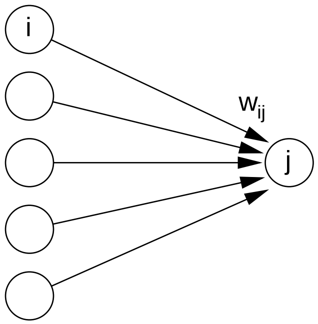
1. **Computational theory**
   What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?
2. **Representation and algorithm**
   How can this computational theory be implemented? What is the representation for the input and output, and what is the algorithm for the transformation?
3. **Hardware implementation**
   How can the representation and algorithm be realized physically?







Notation



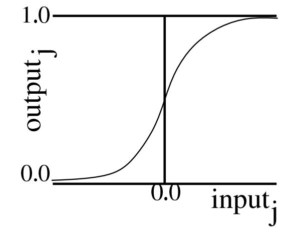
$i,\ j$	indices of units (i sending, j receiving)
a_j	**activation** of unit j
n_j	summed **net input** to unit j
w_{ij}	**weight** on connection from unit i to unit j
θ_j	**threshold** for unit j
e_j	**external input** to unit j
b_j	**bias** (tonic input) to unit j

Linear units

$$a_j = n_j = \sum_i a_i w_{ij}$$

Sigmoidal ("semi-linear") units

$$a_j = \sigma(n_j) = \frac{1}{1+\exp(-n_j)}$$



Binary stochastic units

$$p(a_j = 1) = \frac{1}{1+\exp(-n_j)}$$

Continuous time-averaged (cascaded) units

$$n_j^{[t]} = \tau\, n_j^{[t-1]} + (1-\tau) \sum_i a_i^{[t-1]} w_{ij}$$

OR

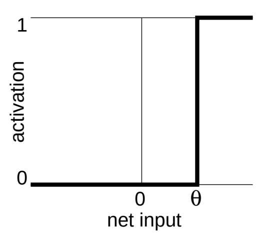
$$a_j^{[t]} = \tau\, a_j^{[t-1]} + (1-\tau)\, \sigma\left(n_j^{[t]}\right)$$

Types of units

Linear threshold unit

$$n_j = \sum_i a_i w_{ij} + e_j$$

$$a_j = \begin{cases} 1 & \text{if } n_j > \theta_j \\ 0 & \text{otherwise} \end{cases}$$



If "bias" $b_j = -\theta_j$, this is the same as

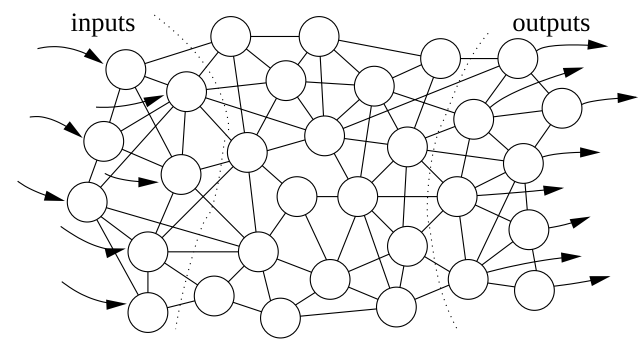
$$n_j = \sum_i a_i w_{ij} + e_j + b_j$$

$$a_j = \begin{cases} 1 & \text{if } n_j > 0 \\ 0 & \text{otherwise} \end{cases}$$

Will generally omit b_j and e_j in equations

- Bias b_j can be treated as weight w_{ij} from special unit with fixed activation $a_i = 1$.
- External input e_j can be treated as incoming activation a_i across connection with fixed weight $w_{ij} = 1$.

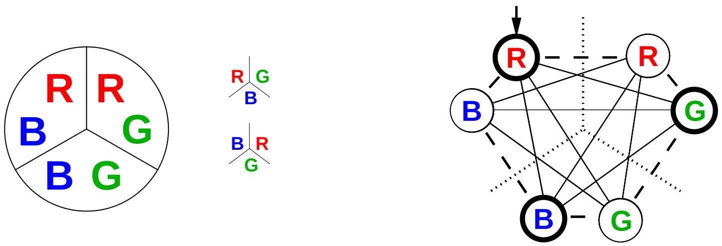
Constraint satisfaction



- Units represent **hypotheses** about parts of a problem
- Weights code **constraints** on how hypotheses can combine (i.e., the degree to which they are consistent or inconsistent)
- Possible **solutions** correspond to particular patterns of active units
- External input introduces **bias** to favor one possible solution over others

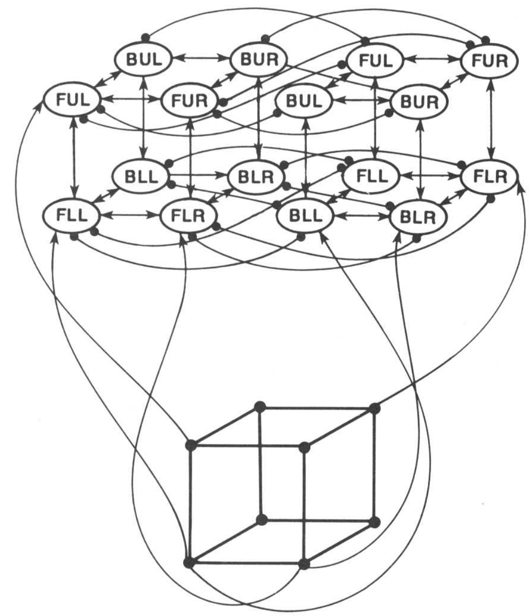
Example: Map coloring

Assign colors to regions so that no two adjacent regions have the same color.



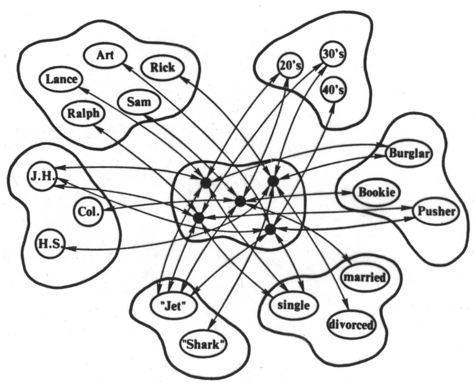
- **Hypotheses**
  - An assignment of a color to a region
- **Constraints**
  - Adjacent regions must be assigned different colors
  - Only one color can be assigned to each region
  - Each region must be assigned a color
- **Biases**
  - Initial color preference for a given region

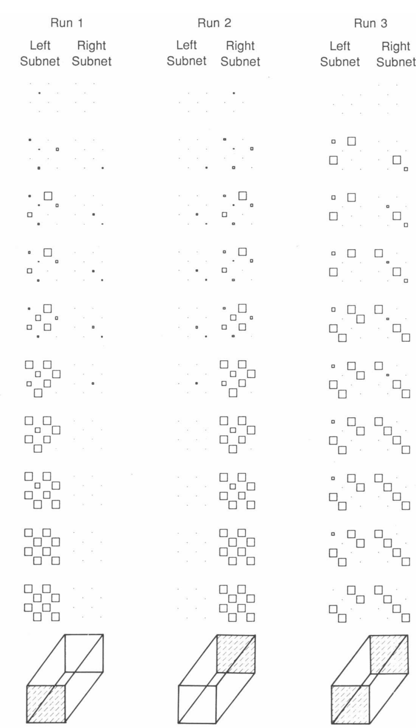
Necker Cube



The Jets and The Sharks

Name	Gang	Age	Edu	Mar	Occupation
Art	Jets	40's	J.H.	Sing.	Pusher
Al	Jets	30's	J.H.	Mar.	Burglar
Sam	Jets	20's	COL.	Sing.	Bookie
Clyde	Jets	40's	J.H.	Sing.	Bookie
Mike	Jets	30's	J.H.	Sing.	Bookie
Jim	Jets	20's	J.H.	Div.	Burglar
Greg	Jets	20's	H.S.	Mar.	Pusher
John	Jets	20's	J.H.	Mar.	Burglar
Doug	Jets	30's	H.S.	Sing.	Bookie
Lance	Jets	20's	J.H.	Mar.	Burglar
George	Jets	20's	J.H.	Div.	Burglar
Pete	Jets	20's	H.S.	Sing.	Bookie
Fred	Jets	20's	H.S.	Sing.	Pusher
Gene	Jets	20's	COL.	Sing.	Pusher
Ralph	Jets	30's	J.H.	Sing.	Pusher
Phil	Sharks	30's	COL.	Mar.	Pusher
Ike	Sharks	30's	J.H.	Sing.	Bookie
Nick	Sharks	30's	H.S.	Sing.	Pusher
Don	Sharks	30's	COL.	Mar.	Burglar
Ned	Sharks	30's	COL.	Mar.	Bookie
Karl	Sharks	40's	H.S.	Mar.	Bookie
Ken	Sharks	20's	H.S.	Sing.	Burglar
Earl	Sharks	40's	H.S.	Mar.	Burglar
Rick	Sharks	30's	H.S.	Div.	Burglar
Ol	Sharks	30's	COL.	Mar.	Pusher
Neal	Sharks	30's	H.S.	Sing.	Bookie
Dave	Sharks	30's	H.S.	Div.	Pusher





Maximizing Goodness (= minimizing Energy; Hopfield, 1982)

Global measure of the degree to which activations satisfy weight constraints

$$G\,(\text{Goodness} = -\text{Energy}) = \sum_{i<j} a_i\, a_j\, w_{ij}$$

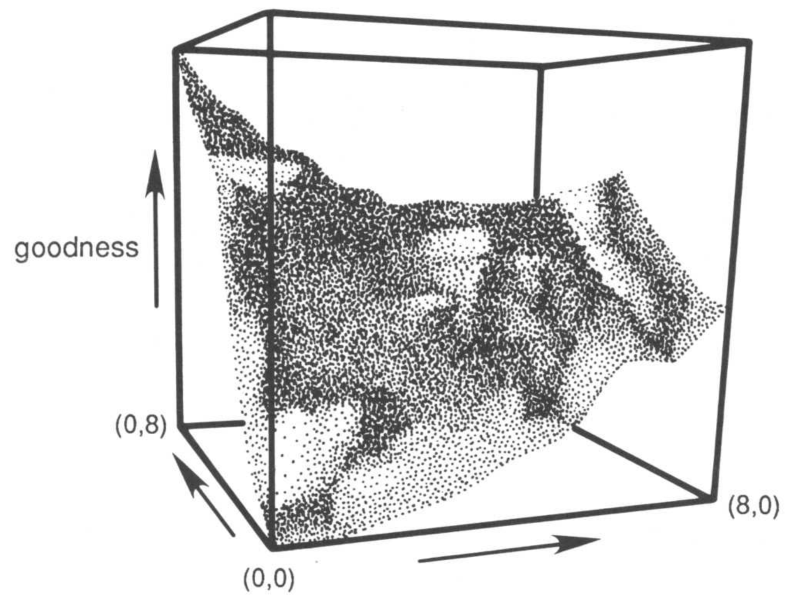
How should unit k behave locally so as to increase global Goodness?

- Set $a_k = 1$ if $G_{a_k=1} > G_{a_k=0}$ (or, equivalently, $G_{a_k=1} - G_{a_k=0} > 0$)

$$\begin{aligned} G_{a_k=1} &= \sum_i a_i\, w_{ik} + \sum_{i<j\neq k} a_i\, a_j\, w_{ij} \\ G_{a_k=0} &= \sum_{i<j\neq k} a_i\, a_j\, w_{ij} \\ G_{a_k=1} - G_{a_k=0} &= \sum_i a_i\, w_{ik} \end{aligned}$$

- Set $a_k = 1$ if $\sum_i a_i\, w_{ik} > 0$ (= **binary threshold unit**)
- *Sigmoid units*: increase activation as net input increases (monotonicity)
- *Binary stochastic units*: increase likelihood of "spike" ($a_k = 1$) as net input increases

Goodness surface with input to front face (left cube)



Goodness surface (Necker cube)