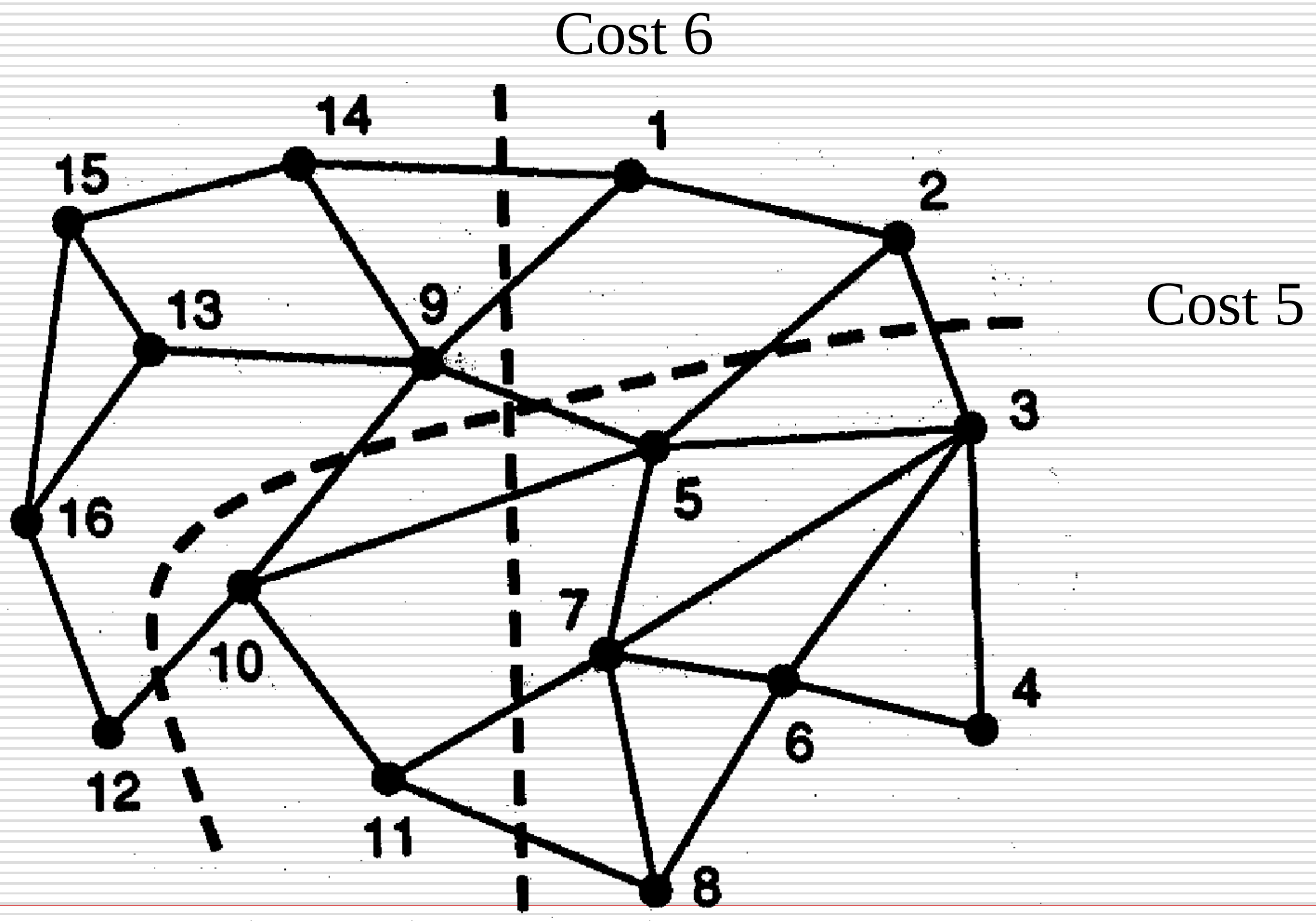
Graph bisection

- Problem statement
- Mathematical framework
- Methods
- Numerical results

Graph generation

```
T = zeros(N,N);
for i = 1:N
for j = 1:N
        if rand(1,1) > 0.5 & j~=i
                T(i,j) = 1;
                T(j,i)=1;
        end
end
end
```

Graph Bipartition



Graph bisection

- Operation: partition all nodes to two sets
- Objective
  - Minimization of cut size
  - Cut size means the number of connections between two sets
- Constraints
  - Two sets have equal size

Representation of Edges and Memberships

$C_{ij} = 1$ if vertices i and j are connected

$C_{ij} = 0$ if they are not

At each vertex define a variable S_i

$+1$ if the site is in one set and -1

if it is in the other

Mathematical framework

MINIMIZE

$$L = -\sum_{\langle ij \rangle} C_{ij} S_i S_j$$

SUBJECT TO

$$\sum_i S_i = 0 .$$

Energy function

$$H = -\sum_{(ij)} C_{ij} S_i S_j + \mu \left(\sum_i S_i \right)^2 .$$

$$H = N\mu - \sum_{(ij)} w_{ij} S_i S_j$$

$$w_{ij} = C_{ij} - 2\mu .$$

Spin model

- Graph bisection

$$\mathrm{E}(S) = -\sum_{i=1}^{N}\sum_{j \neq i}^{N} w_{ij} S_i S_j$$

$$\min_{\{S\}} E(S)$$

A physical-like random system

- Boltzmann assumption
  - S is regarded as a random vector

$$\Pr(S) \propto \exp(-\beta E(S))$$

- Free energy

$$F = \langle E(S) \rangle - \frac{1}{\beta} H(S)$$

Entropy

- Entropy of the whole system

$$H(S) = -\sum_{\{S\}} \Pr(S) \log \Pr(S)$$

- Sum of individual entropies

$$H(S) \approx \sum_{i} \mathrm{H}(\mathrm{S}_{\mathrm{i}})$$

Mean field approximation

- Individual entropy

$$H(S_i) = -\sum_{S_i=\pm 1} \Pr(S_i)\log\Pr(S_i)$$

$$\Pr(S_i) \propto \exp(\beta u_i S_i)$$

$$\Pr(S_i) = \frac{\exp(\beta u_i S_i)}{\exp(\beta u_i) + \exp(-\beta u_i)}$$

$$\Pr(S_i) \propto \exp(\beta u_i S_i)$$

$$\Pr(\mathrm{S_i}) = \frac{\exp(\beta u_i S_i)}{\exp(\beta u_i) + \exp(-\beta u_i)}$$

$$\Downarrow$$

$$\langle S_i \rangle = \frac{\exp(\beta u_i) - \exp(-\beta u_i)}{\exp(\beta u_i) + \exp(-\beta u_i)} = \tanh(\beta u_i)$$

$$\Pr(\mathrm{S_i} = 1) = \frac{\langle S_i \rangle + 1}{2}, \Pr(\mathrm{S_i} = -1) = \frac{1 - \langle S_i \rangle}{2}$$

$$\langle S_i \rangle = \frac{\exp(\beta u_i) - \exp(-\beta u_i)}{\exp(\beta u_i) + \exp(-\beta u_i)}$$

$$\Pr(\mathrm{S_i} = 1) = \frac{\langle S_i \rangle + 1}{2}, \Pr(\mathrm{S_i} = -1) = \frac{1 - \langle S_i \rangle}{2}$$

$$\Downarrow$$

$$\mathrm{H(S_i)} = -\sum_{S_i = \pm 1} \Pr(S_i) \log \Pr(S_i)$$

$$= -\Pr(S_i = 1) \log \Pr(S_i = 1) - \Pr(S_i = -1) \log \Pr(S_i = -1)$$

$$= -\frac{\langle S_i \rangle + 1}{2} \log \frac{\langle S_i \rangle + 1}{2} - \frac{1 - \langle S_i \rangle}{2} \log \frac{1 - \langle S_i \rangle}{2}$$

Mean field approximation

$$\mathrm{F} = \langle \mathrm{E(S)} \rangle - \frac{1}{\beta} H(S)$$

$$\approx \mathrm{E}(\{\langle \mathrm{S_i} \rangle\}) - \frac{1}{\beta} \sum_i H(S_i)$$

$$= -\sum_{i=1}^{N} \sum_{j \neq i}^{N} w_{ij} \langle S_i \rangle \langle S_j \rangle - \frac{1}{\beta} \sum_i H(S_i)$$

$$\mathrm{F} \approx -\sum_{i=1}^{N}\sum_{j\neq i}^{N} w_{ij}\langle S_i\rangle\langle S_j\rangle$$

$$+\frac{1}{\beta}\sum_{i}\left\{\frac{\langle S_i\rangle+1}{2}\log\frac{\langle S_i\rangle+1}{2}+\frac{1-\langle S_i\rangle}{2}\log\frac{1-\langle S_i\rangle}{2}\right\}$$

$$v_i \equiv \langle S_i\rangle$$

$$\frac{\partial \mathrm{F}}{\partial v_i} = -\sum_{j\neq i}^{N} w_{ij}v_j + \tanh^{-1}(v_i) = 0$$

$$v_i = \tanh(\beta\sum_{j\neq i}^{N} w_{ij}v_j)$$

Mean field equation

$$v_i = \tanh(\beta \sum_{j \neq i}^{N} w_{ij} v_j)$$

Mean field equations

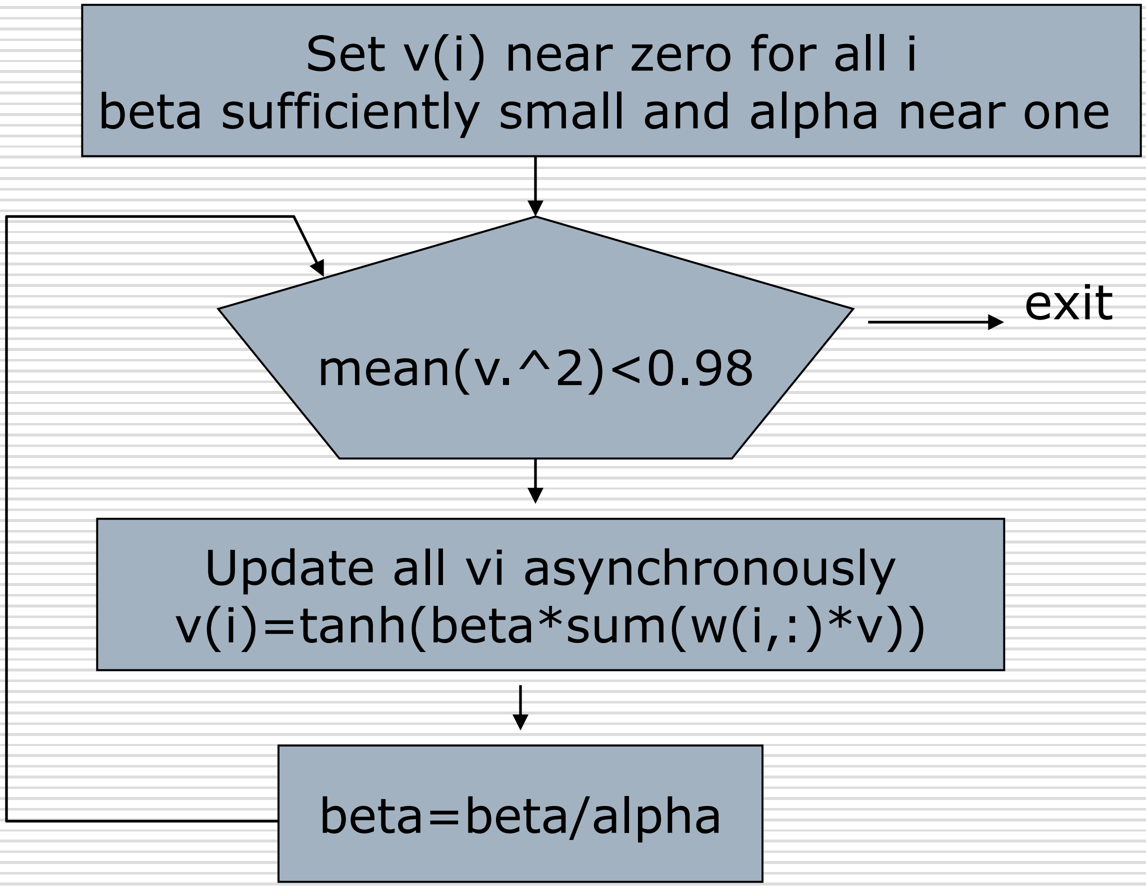
$$\langle h_i \rangle = \sum_j w_{ij} \langle S_j \rangle + h^{\text{ext}} .$$

$$\langle S_i \rangle = \tanh(\beta \langle h_i \rangle) = \tanh\left(\beta \sum_j w_{ij} \langle S_j \rangle + \beta h^{\text{ext}}\right)$$

Procedure of mean field annealing

1. Set $<s_i>$ near zero randomly
2. Set beta to a sufficiently small value
3. Employ mean field equations
   to determine a fixed point
4. If the halting condition holds, halt
5. Increase beta by an annealing schedule
6. Go to step 3

Flow chart


Set v(i) near zero for all i
beta sufficiently small and alpha near one
mean(v.^2)<0.98
exit
Update all vi asynchronously
v(i)=tanh(beta*sum(w(i,:)*v))
beta=beta/alpha

Halting condition

- Mean of squares of all <s_i> exceeds a predetermined threshold value

$$v \text{ is a vector that collects all } \langle s_i \rangle$$

$$\text{halting condition :}$$

$$\text{mean(v.\^{}2)} > 0.99$$

Annealing schedule

- Set beta to a sufficiently small value
- Increase beta carefully

Hopfield neural networks for TSP

- Problem statement
- Mathematical framework
- Mean field annealing
- Numerical results

KL(Kullback-Leberler) Divergence

- Boltzmann distribution

$$P_x(x) \propto \exp(-\beta H_x(x))$$

- Normalization

$$P_x(x) = \frac{\exp(-\beta H_x(x))}{Z_x}$$

Intractable Partition function

$$Z_x = \sum_{\{x\}} \exp(-\beta H_x(x))$$

Factorial distribution

- An approximation

$$Q_x(x) = \prod_i q_i(x_i)$$

KL divergence

- Semi-distance between two pdfs

$$\mathrm{KL}(Q_x \,||\, P_x) = \sum_{\{x\}} Q_x(x) \ln \frac{Q_x(x)}{P_x(x)}$$

Step 1

$$\mathrm{KL}(Q_x \,||\, P_x) = \sum_{\{x\}} Q_x(x) \ln Q_x(x) + \beta \sum_{\{x\}} Q_x(x) H_x(x) + \ln Z_x .$$

Step 2

$$
\begin{aligned}
S[Q_x] &\triangleq -\sum_{\{x\}} Q_x(x) \ln Q_x(x) \\
&= -\sum_i \left(\frac{1+m_i}{2} \ln \frac{1+m_i}{2} + \frac{1-m_i}{2} \ln \frac{1-m_i}{2} \right)
\end{aligned}
$$

Step 3

$$\text{and} E_{Q_x}[H(x)] \triangleq \sum_{\{x\}} Q_x(x) H_x(x) = -\frac{1}{2} \sum_{i,j \neq i} J_{ij} m_i m_j.$$

Step 4

$$
\begin{aligned}
F(m) &\triangleq -S[Q_x] + \beta E_{Q_x}[H(x)] \\
&= \sum_i \left(\frac{1+m_i}{2} \ln \frac{1+m_i}{2} + \frac{1-m_i}{2} \ln \frac{1-m_i}{2} \right) \\
&\quad - \frac{\beta}{2} \sum_{i,j \neq i} J_{ij} m_i m_j. \qquad (10)
\end{aligned}
$$