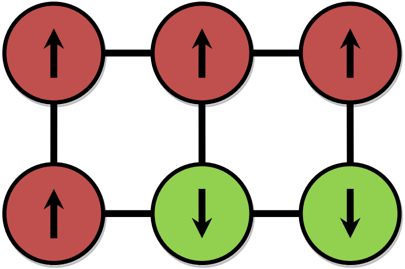
Simulation reductions for the Ising model

Mark Huber

Department of Mathematics and Computer Science, Claremont McKenna College

8 Nov, 2009

The Ising model



Begin with graph $G = (V, E)$

- Each node either “spin up” (+1) or “spin down” (-1)
- Each edge $\{i, j\}$ has strength of interaction β

Spins distribution

Configuration x has weight:

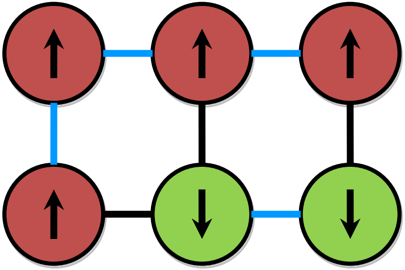
$$w_{\text{spins}}(x) = \prod_{\text{edges have same spin}} \exp(2\beta)$$

Distribution:

$$\pi_{\text{spins}}(x) = \frac{w_{\text{spins}}(x)}{Z_{\text{spins}}}, \quad Z_{\text{spins}} = \sum_{x'} w_{\text{spins}}(x')$$

Call Z the normalizing constant or partition function

The Ising model



Since 4 edges have same spin:

$$\pi_{\text{spins}}(x) = \frac{\exp(2\beta)^4}{Z_{\text{spins}}} = \frac{\exp(8\beta)}{Z_{\text{spins}}}$$

Why study Ising?

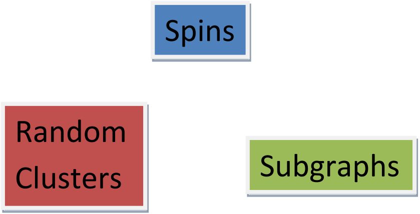
Originated in statistical physics

- Simple model
- Has phase transition on 2-D lattice

Computations

- Spatial model in statistics
- Finding Z_{spins} is a $\#P$ complete problem

The triple scoop: Three flavors of Ising



All of form $\pi(\cdot) = w(\cdot)/Z$

- $Z = \sum_{a \in \Omega} w(a)$
- Random cluster and subgraphs work directly on edges
- $Z_{\text{spins}}, Z_{\text{subgraphs}}, Z_{\text{random cluster}}$ are related by explicit formulae

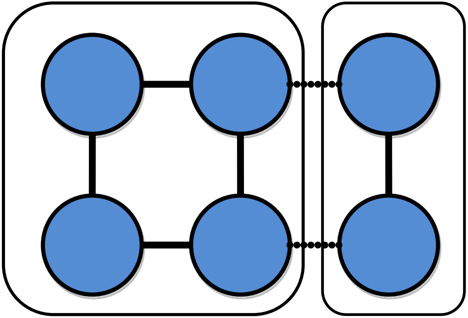
Random cluster distribution

Configuration $y \in \{0,1\}^E$ **has weight:**

$$w_{\text{random cluster}}(y) = \left\{ \prod_{e : y(e) = 1} [\exp(2\beta) - 1] \right\} 2^{c(y)},$$

$$c(y) := \#\text{ of clusters formed by edges with } y(e) = 1$$

Random cluster example



Since 5 edges and two connected components (clusters):

$$\pi_{\text{rc}}(x) = \frac{[\exp(2\beta) - 1]^5 (2)^2}{Z_{\text{rc}}}$$

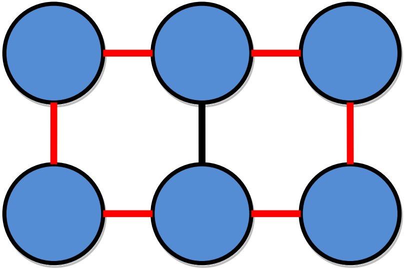
Subgraphs distribution

Configuration $y \in \{0,1\}^E$ has weight:

$$w_{\text{subgraphs}}(y) = \left\{ \prod_{e:y(e)=1} \tanh(\beta) \right\} E(y)$$

$$E(y) := \begin{cases} 1 & \text{degrees of all nodes are even,} \\ 0 & \text{otherwise} \end{cases}$$

Subgraphs example



Since 6 edges and all nodes even

$$\pi_{\text{subgraphs}}(x) = \frac{\tanh(\beta)^6}{Z_{\text{subgraphs}}}$$

Why is the subgraphs view important?

Jerrum and Sinclair [3] gave first fpras for $Z_{\mathbf{spins}}$

- Used a Markov chain approach
- Did not use spins view!
- Operated on subgraphs view

Subgraphs to spins connection not obvious:

> *"Although there is no direct correspondence between configurations in the two domains, and the subgraph configurations have no obvious physical significance..."*

Simulation reductions

Definition

Consider two families of distributions: π indexed by inputs $\mathcal{I}$, and π' indexed by $\mathcal{I}'$. Then π is *simulation reducible* to π' if a draw from π' together with a number of Bernoulli draws, is enough to generate a draw from π

new distribution π'
$+$
extra (unfair) coin flips

$\Rightarrow$ old distribution π

Swendsen-Wang [1]

Two step chain:

- 1st step: Given $X \sim \pi_{\text{spins}}$, generate $Y \sim \pi_{\text{random cluster}}$
- 2nd step: Given $Y \sim \pi_{\text{random cluster}}$, generate $X \sim \pi_{\text{spins}}$

Theorem (Swendsen-Wang 1986: Spins to clusters)

Given a draw $X \sim \pi_{spins}$ and a number of Bernoulli draws at most the number of edges, it is possible in time linear in the size of the graph to generate $Y \sim \pi_{random\ cluster}$.

Theorem (Swendsen-Wang 1986: Clusters to subgraphs)

Given a draw $Y \sim \pi_{random\ clusters}$ and a number of Bernoulli draws at most the number of nodes, it is possible in time linear in the size of the graph to generate $X \sim \pi_{spins}$.

The results [2]

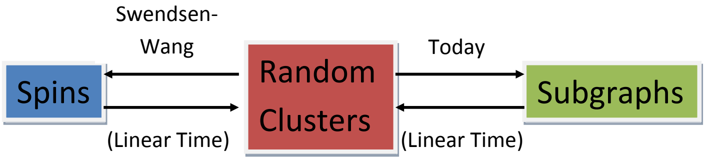
Theorem (H. 2009: Subgraphs to random clusters)

Given a draw $W \sim \pi_{subgraphs}$ and a number of Bernoulli draws at most the number of edges, it is possible in time linear in the size of the graph to generate $Y \sim \pi_{random\ cluster}$.

Theorem (H. 2009: Random clusters to subgraphs)

Given a draw $Y \sim \pi_{subgraphs}$ and a number of Bernoulli draws at most the number of edges, it is possible in time linear in the size of the graph to generate $W \sim \pi_{random\ cluster}$.

Graphically



Spins, subgraphs and random clusters are simulation equivalent

How does it work: preparation

Hyperbolic Tangent:

$$\tanh \beta = \frac{\exp(\beta) - \exp(-\beta)}{\exp(\beta) + \exp(-\beta)} \in [0, 1]$$

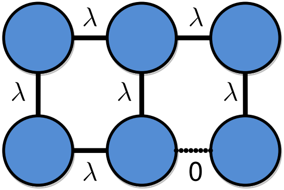
Goal:

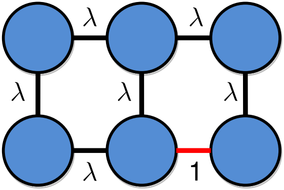
- Transform all β to 0 or ∞, so $\tanh \beta$ either 0 or 1

How does it work: mixtures

Distribution is a mixture of distributions on two graphs:

$$\lambda = \text{tanh}\,\beta$$





Do not know coefficients of mixture: 70/30? 50/50?

Do not need coefficients of mixture!

Start with draw from mixture:

$$X \sim \pi = \alpha_1 \pi_1 + \alpha_2 \pi_2$$

Write distributions using weights:

$$\pi_1 = \frac{w_1}{Z_1}, \ \ \pi_2 = \frac{w_2}{Z_2}, \ \ \pi = \frac{w}{Z} = \frac{c_1 w_1 + c_2 w_2}{Z}$$

Choose component of mixture as:

$$\mathbb{P}(I = 1) = \frac{c_1 w_1(X)}{c_1 w_1(X) + c_2 w_2(X)}, \ \mathbb{P}(I = 2) = \frac{c_1 w_2(X)}{c_1 w_1(X) + c_2 w_2(X)}$$

Procedure gives correct coeffients

Theorem

The procedure on the previous slide yields:

$$\mathbb{P}(I = i) = \alpha_i$$

and

$$[X|I] \sim \pi_I$$

Proof.

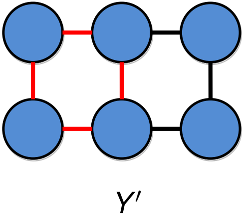
Implicit in auxiliary variable methods. (Explicit proof in [2].) $\square$

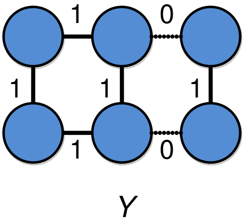
Reducing subgraphs to random clusters

Start with $Y \sim \pi_{\textbf{subgraphs}}$
While an edge has $\lambda \in (0, 1)$

- Choose either $\lambda = 0$ or $\lambda = 1$ using theorem

Example:





End result: $Y' \sim \pi_{\textbf{random cluster}}$

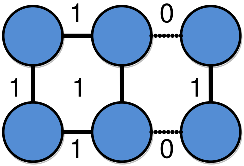
Details

Reducing random clusters to subgraphs

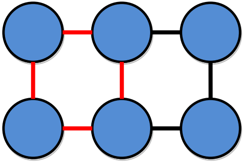
Key Question:

- Given that all $\lambda(e)$ are either 0 or 1, can we draw a subgraphs draw?

Example:



Y



Y'

How to reduce to

Leaves

- Call an edge a *leaf* if degree of one end is 1 (using only edges with $\lambda(e) = 1$ or $Y(e)$ determined)
- Set $Y(e)$ to maintain even requirement

Cycles

- If no leaves, all edges part of a cycle of $\lambda(e) = 1$ edges
- Can "flip" cycle, so $\mathbb{P}(Y(e) = 1) = \mathbb{P}(Y(e) = 0) = 1/2$

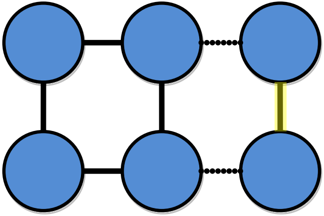
Procedure

1. Deal with leaves until no leaves left
2. Deal with cycles until no cycles left
3. If still undetermined edges, goto 1

Example: leaf

Highlighted edge is a leaf:

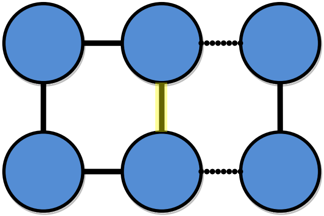
- Must have $Y(e) = 0$



Example: cycle

Highlighted edge part of a cycle

- 50-50 chance of $Y(e) = 1$
- Once $Y(e)$ set, rest of edges are leaves



Reducing random clusters to subgraphs

The procedure:

- Draw $Y' \sim \pi_{\text{random cluster}}$
- All edges with $Y'(e) = 0$ have $Y(e) = 0$
- Find maximal forest F using $Y'(e) = 1$ edges
- All remaining edges are independent with $Y(e) \sim \text{Bern}(1/2)$
- Use even degree restriction to complete $Y(F)$

Applications

Immediate results:

- New proof of formula relating $Z_{\text{subgraphs}}$ to $Z_{\text{random cluster}}$
- New Gibbs sampler for Ising:
  - From subgraphs draw, generate random cluster
  - From random cluster, generate subgraphs draw

Summary: results and future work

What is now known:

- Reduce subgraphs draw to random cluster draw with # of Bernoullis at most # of edges
- Reduce random cluster draw to subgrpahs draw with # of Bernoullis at most # of edges
- New "Swendsen-Wang" style chain for Ising model

Open questions

- What is mixing time of new Markov chain
- In particular, is it faster than Jerrum-Sinclair chain on subgraphs?

References

R. Swendsen and J-S. Wang",
Non-universal critical dynamics in Monte Carlo simulation
Phys. Rev. Lett., **58**, 86–88, 1987

M. Huber
Simulation reductions for the Ising model
arXiv:0908.2151, 2009

M. Jerrum and A. Sinclair,
Polynomial-time approximation algorithms for the Ising model",
SIAM Journal of Computing, 22, 1087–1116, 1993

Details

Recall for mixtures, need:

$$w = c_1 w_{\lambda(e)\leftarrow 1} + c_2 w_{\lambda(e)\leftarrow 0}$$

For $Y(e) = 1$ edges:

- $w_{\lambda(e)\leftarrow 0}(Y) = 0$ (can't remove this edge!)
- $w_{\lambda(e)\leftarrow 1}(Y) = w(Y)/\lambda(e)$ (remove factor of λ)

For $Y(e) = 0$ edges:

- $w_{\lambda(e)\leftarrow 0}(Y) = w(Y)$
- $w_{\lambda(e)\leftarrow 1}(Y) = w(Y)$

Combining:

$$w = \lambda(e) w_{\lambda(e)\leftarrow 1} + (1 - \lambda(e)) w_{\lambda(e)\leftarrow 0}$$

Choosing which mixture

From last page:

$$w = \lambda(e) w_{\lambda(e) \leftarrow 1} + (1 - \lambda(e)) w_{\lambda(e) \leftarrow 0}$$

Recall algorithm from theorem:

- Start with $w = c_1 w_1 + c_2 w_2$
- Choose $X \sim \pi$
- Choose which mixture I by:

$$\mathbb{P}(I = 1) = \frac{c_1 w_1(X)}{c_1 w_1(X) + c_2 w_2(X)}, \ \mathbb{P}(I = 2) = \frac{c_2 w_2(X)}{c_1 w_1(X) + c_2 w_2(X)}.$$

After using theorem, returns following algorithm:

Start with $Y \sim \pi_{\textbf{subgraphs}}$
While an edge has $\lambda \in (0, 1)$

- If $Y(e) = 1$ set $\lambda(e) \leftarrow 1$
- If $Y(e) = 0$ set $\lambda(e) \leftarrow \begin{cases} 1 & \text{with prob } \lambda(e) \\ 0 & \text{with prob } 1 - \lambda(e) \end{cases}$

Return