Asymptotics of the log-Partition Function and New Algorithms for Counting without Sampling

Antar Bandyopadhyay
(Joint work with David Gamarnik)

6$^{\text{th}}$ International Congress on Industrial and Applied Mathematics
Zürich, Switzerland

Theoretical Statistics and Mathematics Unit
Indian Statistical Institute, New Delhi Centre

http://www.isid.ac.in/∼antar

July 16 - 20, 2007

Two Counting Problems

Definition 1 (Independent Set) *Suppose $G := (V, E)$ be a finite graph. We will say a subset $I \subseteq V$ is an* independent set *of G, if for any two vertices $u, v \in I$ there is no edge between u and v.*

We will denote by $\mathcal{I}_G$, the set of all independent sets of G.

Problem 1: Given a finite graph G, count the *number of independent sets* of G.

Definition 2 (Proper q-Coloring) *Fix $q \geq 2$ an integer, and suppose $G := (V, E)$ be a finite graph. A map $C : V \to \{1, 2, \ldots, q\}$ is called a* proper q-coloring *of G, if no two vertices of same color share an edge.*

We will denote by $\mathcal{C}_G(q)$, the set of all proper q-colorings of G.

Problem 2: Given a finite graph G, and $q \geq 2$ an integer, count the *number of proper q-colorings* of G.

Exact/Approximate Counting

Q: Can we do exact counting ?

A:
- Perhaps not !
- The sets are typically exponentially large.
- No polynomial time algorithm [Valiant 1979].

Q: So what do we do ?

A: We can try "approximate" counting.

Q:
- How do we approximate ?
- What kind of approximation ?

A:
- Typical approach is to use a *Markov chain Monte Carlo* (MCMC) sampling scheme.
- One need to prove *rapid mixing* for the chain.

What Do We Propose to Do ?

- We will give *deterministic* approximation schemes, which will not use sampling.

- But we will provide ε-approximation to $\log|\mathcal{I}_G|$ and $\log|\mathcal{C}_G(q)|$. (Unfortunately, this is *obviously* less efficient !)

- Moreover, we will need restrictions on our graphs ! For example, we will need *low degree* graphs, and a "*large girth*" assumption (will be more specific later).

Get Lost !!!!

We are *obviously* doing *less* than what is known !

Then why am I giving this talk ?

- Well well ... I like this work ! ☺
- But there are more reasons than just that ! ☺

Motivation and Achievements

- Our motivation comes from *statistical physics*.

- Computation of $\log |\mathcal{I}_G|$ or $\log |\mathcal{C}_G(q)|$ are interesting, because they correspond to the *free energy* for certain models in statistical physics (the models will be given later).

- We can achieve (new) explicit results for regular graphs, which are not possible to derive using the MCMC methods. To give some example:
  - ▶ We can show that for every 4-regular graph of n vertices and large girth, the number of independent sets is approximately $(1.494\ldots)^n$.
  - ▶ We can also show that if $q \geq r+1$ then for every r-regular graph with large girth, the number of proper q-coloring is approximately
$$\left[q\left(1-\frac{1}{q}\right)^{\frac{r}{2}}\right]^n .$$

- We can drop the "large girth" assumption and work with random regular graphs to get concentration results.

A Statistical Physics Models

Hard-Core Model: Given a finite graph G and a real number $\lambda > 0$, consider a (discrete) probability distribution on $\mathcal{I}_G$ given by

$$\mathbb{P}(I) \propto \lambda^{|I|} \;\; \Leftrightarrow \;\; \mathbb{P}(I) = \frac{\lambda^{|I|}}{Z(\lambda, G)}, \quad I \in \mathcal{I}_G\,,$$

where

$$Z(\lambda, G) := \sum_{I \in \mathcal{I}_G} \lambda^{|I|}\,.$$

Remarks:

- $\mathbb{P}$ is called the *Gibbs distribution* on $\mathcal{I}_G$.
- $Z(\lambda, G)$ is called the *partition function*.
- λ is called the *activity parameter*.
- Observe $Z(\lambda, G) = |\mathcal{I}_G|$ when $\lambda = 1$, then we are back to the original counting problem.

Some Families of Graphs

- **Large girth:** An infinite family of graphs $\mathcal{G}$ is defined to have *large girth*, if there exists an increasing function $f : \mathbb{N} \to \mathbb{N}$ with $\lim_{s\to\infty} f(s) = \infty$, such that for every $G \in \mathcal{G}$ with n vertices, we have

$$\text{girth}\,(G) \geq f(n)\,.$$

- **[Low degree]:** Let $\mathcal{G}\,(n, r, g)$ be the family of graphs on n vertices, such that the maximum degree of any vertex is bounded by r and each graph has girth at least g.

- **Regular:** Let $\mathcal{G}_{\texttt{reg}}\,(n, r, g)$ be the family of r-regular graphs on n vertices, such that each graph has girth at least g.

The Main Results

Algorithm Result:

Theorem 1 *For every family of graphs $\mathcal{G}$ with maximum degree at most* 4 *and large girth, there is an algorithm $\mathcal{A}$, such that for any $\varepsilon > 0$ and $G \in \mathcal{G}$, $\mathcal{A}$ produces a quantity $\mathfrak{Z}$ in time polynomial in $n := |V|$, such that*

$$(1-\varepsilon)\,\frac{\log|\mathcal{I}_G|}{n} \leq \mathfrak{Z} \leq (1+\varepsilon)\,\frac{\log|\mathcal{I}_G|}{n}\,.$$

Results for the Regular Graphs with Large Girth:

Theorem 2 *Suppose* $\lambda < \lambda_c(r)$ *where* $\lambda_c(r) := (r-1)^{r-1}/(r-2)^r$, *then*

$$\lim_{g\to\infty} \sup_{G\in\mathcal{G}_{\text{reg}}(n,r,g)} \left| \frac{\log Z(\lambda,G)}{n} - \log\left(x^{-\frac{r}{2}}(2-x)^{-\frac{r-2}{2}}\right)\right| = 0\,,$$

where x *is the unique positive solution of*

$$x = 1/(1+\lambda x^{r-1})\,.$$

In particular, if $r = 2,3,4,5$ *and* $\lambda = 1$, *then the corresponding limits for* $\frac{\log|\mathcal{I}_G|}{n}$ *are respectively,* $\log 1.618\ldots$, $\log 1.545\ldots$, $\log 1.494\ldots$ *and* $\log 1.453\ldots$.

Results for the Random Regular Graphs:

Theorem 3 *For every* $r \geq 2$ *and every* $\lambda < \lambda_c(r)$, *the (random) partition function* $Z(\lambda, G_r(n))$, *of a random* r*-regular graph* $G_r(n)$ *satisfies*

$$\frac{\log Z(\lambda, G_r(n))}{n} \to \log\left[x^{-\frac{r}{2}}(2-x)^{-\frac{r-2}{2}}\right],$$

with high probability (w.h.p.), as $n\to\infty$, *where* x *is the unique positive solution of* $x = 1/(1+\lambda x^{r-1})$.

Two Main Steps of the Algorithm (Illustrated only for the Independent Sets)

STEP 1 (The Cavity Equation):

- In this step we relate the computation of the partition function to the computation of the marginal probabilities.
- This is done by creating a *cavity* in the original graph.
- Let G_1 be the original graph G with one vertex, say v_1, removed.
- By definition

$$Z(\lambda, G_1) = \sum_{I \in \mathcal{I}_{G_1}} \lambda^{|I|} = \sum_{\substack{I \in \mathcal{I}_G \\ v_1 \notin I}} \lambda^{|I|}.$$

- Thus

$$\frac{Z(\lambda, G_1)}{Z(\lambda, G)} = \mathbb{P}_G(v_1 \notin \mathbf{I}) .$$

STEP 2 (Computation on Trees):

- Note our *large girth* assumption makes our graphs "*locally*" tree like !
- So in this step we only make computation for the marginal probabilities when the graph is a finite tree.
- This can be done easily by a recursive method.

Proposition 4 *Suppose T be a finite rooted tree with root v_0, and let $\{v_1, v_2, \ldots, v_k\}$ be $k \geq 0$ children of v_0. For each $1 \leq j \leq k$, let $T(v_j)$ denote the tree rooted at v_j consisting only the descendants of v_j (if any). Then the following recursion holds*

$$\mathbb{P}_T\left(v_0 \notin \mathbf{I}\right) = \frac{1}{1 + \lambda \prod_{1 \leq j \leq k} \mathbb{P}_{T(v_j)}\left(v_j \notin \mathbf{I}\right)}.$$

The Algorithm

INPUT: A graph G with vertex set $\{v_1, v_2, \ldots, v_n\}$, and a number $\varepsilon > 0$.

BEGIN

1. Compute the girth $g = g(G)$.

2. If $(0.9)^{\frac{g}{2}-2} \geq \varepsilon$ then find $|\mathcal{I}_G|$ by enumeration and STOP.

If not then

3. Set $Z \leftarrow 1$, $t \leftarrow \lfloor g/2 \rfloor$ and $k \leftarrow 1$.

4. Find the t-depth neighborhood $T(v_k)$ of v_k.

5. Compute the marginal probability $p = \mathbb{P}_{T(v_k)}(v_k \notin \mathbf{I})$ for the finite tree $T(v_k)$.

6. Set $Z \leftarrow Z/p$, $G \leftarrow G \setminus \{v_k\}$, $k \leftarrow k+1$.

7. If $k \leq n$ then goto Step 4, otherwise STOP.

END

OUTPUT: Z

Why Does It Works ?

⋆ **Strong Correlation Decay:**

- This is the crucial part !
- We prove that under certain assumptions [e.g. $\lambda < \lambda_c(r) = (r-1)^{r-1}/(r-2)^r$ (for the r-regular case), or $r \leq 4$ (for deriving the algorithm), etc.], the *influence* of the boundary at the root decreases exponentially fast.
- A statistical physics consequence of this is the Gibbs measure on the "limiting infinite graph" (if any !) is unique, that is there is no *phase transition*.
- For the infinite r-regular trees it was shown by Kelly (1985), that there is no *phase transition* for the hard-core model if and only if $\lambda \leq \lambda_c(r)$.
- For counting independent sets we extend this result to the class of finite trees with maximum degree at most 4, which is the most crucial result for our algorithm to succeed.

$\star$ **From the Tree to the Original Graph:**

- In this step we show that the error we make by taking a *local tree* around a vertex is *small*.

  Note: The *local tree* comes from the *large girth* assumption.

- This is again done by using the strong correlation decay property and the (spacial) Markovian nature of the Gibbs distribution.

Some Final Remarks

- A recent work of Weitz (2006) provides a *fully polynomial approximation scheme* for any finite graph with low degree for the problem of counting the independent sets, but it does not give explicit limit results such as ours for the regular graphs.

- Gamarnik and Katz (2006) (personal communication) have extended the work of Weitz (2006) for other counting problems, e.g. counting colorings, and counting matchings on general finite graphs.

- It seems to me that each of this is a “success story” for making a rigorous argument for a very powerful method of statistical physics, called the **cavity method** ! But the full math picture is yet to be discovered.

Thank You