Empirical Support for Algorithmic Conjectures

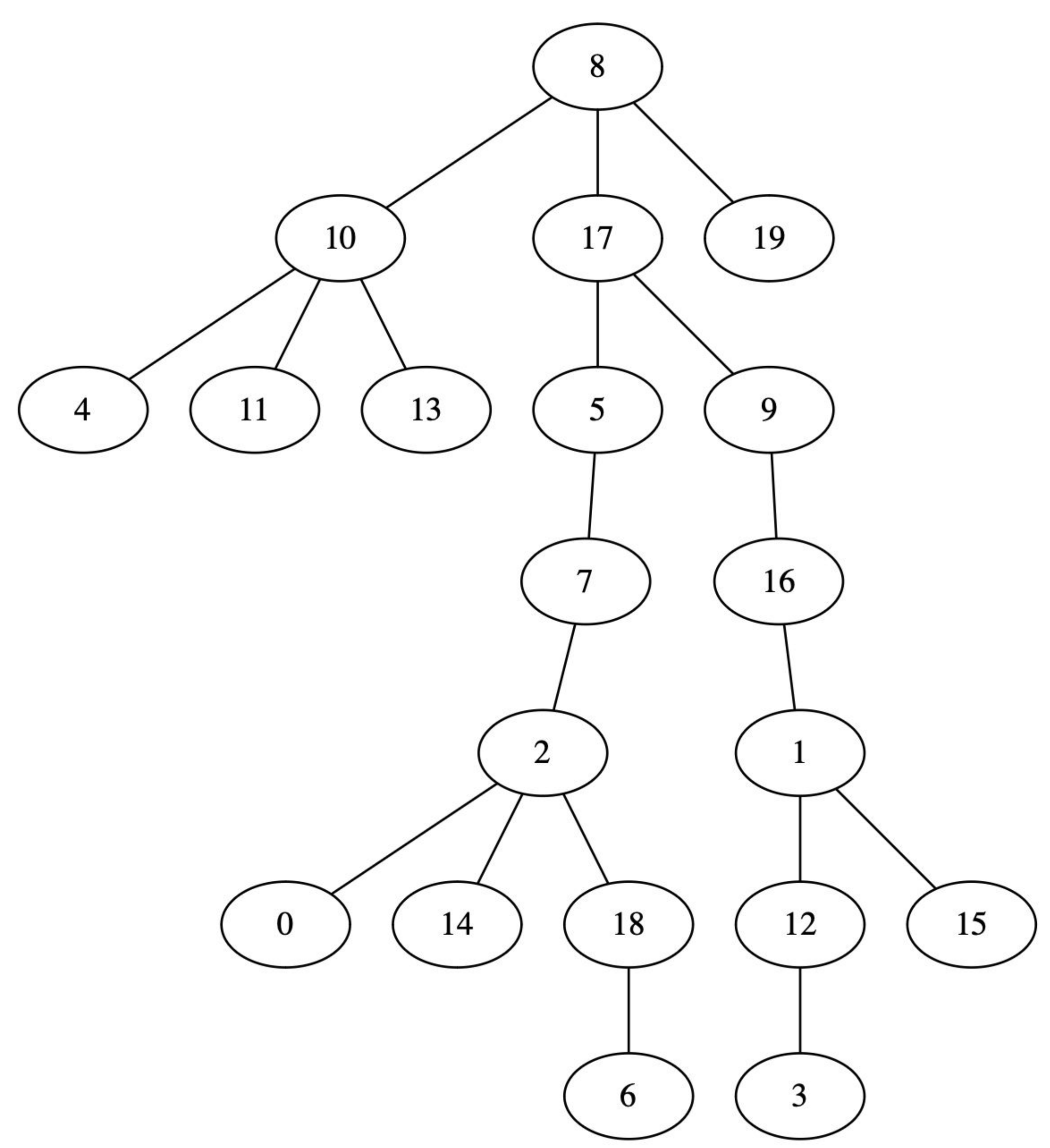
Student Researcher: Shayan Daijavad
Professor: Daniel Frishberg

CAL POLY

Introduction

Markov Chain Monte Carlo methods use Markov chains to sample from distributions. A Markov Chain is a model of a sequence of “states”, where the probability of being at a given state depends only on the state that came before it. Given enough time (known as the mixing time), Markov Chains converge, or “**mix**”, to a stationary distribution, or a distribution that the chain tends to stay at regardless of where it started from.

Our project focuses on a particular Markov Chain Monte Carlo algorithm, with applications in statistical physics, known as hardcore model **Glauber dynamics**. The target distribution of Glauber dynamics is a distribution of all of the independent sets within a graph. An **independent set** is a set of vertices within a graph with no two vertices in the set containing an edge between them.



Research Questions

RQ1: Does the Glauber dynamics for sampling independent sets on trees mix in time O(nlogn)?
RQ2: If we bias the algorithm in favor of larger or smaller sets, how does the mixing time change?

Generally, Glauber dynamics does not mix fast. However, in trees, prior research has shown that it does. [1] Work done by Efthymiou, Hayes, Stefankovic, and Vigoda shows the optimal mixing time in trees of size n to be O(n^2), given a λ value of less than 1.1.

The λ parameter gives more or less weight to different sizes of independent sets.

Glauber Dynamics Pseudocode

Let $G = (V, E)$
for $t = 1 \ldots T$ **do**
  Pick v uniformly at random from V
  if $v \in x_{t-1}$ **then**
    With probability $\frac{1}{2}$, let $x_t = x_{t-1}$ without v
  else
    if $v \notin x_{t-1}$ and $x_{t-1} \cup \{v\}$ is an independent set **then**
      With probability $\frac{1}{2}$, let $x_t = x_{t-1}$ with v
return x_T

Experimental Design

Our experiment design consists of initially generating random trees of size n. Glauber dynamics can then be used as a way to approximately count the number of independent sets within each tree. This is accomplished by taking K samples of independent sets from the tree, except we remove at least one edge from the tree. We count the number of samples from this new tree that would have also been an independent set in the original tree, and then use the ratio between the two to compare against an exact counting algorithm implemented via dynamic programming.

The dynamic programming algorithm gives us the true number of independent sets in a tree, and gives us a way to approximate how well the Glauber dynamics counting algorithm works.

In theory, given enough time and a sufficient number of samples, the Glauber dynamics algorithm will be reasonably accurate. By seeing what values of T and K give us a more accurate estimate, we can measure the mixing time of Glauber dynamics.

Preliminary Data

We implemented the experiments using C++ and the Boost Graph library, running each experiment on a laptop. We ran tests on trees of up to size 1000, trying different values of T, K, and the λ parameter. We found that the counter is more accurate (ratios being off by 0.01-0.1 factor on average) with T being above 1000 and K being around 100, with the λ value being 1. Lower values of all variables, both individually or together, resulted in a far less accurate approximation. This was an unexpected result, as we expected lower λ values to work better.

Next Steps

Our next steps involve running the algorithms on larger instances. This includes running the algorithm with larger values of T and K, in order to see more accurate results from the counter and gather more data for the mixing time of Glauber dynamics.

There are other heuristics that we have not explored yet which may run faster, specifically coupling from the past.

Acknowledgements

This research was funded by the College of Engineering.

REFERENCES

1 Charilaos Efthymiou, Thomas P. Hayes, Daniel Štefankovič, and Eric Vigoda. Optimal Mixing via Tensorization for Random Independent Sets on Arbitrary Trees. In Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2023). Leibniz International Proceedings in Informatics (LIPIcs), Volume 275, pp. 33:1-33:16, Schloss Dagstuhl – Leibniz-Zentrum für Informatik (2023)

https://doi.org/10.4230/LIPIcs.APPROX/RANDOM.2023.33