Report for the winter semester 2001/2002

Dr. Mihyun Kang

Supervisor:	Prof. Dr. Hans Jürgen Prömel
Field of Research:	Random walks,
	random discrete structures
Topic:	Markov chain Monte Carlo methods
Postdoctoral member	of the program since October 2001

Field of research and preview

Random walk is a topic and a method in graph theory, combinatorics, randomized algorithms, probability and statistics. P. Diaconis [3] used group representations to investigate the probabilistic properties of random walks on finite groups. In [6] and [8], I considered random walks on finite graphs, which can not be directly induced by finite groups, but can be decomposed into several finite groups. Investigated were the probability distributions of the first hitting time, which is the minimum number of steps that it takes to reach an element starting from another element. In [1], tested were several pseudorandom numbers using random walks.

The Markov chain Monte Carlo method is an approach of generating or counting random objects, or approximate integration, using random walks. For example, consider a random sampling problem: when one wants to pick an element from a large finite set $S = \{1, \cdots, k\}$ of combinatorial objects according to a probability distribution π on S, one constructs an ergodic Markov chain (X_t) on S with a transition matrix P, i.e., $P_{i,j} = \Pr(X_{t+1} = j|X_t = i)$, and a stationary distribution π, i.e., $\pi P = \pi$, take random walks on the chain for sufficiently long time and output the final state as a sample in that for an ergodic Markov chain, there exists a unique stationary distribution and the transition probability at time t converges to the stationary distribution as t becomes large [7].

In approximate counting and approximate random sampling, it is important to obtain the mixing time of Markov chains, i.e., the number of steps that the Markov chain is very close to its stationary distribution, since it gives an estimation of the running times of such algorithms. Many techniques are developed to obtain upper bounds on the mixing time of Markov chains using spectral properties of the transition matrix and group representations, canonical paths, coupling and path coupling. They are successfully applied for random generations and approximate counting of combinatorial problems, for example, card shuffling, approximate volume estimation for convex bodies, generating random spanning trees, Hamiltonian cycles and graph coloring [4].

Recently many researches were done on perfect or exact sampling algorithms, which generate random objects whose distribution is exactly the desired probability distribution. Among them are Prop-Wilson's coupling from the past (CFTP) [9] and Fill's perfect rejection sampling algorithms [5]. These algorithms are based on coupling methods which run independently and in parallel the copies of a Markov chain starting from each initial state and stop when all copies have coalesced.

Unfortunately it can be very difficult to obtain the mixing time of Markov chains for a certain combinatorial problem. Unknown is as well the mixing time of the Markov chain for random planar graphs, whose stationary distribution is the uniform distribution on the set of all planar graphs with n vertices [2]. Two questions arise on random planar graphs. One is whether this Markov chain is rapidly mixing, i.e., whether its mixing time is polynomial in the logarithm of the size of the state space. The other question is whether it is possible to apply CFTPs and Fill's algorithms to generate random planar graphs. One of the advantages of these algorithms is that the algorithms decide when to stop so that one does not need to compute the mixing times in advance even though the expected running times of the algorithms are about the mixing times.

Similar questions arise on random outerplanar graphs, i.e., whether one can design a rapidly mixing Markov chain for random outerplanar graphs and whether one can approximately or perfectly generate random outerplanar graphs using Markov chains. This is a joint work with M. Bodirsky, a Ph. D. student of the program.

Activities

- Research seminar of the group Algorithms and Complexity at Humboldt University of Berlin with talks "Random walks on a finite group" and "Markov chain Monte Carlo algorithms".

- Lectures and colloquia of the program CGC.

- Fall school on Discrete Geometry: Triangulations from various points of view, October 4 to 6, 2001, Berlin.
- 28. Berliner Algorithmentag, February 15, 2002, Technical University Berlin.

References

[1] G. H. Choe, M. Kang and C. Kim, Efficiency of Knuth's pseudorandom number generators, draft.

[2] A. Denise, M. Vasconcellos and D. Welsh, The random planar graph, *Congressus Numerantium* 113:61–79, 1996.

[3] P. Diaconis, *Group Representations in Probability and Statistics*, Inst. Math. Stat., Hayward, 1988.

[4] M. Dyer and C. Greenhill, Random walks on combinatorial objects, *Surveys in Combinatorics* (J. D. Lamb and D. A. Preece, eds.), London Mathematical Society Lecture Note Series 267:101–136, Cambridge University Press, Cambridge, 1999.

[5] J. A. Fill, An interruptible algorithm for perfect sampling via Markov chains, *The Annals of Applied Probability* 8:131–162, 1998.

[6] M. Kang, First hitting times of simple random walks on several graphs, submitted.

[7] M. Kang, Markov chain Monte Carlo algorithms, draft.

[8] M. Kang, Random walks on a union of two finite groups with one common element, submitted.

[9] J. G. Propp and D. B. Wilson, Exact sampling with coupled Markov chains and applications to statistical mechanics, *Random Structures and Algorithms* 9:223–252, 1996.