Better than Random: Quasirandomness for Discrete Stochastic Systems

James Propp (U. Wisconsin)

September 21, 2004

(based on articles in progress with Ander Holroyd and Lionel Levine;

with thanks also to Hal Canary, Matt Cook, Dan Hoey, Michael Kleber, Yuval Peres, and Oded Schramm)

Overview, with an example

Consider a discrete random system with some numerical property X whose average-case behavior $E(X)$ we want to determine.

One way to estimate $E(X)$ is to generate N independent samples of the random variable X and take the average $(X_1 + \ldots + X_N)/N$. This estimate is typically within $O(1/\sqrt{N})$ of $E(X)$.

The theme of this talk is that you can sometimes get even better estimates by using $(x_1 + \ldots + x_N)/N$ where $x_1, \ldots, x_N$ are properly chosen deterministic samples.

Key idea: Replace randomness by a low-discrepancy property.

Recall that the archetype for discrete randomness is an "unpredictable" fair coin.

The archetype for discrete *quasirandomness* is the deterministic sequence H, T, H, T, H, T, ... (or, equally good, T, H, T, H, T, H, ...).

(This is no longer unpredictable, but it's still "fair"!)

After N tosses of a "quasirandom coin", the number of heads is $N/2 + O(1)$; i.e., the empirical estimate of the bias is $1/2 + O(1/N)$.

Quasirandom walk on finite graphs

Consider a strongly connected finite directed graph in which each vertex has outdegree 2. A bug moving through the directed graph chooses which way to go at each vertex by using a quasirandom coin sitting at that vertex.

For some fixed vertex v, let s_N be the number of times the bug visits vertex v during the first N steps of its quasirandom walk.

FACT: $|s_N - Np| = O(1)$, where p is the steady-state probability associated with v under random walk. That is, $|s_N/N - p| = O(1/N)$.

Note that $O(1/N)$ is the best one could hope for.

Quasirandom approaches tend to give estimates with error deterministically bounded by C/N or $C(\log N)/N$, where the straightforward random approach gives estimates bounded in expectation by $C/\sqrt{N}$.

(Sometimes this is rigorously known; sometimes this is only conjectural but is strikingly well-supported by data.)

More generally, the simplest kind of quasirandom variable with m different equally probable values is a sequence that rotates through the m allowed values in some fixed order:

$$1, 2, 3, \ldots, m-1, m,$$
$$1, 2, 3, \ldots, m-1, m,$$
$$1, 2, 3, \ldots, m-1, m,$$
$$\ldots$$

We call this a rotor.

If these m values are the m arcs emanating from v, we call this quasirandom variable a **rotor-router**, and we picture it as an arrow that points at the neighbors of v in some fixed cyclic order.

For reasons that will become clearer during tomorrow's talk, we advance the rotor at a vertex *before* we move the bug.

Thus, the rotor at an unoccupied site that has been visited before always points in the direction in which the bug left the vertex on its most recent visit.

FACT: Quasirandom walk on any strongly connected finite directed graph gives discrepancy

$$|s_N - Np| = O(1)$$

where s_N is the number of times the bug visits vertex v during the first N steps of its quasirandom walk, and p is the steady-state probability associated with v under random walk.

A quasirandom walk in an infinite graph

Start by considering a random walk.

States: $-1, 0; 1, 2, 3, ...$

Transitions: 1 step to the right or 2 to the left (equally likely)

Start bug at 1

Absorb bug at -1 and 0

If we put a bug at 1, what's the probability that it wanders off to the right without ever getting absorbed?

Zero.

If we put a bug at 1, what's the probability that it gets absorbed at -1?

Approximately .618034.

Let S_N be the number of successes after N trials.

Then S_N/N goes to $1/\phi$ where $\phi = (1+\sqrt{5})/2$.

We expect $|S_N - N/\phi|$ to be on the order of $\sqrt{N}$ for large N.

Now switch to quasirandom walk, where all but finitely many rotors originally point to the right.

We add a bug at 1. It can be shown that the bug will be absorbed at either -1 or 0 after a finite number of steps.

We can continue to add more bugs, each of which will get absorbed at -1 or 0.

Let x_N be 1 if the Nth bug gets absorbed at -1, and 0 otherwise.

Let $s_N = x_1 + \ldots + x_N$ be number of successes (bugs absorbed at -1) after N trials.

Fact: $|s_N - N/\phi| < 1/\phi$ (which implies $s_N/N = 1/\phi + O(1/N)$).

Proof idea: The states of the rotors encode a number between 0 and 1 in base $r = -\phi$, using the digits

0 (Outward arrow)

and

1 (Inward arrow).

An Inward arrow at position i has value $1/r^i$.

The bug itself is the digit $1/r$; its value when at position i is $1/r^{i+1}$.

The value of the rotors plus the value of the bug is invariant under the operation of updating the rotor and sliding the bug.

Further properties:

1. When there is no bug in the system, the value of the system lies between -1 and $1/\phi$.
2. When a bug is added at 1 and wanders before exiting at 0, the overall change in the value of the system is $+1/\phi^2 + 1/\phi = +1$.
3. When a bug is added at 1 and wanders before exiting at -1, the overall change in the value of the system is $+1/\phi^2 - 1 = -1/\phi$.

So the value of the system always increases by 1 mod ϕ, and this nearly always uniquely specifies its new value, since the value of the system stays in $[-1, 1/\phi]$, an interval of length ϕ.

The boundedness of the discrepancy

$$|s_N - N/\phi|$$

follows immediately from the fact that the value of the system always stays in a bounded interval.

Recall that $s_N = x_1 + \ldots + x_N$, where x_k is 1 if the kth bug ends up at -1 and is 0 if the kth bug ends up at 0. The x_k's are quasirandomized versions of independent trials X_k.

Estimating $E(X)$ using $x_1, x_2, \ldots$ instead of $X_1, X_2, \ldots$ corresponds to the (classical) quasirandom method of computing the integral of f on I, where I is the interval $[-1, 1/\phi]$, $f : I \to R$ is the indicator function of $[0, 1/\phi]$, and the sample points $t_1, t_2, \ldots$ satisfy $t_N \equiv N \pmod{\phi}$.

Specifically, x_N equals $f(t_N)$.

A quasirandom walk in two dimensions

If a bug does random walk in Z^2 starting from (0,0), the chance that it will arrive at (1,1) before it ever returns to (0,0) is $p = \pi/8$.

If we do N trials, the number of successes divided by the number of trials should be close to $\pi/8$, with an error on the order of $1/\sqrt{N}$.

Equivalently, the number of successes minus $\pi/8$ times the number of trials (the discrepancy) should be on the order of $\sqrt{N}$ if we do independent random trials.

With rotor-routers, the discrepancy seems to be bounded.

The discrepancy $s_N - pN$ stays in the interval $[-.22, +3.39]$ for all N between 1 and 10,000.

This is much smaller than the discrepancy one would see for true random walk, which would be on the order of 50.

The interval $[-.22, +3.39]$ seems even more surprisingly narrow when one considers that for N between 1 and 5, the discrepancy $s_N - pN$ ranges between 0 and 2.963495. So the discrepancy interval doesn't grow much during the bug's next 9,995 visits to (0,0).

No proof of boundedness of $s_N - pN$ is known.

The Cooper-Spencer theorem

Put some bugs at even sites in Z^d (i.e., sites for which the sum of the coordinates is even), where the sites are equipped with rotors. Let each bug do one step of rotor-router walk (this is well-defined). Do this a total of T times.

Cooper and Spencer show that the difference between (1) the number of bugs at a site after T rounds of rotor-router walk, and (2) the expected number of bugs at a site after T rounds of random walk, is bounded by a constant C that doesn't depend on T, or on what the original distribution of bugs was, or which way the rotors were originally pointing. All it depends on is d, the dimension of the lattice.

See "Simulating a random walk with constant error", by Joshua Cooper and Joel Spencer:

arXiv:`math.CO/0402323`.

Cooper and Spencer also showed that in Z^d, the error over the hypercube $[1, L]^d$ is $O(L^{d-1} \ln L)$.

In particular, in Z^1 the error over an interval of length L is $O(\ln L)$.

So far, we've seen examples where quasirandomness gives sharp estimates for random walk: steady-state distribution, absorption probability, and distribution after T steps.

The same is true for expected time until absorption.

But now let's change gears and think about random aggregation instead of random walk.

Quasirandom aggregation in one dimension

Internal Diffusion-Limited Aggregation (IDLA): To add a new bug to the (initially empty) blob, put the bug at the origin and let it do random walk until it hits an unoccupied site. Adjoin this site to the blob. Repeat.

Modify the rule so that when a bug finds a vacancy at location $x < 0$, both sites x and $x - 1$ get adjoined to the blob.

Say the blob at time t is $[-x(t), y(t)]$.

True randomness: $x/y \to \sqrt{2}$.

Quasirandomness with rotors:

Theorem (Levine): $|x_n - y_n\sqrt{2}|$ is bounded.

Quasirandom aggregation in two dimensions

Theorem (Bramson, Griffeath, and Lawler): The N-bug IDLA blob in two dimensions is round to within radial fluctuation that are $O(N^{1/3})$.

Theorem (Blachère): Can replace $O(N^{1/3})$ by $O(\ln N)$.

Rotor version: After a million bugs have been added to the system, the inradius is 563.5 and the outradius is 565.1: they differ by 1.6 (about three tenths of one percent).

Difference between inradius and circumradius seems to be bounded.

No proof that **ratio** between inradius and circumradius is bounded.

Related work

- Derandomization of Monte Carlo integration via quasirandomness (1960s?)
- Arthur Engel's "probabilistic abacus" (1975, 1976)
- Sandpiles and avalanches (Bak, Tang, and Wiesenfeld, 1988; Dhar, 1990)
- Chip-firing (Anderson, Lovász, Shor, Spencer, Tardos, and Winograd, 1989)
- Eulerian walkers (Priezzhev, Dhar, Dhar, and Krishnamurthy, 1996)
- Balancers and balancing circuits (Rabani, Sinclair, and Wanka, 1998)

Problems

Theory lags behind observation (sometimes quasirandomness appears to be extremely good but we can't prove it yet)

"Quasirandomization" is not a clearly defined procedure

Current methods apply only to first-order estimates (e.g., mean, not variance)

Fancier forms of quasirandomness than $+1, -1, +1, -1, \ldots$?:

Does there exist an infinite sequence $(a_k)_{k=1}^{\infty}$ of $+1$'s and -1's such that the partial sums

$$\sum_{k=1}^{N} a_k$$

are bounded and such that for every positive integer r, the partial sums

$$\sum_{k=1}^{N} a_k a_{k+r}$$

are all bounded? (The bound may depend on r, but not on N.)

Does there exist an infinite sequence $(a_k)_{k=1}^{\infty}$ of $+1$'s and -1's such that the partial sums

$$\sum_{k=1}^{N} a_k$$

are bounded and such that for every complex number z with $|z| = 1$ (or maybe just every root of unity), the partial sums

$$\sum_{k=1}^{N} a_k z^k$$

are all bounded? (The bound may depend on z, but not on N.)

Quasirandomness with different orders of discrepancy:

Instead of rotors (whose discrepancy is $O(1)$), we could use deterministic sequences whose discrepancy is $O(N^\alpha)$ or $o(N^\alpha)$ for some particular α between 0 and 1.

($\alpha = 1/2$ is a natural choice, if one is hoping to find some still-unknown kind of quasirandom bit string that will satisfy the central limit theorem.)

For each such notion of discrepancy, there are theorems that say that if you use low-discrepancy quasirandomness at a local level, the macroscopic effects that are produced by the model will agree (to within low discrepancy) with the average-case behavior for true randomness.

Summary

Computer random-number generators are non-random anyway!

Why settle for one-size-fits-all pseudo-randomness, if you can have the perfect kind of quasirandomness for your model?

When available, quasirandom methods typically:

- are fast;
- give small errors;
- give deterministic bounds on errors; and
- can be *proved* to have those three properties.

For more information:

Email: propp@math.wisc.edu

Video etc.: http://murl.microsoft.com/LectureDetails.asp?1050

Applets: http://www.math.wisc.edu/~propp/rotor-router-1.0/

Probability seminar: 9/22 and 9/29, 334 Evans, 3:10 – 4:00 pm

Michael Kleber's upcoming article in the Winter 2005 issue of The Mathematical Intelligencer