The Monte Carlo Method

An Introduction to Markov Chain MC

Sachin Shanbhag

Department of Scientific Computing, Florida State University

Sources

1. “Introduction to Monte Carlo Algorithms”: Werner Krauth. Available as a pdf download at http://lps.ens.fr/~krauth. This is a great introduction to MC methods, geared towards applications in physical sciences, through puzzles and simple non-technical problems.
2. David Kofke’s page on Molecular Simulation.

Markov Chains

Transition-Probability Matrix

Example

–system with three states

If in state 1, will stay in state 1 with probability 0.1

If in state 1, will move to state 3 with probability 0.4

$$\Pi \equiv \begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \\ \pi_{31} & \pi_{32} & \pi_{33} \end{pmatrix} = \begin{pmatrix} 0.1 & 0.5 & 0.4 \\ 0.9 & 0.1 & 0.0 \\ 0.3 & 0.3 & 0.4 \end{pmatrix}$$

Never go to state 3 from state 2

Requirements of transition-probability matrix

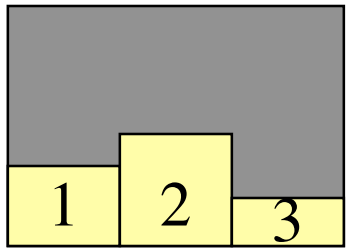
–all probabilities non-negative, and no greater than unity

–sum of each row is unity

–probability of staying in present state may be non-zero

Distribution of State Occupancies

- Consider process of repeatedly moving from one state to the next, choosing each subsequent state according to Π
  - $1 \rightarrow 2 \rightarrow 2 \rightarrow 1 \rightarrow 3 \rightarrow 2 \rightarrow 2 \rightarrow 3 \rightarrow 3 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow$ etc.
- Histogram the occupancy number for each state
  - $n_1 = 3$ $p_1 = 0.33$
  - $n_2 = 5$ $p_2 = 0.42$
  - $n_3 = 4$ $p_3 = 0.25$



- After very many steps, a limiting distribution emerges

The Limiting Distribution

- Consider the product of Π with itself

$$\Pi^2 \equiv \begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \\ \pi_{31} & \pi_{32} & \pi_{33} \end{pmatrix} \times \begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \\ \pi_{31} & \pi_{32} & \pi_{33} \end{pmatrix}$$

All ways of going from state 1 to state 2 in two steps

$$= \begin{pmatrix} \pi_{11}\pi_{11} + \pi_{12}\pi_{21} + \pi_{13}\pi_{31} & \pi_{11}\pi_{12} + \pi_{12}\pi_{22} + \pi_{13}\pi_{32} & etc. \\ \pi_{21}\pi_{11} + \pi_{22}\pi_{21} + \pi_{23}\pi_{31} & \pi_{21}\pi_{12} + \pi_{22}\pi_{22} + \pi_{23}\pi_{32} & etc. \\ \pi_{31}\pi_{11} + \pi_{32}\pi_{21} + \pi_{33}\pi_{31} & \pi_{31}\pi_{12} + \pi_{32}\pi_{22} + \pi_{33}\pi_{32} & etc. \end{pmatrix}$$

Probability of going from state 3 to state 2 in two steps

- In general Π^n is the n-step transition probability matrix
  - probabilities of going from state *i* to *j* in exactly n steps

$$\Pi^n \equiv \begin{pmatrix} \pi_{11}^{(n)} & \pi_{12}^{(n)} & \pi_{13}^{(n)} \\ \pi_{21}^{(n)} & \pi_{22}^{(n)} & \pi_{23}^{(n)} \\ \pi_{31}^{(n)} & \pi_{32}^{(n)} & \pi_{33}^{(n)} \end{pmatrix} \quad \text{defines } \pi_{ij}^{(n)}$$

The Limiting Distribution

- Define $\pi_i^{(0)}$ as a unit state vector

$$\pi_1^{(0)} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \quad \pi_2^{(0)} = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \quad \pi_3^{(0)} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$$

- Then $\pi_i^{(n)} \equiv \pi_i^{(0)} \Pi^n$ is a vector of probabilities for ending at each state after *n* steps if beginning at state *i*

$$\pi_1^{(n)} = \pi_1^{(0)} \Pi^n \equiv \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \pi_{11}^{(n)} & \pi_{12}^{(n)} & \pi_{13}^{(n)} \\ \pi_{21}^{(n)} & \pi_{22}^{(n)} & \pi_{23}^{(n)} \\ \pi_{31}^{(n)} & \pi_{32}^{(n)} & \pi_{33}^{(n)} \end{pmatrix} = \begin{pmatrix} \pi_{11}^{(n)} & \pi_{12}^{(n)} & \pi_{13}^{(n)} \end{pmatrix}$$

- The limiting distribution corresponds to $n \rightarrow \infty$
  - independent of initial state $\pi_1^{(\infty)} = \pi_2^{(\infty)} = \pi_3^{(\infty)} \equiv \pi$

The Limiting Distribution

- Stationary property of π

$$\begin{aligned}\pi &= \lim_{n\to\infty}\left[\pi_i^{(0)}\Pi^n\right] \\ &= \left(\lim_{n\to\infty}\left[\pi_i^{(0)}\Pi^{n-1}\right]\right)\Pi \\ &= \pi\Pi\end{aligned}$$

- π is a left eigenvector of Π with unit eigenvalue
  - such an eigenvector is guaranteed to exist for matrices with rows that each sum to unity
- Equation for elements of limiting distribution π

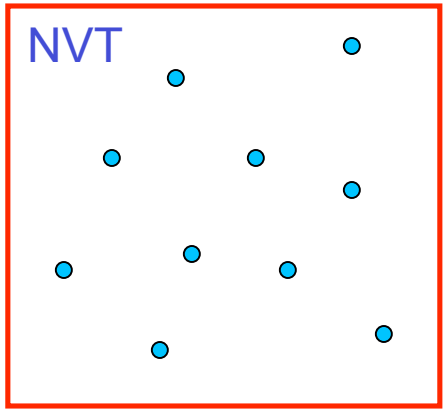
$$\pi_i = \sum_j \pi_j \pi_{ji}$$

$$e.g.\quad \Pi = \begin{pmatrix} 0.1 & 0.5 & 0.4 \\ 0.9 & 0.1 & 0.0 \\ 0.3 & 0.3 & 0.4 \end{pmatrix}$$

$$\begin{aligned}\pi_1 &= 0.1\pi_1 + 0.9\pi_2 + 0.3\pi_3 \\ \pi_2 &= 0.5\pi_1 + 0.1\pi_2 + 0.3\pi_3 \\ \pi_3 &= 0.4\pi_1 + 0.0\pi_2 + 0.4\pi_3 \\ \pi_1 + \pi_2 + \pi_3 &= \pi_1 + \pi_2 + \pi_3\end{aligned}$$

not independent

Statistical Mechanics



N particles,
position $\mathbf{r}_i = (x,y,z)$
momentum $\mathbf{p}_i = (p_x, p_y, p_z)$

$\{\mathbf{r}_i, \mathbf{p}_i\}$ = 6N degrees of freedom

A particular snapshot of $\{\mathbf{r}_i, \mathbf{p}_i\}$ is called a *microstate*

Total energy of the system

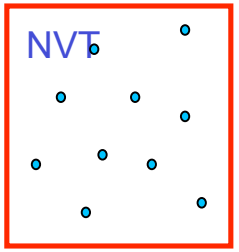
$H(\{\mathbf{r}_i, \mathbf{p}_i\}) = U(\{\mathbf{r}_i\}) + K(\{\mathbf{p}_i\})$

$U = U(\{\mathbf{r}_i\})$ potential energy (often pairwise summation)

$K = \Sigma\, \mathbf{p}_i^2/2m_i$ kinetic energy

Statistical Mechanics

In an NVT ensemble,



$$\pi(\mathbf{r}_i,\mathbf{p}_i) = \frac{1}{Q}\exp(-\frac{H(\mathbf{r}_i,\mathbf{p}_i)}{kT})$$

where Q = classical partition function, defined by

$$Q = \frac{1}{h^{3N}N!}\int d\mathbf{p}_1 \int d\mathbf{p}_2 \cdots \int d\mathbf{p}_N \int d\mathbf{r}_1 \int d\mathbf{r}_2 \cdots \int d\mathbf{r}_N \exp(-\frac{H(\mathbf{r}_i,\mathbf{p}_i)}{kT})$$

this is like a effective volume of the phase space

It is related to the free energy via A/kT = - ln Q(N,V,T)

An average property (e.g. specific heat) may be obtained via,

$$\langle A \rangle = \int d\mathbf{p}^N \int d\mathbf{r}^N A(\mathbf{r}_i,\mathbf{p}_i)\pi(\mathbf{r}_i,\mathbf{p}_i)$$

Statistical Mechanics

The kinetic energy factors can be integrated out,

$$Q = \frac{1}{h^{3N} N!} \int d\mathbf{p}^N \int d\mathbf{r}^N \exp\left(-\frac{U(\{\mathbf{r}_i\}) + K(\{\mathbf{p}_i\})}{kT}\right)$$

$$Q = \frac{1}{h^{3N} N!} \int d\mathbf{r}^N \exp\left(-\frac{U(\{\mathbf{r}_i\})}{kT}\right) \int d\mathbf{p}^N \sum \frac{p_i^2}{2m_i}$$

Z_N = configurational integral

$$Q = \frac{Z_N}{\Lambda^{3N}}$$

It is easy to show that,

$$\langle A \rangle = \frac{1}{Z_N N!} \int d\mathbf{r}^N A(\mathbf{r}_i) \exp\left(-\frac{U}{kT}\right)$$

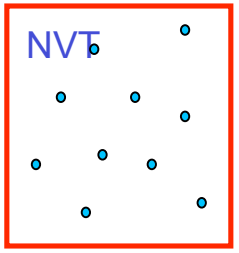
Why do we care?

In MC, unlike MD, we don't typically have access to momenta (only positions).

Statistical Mechanics

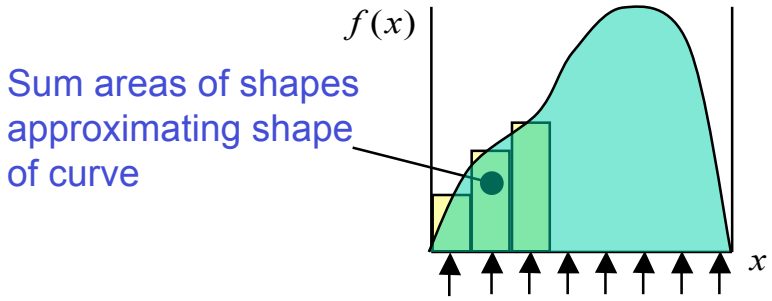
This is an integration problem, I can use quadrature!

$$\langle A \rangle = \frac{1}{Z_N N!} \int d\mathbf{r}^N A(\mathbf{r}_i) \exp(-\frac{U}{kT})$$



One dimensional integrals

$$I = \int_a^b f(x)dx$$



$$I \approx \Delta x \sum_{i=1}^{n} f(x_i) = \frac{b-a}{n} \sum_{i=1}^{n} f(x_i)$$

Error $\delta I \sim 1/n^2$ (for 1D)

$\delta I \sim 1/n^{2/D}$ (for dD)

Where is Monte Carlo?

This is an integration problem, I can use quadrature!

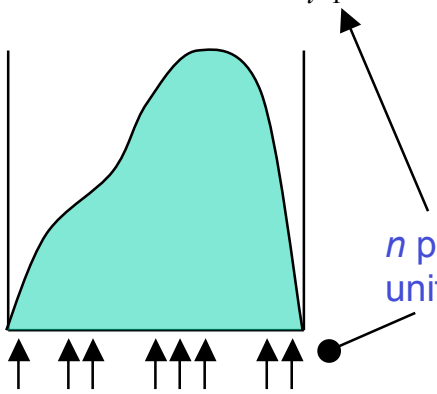
$$\langle A \rangle = \frac{1}{Z_N N!} \int d\mathbf{r}^N A(\mathbf{r}_i) \exp(-\frac{U}{kT})$$

One dimensional integrals

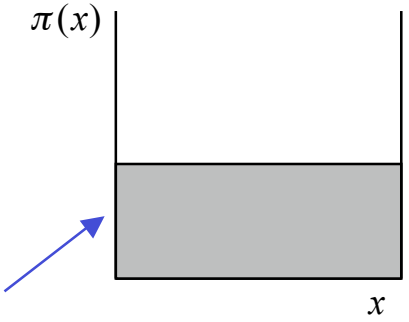
$$I = \int_a^b f(x)dx$$

Error $\delta I \sim 1/n^{0.5}$ (for 1D)
$\delta I \sim 1/n^{0.5}$ (for dD)

$$I \approx \frac{b-a}{n} \sum_{i=1}^{n} f(x_i)$$



n points selected from uniform distribution p(x)



For large dimensions, MC has better convergence

Where is Monte Carlo?

$$\langle A \rangle = \frac{1}{Z_N N!} \int d\mathbf{r}^N A(\mathbf{r}_i) \exp(-\frac{U}{kT})$$

If each function evaluation is 1 flop (lower-bound), need 10^{300} function evaluations to get integral.

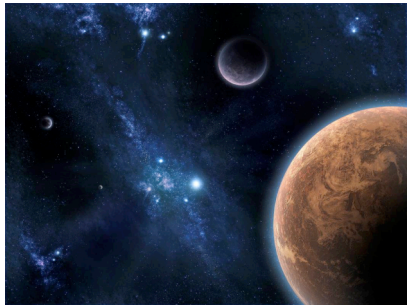
Typically, 3D space, say

N=100 particles, DOF = 3N = 300

n=10 points between (-L, +L) (dx_1, for example)

Fastest computers (IBM BlueGene/L ~ 500 Tflops)

Time required = $10^{300}/5*10^{14} \sim 10^{278}$ years



age of the universe 13 billion years

IMPOSSIBLE

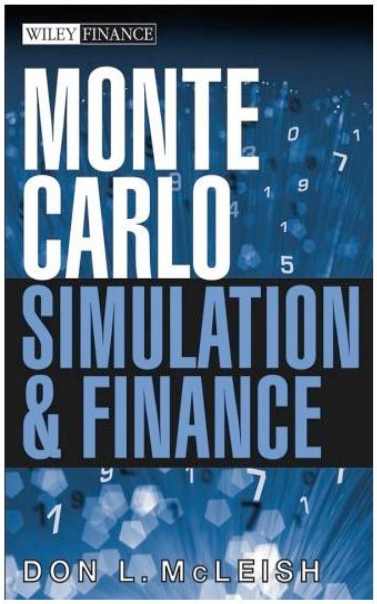
The Etymology of “Monte Carlo”

Principality of Monaco (~ 1 mile^2 - UM)

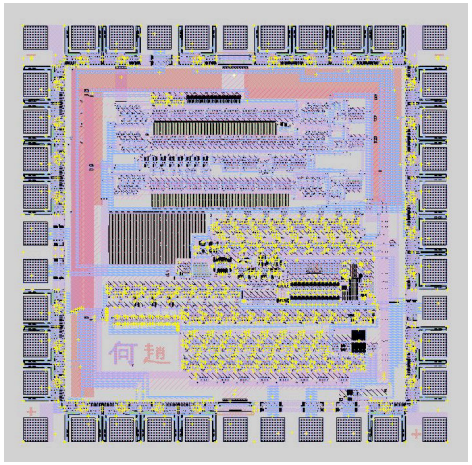
Famous for casinos, Formula-1 races, royalty

Random numbers

The Ubiquity of Monte Carlo… Random Examples

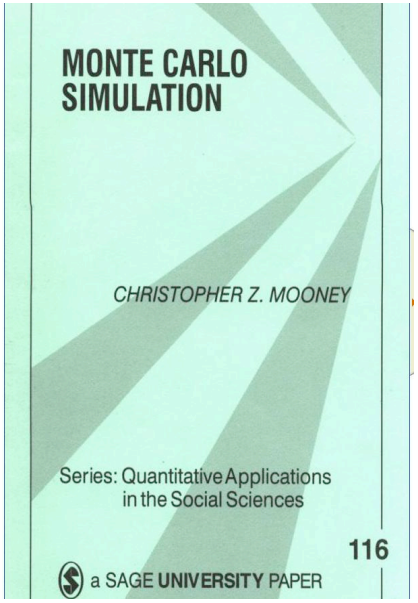


Finance and Social Sciences

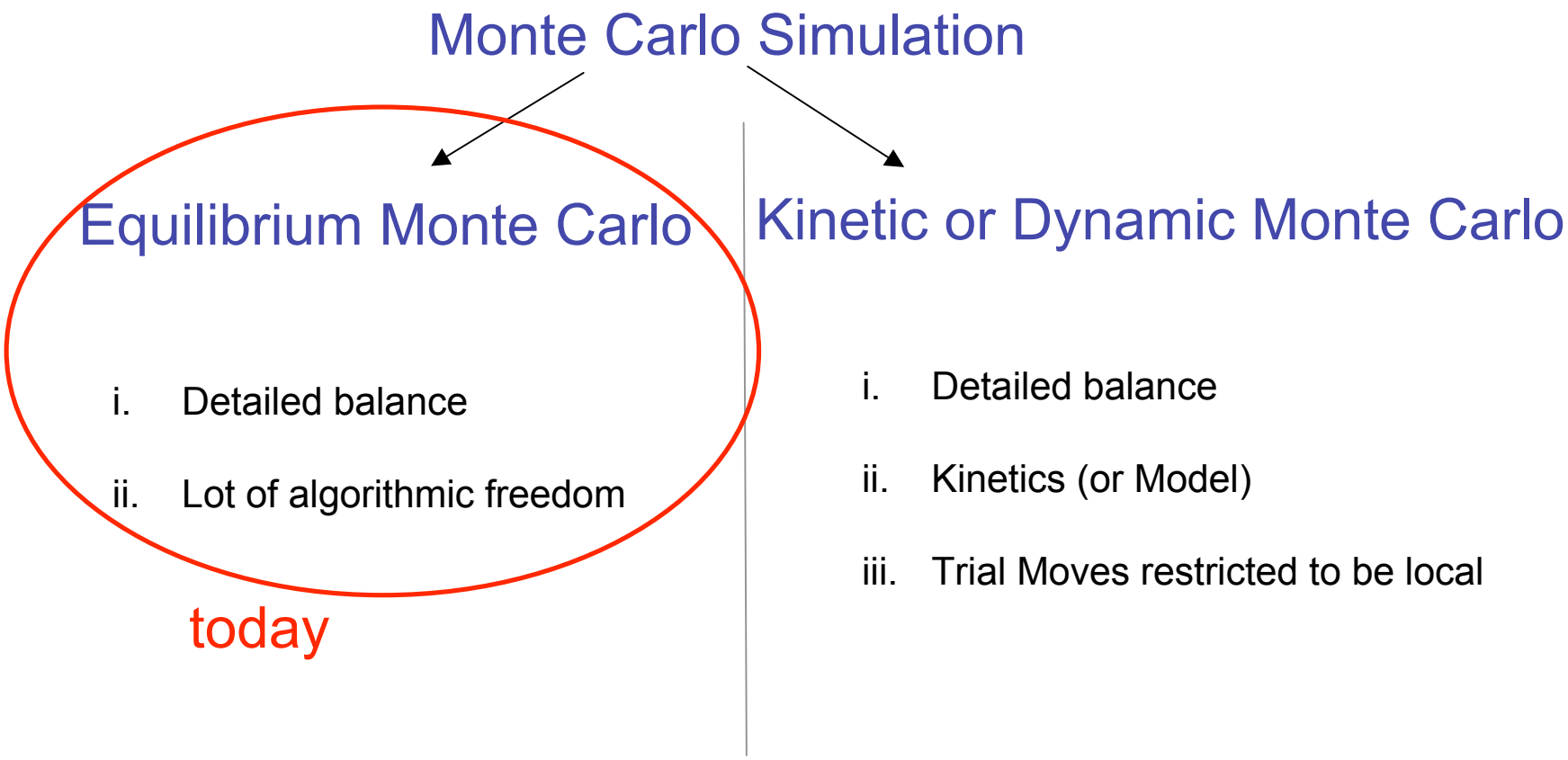


Chip Design and Newpaper Panels

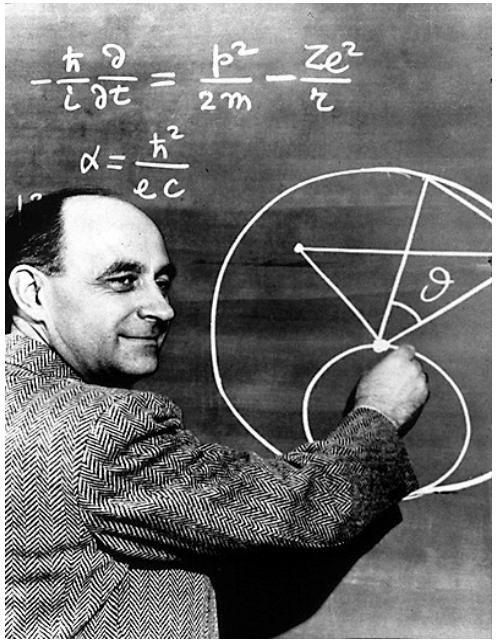
"floorplanning"



Broadly Speaking…



The Tradition of Monte Carlo

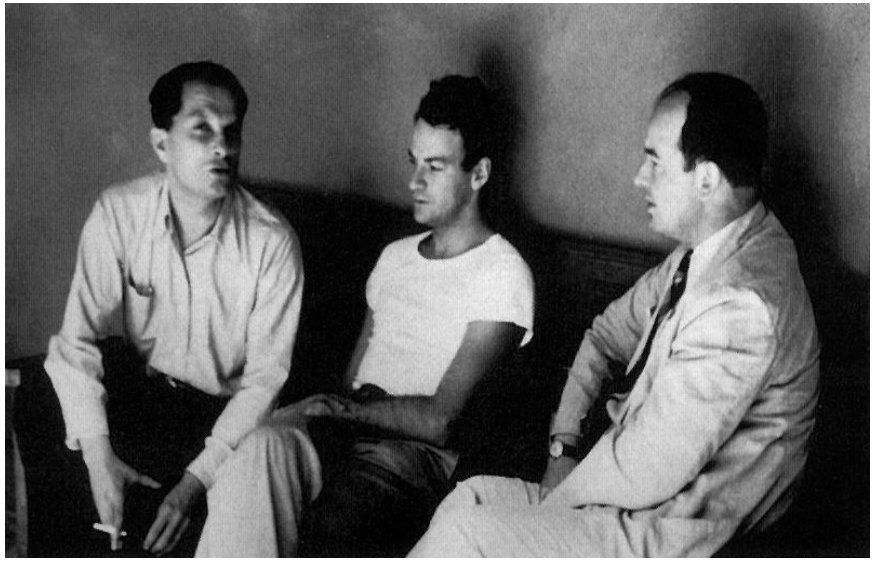


Fermi

(1930-1950) various attempts

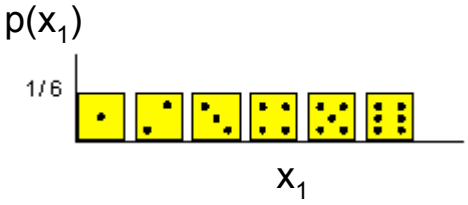
The Manhattan Project

Ulam, Feynman, von Neumann

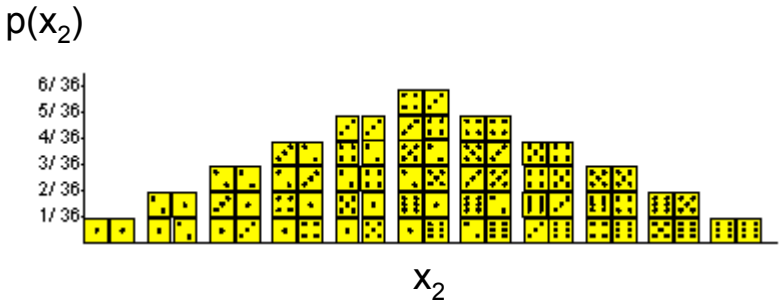


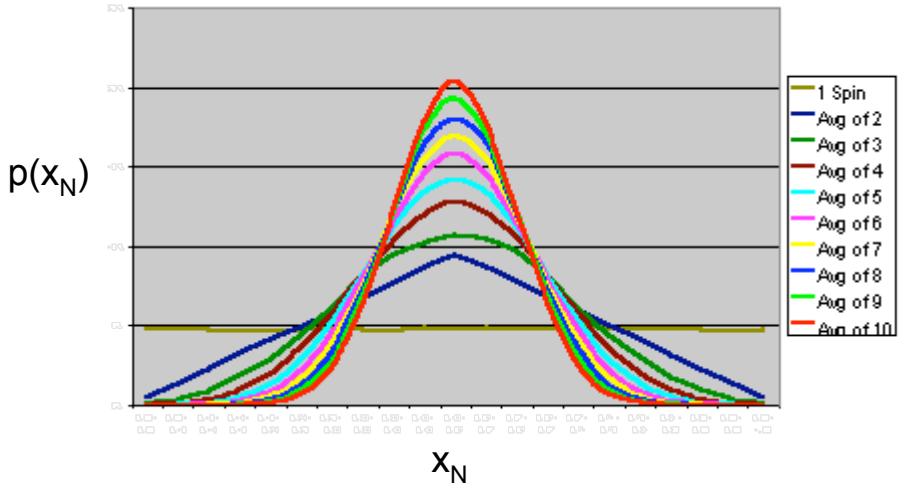
The Miracle of the Central Limit Theorem

$x_1 =$ 



$x_2 =$  $\div 2$





The probability distribution of the average of independent random variables becomes increasingly Gaussian and increasingly narrower

sigma ~ 1/sqrt(N)

Children's game in Monaco

Mathematically oriented kids of Monaco want to find the "pi"

$N_{red}/(N_{red}+N_{blue})$

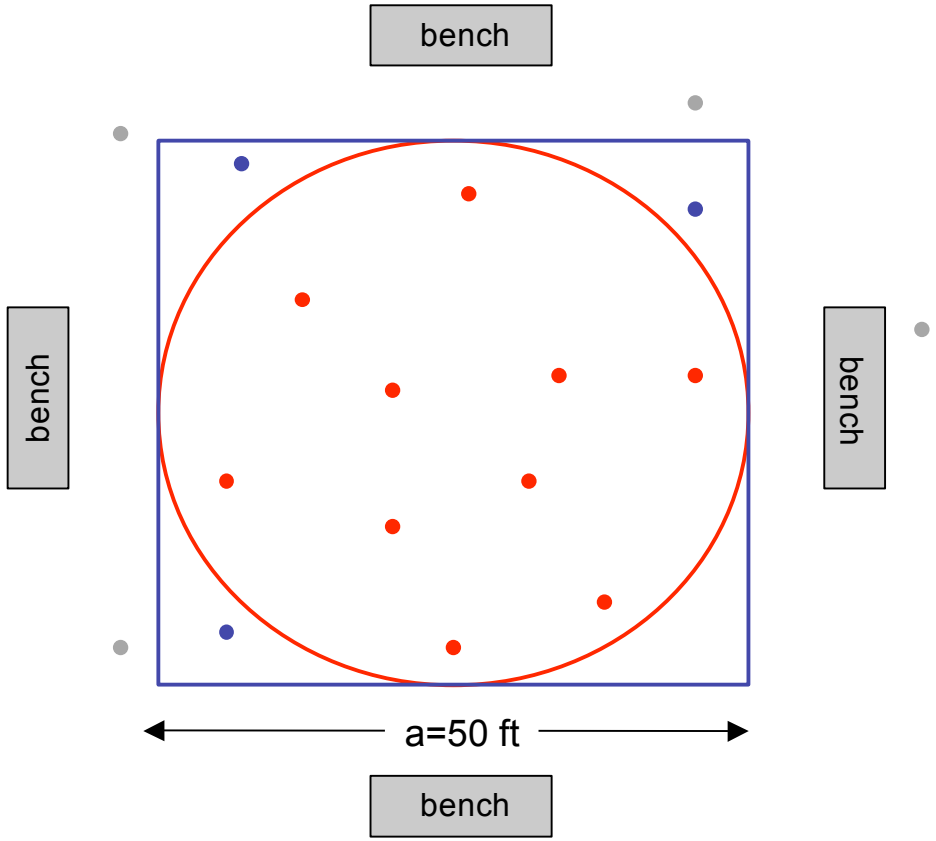
$= A_{cir}/A_{sq}$

$=pi*(a/2)^2/a^2$

$= pi/4$

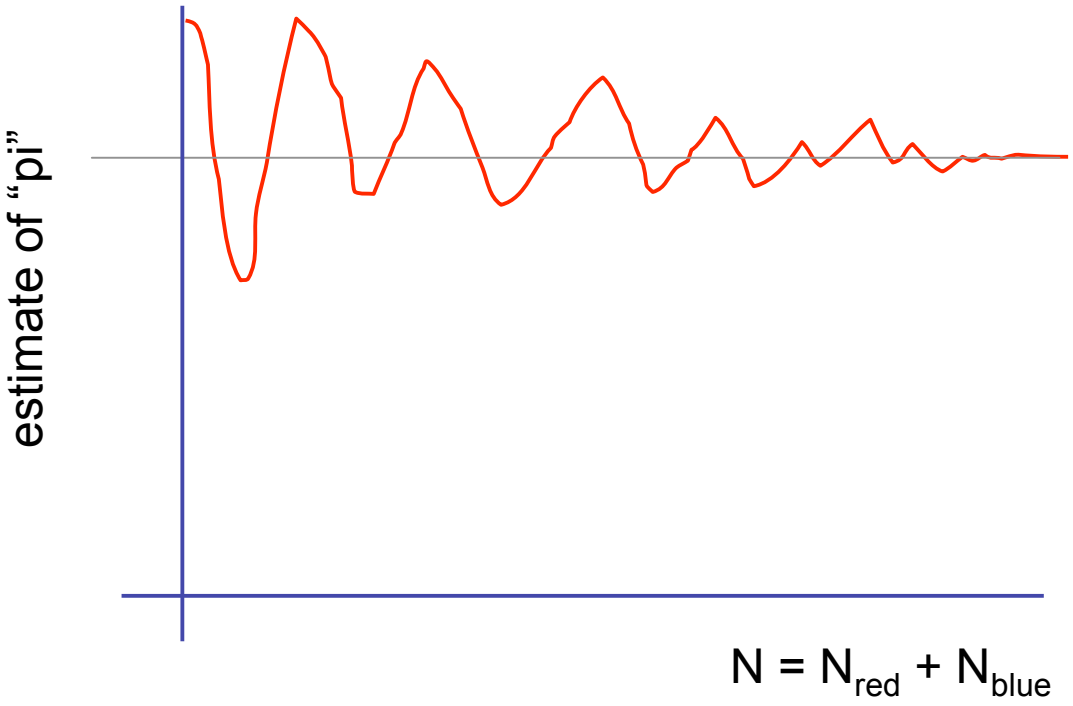
$pi = 4 * N_{red}/(N_{red}+N_{blue})$

$N = N_{red} + N_{blue}$



*Here, pi = 4 *10/13 ~ 3.08*

Children's game in Monaco



The invisible hand of the CLT operates to ensure that we converge to the true value of "pi" as the number of rocks thrown is increased. "Error" vanishes as 1/sqrt(N)

Older people's game in Monaco

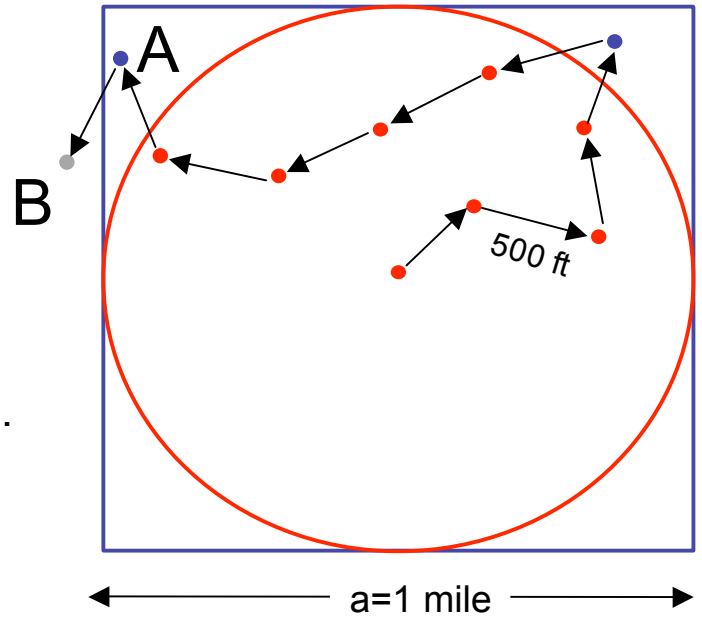
"Time" spent planting flags in the circular region is proportional to the area of the circle

CHOICES

(i) Move to "B". Plant a gray flag. Continue throwing.
(ii) Stay at "A". **Don't** plant flag. Throw again.
(iii) Stay at "A". Plant a flag. Continue throwing.

Many MC simulations are wrong because they make the wrong choice.

plant flags on a helipad



$pi = 4 * N_{red}/(N_{red}+N_{blue})$

Condition of Detailed Balance will give us sound advice.

Some Concepts

KID'S GAME

Each sample (place where throw lands) independent

Direct Sampling

Converges faster (Think CLT, think independent samples)

May not be feasible for realistic problems!

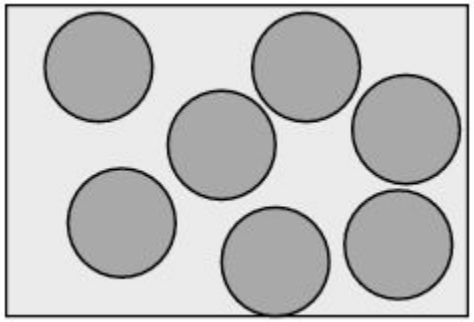
OLDER PEOPLE'S VERSION

Sample dependent on previous position (is within radius of 500ft)

Markov-chain Sampling

Need more throws (samples) to attain same level of accuracy

Coins in a Shoebox Problem



Q: How do you generate random configurations of coins such that they don't overlap?

Direct Sampling

Random Sequential Depositon (?)

WRONG!

$phi_{RSD} < phi_{CP}$

Physicists (since 1953) haven't found a direct sampling algorithm

Markov-chain Sampling

Start with a legit configuration

Trial: Move a randomly chosen coin just a little in a randomly chosen direction

Overlap?

yes

no

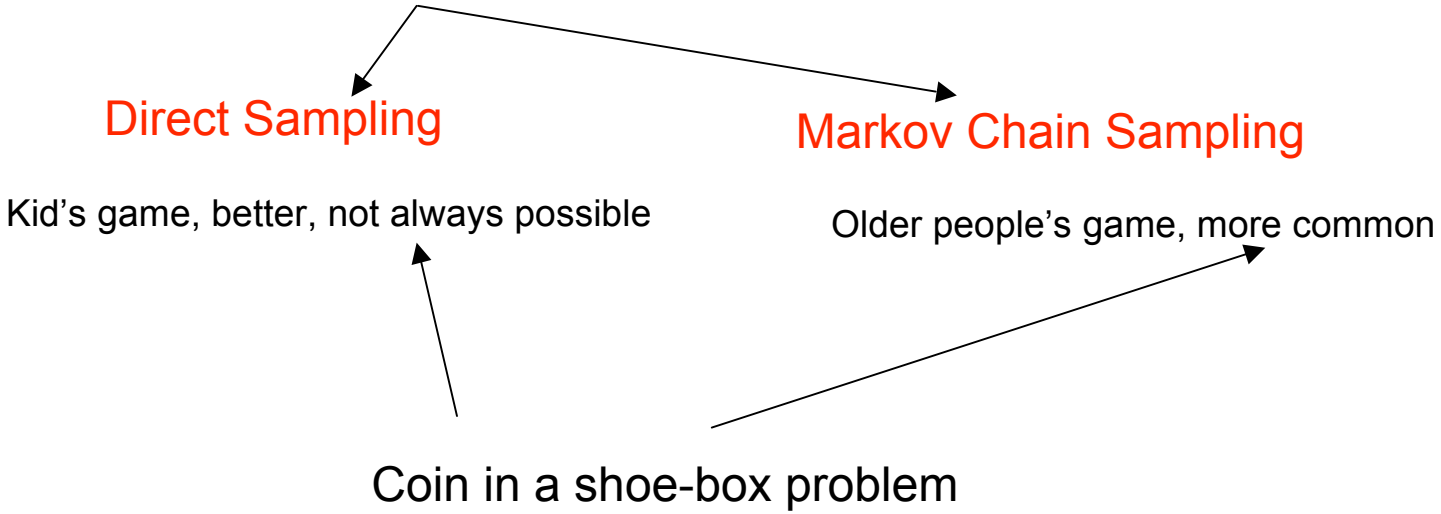
Reject move

Enact move

So far…

Central Limit Theorem: The Engine

Sampling: How do I generate trials?



Unsolved Mystery: What should you do if you throw the stone out of bounds in the adults' game? Tricky question, full of consequences for world peace and prosperity. We said upcoming "condition of detailed balance" will resolve it.

Scrambling Game

4	2	8
1		7
6	3	5

Objective: *Create a perfect scrambling algorithm such that any possible configuration of the puzzle is generated with equal probability*

"equilibrium" where each state is equi-probable.

Toddler's algorithm

Rip the whole thing apart

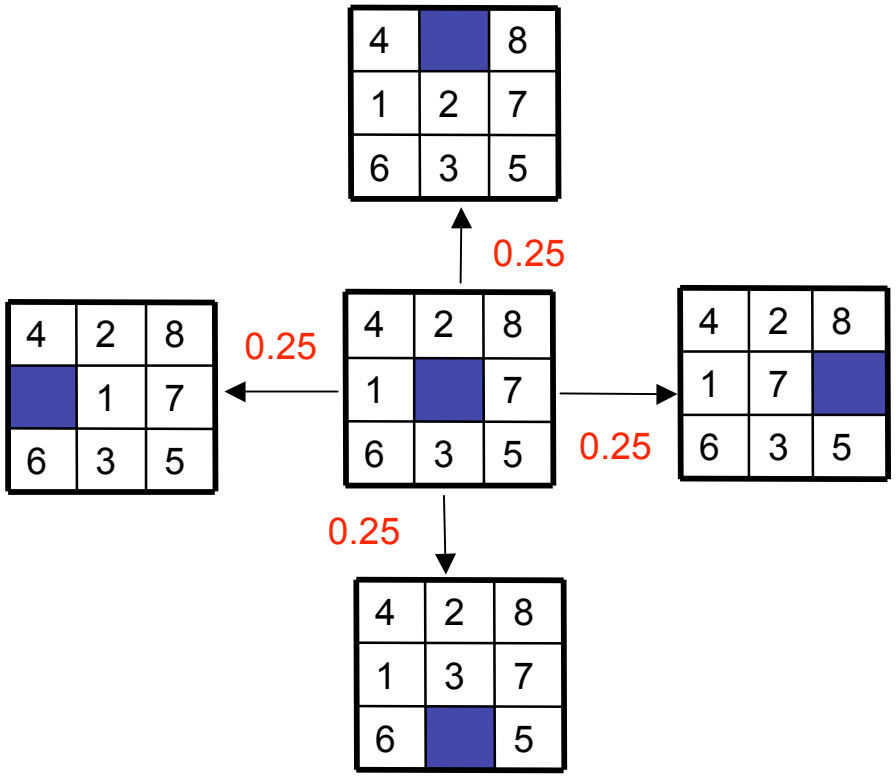
Put it back together

Scrambling Game

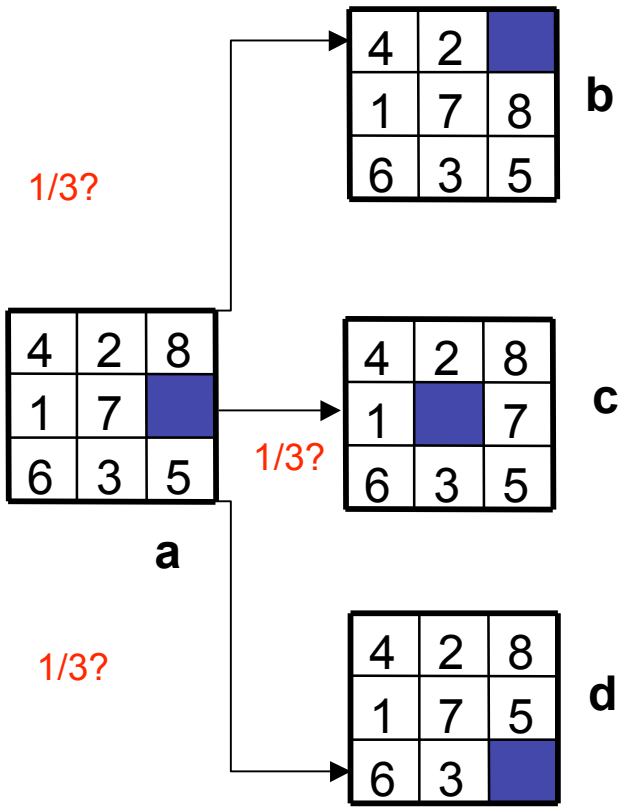
Older People Version

Move the empty square in a random direction, each timestep

Choose central square to go (N, E, W, S) each with probability ¼ (*which is indeed correct*)



Scrambling Game



$F(a) = F(b) = F(c) = F(d)$: at "equilibrium"

$$F(a) = F(a)\,p(a\to a) + F(b)\,p(b\to a) + F(c)\,p(c\to a) + F(d)\,p(d\to a)$$

$$[1 - p(a\to a)]F(a) = F(b)\,p(b\to a) + F(c)\,p(c\to a) + F(d)\,p(d\to a)$$

$$p(a\to a) + p(a\to b) + p(a\to c) + p(a\to d) = 1$$

$$p(a\to b) + p(a\to c) + p(a\to d) = 1 - p(a\to a)$$

$$[p(a\to b) + p(a\to c) + p(a\to d)]F(a) = F(b)p(b\to a) + F(c)\,p(c\to a) + F(d)\,p(d\to a)$$

Scrambling Game

$$[p(a\to b) + p(a\to c) + p(a\to d)\,]F(a) = F(b)p(b\to a) + F(c)\,p(c\to a) + F(d)\,p(d\to a)$$

$$\begin{aligned} &F(a)\,p(a\to b) + && F(b)\,p(b\to a) + \\ &F(a)\,p(a\to c) + = && F(c)\,p(c\to a) + \\ &F(a)\,p(a\to d) && F(d)\,p(d\to a) \end{aligned} \qquad \text{(A)}$$

Impose a condition that terms of the same color on either side of the "=" are equal

$$\begin{aligned} F(a)\,p(a\to b) &= F(b)\,p(b\to a) \\ F(a)\,p(a\to c) &= F(c)\,p(c\to a) \\ F(a)\,p(a\to d) &= F(d)\,p(d\to a) \end{aligned}$$

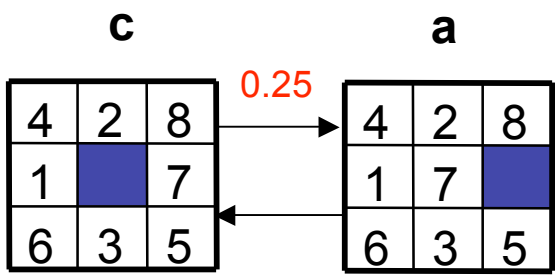
This "recipe" for satisfying equation (A) is called the ***condition of detailed balance***

Sufficient, but not necessary condition

Scrambling Game

Look at one of these carefully,

$p(a \rightarrow c)/p(c \rightarrow a) = F(c)/F(a)$



Given: F(a) = F(c)

$p(a \rightarrow c) = p(c \rightarrow a)$
$= 0.25$ (not 1/3)

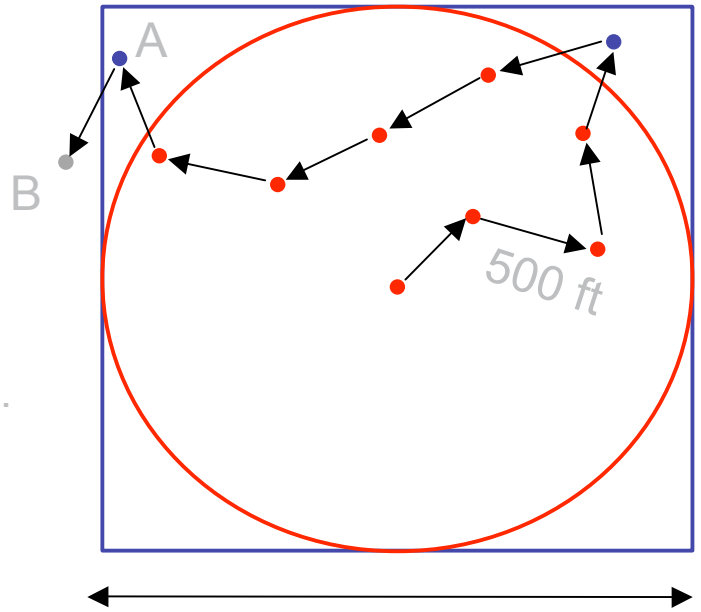
Therefore algorithm is:

(i) Pick {N, E, W, S} with equal probability
(ii) Move blank square to corresponding direction, if possible. Otherwise, stay where you are, *reject* the move, and *advance* the clock.

Older people's game in Monaco

"Time" spent planting flags in the circular region is proportional to the area of the circle

plant flags on a helipad



CHOICES

(i) Move to "B". Plant a gray flag. Continue throwing.
(ii) Stay at "A". **Don't** plant flag. Throw again.
(iii) Stay at "A". Plant a flag. Continue throwing.

Many MC simulations are wrong because they make the wrong choice.

Condition of Detailed Balance will give us wise advice.

$pi = 4 * N_{red}/(N_{red}+N_{blue})$

Some More Concepts

Concept of **REJECTION** arose from the condition of detailed balance

$$p(a \rightarrow c)/p(c \rightarrow a) = F(c)/F(a)$$

This method of rejection has been enshrined into the "Metropolis Algorithm"

$$p(a \rightarrow c) = \min[\, 1, F(c)/F(a)]$$

Connection with Statistical Mechanics

$$F(c) = \exp(-U_c/k_B T)$$

$$p(a \rightarrow c) = \min[\, 1, \exp[-(U_c - U_a)/k_B T]]$$

Boltzmann

The Metropolis Method

(over 7500 citations between 1988-2003)

THE JOURNAL OF CHEMICAL PHYSICS VOLUME 21, NUMBER 6 JUNE, 1953

Equation of State Calculations by Fast Computing Machines

NICHOLAS METROPOLIS, ARIANNA W. ROSENBLUTH, MARSHALL N. ROSENBLUTH, AND AUGUSTA H. TELLER,

Los Alamos Scientific Laboratory, Los Alamos, New Mexico

AND

EDWARD TELLER, * *Department of Physics, University of Chicago, Chicag*

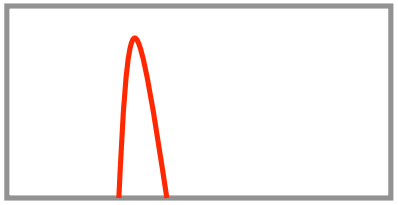
(Received March 6, 1953)

A general method, suitable for fast computing machines, for investigating such p
of state for substances consisting of interacting individual molecules is described.
of a modified Monte Carlo integration over configuration space. Results for the two
sphere system have been obtained on the Los Alamos MANIAC and are presente
are compared to the free volume equation of state and to a four-term virial coefficie

Nicholas Metropolis with the MANIAC

Back to Statistical Mechanics

$$\langle A \rangle = \frac{1}{Z_N N!} \int d\mathbf{r}^N A(\mathbf{r}_i) \exp(-\frac{U}{kT})$$



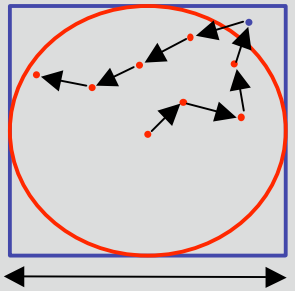
This function is sharply peaked.

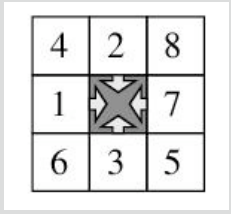
Observation 1

Observation 2

It may be possible to compute <A> without evaluating Z_N?

In the area of the circle problem, we evaluated an area (like a partition function), and the probability distribution was uniform (obs 1 and 2 inapplicable).



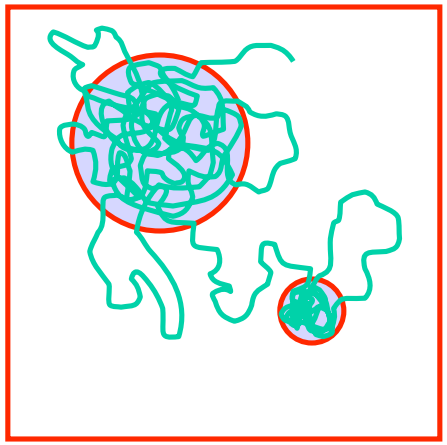


In the scramble problem, the probability distribution was uniform (obs 1 inapplicable)

Back to Statistical Mechanics

$$\langle A \rangle = \frac{1}{Z_N N!} \int d\mathbf{r}^N A(\mathbf{r}_i) \exp(-\frac{U}{kT})$$

Metropolis helps us generate new microstates with the appropriate probability. Thus a simple average over all the microstates gives us <A>



Samples the important region of phase space

Detailed Balance

- Eigenvector equation for limiting distribution
  - $\pi_i = \sum_j \pi_j \pi_{ji}$
- A sufficient (but not necessary) condition for solution is
  - $\pi_i \pi_{ij} = \pi_j \pi_{ji}$
  - "detailed balance" or "microscopic reversibility"
- Thus
  - $$\pi_i = \sum_j \pi_j \pi_{ji} = \sum_j \pi_i \pi_{ij} = \pi_i \sum_j \pi_{ij} = \pi_i$$

$$\Pi = \begin{pmatrix} 0.1 & 0.5 & 0.4 \\ 0.9 & 0.1 & 0.0 \\ 0.3 & 0.3 & 0.4 \end{pmatrix}$$

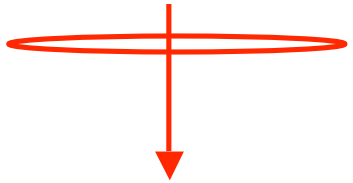
For a given Π, it is not always possible to satisfy detailed balance; e.g. for this Π $\quad \pi_3 \pi_{32} \neq \pi_2 \pi_{23}$

zero

Tutorial: Ising Model

1924: Ising solved 1D
1944: Onsager solved 2D

Simple model for understanding magnetism



Unpaired electrons = origin of magnetism

Ferromagnets: Fe, Ni - magnetic even in the absence of an electric field

Paramagnets: Al, Ba - magnetic only in the presence of an electric field

Up-state

Down state

Tutorial: Ising Model

(a) Energy

$$E(\{\sigma\}) = -J\sum \sigma_i \sigma_j$$

(b) Magnetization

$$M(\{\sigma\}) = \frac{\sum \sigma_i}{N}$$

Tutorial: Markov Chains

Tutorial: Markov Chains