الگوریتم متروپلیس (Metropolis)

الگوریتم متروپلیس در سال 1953 توسط متروپلیس (Metropolis)، روزنبلوث (Rosenbluth) و تلر (Teller) معرفی شد.
این الگوریتم جزء 10 الگوریتم برتر قرن ۲۰ در حیطه علمی و مهندسی است.
آلن دریک جمله بخواهیم الگوریتم متروپلیس را توضیح دهیم اینگونه توضیح می دهیم:

" الگوریتم متروپلیس روش پیمایشی نمونه گیری با اهمیت در کامپیوتر است "

↳ Implementation

برای اینکه بیشتر با این الگوریتم آشنا شویم، بیایید با مثال قبلی برویم، یعنی انتگرال زیر

$$F = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-x^2/2} dx = 1$$

در این انتگرال $p(x)$ را به صورت زیر معرفی کردیم:

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \propto e^{-x^2/2}$$

و تابع $g(x)$:

$$g(x) = x^2$$

به این انتگرال به صورت زیر خواهیم شد:

$$F = \int g(x) p(x) dx$$

که طبق روش نمونه گیری با اهمیت

$$F \simeq \frac{1}{n} \sum_{i=1}^{n} g(x_i)$$

که x_i ها طبق توزیع $P(x)$ انتخاب می‌شوند.

در واقع آنچه در بخش قبل گفته شد، استفاده از توزیع آماده $P(x) \propto e^{-\frac{x^2}{2}}$ برای محاسبه انتگرال بود. حال بیایم خود این توزیع را به روش زیر (که همان الگوریتم متروپلیس است) بدست آوریم:

اعداد تصادفی $\{x_i \text{ و } i=0, \dots, n-1\}$ را که می‌خواهیم از توزیع $P(x)$ تبعیت کنند به صورت زیر تولید می‌کنیم:

۱- یک نقطه شروع انتخاب می‌کنیم (x_0)

۲- یک بیشینه گام انتخاب می‌کنیم (δ)

۳- نقطه بعدی تصادفی یعنی x_{i+1} را به صورت زیر تولید می‌کنیم:

۳-۱

$x_{trial} = x_i + \delta\xi$ (نقطه جدید تصادفی (آزمایشی))

↓ trial

که ξ یک عدد تصادفی بین ۱- و ۱ است. بنابراین

$$x_{trial} \in [x_i - \delta, x_i + \delta]$$

۳-۲ مقدار نسبت

$$w = \frac{P(x_{trial})}{P(x_i)}$$

را محاسبه می‌کنیم

۳-۳ - اگر $W > 1$ ، گام آزمایشی (تصادفی) در جهت درستی برداشته شده است!
(نقطهٔ جدید به خواف احتمال بیشتر حرکت کرده است) بنابراین این نقطه را
می پذیریم :

$$x_{i+1} = x_{trial}$$

۳-۴ - اگر $W < 1$ نقطهٔ آزمایشی تصادفی را با احتمال W می پذیریم (کاملاً رد نمی‌کنیم):
برای این کار چنین عمل می‌کنیم :

- یک عدد تصادفی بین ۰ و ۱ تولید می‌کنیم (r).
- اگر $W > r$ بود، این نقطهٔ جدید را می‌پذیریم ($x_{i+1} = x_{trial}$)
- اگر $W < r$ بود، این گام جدید را رد می‌کنیم و $x_{i+1} = x_i$

۳-۵ به گام ۳ برای تولید نقطهٔ بعدی برمی‌گردیم.

نکته‌ای که باید در اینجا در نظر گرفت، انتخاب اندازهٔ بیشینهٔ گام یعنی δ است.
می‌توان به راحتی مشاهده کرد که با افزایش δ تعداد گام‌هایی که رد می‌شوند زیاد می‌شود
و عملاً نمونه‌برداری درست انجام نمی‌شود. یک قاعدهٔ ساده برای تشخیص مقدار δ،
بررسی تعداد گام‌هایی پذیرفته شده به کل گام‌ها است. موقعی کار روی مناسب (ایده‌آل)
است که نسبت بین ۵۰٪ تا ۷۰٪ باشد.

توضیح مختصری درمورد مقاله متروپلیس 1953 :

عنوان مقاله:

Equation of state calculations by Fast computing Machines

مجله ... :

J. chem. phys. 21, 1087 (1953)

تعداد ارجاعات تا آذر 1401 : 49047

نویسندگان :

Nicholos Metropolis ,

زوج { Arinna W. Rosenbluth
Marshall N. Rosenbluth

زوج { Augusta H. Teller
Edward Teller

این نویسندگان در آزمایشگاه ملی لوس آلاموس برروی توسعه ریاضیات کار می کردند، ولی هرکدام حل یک مسئله که در ادامه درمورد آن صحبت خواهد شد، علاقمند شدند.

اهمیت مقاله:

۱- اولین شبیه‌سازی کامپیوتری یک سیستم فیزیکی

۲- شروع شاخه جدیدی در علم یعنی علم محاسباتی (computational science)

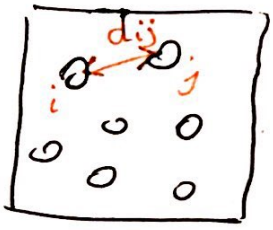
شاخه‌های فیزیک: تئوری، تجربی، فیزیک محاسباتی (1953)

موضوع مقاله:

تمرکز مقاله بر روی بدست آوردن معادله حالت برای یک سیستم دوبعدی از ذرات در یک جعبه دربسته، با استفاده از روش مونت‌کارلو بود (در ادامه توضیح داده شده است که چرا مونت‌کارلو انتخاب شد).

انرژی پتانسیل یک سیستم دوبعدی از ذرات:

$$E = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} V(d_{ij})$$



d_{ij} فاصله بین ذره i ام و ذره j ام است و

$V(d_{ij})$ پتانسیل (انرژی) بین این دو ذره را نشان می‌دهد.

با توجه به اینکه هر خاصیت (یا هر کمیتی) سیستم در آنسامبل کانونیک بصورت زیر است:

$$\bar{F} = \frac{\left[\int F e^{-E/k_B T} d^{2N}p\, d^{2N}q\right]}{\left[\int e^{-E/k_B T} d^{2N}p\, d^{2N}q\right]}$$

مختصه کمان ، مختصه کمانه

بنابراین ابعاد انتگرال گیری بسیار بزرگ خواهد شد، پس می‌توان این انتگرال گیری را با استفاده از روش مونت کارلو (ساده) بدست آورد ولی تعداد حالت های مختلف بسیار زیاد است و حالت هایی که مقدار E در آنها بسیار زیاد است عملاً در انتگرال نقشی ندارند. بنابراین متروپلیس و همکاران الگوریتمی مناسب برای محاسبهٔ $\bar{F}$ پیشنهاد کردند، که در ادامه به آن می‌پردازیم.

(البته باید به این نکته نیز اشاره کرد که طبق لقمه‌های روز نیولو عملاً او و همسرش آرنا الگوریتم را توسعه داده‌اند و در واقع روز نیولو ایده‌هایی از تیلر و نیومن برای متوسط گیری آماری گرفته است و این ایده (یعنی الگوریتمی که در ادامه گفته می‌شود) توسط همسر او بروی دستگاه های محاسباتی در لوس آلاموس پیاده سازی شده است. این فرانسه در حدود پنج سال حول کشیده است و طبق گفته‌های روز نیولو، متروپلیس عملاً زمان محاسبه برای استفاده از ماشین‌ها در لوس آلاموس برای این محاسبات برای آنها گرفته است).

برای اطلاعات بیشتر در مورد الگوریتم متروپلیس می‌توانید به مراجع زیر مراجعه کنید:

1 - Metropolis Algorithm wiki

2 - "Genesis of the Monte carlo Algorithm for statistical Mechanics" by Marshall N. Rosenbluth

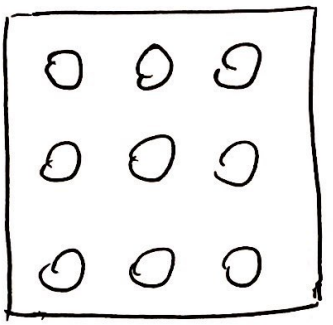
(ژنسیس، پیدایش)

3 - "Marshall Rosenbluth and Metropolis Algorithm" by J.E. Gubernatis

مرجع 2 و 3 در مرجع 1 معرفی شده‌اند.

الگوریتمی که متروپلیس و همکاران برای محاسبه F پیشنهاد کردند به شرح زیر است:

۱- ابتدا N ذره در داخل یک جعبه مربعی قرار می‌دهیم (به صورت دلخواه و بدون همپوشانی)، مثلاً می‌توان به صورت منظم قرار داد:



۲- سپس یک ذره را انتخاب و به صورت زیر جابجا می‌کنیم:

$$\text{مکان ذره} \begin{cases} X \longrightarrow X+\alpha\xi_1 \\ Y \longrightarrow Y+\alpha\xi_2 \end{cases} \longrightarrow (\text{ksi} \text{ تلفظ})$$

که α بیشینه جابجایی در الگوریتم ما را نشان می‌دهد و ξ_1 و ξ_2 نیز اعداد تصادفی بین ۱- تا ۱ هستند.

۳- بعد از جابجایی هر ذره، مقدار تغییر انرژی سیستم را محاسبه می‌کنیم (ΔE)

۴- اگر $\Delta E<0$ بود، این جابجایی ذره را می‌پذیریم و ذره را جابجا می‌کنیم.

۵- اگر $\Delta E>0$، با احتمال $e^{-\Delta E/k_BT}$ این جابجایی را می‌پذیریم.

برای اینکار یک عدد تصادفی بین ۰ و ۱ تولید می‌کنیم ($1>\xi_3>0$)، اگر $e^{-\Delta E/k_BT}>\xi_3$ بود، این جابجایی را می‌پذیریم و ذره را جابجا می‌کنیم، در غیر این صورت این جابجایی را نمی‌پذیریم و ذره در حالت قبلی خود باقی می‌ماند.

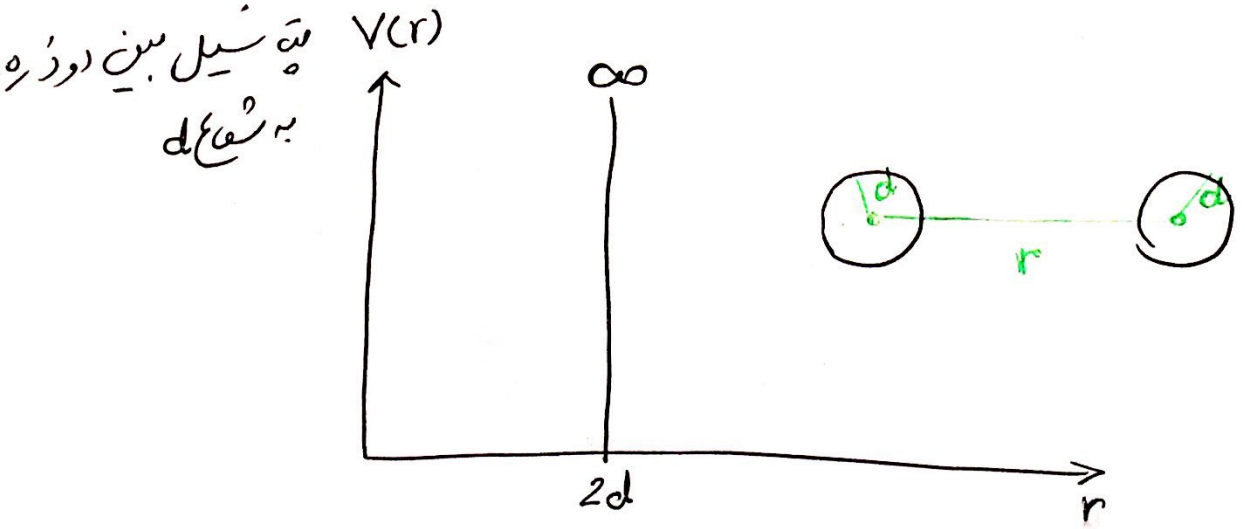
۶- در هر گام (چه پذیرش انجام شود و چه نشود) مقدار کمیت F را می‌سنجیم.

۷- با تکرار مراحل ۴ تا ۶، مقدار متوسط F را به صورت زیر به دست می‌آوریم:

$$\bar{F} = \frac{1}{M} \sum_{j}^{M} F_j$$

که از یعنی حالتی که در تعادل ترمودینامیکی به دست آمده است (چه حالتی که پذیرفته شده و چه حالتی که سیستم در حالت قبلی باقی مانده، بشمار).

مثالی که در مقاله سال 1953 با این روش به آن پرداخته شده، مسئله دیسک‌های سخت (hard-disk) بود. در این مسئله خاص، پتانسیل بین ذرات تا زمانی که به هم برخورد نکنند وجود ندارد و در زمان برخورد پتانسیل بی‌نهایت می‌شود:



می‌توان نشان داد که معادله حالت حاکم بر سیستم‌های ساخت، دقیقاً معادله‌ای نیست که برای گاز ایده‌آل در نظر می‌گیریم. در واقع می‌توان معادله حالت را طبق معادلات زیر بدست آورد:

$$F^{(D)} = -k_B T \ln Z^{(D)}$$

$Z^{(D)}$ ← تابع پارتیشن

$F^{(D)}$ ← انرژی آزاد هلمهولتز (Helmholtz)

$D = 1, 2, 3$ ابعاد

و فشار به صورت زیر خواهد بود:

$$P^{(D)} = -\left[\frac{\partial F^{(D)}}{\partial L^{(D)}}\right]_T = k_B T \frac{\partial \ln Z^{(D)}}{\partial L^{(D)}}$$

$L^{(D)} = L$ یک بُعد

$L^{(D)} = A$ دو بُعد

$L^{(D)} = V$ سه بُعد

برای ذرات کاملاً بدون برهمکنش و در اینجا ذراتی با شعاع صفر) معادله بالا به صورت زیر خواهد شد:

$$Z^{(D)} = \left(L^{(D)}\right)^N$$ ← تعداد ذرات

(کراوث Krauth)

بنابراین

$$P^{(D)} = k_B T \frac{\partial (N \ln L^{(D)})}{\partial L^{(D)}}$$

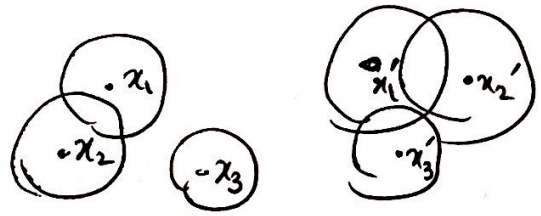
$$\rightarrow P^{(D)} = k_B T \frac{N}{L^{(D)}} \longrightarrow P^{(D)} L^{(D)} = N k_B T$$

بنابراین معادله حالت برای ۳ بُعد و ۲ بُعد به صورت زیر خواهد شد:

۳ بُعد $\quad PV = N k_B T$

۲ بُعد $\quad P^{(2)} A = N k_B T$

ولی این رابطه از کجا بدست آمده که ذرات می‌توانند در هر جایی قرار گیرند:



حالت‌های ممکن (قابل قبول)

$\longrightarrow$

سه بُعد $\quad Z = V^N$ حجم یا

دو بُعد $\quad Z = A^N$ مساحت یا

$\downarrow$

$\{x_1, x_2, \cdots x_N\}$

در هر کجا فضایی می‌توانند قرار گیرند

یک بُعد

$$Z = \int dx_1 dx_2 \cdots dx_N = L^N$$

ولی وقتی فرض دیسک سخت (hard -disk) را در نظر بگیریم، آنگاه

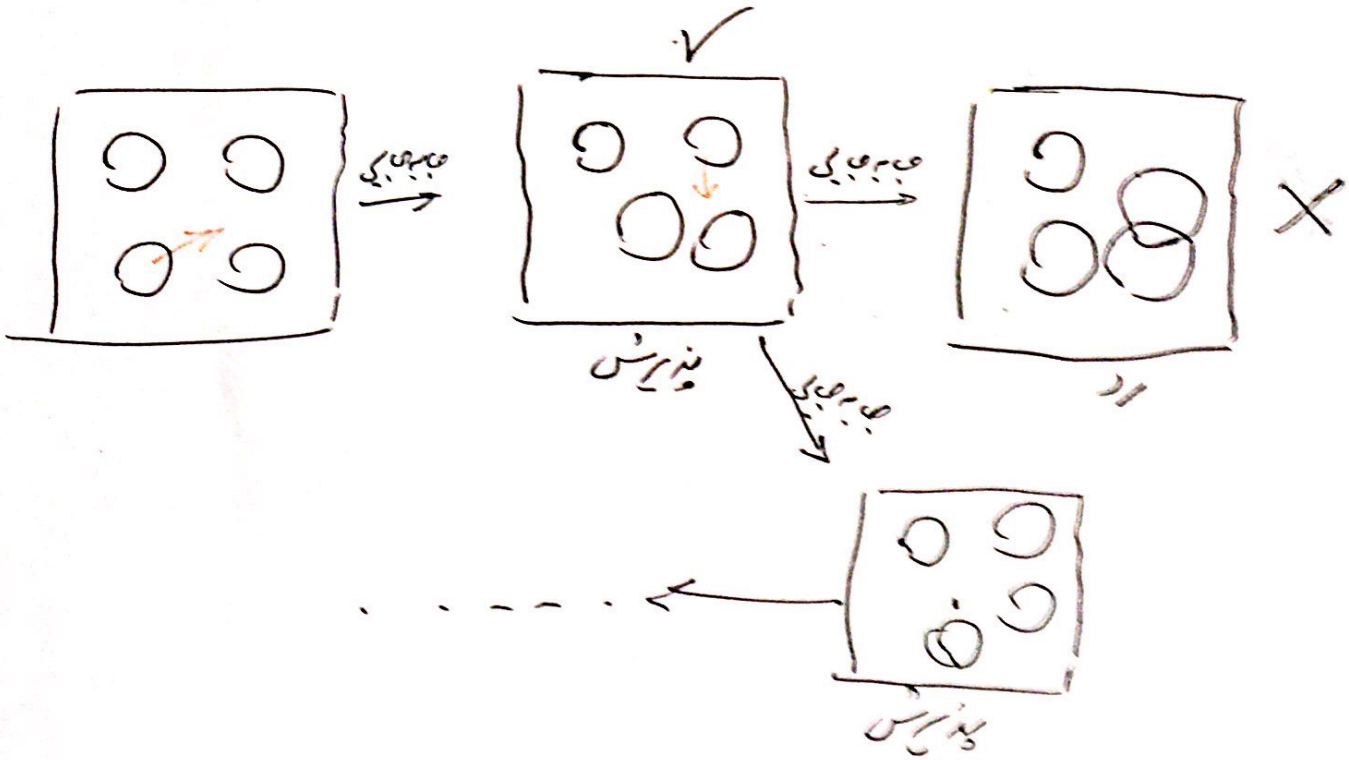
$$Z \neq V^N$$

و به صورت درست‌تر:

$$Z < V^N$$

حال اینکه Z به چه صورت خواهد شد و معادله حالت به چه شکل در می‌آید ($P^{(12)}A = ?$) کاری است که در مقاله برای حل آن انجام شده است.

برای اینکه شبیه‌سازی مونت کارلو بر روی این سیستم انجام شود، نویسندگان دقیقاً طبق الگوریتمی که گفته شد، شبیه‌سازی را انجام داده‌اند:



$P = e^{-\Delta E / k_B T} = 0$ حالتی که همپوشانی بوجود می آید، پذیرش با احتمال صفر انجام می شود

بنابراین جابجایی هایی که منجر به همپوشانی می شوند، پذیرفته نمی شوند (ردی شوند)

rejection

نکته ای که در اینجا باید به آن اشاره کرد، این است که در این شبیه سازی فقط هدف تولید حالت های مکانی مختلف برای ذرات بوده است. با این کار می توان Z (تابع پارش) را حدس زد و بنابراین می توان معادله حالت را بدست آورد.

برای جزئیات بیشتر به سخنرانی های Krauth هفته دوم و هفته سوم از مجموعه درس های coursera به کتاب Krauth مراجعه شود.