

極めてまれな事象を探る計算機シミュレーション

物理・情報・工学のさまざまな分野で、現象を確率分布でモデル化して、それを乱数を用いた計算機シミュレーション（広義のモンテカルロ法）で探ることはもはや常套手段といえる。しかし、興味がある現象が起きる確率が与えられたモデルのもとできわめて小さかったとしたらどうだろう。それが1万回に1回、あるいは1億回に1回であっても、現在の計算機はとても速いので、まだまだ直接計算できる。ところが、それが10の20乗回に1回とか、10の100乗回に1回とかになると、もう普通にやっていると宇宙の終わりまでかかっても計算が終わらない。最近、統計数理研究所の伊庭幸人氏と東京大学の福島孝治氏は、そのようなきわめてまれな事象を探るための計算機シミュレーション法を構築した。この研究は、日本物理学会発行の学術誌 *Journal of the Physical Society of Japan* (JPSJ)の2008年10月号に掲載される。

「珍しい現象を選び出すシミュレーション」は、物理現象の理論的解析を含むさまざまな理由で必要になるが、本研究では、あるアルゴリズムや手法の弱点を知る、という文脈でこれを考えている。論文で提案された手法には一般性があり、さまざまな応用が期待されるが、論文では特に誤り訂正符号の問題が扱われている。

誤り訂正符号とは、信号を冗長化することで信号が通信路で蒙る雑音に対して情報の耐性を保証する情報技術である。送り手は送りたい情報がある規則に従って冗長化し（encode とよぶ）、これを発信する。受信した受け手は冗長化された信号を規則にしたがって元に戻し、送り手の意図した真の情報を得る（decode とよぶ）。誤り訂正符号の問題はスピン系の統計物理と数理的に等価であり、1対1で翻訳できることが知られている。統計物理の言葉でいえば、encoding とはあるスピン配列（ビットパターン）を基底状態として持つようにスピン系の相互作用を定めることに対応し、decoding は受け取った相互作用をもつスピンモデルの基底状態となるスピン配列を探すことに対応している。

本研究で、伊庭・福島両氏が問題にしたのは次のような状況である。雑音の強さがそれほど強くなくても、非常にまれな雑音の並びが「急所」に当たれば、冗長化された情報といえども、decoding で真の情報が回復できないほど破壊されてしまう可能性がある。これは統計物理の例でいえば、強磁性状態が基底状態であるときに、少数の結合をうまく選んで負にすることで強磁性状態を不安定化させることに相当する。本研究の例題では、図1に示す通り、10の20乗回に1回程度起こる現象まで確率が算出されている。ここで、雑音が独立であるという理想化された仮定のもとでは低い確率でしか起こらない事象でも、実世界では相関などの効果で簡単に実現する可能性があることに注意したい。このような小さな確率の評価は理論・応用の両面で有用と思われる。

図1の結果を単純に考えると、少なくとも10の20乗個以上の雑音パターンを調べたことに相当するが、数百個の変数をもつ系に関して文字通りそれを行えば、10ペタ

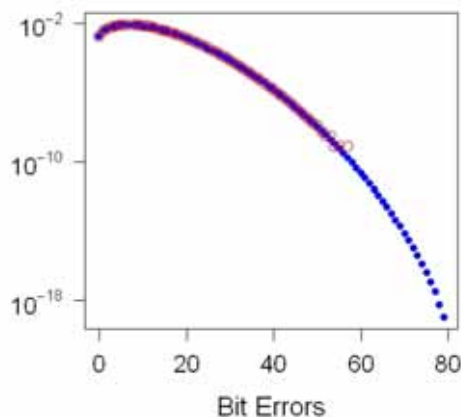


図 1. 簡単な畳み込み符号での例示—入力ビット数 200、符号語の長さ(スピン相互作用の数)396 の場合について、本研究で提案した手法で計算した結果を青点で示した。縦軸は与えられたモデルにおける生起確率（対数目盛り）、横軸はビット誤りの数。提案した方法で求めた確率は素朴なシミュレーションで求めた確率（赤丸）と生起確率の大きいところでは正確に一致する。

フロップス級の次世代スパコンでも年のオーダーを要する計算である。それを可能にしたのは、統計物理学の分野で使われているマルチカノニカル・モンテカルロ法の考え方をこの問題に応用したことによる。マルチカノニカル・モンテカルロ法は、メトロポリス以来半世紀にわたって物理学で利用されてきた動的モンテカルロ法（マルコフ連鎖モンテカルロ法）の一種である。動的モンテカルロ法の応用範囲は物理学の外に拡大しつつあり、最近では統計データ解析でも盛んに利用されているが、そのなかでもマルチカノニカル法は分布にローカルな極大のある問題で強みを発揮する。本研究で提案された手法は、従来の「珍しい現象を選び出す」手法の多くがうまく働かない場合でも定量的な結果を出しうるのが特徴である。

動的モンテカルロ法の応用という観点からすると、本研究で特筆すべきは、雑音（符号反転）の位置についてモンテカルロ法を用いた探索・サンプリングを考えている点である。スピン系の言葉に翻訳すると、クエンチされたランダムネスに関するサンプリング、すなわち、あるアンサンブルから沢山のスピンモデルを選び出したとして、それに伴う諸量の確率的揺らぎをモンテカルロ法で探索していることになる。これは、従来から統計物理学で用いられている動的モンテカルロ法が、1つのスピンモデルにおける熱揺らぎに伴う諸量の確率的振舞いを見るために使われているのと対照的である。

本研究で提起された、マルチカノニカル・モンテカルロ法に基づく手法は、誤り訂正符号の問題に役立つのみならず、「急所」のサンプリングという観点から情報処理の分野や、極めてまれな事象が重要になる物理現象への広い応用展開が今後大いに期待される。

論文掲載誌: J.Phys. Soc. Jpn. 77 (2008) 103801

電子版: <http://jpsj.ipap.jp/link?JPSJ/77/103801/>

<情報提供: 伊庭幸人（統計数理研究所）>