

平成 22 年度 国立情報学研究所 市民講座
「計算しない数学－意外と身近な離散数学とは？」
講師： 河原林 健一
(国立情報学研究所 情報学プリンシプル研究系教授)

◆ 講 義 ◆

・スライド1「計算しない数学－意外と身近な離散数学－」

ただ今ご紹介いただいた、国立情報学研究所の河原林です。

本日は「計算しない数学－意外と身近な離散数学とは？」についてお話しします。

まず、暑い中、お越しいただき、ありがとうございます。明日あたり梅雨明けのようで、体調にお気を付けてお過ごし下さい。

本日は「計算しない数学」について話しをしたいと思います。ちょっとだけ計算します。すみません。

(笑) 足し算、引き算くらいはさせてください。そのくらいはさせてもらわないとお話し出来ないのので申し訳ないですが、その辺はお許し下さい。

ただ、難しい計算式はあまり出さないような形でやりたいと思います。

・スライド2「講演者の簡単な紹介」

まず1つ、私のプロフィールを。

詳しいプロフィールは、お手元にあると思います。軽くご紹介します。

専門分野は、数学と情報科学を中心にやっております。

その中でも、数学の中では離散数学といわれる分野、結構新しい分野で、ここ 30～40 年くらいで発展してきた新しい数学分野を中心に研究しています。

情報科学分野では、理論計算機分野、完全に理論を中心に研究しています。

実際、離散数学と理論計算機科学というのは非常に交わりがありまして、両方、密接に関係しています。

離散数学の知識、手法が計算機、例えば、コンピュータなどにもいろいろと利用されています。

今回も、いくつかそういう例を皆さんに紹介したいと思います。

離散数学ですが、私の研究としては離散数学、計算機科学なのですが、数学というのは、実際数学的興味というのがありまして、数学者独特の世界観を持っているのです。数学はきれいな世界で出来ると。数学というのは全てが美しく、全てがまとまっていて、神によって創られているというようなきれいな世界だというような認識があり、そういうことにも私は興味があります。ある意味、人間が到達できる頂きみたいなものを見る感じで、数学者は数学の問題に取り組んでいます。そのようなものにも興味あります。また、いくつか例を出しますが、実用に応用可能な離散数学、又は計算科学にも興味を持っています。

要するに、自分の持つ知識が応用できる分野であれば、実用的なものなら何でも興味があるということです。

ただし、計算は非常に苦手で、苦手というより、よく間違えます。

実際、本日、お配りした資料ですが、最終的にチェックしてみたら、だいぶ、数字が間違っていることに気づきました。申し訳ないですが、後で訂正させていただきたいと思います。

計算が苦手だから間違えるのではなく、逆でもあるのです。間違えるから苦手になるわけです。

計算はあまり好きじゃないんです。

ただ、論理的に考えるのは好きなので、今日はちょっと論理的に、離散数学の論理的に考える部分を皆さんに触れてもらえればと思います。

趣味はスポーツ観戦、ジョギング、寝ることです。

3つ目の寝ること、これは、数学の問題は難しい問題が多いですね～、だから大体考えているうちに寝てしまうと。(笑)大体解けなくてそのまま寝てしまって、夢の中で解けたところで、現実の中では何も出来てなかったということがよくあります。

あと、猫が大好きで猫と遊ぶこと、猫の写真を見ること。飼い猫が3匹います。

最近困ったことは、ワールドカップが一昨日まであり、見事に3時半に起きて試合を見ることになったわけです。そのおかげで寝不足です。実は3時半はワールドカップの試合が始まる時刻でもあったのですが、私の猫が起こしに来る時間でもあったので、都合はよかったけれど、寝不足でしょうがないと・・・。

飼い猫の1匹が私に反抗的なんです。最近、分かったのは「敵の敵は味方」。猫同士だけでもあるんですが、これは人間でも正しいわけです。3匹飼っていて、その1匹がもう1匹と仲が悪い。

仲が悪い猫が私になついています。

なぜかという、反抗的な猫、縄張争いの猫にとってみれば、私は敵の敵なわけです。なのでどうやら「敵の敵は味方」は猫にも当てはまるみたいです。「敵の敵は味方」は正しいのかはよく分からないのですが・・・。

そんな感じで、研究についてはプロフィールに書いてあります。

・スライド3「講演のアウトライン」

では、次に進みます。講演のアウトラインです。

単純にこの3つだけを今日はお話したいと思います。

申し訳ないですが、本日出席されている方のリストというか所属先を事前に、昨日ですが見せていただきました。

その結果、非常にバラエティがあって、数学がものすごく嫌いな人もいらっしやれば、結構その道のプロの方もいる。残念ながらその道のプロの方には私の講演は、たぶん簡単すぎると思います。

申し訳ないですが、そのような方は私の講演を邪魔しないで下さい。いじめることも可能ですので、そのようなことはお控えいただきたいと思います。

実用的な問題を数学的に考えようというのと、もう1つ、実用的問題には離散数学が使われているよということを今日、ちょっと分かっていただければいいと思います。

最後に、離散数学を使って基本的解決法を学ぼうということ、これらを中心に話していきたいと思いません。いくつか例を出しますが、簡単な順番に話しますので、なるべく最初はゆっくり、皆さんに分かっていただければと思っています。

・スライド4「実用的な問題を「数学的」に考えてみよう」

では、実用的な問題をちょっと考えてみましょう。

まず、1個目。単純にこういう問題。

128 人参加のテニストーナメント。

ウィンブルドンとか、全仏オープンとか、全米オープンなどは、128 人きっかり参加します。

これだけ参加するテニストーナメントというのは、何試合ありますか？

トーナメントが書いてありますが、16 チームが参加していて、数えれば皆さん、15 試合だとすぐわかりますね。ただ 128 人になると、ちょっと数えるわけにはいかない。

この解法をご存じの方は簡単だと思いますが、トーナメントは、参加者は全員、優勝者を除いて 1 回しか負けません。

2 回負けることはありませんね。負けたらおしまいですから。だから 1 回しか負けません。

各試合に勝者が 1 人、敗者は 1 人です。各試合に敗者は 1 人ずつ出るわけです。

全部で何人の敗者がいるかという、127 人いることになります。1 人だけ全部勝つわけですから。

1 試合につき 1 人敗者がいて、それで全部で 127 人負けるわけですから、127 試合です。

いいですか。簡単なトリックですよ。

2 番目は、バースデートリックと言われています。

大体 30 人くらい集まれば、90%以上という高い確率で、同じ誕生日の人がいます。本当はもっと高いのですが。これは「計算しない数学」なので端折らせていただきますが、こんなことも数学を使うとしたら離散数学を使う。

3 番目、漢字、間違っています。白鵬です。

今日、勝ったかどうかわかりませんが、だいたい白鵬の勝率は 90%くらいです。

年間 60 番くらいありますが、だいたい 54 勝、55 勝くらいします。

大体 9 割くらいなんですね。その白鵬が、どっかの場所中に 2 連敗する確率は、どれくらいでしょうか。

これは異常に小さいんですね。ちゃんと計算すると 3%よりもっと小さくなると思います。

これは何を示唆しているかという、もし懸賞をかけるなら、負けた翌日に賭けるべきなんです。その時に懸賞を白鵬に与える確率は非常に高い訳ですね。

このような考察もできる。このような考察の後に何を考えるかという、まず 1 問目。

・スライド 5 「テニスボール問題（基本編）」

基本的な問題を考えてみることにします。離散数学を使って。

まず、テニスボールがあります。12 個用意しました。テニスボール、全部同じ形をして、全部同じ色をしています。ただし、欠陥ボールが 1 個だけある。しかもその欠陥ボールは、1 個だけ重いことが分かっています。もちろん見ただけでは判別できません。

ここに幸い、天秤があります。

あなたは、天秤を使ってこの 12 個の中から 1 個だけ重いものを判別して下さい。

ただし、3 回しか天秤を使ってはいけません。さあ、どうやりましょうかと。。これが問題です。

・スライド 6 「テニスボール問題（基本編）の答え」

これは、分かっている人から見れば、すぐに分かっちゃって、今まで見たことない方にとっては結構考えてしまうんですね。

だから、分かっているといるんなところでちょっと自慢が出来たりする。

ちょっとだけこの解法を見てみましょう。こうやると3回ですと。

まず、12個のボールを4個ずつに分けます。これは勝手に分けてもらって結構です。

便宜上色が変わっていますが、これは、実際は色は全く同じだと思って下さい。

単純に4個、4個、4個に分けます。

さて、ここで天秤が登場です。

まず1回目は、A、3つに分けたうちの1個目の4つと、2個目、Bを天秤にかけました。これが1回目。

天秤に掛けた時、何が起るかという、3通りあります。

1通り目は、Aが重い。

2通り目はBが重い。

3通り目は釣り合う。

もし、Aのほうが重かったら、Aのほうに欠陥ボールがあることになります。

なぜかという、1個だけ重いわけですから。

もしAとB、両方とも欠陥ボールがなければ、釣り合わなければいけませんからね。

Aが重ければ、Aに欠陥ボールがあることが分かる。

もし、Bが重かったらBに欠陥ボールがある。3番目、釣り合ったらどうなるか。

釣り合うというのは、Aの4つのボールも、Bの4つのボールと、全部同じ重さだということが分かります。ということは、何を言っているかという、Cに重いボールがなければいけない。

ですから、天秤の1回だけで、候補が4個に絞られてしまう。

一気に4個になる。つまり、8個を除外できてしまう。

・スライド7「テニスボール問題（基本編）の答え 続き」

その次何をするか。単純に、このAをまた2つに分けます。先ほどみたいにAが重かったとしましょう。

次はAを単純に2つに分けます。

でまた、天秤を使って測りましょう。

Aの中に重いのが1個入っていることが分かるので、必ずどっちかに傾きますね。

そうすると最後の2個に絞られて、この最後の2個を3回目に量ることで、どれが重いか判別できるわけです。これで3回で分かりますね。

これが離散数学の1個の考え方で、単純に分けて、大体半分か3分の1を1回で捨てましょうというそういう考え方の元で成り立っています。

・スライド8「テニスボール問題（基本編）の発展問題」

応用問題です。解法は今しませんが、家でちょっと考えてみて下さい。

先ほど使った問題では、重いボールという限定がありました。

だから、最初の段階でAとBというのを4つずつ乗った状況で、Aのほうに重ければ、Aの4つの中に必ず1個重いボールがあると分かったわけです。

でも実際、この問題をちょっと変えて、欠陥ボールがあるのは正しいのですが、仮に重いか軽いかわからなかったらどうしますか。

つまり、最初の段階で、AとBを先ほど比べましたね。

Aが重かったから、Aのほうに欠陥ボールがあったと分かったんですが、重さが軽いか重いかわからな

いと、Aにあるのか、Bにあるのか今のところ分からない。分かるのは、Cにはないということだけです。でも、AとB、どちらにあるかは、まだ分からない。問題は、実は3回ではなく4回でやってみて下さいということ。これが問題1です。今の12個というのは、何か人工的みたいな感じがする訳ですね。では100個やってみましょう、どうなるでしょうか、というのが問題。いま、解法はしませんが、帰りの電車の中でちょっと考えてみて下さい。いろいろ、面白いことが言えると思いますので・・・。

・スライド9「テニスボール問題（応用編）」

さて、これが基本問題でした。では、もうちょっと応用編に進みましょう。今度はまた、ボールの問題です。テニスボールなのですが、全部、今度は重さが異なります。重さが異なるボールが16個あったとしましょう。問題は単純です、見た目からは、どちらが重いかわかりません。ボールは基本的に全部同じ形です。便宜上、色を変えましたが、見た目からはどちらが重いかわかりません。問題は、単純に重い順に並び替えて下さい。ただし、次の作業しかはやってはいけません。やっていい作業は、天秤みたいなのですが、1個ずつ、2個の比較だけしかやってはいけません。つまり、一気に3個取ってきたり、4個取ってきたりしてその4つから1個を取り出したり、重さをみつけたり、3個取り出して、そのうちの1個が一番重い、これが2番目に重い、例えば3つを取ってきて、これが重い、これが重い1番目に重い、2番目に重い3番目に重いとかの判定してはいけません。単純に2個とってきて判定する作業だけでやって下さい。さて、1回の作業では2個のボールしか比べてはいけません。問題は、何回の作業で重い順に並べ替えることができますか、というこれが問題です。これは実は後で出てきますが、計算機の中で、すごく重要な概念に結びついています。それをちょっと一言おきまして、この問題の動機というものを伝えさせていただきました。

・スライド10「テニスボール問題（応用編） 解決法（その1）」

では、これをどうやって解くか。2通りお見せします。1通り目は、誰でも思いつくものです。単純にボールを16個、取って来ました。単純にこの最初の2個を比較します。重いほうが分かりますね。重い方を取ってきます。2番目には、1番目に重かった方と、3番目のボールを取ってきて比較しましょう。で、重い方をとってきて、3番目には単純に4番目のボールを持ってきて、重いのと軽いのを比べましょう。だから3番目の作業までに、何が言いたいかというと、1番目、2番目、3番目、4番目のボールの中で1番重いのが分かるわけです。単純に最初の2個を比較して一番重いのが分かる。その後の3つめを取った時に、最初の2個の内の重いものと、3番目を比較しているわけです。ただし、これを15回繰り返さないと、1番重いボールが判明しないわけです。

・スライド11「テニスボール問題（応用編） 解決法（その1） 続き」

もう少し砕いて言うと、このやり方は何をやっているかというと、

最初の5回やったときは何かというと、最初の6番めの中で1番重いボールは何かというのを判定しているわけです。いいですか？こうやれば、必ず左から順々にこの作業をやっていくと、最後の段階で、いちばん重いボールが分かるわけです。

で、1番重いボールが分かったら、その重いボールを列から除いて同じことを2番目に重いボールをやってみましょう、と。これは確実にこの方法を使えば分かりますね。いいですか？

ところが、これはボールの総当たり戦みたいなものなんです。

全員が全員、比べないといけないから、このやり方を「総当たり戦」と呼びます。総当たり戦方式なので、実は、最悪何回の作業が必要かと言うと、全部でこの足し算だけ(120回)必要になってくるわけです。今の作業は実は総当たり戦と全く同じでなんです。

全員と全員を比べないといけない。こんなに多くの数が必要になってきてしまいます。

実は、こんなに要りませんよ、ということをご案内からみなさんにご紹介したいと思います。これは膨大な作業ですね。実はこんなにいりませんと。

もっと少ない回数でやりたい。頭を使ってやりましょう、というのがこれからの話です。

・スライド12「テニスボール問題(応用編) 解決法(その2)」

では、どうやるか、ちょっとお話ししたいと思います。

先ほどは総当たり戦でした。総当たり戦が出てくると、今度出てくるのは何かというと、今度はトーナメント戦なんです。単純に16個のボールを並べてトーナメントを作りました。

トーナメント戦をやって、それで判定しましょう。

どうやってトーナメント戦をやるか、ちょっとだけお見せしたいと思います。どうやって証明するかというと、単純にトーナメントですね。

1番最初にやることは、この左の2つを見て勝者をみる、第2回戦を見て勝者を見る、第3回戦を見て勝者を見るというように続けていって、次を同様にやる…と繰り返し、1回負けたら、トーナメント戦だからノックアウト方式なので、試合はもうしません。単純に勝者同士だけで試合をするという形で試合をしていきます。

・スライド13「テニスボール問題(応用編) 解決法(その2) 続き

そうするとチャンピオンが決まります。トーナメントは結構いいんです。

・スライド14「テニスボール問題(応用編) 解決法(その2) 続き」

なぜかというと、優勝者は必ず1番なんです、絶対に。いちばん強い人が優勝者なんです。これはいいですよ。2番目がちょっとポイントでちょっと見てみてください。

準優勝者、ワールドカップでいえばオランダです。

*スライド13参照

オランダは準優勝ですが、準優勝のオランダは、何を言っているかということ、オランダはここ(左から2番目)にあたるわけですが、これは丁度16でワールドカップのトーナメントと同じなんです。この山の中(一番左)では、オランダが一番強いということだったのです。

だけど、準優勝者はこの中で1番強い、つまり重いボールであることは間違いないが、だからと言って全体の2番目という保証はないわけですね。

一番いい例は、ワールドカップの例ですが、*スライド13参照

これがスペインだとしたら、こっちはドイツなんです。どうやってドイツとオランダの強い方を比較するか。単純に準優勝者がいちばん強いとは残念ながら言えない。ただし、2番目に強い人を決めるのは、どうやるか。この3番目がポイントです。

2番目に重いボールを決定するためには、準決勝、準々決勝、1回戦で優勝者に負けた相手とだけ比較すればいい。

見てください。ドイツ。スペインには負けましたが、ドイツはなぜここまで来られたかというこの4つの中では1番だったわけです。

だから2番目に強い候補者としては……。この人か、この人。または、この人。またはこの人。

(映像参照)で、なぜ敗者といっぱい決めなきゃいけないかというと、

例えば、ここスペインは初戦相手がブラジルということもあり得たわけです。先ほど申しましたが、優勝者は常にいちばん。

だけど2番目の人は、どこにいるかというと、実は優勝者に負けた人の中にはいるはずですよ。

もし優勝者に負けた人じゃないのが2番目だったら、その人が優勝者に辿り着いているはずですよ。そうですね。他の人、全て倒していないといけないわけですから。

だから、2番目を判定するためには、何をすればいいかというと、優勝者に負けた人だけを見ればいいわけです。いいですか、今の考え方は。それが3番目にあたります。

単純に1回戦、準決勝、準々決勝で優勝者に負けた人達の中で、どれが1番かというのを決めればよいのです。

そうするとその優秀なところは何かというと、先ほどの総当たり戦で一位に決まるのに最悪で29回必要だったわけです。15回と14回。

*スライド13参照

ところが今回の場合は18回で済んだのです。なぜかと言うと、トーナメント戦、全部で何試合あるかというと、16個ボールがあるので、先ほどの論理からいうと15試合ですね。

優勝者に負けた人は何人かというと、3人。4人ですね、すみません。

ですから、この15の中に決勝が含まれているので、15+3で18回だけで済むわけですね。

この時点で11回も得しているわけです。トーナメントは、もう少し良いことが言えます。

・スライド15「テニスボール問題(応用編)解決法(その2)続き

ここからは皆さんにも後で考えてほしいんですが、3位以下はどうやって決めましょうか。

2位が決まりました。3位以下はどう決めましょうか?

ちょっと考えてみて下さい。

もうちょっとこの方法を拡張させると、次のようなことを考えつきます。

見にくいかもしれませんが、ホワイトボードのトーナメント表を使います。

先ほどの、トーナメントの結果が出ているわけです。*映像参照(ホワイトボード)

それぞれの1回戦で全部、敗者がいるわけですね。

実は1番軽いボール、または1番弱いチームを見つけるのはどうすれば良いかというと、敗者同士で決戦すればいいんですね。

今度は、トーナメントが逆になりますが、赤いほうは、弱いほうが進むと。つまり軽いボールが進んで

いく。そういう形でトーナメントをすると、この勝者は一番弱い人になります。同じ論理から。

2番目の弱い人も、同じようにして探せます。そこで大分得するんですね。

さきほどのトーナメントで1番、2番が決まって、もう少し、敗者の弱いトーナメントをやると、1番弱い人、2番目に弱い人がすぐに決まる。こういうことを考える、これがヒントです。

ちょっとここから先を考えて下さい。

ある情報だけを使って頑張ると、64回位でできます。

ですからポイントは、ここにある情報は最大限利用しましょうということです。

今皆さんにお話しましたこの問題、一番最初に申し上げましたが、非常にコンピュータに対して重要だと申しました。

・スライド16「テニスボールからソーティングへ」

もし皆さんが専門家だとすぐにピンと来ると思うんですが、実はコンピュータのソーティングという問題と全く同じ問題でもあります。

ソーティングとは何かというと、MicrosoftのExcelなどを使われている人はご存じだと思いますが、単純に数字がいっぱい並べられていたときに、大きい順・小さい順に並び替える。

今のテニスボール問題も、単純にテニスボールを数字だと思うと全く同じなんですね。

重い順に並び替える、軽い順に並び替えると。

問題設定として先ほど、作業として2個を比較するしかできないといたしました。

これはなぜかということコンピュータそのままなんですね。

まず青いところを見て下さい。コンピュータは入力してもらいます。

数字ではなく2進法でもらうのですが、数字でコンピュータはもらいます。

コンピュータはデータ、数字を仮にもらったとしても、人間みたいに瞬時にどちらが大きいかわ判断できません。

単純に情報としてもらっているだけなので、数として認識してもらえない。

コンピュータというのは全く認識できないわけですね。

で、コンピュータがどういう動作をするのかということ、残念ながら、1回の動作で1つしかできない。

先ほど言いました1回の動作で大小を比較するというのは、

コンピュータの動作では、1回の動作で単純に引き算しかできない。

実際コンピュータというのはかけ算すら、本当はできないんです。

かけ算は足し算を繰り返してかけ算にしなければいけないくらいなんで、本当は足し算と引き算しかできないんです。

だからコンピュータというのは、1回の動作で単純に3つの数字をもらってどれが一番大きいかなんてできない。

単純に2個数字もらって、プラスかマイナスかぐらいしか分からない。

だから先ほど言った、1回の作業で2個しか比較できないというのは、単純にコンピュータと同じことしかやってはいけませんよという条件と全く同じだったわけです。

ここで問題なのは、コンピュータに何回作業させれば、先ほどの問題が解けるかということなんですね。

そこがコンピュータで問題を解く時の速度を決定する大きな要因の1つになります。

・スライド17「アルゴリズムの計算量」

もうちょっとだけ専門的な話になります。

先ほど申しました、速度を決定する大きな要因が計算量というものに関わってきます。

その前に、アルゴリズムというのは単純に解法のやり方です。

先ほどの総当たり戦だとか、トーナメントを使ったやり方というのが、アルゴリズムだと思って下さい。

アルゴリズムの解決に要する時間を計算量と呼ぶのですが、コンピュータの速度を決定する大きな要因は、計算量と呼びます。計算量というのはどういうものか。

単純に入力数の関数。すみません。関数というのが出てきてしまいました

関数というのは単純に、 X の2乗だとか、2の X 乗とかになります。

そこだけは、皆さんご存じだということで、行かせて下さい。

単純に計算量というのは、ここで言うと入力数。

今の例だと、テニスボールの数の関数になります。

先ほどのやり方、テニスボール問題の解法を、さっきは16個でしたが、 N 個でやるとどのくらいかかるか、どれくらいの回数が必要か。

評価すると、総当たり戦では $N(N-1)/2$ くらい。

これは単純に、 N 人と N 人がいた時、どれだけペアがあるか、というのと全く一緒です。

トーナメント方式は、数学的にちゃんと評価してあげると、 $N \times \text{LOG } N$ 。

LOGとは何か。トーナメント戦で、優勝するために何回勝てばいいか。

もうちょっと専門的に言うと、この下に2とか10とか付くんですが、今ここでは2だと思ってください。単純に $\text{LOG } N$ というのは、 N チームが参加したトーナメントにおいて、優勝するためには何勝すればいいか、その数です。

計算量を数学的に評価すると、だいたいこんな感じ（スクリーン）になります。

これはなぜ重要かといいますと、この2つの数字、数の大きさが相当違ってくる。

・スライド18「アルゴリズムの計算量」

この2番目に行きます。

例えば、現実的じゃないですが、ボールを1万個だと思ひましよう、いや、1億個だと思ひましよう。

一億個というのは、本当は現実的じゃないですが、ソーティングという概念から言うと例えばアメリカ合衆国でお金を稼いでいる人は大体1億人位はいるので、その人たちがどの位稼いでるかをソーティングするとなると、やはりそれくらいのデータは必要になってくるんです。それくらいのデータがあると、さっきの総当たり戦で1億個のものを解くとすると、どんなに速いコンピュータを使っても総当たり戦だと100日位かかってしまいます。

でも、トーナメント戦を使うと、だいぶ速くなるんです。 $N \times \text{LOG } N$ だとすると大体1000分くらい。1000分というのは、どのくらいかという……え〜と…大体1日かからないくらいできてしまう。

なんでこんなに違って来るかという、単純に1億という数字を、ここ* $N(N-1)/2$ の N に入れてみて下さい。

どれだけ違うか、よく分かると思います。単純に0の数を数えただけでも、どれだけ違うのか分かると思います。これだいぶ違うんですね。大きなデータを考えると。

ここで言いたいのは、高速化の必要性が絶対にある。

時間がかかりすぎては意味がないので、なるべく速く解けるようにしましょう。
それがここで言いたかったことです。

因みに、世の中にはさっきのような $N(N-1)/2$ みたいなものよりもっとひどいアルゴリズムがあります。
 N の 3 乗とか、あるいは 2 の N 乗とか 3 の N 乗・・・みたいな回数必要なアルゴリズムは世の中にいっぱいあるんですね。これは、どれだけひどいかというのを次にお見せしたいと思います。

・スライド19「データ量 N の問題を解く時間量」

これ、見てください。 N の 2 乗だと、まだこれくらいなら解けます。60 くらいなら解けるんです。
 N の 3 乗もこれもまだ許せますね。これへんまでは。

ところが、50 を入れても、0.0025 秒と 5.2 分ではだいぶ違いますね。

で、もっと言うと、この 2 の N 乗と N の 2 乗はどっちが大きいかというと、小さいところでは 2 の N 乗のほうが大きいかも知れませんが、50 とか入れると、2 の 50 乗というとすごい数になります。絶対に電卓の中には入りません。

電卓に入るのは 2 の 10 乗くらいまでです。恐らく。それだけでもでかいですね。こんなにかかっちゃうんです。それで 3 の N 乗というとは全然違う単位になってるわけです。

・スライド20「離散数学が実用面で役立つこと」

なので、ここで重要なのは、巨大データを扱わなくてはならない問題が、今多々発生しているんですね。例えば先ほどのソーティングみたいなもの。

現在地球上で、1 億ぐらいのデータを扱わなければならない問題というのはいくつかあります。

1 億でなくても、1000 万ぐらいのデータの数を扱わなければならない問題もたくさんあります。

1000 万ぐらいなら、日本の人口の中で、5000 万人ぐらいの人がお金を稼いでいると思います。

その年収を多い順に並べなさいというような時、税金をいくら取るかなどを考える時、必要になってきますよね。

それぐらい大きいデータを扱うとなると、計算量、どれだけのステップが必要かというのは、だいたい N とか $N \times \text{LOG } N$ くらいじゃないと解けません。

これより大きくなるとどうしようもなく、時間がかかりすぎて解けなくなってしまいます。

先ほどのテニスボール問題の解法、あれはソーティングと言いますが、現状で、1 億はでかすぎますが、1000 万というぐらいの巨大データを扱う問題に対して、実用的に解ける限界のアルゴリズムがソーティングなのです。

あれぐらいしか、実は今のところ、大きなデータは扱えないわけです。このトーナメント方式 ($N \times \text{LOG } N$)、これぐらいしか扱えない。これがほぼ限界なんですね。今の大きなデータを扱わなければならない世界では。

従って、巨大データを扱う時は、トーナメント式みたいなものじゃないと採用できない。

トーナメント式は離散数学のエッセンス、アイデアをたくさん使ったものなので、単純に離散数学を使うことによって、誰でも思いつきそうな総当たり戦からトーナメント方式に発展して、しかも、それで大きなデータも扱えるようになるのですね。

ソーティングみたいなことは、離散数学の力がないと解けません。

・スライド21「パ・リーグの試合日程」

さて、もうちょっと実用的な話題にちょっとだけ入りたいと思います。

単純にプロ野球のパ・リーグの日程を考えましょう。何をやりたいかということ、単純に日程を組みたいと。

パリーグ、6球団ありますね、北海道、仙台、千葉、さいたま、大阪、福岡。

彼らのために、何とかして各球団が移動距離を短くする、飛行機で飛ぶ距離を短くするような日程を組みましょうという、それが問題です。

これが、離散数学を使うといろいろなことができるのです。

ちなみになぜパシフィック・リーグのほうがいいのかということ、パ・リーグの方が、各ホーム球場を見ると、仙台の楽天と、千葉ロッテ以外はドーム球場なんですね。だから雨で中止というのがあまり無いのです。ですから、スケジュールを作ったときに、雨天中止のために後ろにスケジュールを組まなくてはならないなどということがないので、問題設定としては簡単です。

・スライド23「プロ野球の球団本拠地」

それからもう1つ、問題が数学的に面白い。なぜかということ、パ・リーグのほうが日本全国に散っているんですね。 *スライド23 参照

セ・リーグは在京セ・リーグと言われ、3チームはほとんど移動距離がないに等しく、そこを解析しても面白くない。だから学問的にはパ・リーグのほうがちょっと面白いのです。

・スライド22「パ・リーグの試合日程」

さて、問題設定です。各チーム120試合。

単純に同一カード24試合、年間24試合で、ホーム12試合、アウェイ12試合を考えます。

24×5で120試合ですね。

今、実際プロ野球は、144試合なんですけど、なぜ120試合にしたかということ、残り24試合は交流戦です。

ちょっと交流戦は考えるのを止めましょう。

120試合の日程を作りなさい。移動距離を短くするように作りなさいと。

ただし、このような条件をつけましょうと。

まず1つのカードはもちろん3連戦までしかできません。

例えば巨人-阪神で12連戦したら、どっちのファンも食傷気味になりますね。

その後、2番目の条件として、パ・リーグは4カード連続ビジターで、アウェイに行くのは無しにしましょうと。セ・リーグはどうしてもこれが発生します。

阪神が甲子園球場を高校野球期間中、使えませんから。

そういう意味で、セ・リーグよりパ・リーグのほうが面白いんですけどね。

単純にホームゲームもアウェイゲームも9試合までしかできません。

あとは、なるべくひと月の間に、ホームの試合の数と、アウェイの試合の数を均等にしましょうという条件を付けました。この条件の下でなるべく移動距離を短くしましょうというのが問題です。120試合の。

・スライド24「パ・リーグの試合日程」

さっきの3連戦まで等の条件がなければ、最適回は簡単なのです。単純に12試合連続、日本ハムと仙

台を札幌ドームで戦わせて、楽天は札幌ドームでおしまい、それで移動するという、これが一番単純でいいんです。なぜかという、もちろん移動回数がなくなりますからいいに決まっています。でもこれはやめましょう。こういうことを許すとフェアでなくなるし、ファンもずっと同じ試合を見なければいけないわけですからね、こういのはダメですと。

・スライド25「新しいスケジュールの提案」

そこで、新しいカードスケジュールをこれから提案するんですけども。

これは私のポストクと一緒にやっているんですが、どういうふうな形で提案するかというと、メジャーリーグ方式を採用します。メジャーリーグの基本方式は9連戦、9連戦なんですね。

9試合まず連続でホームでやって、その後、9試合アウェイに行つて、今度9試合またホームに戻つて、また9試合アウェイに行く。このような方式を基本にしているわけです。

とにかくどのチームも、ホームとアウェイの試合の数を消化した時点で、6試合以内にしましょうと。ちょっとテクニカルな部分なんですけど、何を言っているかということ、最初にずっとアウェイ、アウェイ、アウェイ・・・ビジター、ビジター、ビジター・・・で、最後のほうにホーム、ホーム、ホーム・・・みたいなことになるのはやめましょうと。

そうすると、最後のほうで例えば優勝争いをしている巨人ー阪神、全部6試合連続甲子園みたいなこともありえますので、そのようなことはやめましょうと。

あとは、どのカードも消化試合はなるべく同じにしましょう。

例えば阪神の例で言うと、阪神ー巨人、阪神ー横浜、阪神ー広島、そういうのが5カードありますけど、例えば、阪神ーヤクルトは20試合消化しているけれども、阪神ー巨人は14試合しか消化していないというようなのはやめましょうと。

なるべく同時に消化していくようにしましょう。これが提案したい方法なわけです。

実は、これがちょっとネックです。

結構難しくしています。現状のスケジュールだと、6試合差が付くことがあるんですね。

・スライド26「今シーズンのパ・リーグの対戦スケジュール」

現状のスケジュールをお見せします。小さくなって申し訳ないのですが。

最後の5試合、38から42までをご覧ください。すみません、37からですね。

それを見ると、楽天は、37の状況で福岡と終わってるんですが、その後に北海道日本ハムと6試合残してるんですね。こうすると、いろいろな面で不利なことが起こります。

営業面で言えば、日本ハムと楽天が優勝争いをしていけば良いのですが、両方とも下位に沈んでいたら単純に消化試合でしかないわけです。

それを最後のほうでたくさんやってもあまり意味がない。

もうひとつは、これが、カードの妙で、北海道と楽天がこんな短期間で2回も対戦してしまうと、日本ハムが大分損なわけですね。

例えば、田中ー岩隈が6試合のうちに4つ出てきてしまったりするとそれで、だいぶ不利になってしまうわけです。こういうスケジュールはなるべく避けようというものです。

新しく試合を立てましょうと。基本は、3 試合、3 試合。最初の 10 カード、赤いのがホームです。

1 のところは、千葉と埼玉西武が対戦しているんですが、1 カードに 3 試合あります。

単純に提案する手法というのは、大体 3 つ、3 つでやってるカードです。

これを中心としてカードを提案して、後は先ほど言いましたように、なるべく 1 月でのホームとアウェイの消化試合を一緒にしたり、同一カードはなるべく一緒に進行したりということを考慮したカードです。

何がいいかということ、結構削減できるんですね。大体平均して 20% ちょっとくらい削減できる。

単純に何をやったのかということ、 *スライド 26 参照

今までの現行のカードは、お見せしますが、2 連戦とか、2 連戦、2 連戦とか、2 カード連続ホーム、2 カード連続アウェイ。

または東北楽天に至ってはひどく、アウェイ、ホーム、アウェイ、ホームと繰り返しているんです。

東北はかわいそうに、ここ全部移動しないとイケないんです。

こういうことを、なるべく避ける方法でスケジュールを組んだのが、このようになりました。

*スライド 27 参照

9 連戦、3 カード連続ホーム、3 カード連続アウェイを中心にするので、だいぶ削減できることが分かりました。

因みに、このスケジュールは結構良いかということ、ちょっと実は問題があります。

いくつか問題点をお話したいと思います。

何が問題点か。

千葉を見て下さい。最初の 5 個で、残り 5 チームと当たっています。

最初の 3 連戦はホームでそのあとアウェイ。全部と当たってるわけです。

次の 5 つのカードも、全部と当たっています。

いいんですけど、問題は何が問題かといいますと、千葉対東北楽天を見ると、両方ともアウェイになっちゃうんですね。

両方ともアウェイになってしまう、ということは、どこかに歪みが出て、両方、ホームというのが出てきます。

つまり、このやり方だとどのカードも大体、並行して進んでいくが、ホームとアウェイのバランスがちょっと崩れてしまう。

もうちょっと具体的に言うと、千葉とさいたま、千葉とオリックスは消化するゲームは同じくらいで進んでいきますが、残念ながら、千葉と東北は最初の 6 試合、全部仙台で試合をしないといけなくて、千葉とさいたまは、最初の 6 試合、全部千葉で試合をしなければならないという、そういうずれが出てきてしまう。

そういう意味で、このスケジュールはある意味、完璧ではないんです。

これが営業面では相当響くと思います。

最初と最後、開幕すぐと終盤は、ある意味かき入れ時なんですよ。

始まる時は、注目度が高いし、終わりは優勝争いで人が集まる。それを考えると、

最初の時にホームが集中してくれたほうが、いっぱいお金が入る。

ある意味、営業側としては嬉しいわけですね。

後ろのほうにいっぱい、ホームが集中すれば嬉しいわけです。

そういう意味でこの新しいスケジュールによって、移動距離は削減されるし、エコにもいいんだけど、営業的にこれがいいかどうか、パリーグのチームたちが採用するかは、ちょっと今のところまだ分からない。残念ながら営業面を優先させると、まだ完璧じゃないように思います。

さて、私の研究の一部を紹介するのは、時間の関係で申し訳ないですが端折ることにさせていただきます。最後にまとめに入らせていただきたいと思います。

* スライド 29~33 割愛

・スライド34「これからの離散数学ー離散数学が向かう未来ー」

今日、本当に言いたかったのは、ここに書いてあります。

離散数学というのは、実は学問としていろいろピンと来るようなアイデアを使って、数字じゃなく、トーナメントみたいなアイデアを使って問題を解く。

ある意味パズル的感覚があって面白いですが、離散数学をやることによって、コンピュータのさらなる高速化、コンピュータ自体ではなく、コンピュータの動作の高速化に貢献できたりする。さきほどのソーティングなどの例ですね。それから営業面を無視すれば、スケジューリングもかなり効率よくできて、25%、30%のコストダウンにつながりますよと。

あるいは、私の今やっている研究では、どちらかというたとえばGPSやカーナビなどの情報のアップデートにも貢献できます。

あとは、皆さんインターネットが使われると思いますが、Googleなどの検索の仕方も、離散数学が非常にたくさん使われています。

他にも沢山あります、離散数学の使われている例が。

今日はほんとに氷山の一角しか話せませんでした、少しでも興味を持っていただければと思います。残念ながら離散数学の入門書は、あまり多くありません。

もし興味がある方は、最初のページに私の e-mail アドレスが書いてありますので、連絡してくれば個別に対応したいと思います。

長い間、どうもありがとうございました。

(拍手)