

いや、実のところはねえ

宇野 毅明

共同研究は狐と狸の化かし合い。というのは嘘。でも、知らない人が集まり、互いに自分の利益を得るよう行動した場合、コミュニケーションが悪ければ自然とそうなるであろう。お互いに Win-Win の関係を構築し、互いに意義のある成果を得られるような問題設定をし、負担の大幅な偏りがないよう研究を進める。理想的な話ではあるが、このような共同研究を目指したい。とはいえ、目指すだけ、志を持つだけではうまくいかないこともまた事実。本稿では、筆者の共同研究体験談を通じて、共同研究でのより良い関係を作っていくためのコツを紹介したい。

キーワード：共同研究，応用事例，問題設定，研究打合せ，出張

時は 2002 年、ヒトゲノムの読み取りは完了し、まさに大規模データ時代の幕開けが近づいていた。ヒトゲノムの大きさはおよそ 3GB、このような大規模データでは、アルゴリズムによる高速化は 1 万倍以上の高速化をもたらし、革命的な技術革新を起こす可能性を秘めている。列挙アルゴリズム高速化で質の良い問題を見つけることができずにいた私にとって、ゲノムの世界は非常に魅力的に見えた。文字列アルゴリズム、ネットワーク、類似性、...。ゲノムの持つある種異質な構造が、さまざまな問題を導き出すはずである。アルゴリズム分野にも、私の所属する情報学研究所にもゲノムの研究者は点在し、ゲノム関係の研究集会も頻繁に開催されていた。

定理 1：目標は実現できない

ゲノム問題を手に入れるためには、まずゲノムの人と話をしなければならない。幸運なことに、私は第一線で活躍する研究者と談義する機会を何回か得ることができた。見知らぬ研究者が研究について議論を行うとき、双方合意できる談義の目的は、良い研究テーマを見つけ、共同研究の可能性を探ることである。そのため、お互いの研究内容と目的意識を交換するところから話が始まる。話してみるとわかるゲノム研究者の根源的な目標は、生命の真理に迫ること。ヒトゲノムに限定すれば、人間の生命活動のシステムを、ゲノム配列の面から理解することである。しかし、この目標はあまりにも遠い。付き合っていたら来世までかかるであろう。いや、来世でも無理である。世界を征服し

たい、のほうがよほど現実的だ。そのあたりは適宜相づちを打ち話を進め、より具体的な目標に絞り込んでもらう。すると、タンパク質の働きを理解したい、遺伝子が発現する条件が知りたい、ゲノム配列から働いている遺伝子の部位を特定したい、などが出てくる。ここまで具体化されると、ある程度アプローチできる気分になる。しかし、研究計画に思いを巡らしていると、はたと気づく。情報科学的な技術でアプローチできるものはほとんどないのである。彼らの目標は、彼らの研究者としての興味から出てきたものであり、情報に関わるかどうかなど関係ない。われわれにはやはり距離のあるものなのである。

「で、実のところはどうなんですか？」このような話は適当に聞き流し、さっさと具体的な話に入る。これが話し合いを迅速に進めるには不可欠である。しかし、これはお勧めできない。共同研究は、共同でタスクを設計することが重要であり、そのためにはお互いの状況をよく知る必要がある。最終的なゴールや、研究の価値観も理解せず、短期的な方法論だけでタスク設計を行って、はたして意義のある共同研究ができるであろうか？ 一見無意味に思えるこの大目標の設定と共有。これはエンジニアや数理をはじめとする OR の研究者に欠けていることが多く、ついつい短期的、即物的な理解で課題を設定してしまう。まずは相手の問題意識を大局的な観点から理解し、その上で問題の設定を行うことが重要なのである。

では振り返って、自身に大局的な目標や技術体系があるだろうか？ これがなければ、相手は自分の芯の考えを理解したパートナーとはなり得ない、つまりは良い問題設定のための情報を提供できないということである。常日頃から自分の大目標と世界観を醸成し、他

うの たけあき
国立情報学研究所情報学プリンシプル研究系
〒101-8430 東京都千代田区一ツ橋 2-1-2

人の大目標や価値観を理解するよう努めることが重要なのである。

定理 2：求められることはできない，求めることもできない

さて，ある程度の相互理解はできたでしょう。次に具体的な問題の検討に入る。ここまでくると，データは何がある，技術は何がある，何ができるとうれい，というようなパーツがそろってきており，議論がしやすい。さまざまなアイデアを出してその可能性を検討する。うまく回れば，非常に楽しい時間が過ごせる。しかし，可能性の議論は空回りすることが多く，なかなかピシッとままることはない。ここには実働と可能性に関する感覚の大きなギャップが存在する。

計算機科学の外側にいる人間は，計算機の研究をしている人はなんでも計算できると思っている。例えば，必要と思われるデータがそろえば，結果はちゃんと出せると考える。ある意味で正しいが，正しくない。まず，世の中には計算できないことがある。最適な政策，遺伝子と生命活動の因果関係，こういったものはまず計算できない。どんなにデータがそろっても評価値やメカニズムがわからない以上無理なのである。この，数理的なモデリングの可否の部分ですっ飛ばした要求は意外と多く，計算とモデル化がごっちゃになっている現状を伺わせる。また，手法を研究する研究者というのは，新手法の開発と実装がコスト 0 でできる，あるいはすでにできたものを持っていると思われる。ちょっと持ってくればすむでしょ，という感覚だ。しかし実働するほうとしては，必要となる労働力に辟易する。しかも，労力を掛けても実にならない可能性が大きいとなればなおさらだ。

こういった場合，相手は労力を惜しむことに首をかしげる。なぜやらない？ 実験系分野では，実験は失敗して当たり前。積み重ねで結果を得る。そのため，実験をしたこと自体も（簡単な査読を経て）業績となる。実験に対するリスペクトがある。しかし，理論系・工学系は実装に対する理解は薄い。単なる実装は研究ではないと思われる。労力は可能性の高いところのみにつき込みたい。その一方で，相手の目標は高い。生命の理解，そのためには 10 年はかかるだろう。骨を埋める覚悟なしではいい研究などできやしない，と来る。ここで，双方の大目標の理解は重要である。大目標が異なる人達が集まっている，とわかっていれば，導く結論も自然と落ち着いたものになるであろう。

逆に，情報系の研究者は現場の実態を知らない。デー

タがいくらするか，実験がどれほどの労力を伴うか，ぼつとでの研究者が出したアイデア，価値基準（モデル）がどの程度の説得力を持つか，想像がつかない。情報科学的に美しく，価値のある問題設定でも，彼らの興味を引くとは限らない。そのために高額（ときに 1,000 万！）なデータや機械を買い，多大な労力（1 回の実験に 1 カ月！）をかける意味はなかなか見いだせないであろう。「実のところは，そんなこと馬鹿馬鹿しくてできない」のである。われわれも，自分たちの価値観で相手に無理強いをしてはいけないのである。

定理 3：小さい研究は大きくならない

議論の末，具体的な問題が決まる。このとき，たいていのっているのは相手方だ。テーマは相手方の本流であることが多く，タスクは非常に具体的だ。1 年後にどのような結果が出るか，だいたい想像できる。しかし，具体的であるタスク，というものは，往々にして OR 系の人間には「雑用」であることが多い。研究課題を OR 的な視点から見てみると，単なる応用事例にすぎず，作業は単なる実装であることもしばしばだ。それでも，計算機科学的になんらかの困難があるから，彼らは解けなかったのである。

このような関係は，正直長続きしない。そのうち関係がドライになり，やがて飽きが来て消滅する。利害関係，しかも偏りのあるもので結びついた人間関係は壊れやすい。強いつながりは共通の目標を持ったときのみ生まれる（その意味では，雑用的な研究でも高い共感が得られるのであればそれは幸せである）。しかし，両者に意味のある成果が期待できる課題設定が重要ではあるが，そのような「幸運」に巡り会えることは少ない。

このようなときは，課題を完全に解決しようとするのではなく，上質な部分のみを解くようにするほうが良い。しかし，小さな課題のさらに一部では，インパクトの大きい研究は期待できない。一步下がって部分課題を大局的な視点で眺めることが大事である。一步下がり方にはいろいろあり，実応用や実装でユーザ満足度を高める方向，構造を持つ部分のみを抽出しての定理証明を行う，など多岐にわたるが，私は多くの研究者にインタビューを行い，多くの研究課題で部分問題として現れる，あるいは多くの課題に関連している課題を探した。

構造推定や予測など，数多くの情報処理課題が含有されたが，一番計算資源を消費していたものは類似検索であった。ゲノムの一部の機能推定を行うため，ほ

かの種のゲノムの類似する部分を調べる、意味がないと思われるウィルスの痕跡を取り除くため、ウィルスの配列と似たものを取り除く、シーケンサーで読まれた短いゲノム配列を重ね合わせて伸張し、ゲノム全体を解読する（アセンブリング）といった処理すべてに類似検索が関わっている。類似検索は一般に困難な問題であり、決定打は今のところ存在しない。しかし、ゲノム配列は4種の文字しか含まない、類似性が文字の表現で機械的に定義される、多量の質問が一度に来る、など、この分野ならではの計算上の利点もある。聞いてみると反応がよい。これができれば計算が軽くなるし、多種の生物をすぐに比較できる。きっといろんなことがわかるはず。それこそ1,000倍速くなったら革命が起きるんじゃないかな、と言われる。一般的な問題で、根底に共通するものを見つけると、直接的に課題を解決することはなくとも、これは使えそうだな、という感覚は持ってもらえるものである。

定理4：素晴らしいものは役に立たない

類似検索は確かに難しいが、有利な条件もある。アルゴリズムの研究者にかかれば、ほら、このとおり。かくしてアルゴリズムは順調に開発され、染色体同士の比較など大規模な比較問題は1,000倍どころか1万倍近く速くなった。これは素晴らしい、多くの研究に役立つに違いない、うん。心をときめかせて、ゲノム研究者のところに見せに行く。さてさて、反応は、と期待していたのだが、しかし実際に見てもらったときの反応は「いや、実のところはねえ、ここが速くなくてもしょうがないんですよ」と来た。はあ、すごい研究ができるんじゃないかなかったですか？アルゴリズムが速くなるとハードが売れなくなるってどういうことですか？

さて、何が悪かったのか。彼らに話をよく聞いてみると、結果が受け入れがたいところにはいくつかの原因があった。一つは、デファクトスタンダードの類似検索と、類似性のモデルが微妙に異なっていること。これは、エンジニアリング的には「結果が妥当ならいいじゃないか」と考えられるのだが、実際には引っかかるところがある。例えば実験を行った際、過去の結果を参照・比較しようとする、どうしても検索方法を同一にする必要が出てくる（さもなければ比較不可能となる）。そのためには「同じ方法で検索」する必要がある。再現性や共通性を必要とする課題では、受け入れられない。

また、類似検索は基礎すぎて、自分のやりたいこと

から距離がある、という点がある。類似検索を研究で使えるようにするには、類似検索をデータベースに組み込み、使い勝手を向上させる必要があるのだが、これは非常にコストがかかる。データベースと独立した課題、例えば染色体の解析などでも、やりたいことは「検索を使って行う何か」なのである。それを一から構築するだけの技量と余力はないのである。

では、どうすれば良かったのだろうか。基礎的で共通する課題を選んだところは良かった。その次、新しいことを始める前に、その具体的な利用シーンまでも考え、何が受け入れられるか、何がだめなのか判断し、有効な場面に対する簡単なアプローチの手段を提供すべきであったのだろう。つまり、問題のモデルを変える必要があるのであれば、モデルを変えても受け入れられるところはどこだろうか、という思慮が足りなかったのである。実際、アセンブリング、アライメント、断片配列のマッピングなどの課題では、既存手法のモデルがどうであるかは隠蔽され、その変化は気にならない。このようなアプリケーションを前面に出し、簡単なプロトタイププログラムを作って実行結果が出るようにすれば、違う受け止め方をされただろうと思われる。

さらには、相手の要求をそのまま考えた、というのもよろしくなかった。研究分野、個人研究の目標から価値基準や総合的な問題意識まで理解し、そのうえで相手が必要とする課題をこちらがデザインするようなレベルまで話をすべきであった。われわれは、ORの人間である。つまり、相手の持っていない視点を持ち、どのような問題が解けるか、解くべきであるかを知っているのである。相手のリクエストの共通部分を見ただけでなく、相手が本当にやりたいことをOR的な視点で捉え、その共通部分を取ることが必要だったのである。

定理5：相手は本当の課題を話さない

さて、このようにしてみると、課題設定はどのようなものか、なんとなくご理解いただけたと思う。つまり、共同研究を行うとき、われわれはコンサルタントなのである。コンサルタントは、相手のリクエストをそのまま受け止めてはいけぬ。相手の心の底にある本当の要求をくみ取り、それに合わせて問題設定を行うのである。相手は自身の問題がわかっていない、と考えて話をすべきであろう（自分の視点から見て、であるが）。

家の建築を例に見てみよう。これは、家を建てる依

頼主と建築家の共同作業である。依頼主は、自身が建てようとしている家に対してなんらかのイメージを持っている。リビングは広く、窓は大きく、階段は登りやすく、キッチンはキレイに…、間取りに関してもイメージがあるに違いない。しかし、それをそのまま「はい、そうですか」と実現してしまつては、実際に住んだときに大きな問題が発生する。依頼主は建築のすべを知っているわけではないので、依頼主の設計した間取りが実際に生活するうえでの利便性を兼ね備えているとは限らない。建築家は、依頼主の要求を整理し、優先順位をつけ、あるいは組合せ、それに依頼主の生活状況を加え、生活動線の整理や収納の設計、室内環境の向上などを鑑みて、家具の選択と間取りの提案をする。依頼主から見れば、自身の要求と異なる物を提案されるわけであり、受け入れるには抵抗がある。しかし、ここで建築家からなぜこのような間取りにしたかについて、生活にフィットした建築の観点から説明すれば、依頼主は自身の考えに欠けていた物を発見し、提案を受け入れ、あるいは代案を提案するだろう。共同研究における OR の研究者の役割も、建築家と似ている部分がある。つまり、共同研究の相手が心の底に持つ問題意識に対して、OR 手法の観点から適切なモデルを提案するのである。

このような芸当は、一朝一夕でできるようになるものではない。が、しかし、習得を諦めるべきものでもない。これらのことはまさに問題意識から問題定義を行うプロセスそのものであり、読者の皆さんも OR の

研究者を目指したその瞬間から、このようなことを心がけてきたはずなのである。あとは今ひとつの意識の問題であろう。常日頃からより広く多角的な視野を持ち、他分野の問題意識を理解するように努めれば、自然と身についていくものと考えている。

定理 6：ヒットは足で稼げ

さて、ではこのように共同研究を通じて自分の研究を展開していくにはどのようにすればよいか。一番手取り早い方法は「積極的にいろいろな場所を訪問する」である。シンポジウム、研究会、展示会、興味を持つ他分野の会合に積極的に顔を出す。そこで知り合った人を捕まえ、議論する。懇親会があるなら、その場で話をする。信頼関係が構築できたなら、実際に研究室や企業を訪問し、より深い研究打合せを行う。このようなプロセスを続けていくことで、自然とコンサルタントに必要な素養が養われ、共同研究の実績も増えていくことになる。

幸いにも、OR の研究者は大がかりな実験設備につきっきりになる必要がない。時間の都合さえつけければ、機動力で勝負できる。逆に言えば、実験設備や高価なデータという資産を持たないわれわれには、頭の回転と機動力を活かすしか生きる道はないのである。共同研究という視点に立てば、われわれは自分自身の評価を OR の中で閉じた尺度で行うことはできない。比較相手は常に他分野である。OR の枠を超えた大きな研究者を目指して精進していきたいものである。