

タイトル	北海学園大学人文学会・人文学部・文学研究科主催パネルシンポジウム AIがヒトを超えるとき：相剋から共生に向かうために
著者	森，洋久；MORI, Hirohisa
引用	北海学園大学人文論集(64)：141-169
発行日	2018-03-31

AI がヒトを超えるとき —— 相剋から共生に向かうために ——

講演者 国際日本文化研究センター 森 洋 久 氏

○司会 皆さん、こんにちは。

時間が参りましたので、2016 年度人文学会「AI がヒトを超えるとき」を始めたいと思います。

きょうはお忙しい中お集まりいただき、まことにありがとうございます。

それでは、まず最初に、上野人文学会長から挨拶をもらいたいと思います。

○上野会長 皆さん、こんにちは。

今年度から人文学会長になりました上野と申します。どうぞよろしくお願ひいたします。

今年度のこの人文学会は、ゲストとして国際日本文化研究センターの森洋久先生をお招きして御講演をいただく企画になっております。その後、パネルシンポジウムの時間といたしまして、学部を超えた本学の先生方にも御登壇をいただいて、フロアの皆様とともに人工知能とヒトの関係について議論をしていまいりたいと思っております。

ポスターにもありましたけれども、最近、あちらこちらで「2045 年問題」という言葉をよく耳にするようになってきました。新聞報道などによりますと、今の技術革新がこのまま続いていくと、コンピューターは、あと 30 年もすると、全人類の情報処理能力を超えてしまうと、そういう予想をしたものが「2045 年問題」だそうです。それによって、人類の仕事の大半が機械によって取ってかわられるのではないかと、そういう心配がされてい

るところだと思います。

ただ、個人的には、そのときでも、やはり人間にしかできないような仕事というのが残るのではないのかなという、そういう楽観的な見方もちょっと私の中にはございます。

また、人間が機械に取ってかわってくれるということになれば、我々には、余裕ができるわけですし、余暇といいますか、暇な時間ができるわけです。それをどう生かすのかということが、そのとき問題になって、そのとき我々人間はどうやって生きていくのか、あるいはそういう機械に囲まれた中で、我々はどうやって生きていくべきなのか、そういったことを考える時代が来るのではないかというふうにも思います。

そうすると、まさに人文学の、そこで出番になるのではないかと。人文学の時代がやってくるのではないかと、そういうふうにも思ったりしています。そんなふうによくいくかどうかわかりませんが、ちょっと楽観的な感じを私は持っております。2045年、私が生きているかどうかわかりませんので、そんな適当なことも言えるのかもしれませんが、そんなふうに思っています。

今、現場では、実は2018年問題という、同じ問題でもちょっと違う問題で、受験生がどんどん減っていくという問題を間近に控えていまして、戦々恐々としているのですが、きょうは、そういうことはちょっと置いておいて、ちょっと未来の話、2045年に向けて、我々はどう考えていくべきなのかということについて、考えていく時間になればなというふうに思っております。

この後、森先生に御講演をいただくわけですが、どういうお話になるのかわかりませんが、5時までの短い時間、限られてはいますけれども、皆さんと一緒に有意義なひとときとなりますよう祈念して、挨拶とさせていただきます。

どうもありがとうございました。(拍手)

それでは、この後、柴田先生が企画司会をされますので、柴田先生にバトンタッチということにさせていただきます。

○柴田氏 本日、司会を務めます人文学部の柴田と申します。

森先生の御講演に先立ちまして、簡単に趣旨の説明をいたします。

人文学会、これは2013年に正式に発足いたしました。人文学部の教員を主な構成員とする集まりであります。そのときに開催いたしました記念シンポジウム「人文学の新しい可能性」から数えますと、今回は、4回目のイベントです。人文学会、実は、発足以前も毎年、人文学部教員による講演、シンポジウムを開いて、向学の間、これを設けてまいりました。これは、人文学部の非常によき伝統だと思います。

今回は学部を超えて、さらに大学を超えて、テーマにふさわしい先生をお招きし、より広く、より深く、学ぶ場にしたいと思います。本日のテーマ「AIがヒトを超えるととき——相剋から共生に向かうために——」。幾つか、キーワードについて、簡単に御説明いたします。

まず、「2045年問題」ですが、これは、未来学者であり、発明家であるレイ・カーツワイル、この方が大々的に使っている言葉というふうに言われていますが、例えば2007年、NHK出版から出ました『加速するテクノロジー』という本には、次のような定義があります。

2020年には、1台のコンピューターが1人の人間の知性を凌駕する。そして2045年には、人間の知能の10億倍の能力を持った人工頭脳が登場し、人類とテクノロジーの関係が特異点、人間の能力が根底から覆り、変動するときのことを特異点と言いますが、特異点に達すると、こういうことを言っています。特異点は、英語で言いますとシンギュラリティですが、特異点以後の世界では、人間と機械、物理的な現実とバーチャルリアルティーの間に区別がつかなくなるということを述べています。

このアイデア自体は、さらに前の2006年の、日本語の題で言いますと『ポスト・ヒューマン』という本に出てきます。原題が『The Singularity is Near』という本です。特異点を理解しますには、例えば、幾つかキーワードを確認しなければいけません。その一つが、収穫加速の法則というものがあります。

これは、いかなる法則かと言いますと、ムーアの法則を前提にしないと

理解できない。これは、後ほど森先生から詳しい御説明がありますが、ムーアの法則と言いますのは、CPUのインテルの共同創業者のゴードン・ムーアが、1965年の論文に書いた法則で、およそ18カ月で半導体の性能が倍になるという、こういう法則です。収穫加速の法則と言いますのは、ムーアの法則を超える速度で加速度的に技術が進歩する。つまり、進歩の速度はだんだん早くなっていくと、こういう法則です。これがまず前提になっております。

でも、私の専門は、ポスターですとメディア論、メディア思想というふうになっておりますが、広く言いますと、技術思想というのを専門にしております。実は、古来から人間と道具、あるいは機械との分業というのは、文明を越えて議論されてきた問題であります。古くは、人間の記憶を代替する発明である文字について、プラトンが語ったものにさかのぼりますし、20世紀に入りますと、人間とコンピューターの分業の議論が大体1940年ぐらいから始まります。さらに、人間の器官を機械で置きかえるサイボーグ論など、特に20世紀に入ってから、この分野の議論は百花繚乱の感があります。

さて、私、そのような専門なのですが、「2045年問題」というものについて、どのような認識を持っているかと言いますと、恐らくこういった分業の延長線上にある21世紀の危機であろうという予感はあるのですが、実は、まだフィクションと現実が混在すると言わざるを得ない状況です。一体これは単なる想像なのか、あるいは統計的な外挿による予測と言えるのか。

さらに、昨今、新聞等でにぎわっておりますビックデータ、これが一体どれほど利用可能か、ディープラーニングがどれほど画期的か。そもそも、知性とは何かというのは、実は厳密にはわからないというふうと言わざるを得ません。ただし、一つ確かなのは、特異点なるものが実際に2045年に来るか否か、それにかかわらず、自動運転、さらに人間の棋士を負かすようなAI、これが現在、登場しつつあります。現実人間と機械の関係を大きく変えていく、これだけは間違いのないだろうというふうに考えます。

かつて、産業革命の時期に、機械化にあらがって機械を打ち壊した運動がありました。いわゆるラッドイト運動、イギリスで起こりました。職を失った職工が織機を壊して回るといようなことをやったわけですね。しかしながら、歴史を振り返ってみますと、この種の運動が成功した例というのはありません。必ず人間は敗北します。しかし我々、ラッドイツになることなく、さりとて、空想上の AI や AL に憧れたり、あるいはいたずらに恐怖するのではなくて、本日のテーマにもありますように、機械との共生を考えていきたい、このような立場をとっております。

その議論の基礎を共有するために、まず第1部で、情報工学を御専門にされていて、なおかつ AI について非常に造詣の深い森先生に御登壇いただいて、現在の技術水準、state of the art を踏まえた講演をいただきたいと思います。

簡単に、御経歴を紹介いたしますと、現在、国際日本文化研究センターの准教授の職にあられます。東京大学在籍時に、日本初の OS であるトロンの開発に携わるとい御経歴があります。現在、国際日本文化研究センターで資料のデジタルアーカイブ化からデータベースの構築、活用を担当していらっしゃいます。音楽にも造詣が深く、オペラを企画、作曲されるということです。

御講演を、この後 14 時 15 分から 15 時までの 45 分間いただきます。その後、10 分の休憩を挟みまして、第2部、ここではパネルディスカッションを行います。森先生のお話を受けまして、学部を横断して御参集いただきました4分野の先生に、それぞれの御関心から「2045 年問題」が投げかける課題と展望について御提題いただきます。

本日、御登壇いただきますのは、もう既に会場にいらっしゃいますが、お名前のみ御紹介いたしますと、立っていただきますかね。お話しいただく順番ですけれども、まず、経済学部の水野谷武志先生、どうぞ拍手を（拍手）、それから法学部の山本健太郎先生（拍手）、工学部の竹内潔先生（拍手）、そして人文学部の佐藤貴史先生（拍手）です。

今、御紹介いたしました順にお話をしていただいた後に、登壇者による

パネルディスカッションをまず行います。このディスカッションを経ました後、会場の皆様と質疑応答の時間を設けます。この場にいらっしゃる全員が参加して、この問題を議論し、理解を深める場になるように進行してまいりたいと思います。どうぞよろしくお願いいたします。

森先生の講演に先立ちまして、幾つか事務的な御連絡があります。

本日のイベントの内容ですが、本学部が発行いたします『人文論集』に掲載する予定です。登壇の先生方には、既に御了解いただいておりますが、質問される際には、その点を了解して、つまり、活字になるということをお御了解いただいて、質疑に加わっていただければと思います。

それから、17時にイベントを終了いたしますが、終了後、17時20分から本学生協で懇親会があります。参加費等は、これはいたしません。もう気軽に議論の続きをぜひやりたいという方には立ち寄っていただいて、立食しながら気軽に議論の続きをしていただきたい。ぜひお立ち寄りください。これはとくに出席等はとりません。

最後ですけれども、この企画に関しまして、ESRI ジャパン株式会社に企画協力と後援をいただきました。同社札幌オフィスの福田様、福田様に感謝申し上げます。(拍手)

では、森先生、御講演をお願いいたします。

○森氏 御紹介にあずかりました森洋久です。今回は、この講演にお招きくださった関係者の方々に感謝を申し上げます。ありがとうございました。

それでは、講演を始めたいと思いますが、きょうのテーマとして、講演のやり方ですが、お手元に2016年度人文学会「AIがヒトを超えるとき」というレジュメがございます。前回は、OHPだけを使って行ったのですが、今回はちょっと、いろいろ詳しい内容も書きたいというふうに思いまして、レジュメを用意させていただきました(添付資料参照)。グラフなどはレジュメのほうのが白黒なので見にくい部分もあるかなと思ひまして、グラフなどはここに表示したり、両方を使ってやっていきたいと思ひます。

今回のテーマが「2045年問題」と、前回の場合も「2045年問題」という

ものも想定しまして、人間の感覚と AI、クオリアの問題とか、あと中国人の部屋という、AI というものが本当に知能なのか、それとも、単に機械的な変換を行っているだけのものなのかというようなことを一つの問題提起として、いろいろそれに対して、さまざまな方面から考察を入れるということをやってまいりました。

きょうは、今度は、カーツワイルが言っている、先ほど柴田先生から提起がありましたけれども、2045 年に向けて、2045 年にコンピューターの能力が人類を超えるという、そういう予測という特異点、シンギュラリティというものを、私のちょっと独自の方法で批判していこうというのがテーマです。

私自身の専門、日ごろなりわいとしていっているところですが、ちょっとその辺について、多少詳しくつけ加えたいと思います。日ごろは、AI にプロパーな仕事をやっているわけではなくて、先ほども紹介にあずかりましたように、日文研というところで、人文系的なさまざまなデータベースを構築しているという仕事が主軸にありまして、私自身のプロパーなところというのは、ネットワークシステムなのです。TCP/IP ですか、あるいは Web の HTTP とか、そういったようなプロトコルを開発したり、ブラッシュアップしたりして、それをそういったデータベースなどに応用していくというところが、私の仕事です。

そういった中で、データベースといっても、最近では、データベース＝インターネットという時代です。そうなってくると、単純に日文研だけでデータベースをつくっているだけでは、仕事というのは間に合わなくなって、いろいろな同じような研究をやっているような組織ですとか、そういったところのデータベースとか、アクティビティとコネクションを持ちながら、インターネットの中でコネクションを持ちながらデータベースを構築していかなければいけないと。ネットワーク上、インターネットを使って、データベース自体もネットワーク化されていくというような時代にだんだんなっているわけです。

そういった中では、非常にネット等と深いつながりがある仕事で、イン

ターネットというのは、実はかなり AI との関連性を持っているわけですね。どういうふうに持っているかという、そもそもインターネットの生みの親と言われるスイスにいます原子力研究所の研究者ティム・バーナーズ=リーという人が Web システムというのを最初に考えたのですけれども、そのときに、既に Web というものがネットワーク化されていくということは、それは知識と知識をつなげることであって、そして、知識をつなげていくという先には、その知識を使った人工知能というのがあり得るのだと、ネットワークオートマタという言い方で、そのときはそういう考え方を提示しております。

実際、その後、その考え方というのは、現実化していこうとする人たちがいまして、いろいろ名前を変えるのですが、例えばセマンティック Web というような言い方であったりですとか、オントロジーというような言い方をして、そういった非常に知識をベースにして、AI チックにいろいろ知識を加工して人間に提供していくという考え方をするようなシステムというのがいろいろ開発されていっています。データベースの中でも、それを検索に応用したりとか、例えば Google の検索エンジンなんかでも、単純にデータベースを引っ張ってきて、データ表示しているというだけではなくて、やはり、そこには、どういう検索語が検索されたかというような履歴をためて、そこに AI 的な手法でデータをブラッシュアップして、検索者に対して提供していくというような考え方があるわけです。

そういった意味で、データベースの仕事をしていると、AI なんて関係ないというふうに言ってはられないというような時代にだんだんなってきたわけで、ネットワークプロパーであるけれども、いろいろ AI について調べていこうというようなところから、AI に対して興味が始まっているというふうなところです。

ちょっと前置きが長くなりましたけれども、実際、きょうはお集まりになっている方々も、もともとは AI プロパーでない方というのはかなりいらっしゃると思います。その中で、現在、AI に何らかかかわっている方も、あるいはこれからかかわろうと思っている方も、あるいはそうでもない方

も、ある種社会的にインターネット的にそういうふうな AI にかかわらざるを得ないと、あるいは隣にそういう AI に、例えば胸ポケットに入っているこういう携帯にもう既に、例えばこれだと Siri という人工知能が入っていたりとか、生活の中にそうやって入ってきてしまっているのか、かわらざるを得ないという方は非常に多くいらっしゃるのではないかと思います。

それでは、まず 2015 年問題批判をする前に、「2045 年問題」というのはどういうものなのかというところを、ざっと見たいと思います。レジメ（文末に掲載）のほうには、 $\lim_{\xi \rightarrow +\infty} e^{\xi} = +\infty$ という式が書かれているのですが、この OHP で、その式を打ち直すことはできなかったため、ポチポチになっています。これは、最後の章と 5 ページのところの 7 番の章と、実は対になっております。

さっき、 ξ をプラス無限大にしたのですけれども、こっちはマイナス無限大になると。そうすると、 e の ξ 乗はゼロに収束していく。これは何を意味しているかというのは、読んでいくうちにわかってくるのではないかと思います。まずはプラス無限大に、これは指数爆発という言い方、一言でいうと、そういうことだと思えるのですけれども、シンギュラリティの言わんとしているところは、そういうことなのです。

まず、そのことを理解するために、レジメにあるムーアの法則というのが、実は背景にあります。これが結構、僕が学生の時代、今から 20 年ぐらい前なのですけれども、1990 年代とか、2000 年にちょっとなるぐらいのとき、御紹介にあずかりましたように、トロンプロジェクトというような日本初の OS をつくろうというプロジェクトがあって、トロンチップという日本初のマイクロプロセッサをつくろうと。インテルチップに対抗してつくろうとか、いろいろ非常に血気盛んな時代がありまして、そういつた中で学生時代を過ごしながらか、あるいは助手の時代はプログラムをつくるという段階、あるいはマイクロプロセッサをつくるという段階です。

例えば、あるプログラムをつくろうと思ったら、CPU が全然遅くて、動かないというのでは話にならないわけです。そうすると、CPU、やっぱり

パワーがある CPU は欲しいわけなのですが、プログラムを開発すると、それだけ時間がかかるのです。例えばプログラム開発に1年かかります、といったときに、現在のプロセッサで動くプログラムを開発するのではなくて、1年後のプロセッサで動くプログラムを開発するというと非常に効率がいいですね。現在、余り動かないけれども、1年後発表するときにはさくさく動くというように開発できたらいいわけです。

そのためにはどうするかというと、1年後のマイクロプロセッサというのはどれくらいの能力があるのだろうというのをあらかじめ予測しておいて、その予測された能力の上で動くプログラムを開発するというようなことをよく行ってきたわけです。そのために、このムーアの法則というのは、結構、開発段階とかでよく参照された議論です。

それは何かというと、ある意味単純な話で、1年後大体プロセッサのスピードは2倍になっているというふうに覚えておけば、現在、プログラムを開発するとき、いろいろなレスポンスが2倍くらい遅くても大丈夫であると。1年後には、それはさくさく動くと考えていいわけです。そのためのムーアの法則なのですけれども、正確には、1年後であって2年後、1年後そんな早くはないと言っているわけです、ムーアは。正確には、18カ月、メモリの容量とCPUの速度と、いろいろなパターンでカーブがちょっと違うのですけれども、基本的には、例えば簡単に2年後とすると、2年後にどんどん倍々倍々になっていくと。

何で直線になっているかというと、こっち側が見てわかるとおり、10倍、10倍、10倍、10倍というふうに、対数直線で書かれているから直線になるのですね。これが対数ではなくて、普通のトランジスタ数にすると、カーブとしてはこういうぐつという、指数関数という、右に行けば行くほど、急激に上がってくるカーブになります。その急激に上がっていく指数関数というのは、タイトルにあらわれている e の x 乗という関数なのですから、ムーアはこのように2年で倍になっていくと。

実際、CPUのトランジスタ数、いろいろな評価方法がある、トランジスタ数ですとか、クロックのスピードとか、そういういろいろな評価軸はあ

るのですけれども、こういう直線の上に乗っかっていくということをムーアは言っているわけです。これは、非常にある意味、今、申し上げたとおり真面目な研究でして、開発なんかに応用できる非常に真面目な研究として、ムーアはこういう研究をやっていたわけです。

これも、例えばシリコンという物質を使って実際にCPUをつくっているわけなので、どこかに限界があるだろうと、みんなは言っていたわけです。実際、ここのところを見ていると、ずっと線をペンティアムⅡとかⅢとか、この辺で「さちって」(＝飽和して：saturate)きていたりしますね。現在、ペンティアムⅥぐらいですね、2004年、2016年ですから、こちら辺にいますのかな。現在もちょっと「さちって」いたりするのですけれども、結構、でも線の上に乗っかっているのです。なかなか、こういう時代にしても、そろそろ限界ではないかと言われるのですけれども、ペンティアムⅣとかで持ち直してきたり、いろいろしながら現在まできております。

工業製品の性能に関して、いろいろこういう調査をやってみると、確かに指数関数の上に乗っかっているということはよくあるのですけれども、レイ・カーツワイルは、これは工業製品だけではないということを主張されているわけです。例えば人間の能力ですとか、生物学的な生物の能力ですとか、人間の歴史ですとか、あるいは技術革新、そういったいろいろな、時間によって発展していくアクティビティというものが、こういう指数関数的なところに乗っかってくると。そういうふうにカーツワイルは主張しているわけです。

ところが、このカーブというのは、対象によって実は傾きがいろいろ違うわけです。余り伸びていかないものもあれば、急激に、非常に急勾配なもの、いろいろあるわけです。ところが、勾配を考えていったときに、人間の技術革新のカーブ、勾配と、例えばトランジスタ数の勾配というのを比較したときに、工業製品のトランジスタ数の勾配のほうが実は急勾配であるということに目をつけたわけです。そうすると、急勾配なものが非常に小さいところからスタートするわけですが、いつか、弱い勾配の人間の技術革新を追い抜かすだろうということをレイ・カーツワイルは考

えまして、その高低のポイントを計算したところ、それがおよそ2045年近辺のところに来るというわけです。

さて、非常にある意味、工業製品の議論から、人間や生物の進化とか進歩というところへ、普通の人がぱっと見ると、非常に議論が飛躍しているというイメージというのを受ける人というのも多いだろうし、実際、それは飛躍なのか飛躍でないのかというようなことを、ちょっときょう、いろいろな観点から見ていきたいと思います。レイ・カーツワイルは、これ以外にもいろいろなことを言っているのですが、きょうは、勾配に関するいろいろな考察をちょっとしてみたいと思います。

こういうグラフの上に点を打って、点を打つと、点というのはばらばらとばらつくので、実際に直線ではないのですが、そこに定規を当てて、ぴっと直線を引くのです。こういう作業を、皆さん、統計などの授業なんかでもやったことがあると思います。大ざっぱに見ると直線に見えるというということで、一つの直線というモデルで考えてみようというわけです。

そうすると、大体この直線をずっと向こうへ延ばしたくなると。やはり、直線を引く人のある種の誘惑として、直線を延ばしたくなる。延ばすだけではなくて、386というCPUと486というCPUがあります。この間に、CPUがないわけですが、もしここにCPUがあったとしたら、どうなるだろう。多分、直線の上に乗っているとすると、そのCPUの能力というのは、大体こちら辺に来ているだろうと。そういう考え方でもできるわけです。

実際、そういうことは科学の研究のあらゆるところで出てくるわけですね。こういうところに線を引いて、ここにあるものを予測するのを内挿という言い方をします。あるいは、こういう直線が引けるのだからということで、直線をこっちへ延ばしたりとか、こっちへ延ばしたり、要するに、サンプルのある領域から外側へ延ばすことを外挿という言い方をするわけです。

その外挿や内挿という方法は、非常に直感的に強力な手法でありまして、

ある種の予測に対する根拠を与える方法でもあることは事実なのですが、それを多用し始めると、実はいろいろな問題が起きてくるというのが、2番目のトピックスです。

1ページの下から2ページ目の裏側のほうへ幾つかの例を挙げているのですけれども、人間の感覚的なところに外挿をしたい、内挿をしたいという感覚というのは、誰しもあるものです。その一つのいい例というのは、まず一番最初に出てくる地球は平面か、丸いかと。古来から言う、言わずと知れた、非常に天文学のメジャーな話題ですけれども、日常生活をしている分には、地球の表面は平面であると考えて何の差し支えもないわけです。球形であることを毎日のように意識して生活をされている方というのは、この中には恐らくいらっしゃるだろうと思います。

しかし、その日常的な感覚を地の果てまで突き延ばしていくと、どうもやはり、様子が変わってくるわけです。地球は実際、丸かったのですね。現在の人はそのことを知っています。だけれども、そういうふうにはまさに、地球の表面というモデルを線形的に外挿してきた結果、地球が丸いはずなのに平面だという別の結論に至ってしまったという事例ではないでしょうか。

物理学的な話でいくと、物理学的な世界というのは、そういう話というのはよくあります。例えば相対論とニュートン力学の関係というの、そうですね。ニュートン力学というのは線形的な力学で、それをずっと高エネルギーの状態とか、あるいは非常にスピードが光に近い世界の乗り物の話とか、そういうところへ議論をニュートン力学を延長していってしまつと、実は、誤差が大量に生まれて、矛盾が起きてしまう。その矛盾を解決したのは、相対性理論であるというふうに考えることができます。

そういう意味でいったら、アインシュタインが相対性理論を発表したときに、かつてはリングは木からぼとんと落ちたのに、相対性理論が生まれてから、リングは地面から木の上にぼとんと飛び上がるようになったというようなことを、冗談交じりに言った方がいらっしゃるの、そうですね、実はそういう大きな変化ではなくって、相対性論というのは、エネ

ルギーが非常に小さい段階であったりとか、あるいは非常にスピードがゼロに近い、人間の日常生活で出合うようなスピードのレベルの世界では、実は相対性理論というのは、ニュートン力学とほぼ一緒であると考えているのですね。要するに、こういう日常サンプルの単位で議論しているときは一緒であるというふうに考えていいわけなのですが、それを外側に延長したときに、実はいろいろな齟齬が出てくると。

だんだん、ちょっとAIの話に近い話になってきますけれども、2ページ目の上に、小林秀雄の随筆の中に、将棋の神様という話があります。これを読んだときに、ゲーム理論の話を、そのとき知らなかったところに、これを読んで、僕も非常に驚いたのですけれども、小林秀雄は、将棋がやはり大好きなようで、素人が素人同士でおぼつかない対戦をやっているときというのは、回り道でも、もうそこはこう打てばいいのにみたいな感じで、まどろっこしい将棋なわけですね。それが、プロの棋士同士が対戦するようになってくると、非常に美しい手で相手を追い込めていくという、だんだん迫力がある対戦になっていくと。だけれども、完全に勝てる将棋の神様みたいな人はいないわけです。もし、そういう将棋の神様みたいな人が、完全に将棋のことを知り尽くした神様が2人いて、その将棋の神様同士が対戦をしたらどうなるのだと。物すごい迫力のあるすごい対戦になっている。これは見応えがあるのではないかと、そう小林秀雄は考えたのですよね。

それをある友達の、工学系の研究者だったらいいのですが、その人に話をしたところ、思いもよらない答えが返ってきたと。多分そんなことをやったら、最もつまらない対戦になってしまうだろうと。なぜなら、先手、後手を決めた瞬間に、もうどっちが勝つか決まってしまうと。始める前に、将棋の神様がやる気を失ってしまうわけです。負けるとわかったほうは、あとは非常にだらだらとルーチンワークを2人こなしているわけで、あなた負けたね、終わり、みたいな、そういう将棋になってしまうよと。

つまり、これはゲーム理論において、勝ち負けがはっきりわかっているとき、引き分けがないという、そういうルールの中で、確率的な過程が

入っていない、サイコロを振るとか、そういう確率的な過程が入っていない。そして、先手か、先手後手、先手後手という形で進んでいく。こういうゲームの場合、先手必勝、先手が必ず勝つというやり方があるか、あるいは後手必勝、後手が必ず勝つというやり方があるか、そのどちらかであると。ゲームのルールが決まれば、どちらかに必ず決まってしまうと。そのやり方を知っている人は、必ず勝つというふうになっているのだというのが、ゲーム理論の一つの法則としてあるからです。それを実際に、数学的に証明することも可能であると。これも人間の感覚の外挿の一つの落とし穴だったというふうに言えると思います。

ちょっといろいろ、ここまでの話で30分ぐらい使ってしまったので、飛ばしますけれども、バブル期のグラフというのは、これはバブルを経験した人はある種の失敗案として、当時は青天井とか、右肩上がりとかという言葉がはやって、いろいろな巨大開発がぼんぼん打ち上げられて、お金ががんがん回って、そのときの開発のプレゼンには、必ず右方向に上がったグラフが出てきて、このまんま発展していくと、いつごろにはどれくらい需要があって、その需要を目指すために、こういう商業施設はどうか、こういうプランはどうかという感じで、どんどんどんどんいろいろな巨大プロジェクトが動いたという時代です。しかし、それもバブル崩壊という一種の特異点があって、そこで全てがそういう外挿が崩れてしまったということですね。

ここまで見れば、「2045年問題」というのも、本当にこういう外挿という方法だけでいけるのか、と。どんなに情報を集めても、あるいはサンプルを集めても、現在、2016年ですから、2016年以降のデータというのはないわけです。全てのサンプルが2016年よりも古いものになっていて、どんなに頑張ったとしても、2045年は外挿でしか予測できないわけで、そうすると、こういう外挿による落とし穴というのがあるのではないのかと。

3番目の進化論から見た2045年。では、外挿ができないとするならば、一体この曲線というのはどうなっているのだろうかというのを、僕なりに考えた一つの考え方なのですからけれども、こういう工業製品も一種の進化論的

な発展をしているというふうに考えたり、あるいは、当然、生物であったら進化論的に発展していくと。進化というのは、単純に考えれば樹形図になるわけですね。木があって、分岐があるわけです。分岐から分岐までは、ある種直線的な発展をするのだけれども、時々そうやって分岐するのです。つまり、そういう非常に非線形的な動きが、要するに、突然変異と言われるようなものであったりとか、そういう非線形的なものが時々起こって、しょっちゅう起こるわけではないのですよね、時々起こって、そうやって分岐を始めると。

そこまで来ると、4ページのところの上に図が書いてありますけれども、これはちょっと進化の樹形図そのものではないのですけれども、いろいろなものが発達していく、あるいは、例えば動物の祖先図とか、進化の系統図から交配して、誰がここから、子供が2人生まれて、それが誰さんと結婚して、結婚すると合流するとか、そういう一種の祖先図であったりとか、工業製品のそういう関係図とかを見たり、そういうふうに思っただけだとありがたいですけども、多分、直線的に発達する部分と、こうやって分かれる部分というのがあるとすると、こういうような図になっていく。こうして発展して、一つの大きな流れが、また直線に見えたりする場合があります。

ところが、直線に見える大きなまとまりの流れも、またこうして分岐する場合もあると。そうすると、この大きな流れをもっとマクロ的に、あるいは目を細めてみると、これが一本の直線に見えて、もっと広い範囲が見えてくると、実はこういう大きな流れは、ここの部分を見ているのだったのだと。この大きいところは、実はここのところが、だんと、ここをずっとこうして延ばしていくと、実はまた、こういう大きな流れが見えてくるというようなことというのは、よくあるわけです。

例えば、ここがいわゆる工業製品の先ほど言ったムーアの法則というのは、この赤い線のここのところ、本当のここのところを見ているにしか過ぎないかもしれない。カーツワイルが議論している進化も、実は大きな直線的な流れがある、それは、緑の直線のところを見ているのかもしれない。

この傾きとこの傾きが違いますと言って、どこかで交差するという話をしたとすると、レベルが全然違うわけです。これとこれが交差するということは、ほとんどナンセンスな話になってくるわけです、そうなってくると。

そういうふうに見てくると、いわゆる直線的なものと、分岐する偶然的なものと、直線的なところというのは必然的に進んでいくわけですが、ある種の偶然的な因子によって枝分かれするという、こういう進化論的なモデルを考えると、レイ・カーツワイルの直線だけからなる議論というのは、実は2045年ぐらい、今から30年後ぐらいの予測であれば、ひょっとすると当たるかもしれない。

だけれども、それを100年、1000年、1万年というふうに延ばしていくと、30年の間に偶然的なことが起こるというのは、確率的に非常に小さいわけです。だけれども、時間をどんどんどんどん延ばしていくと、ポキッと折れるところにぶつかってしまう。そうなってくると、その予測というのは信憑性をがくがくっと失ってしまうというふうに言えるわけです。こういう予測不能性というものを、レイ・カーツワイルはどこまで考えていたのかなというところが一つは出てくると思います。

次に、4番目、これもその一つの進化の例なのですが、これは人類の一つの進化として、現在の人類というのは、クロマニヨン人からだんだんと進化してきたと言われています。クロマニヨン人が出てくる以前に、人類というのは分岐をしていて、その一つの別のほうにネアンデルタール人というのがいて、ネアンデルタール人というのもある種の文化を持っていたりして、非常に興味深い人類なのですが、いつの間にか2万年くらい前に絶滅してしまう。それが、何で絶滅してしまったのかということがまだよくわかっていないのですが、知能的にクロマニヨン人とネアンデルタール人と見ると、クロマニヨン人のほうが賢い面も結構ありまして、知能的にクロマニヨン人に負けたのだというふうな言い方もされたりしています。

ネアンデルタール人は、現代人のクロマニヨン人のように祖先ではないのですが、ごくごく最近、実は現代人の遺伝子の型に数%、ネアン

デルタール人と共通する遺伝子があるということがわかってきております。それが地球上の、例えば黄色人種とか、白人とか、いろいろな、どこに住んでいるかによって割合がいろいろ違うらしいのですけれども、少し入っているということは、何らかの混血がここの間であったのではないかと。その結果として、現代人が生まれているという予測ができるわけです。

これをこういうふうに置きかえるわけですね。クロマニヨン人をいわゆるコンピューター、僕たちをネアンデルタール人に考えると。これは、コンピューターのほうが偉いという仮定のもとに立っているわけですね。今後、カーツワイルの言うように、コンピューターは人間を凌駕していくと考えるのだったら、こういう図式というのはいいのではないかと。対応するとよく似ているよねと。2045年迎えて、コンピューターが次なる段階に進んでいくと考えると、コンピューターだけが本当に進化していくのだろうか。現代人と何らかのコンタクトがなければ、やっぱり未来のコンピューターというのも進化していけないと考えると、全くコンピューターと現代人という対立的な関係性、きょうのテーマにもありますけれども、相剋的な関係で考えるということは、できなくなってくるのではないかというふうに言えるわけです。

5番目はちょっと読んでおいてもらいましょうか。ちょっとやりましょう。5番目、未来のコンピューターになる人間を凌駕する、あるいは凌駕するようなコンピューターを生んだ現代人というものを、ある種カーツワイルなどは神としてあがめると、あるいは神と呼べるような存在ではないかとまで言っているのですけれども、現代人、未来のコンピューターというものののはたから見ていると、それはどう見えるだろうというのが5番目ですね。

地球上に住んでいる生命の中には、人間やコンピューターだけではないわけです。生物学者というのは、結構おもしろい人たちが多くて、大体自分の研究対象の生物に思い入れるわけですね。植物学者は植物とか、ここではタコ・イカ学者は、タコ・イカが絶対的に生物界の中で上なのだというのを言うわけです。ここでは、タコ・イカの例ですけれども、人間は

食物連鎖の頂点にいます。人間を食う動物というのはいないわけなのです。タコ・イカというのは、食物連鎖のピラミッドの中間層にいるわけです。

実は、中間層というのは非常に大切なのだということを言うわけです。これはなぜかという、中間層の下動物を捕食して、上の動物へそれを伝えるという役割をしているわけです。人間が食物連鎖の中で絶滅してしまっても、ほかの生物にしてみたら帽子がなくなった程度で、ああ、いなくなっちゃったという感じで、何の影響もない。だけれども、タコやイカがもしいなくなると、要するに、タコ・イカで結ばれている、例えば深海の生物と陸上の生物、海の表面の生物から陸上の生物というもの、あるいは非常に小さい微生物とタコ・イカを食べるような大きな動物、こういった大きく分けて2種類の動物たちが食物連鎖的に分断されてしまうのです。

分断された瞬間に、大きいほうの動物というのは食糧難に陥ってしまって、タコ・イカが絶滅と同時に、絶滅してしまう。それが絶滅してしまうと、そういう動物たちがする糞とかもなくなってしまうので、深海生物や小さい生物たちも絶滅して、生命系全体が壊滅状態になってしまいます。多分、そういう重要な動物というものは、生態系は一生懸命守ろうとするけれども、食物連鎖の頂点にいる人間なんていうのは、大して目もくれないというのが実際ではないかと。

最後に、ロマンチストの豚という、空を飛びたいと思っている豚がいて、一生懸命精進するのだけれども、最後、羽が生えて空に飛んでいっちゃって、その後、その豚からは便りも何にもない。どこへ行っちゃったのだろうなど。タコ・イカから見たら、多分、未来のコンピューターを見て、そういうふうに見えるのではないかとということです。

6番目として、これはある意味、一つのAIというものを考える上で、AIというものがある種擁護する方面、例えばここに出てきた西垣通という情報哲学者といたらいいのかな、が絶対的にコンピューターと生命というのは、別物だという立場をとっているわけです。AIがどんなに発展

しても、それは機械でしかないという立場をとっているわけです。

この立場というのは、間違っているとは言いませんが、しかし、そういう立場を掲げる背景には、西垣通がもともと情報関係のエンジニアであったと、コンピュータープログラムをつくっていたという過去があるから、というのは、僕は非常に関係しているのではないかというふうに考えています。彼にとってコンピューターというのは、まさにつくっていた対象なので、既知のもの、全てどういうふうに動いているかわかっているというもののなのです。そういう人にとって、そういう既知のものを、それが生命だと、あるいは知能だというふうにいることはなかなかできないわけです。

というわけで、生命というのは一体何なのだろうという一つの考え方として、生命というものが、ここに何か動物がいる、あるいはここに機械が置いてある。これを生命として名づけようという、ある意味、生物学、分子生物学は、生命とは何かという定義を持っていて、その定義に合致したものを生物と考えるわけなのですけれども、我々が普通に考えている生命というのは、そういう対象を特定して、それに条件を与えて定義を与えるというのが、形で捉えられるものではないというふうに考えられるのではないかと。そうしたら、どういうふうに考えなければいけないかというと、生命だと思われる対象と自分との関係において、生命というものを捉えるべきではないかと。

ここに未知=生命というふうに書かれているわけですが、要するに、自分から見て、これは理解不能だと、わからないというものがあるもの、それが生命の一つの条件として、そういう、要するにわからないものであるということが、一つの生命の条件としてあるのではないか。それは完全に自分と対象との関係ですね。

西垣にしてみたら、コンピューターは完全にわかっているもの。けれども、Siriを動かしている、例えばある大学生にとってみたら、別にコンピューターが専門ではないので、Siriがどう動いているかわからないと。そうすると、Siriというものは、ある種のそこに生命的な現象を感じるようになってくると。そういう自分と相手の関係の上で、生命というのを捉

えるべきではないかと。

考えてみれば、民俗学におけるさまざまなシャーマニズムや宗教といったこととか、そういったものというのは、山岳信仰、山を一つの神様として信仰している、そういうのがあるわけです。山は単なる山で、それは土の塊だよと言ってしまう人にとって、土の塊として理解してしまう人にとってみたら、それが信仰の対象でも何でもないわけです。だけれども、そこにやはり生命が宿っているというふうに加え、そこに山というものに対して未知なるものを感じる人たちは、それを一つの信仰の対象として、生命をそこに感じるわけですね。

そういった文化的な人間のアクティビティとも、生命というのはかかわっている。それと同じように考えられるのではないかと。人工知能というものも、理解不能なものとして、もし2045年以降、どんどん理解不能なものになっていくとすると、それを生命として人間は感じるようになってくる。

先ほど、この会議の前に、ブラッシュアップの議論をさせていただいたときに、人工知能がノーベル賞をとるという時代が来るのではないかと。だけれども、それというのは、人工知能を発明した人がノーベル賞をとったことになるのではないのかという質問をしたときに、柴田先生が、人工知能が、ある種の発見をしたプロセスというのが、どうやってそのプロセスを得たのかということ周りの人間どもに、あるいは人工知能をつくった人にもわからなくなってくると。そのわからなくなってくるということにおいて、人工知能に発明した者に対する著作権、あるいは初期発見者としての栄誉を与えざるを得なくなってくるのではないかと。それも、この話と非常に符合してくるわけです。

一方で、分子生物学という学問があります。これは、例えばiPS細胞で細胞をリセットするというようなことができるようにだんだんできてきているわけですが、そこには、生命に対して、細胞というものが分裂してふえるという機能を持っている、それから、代謝を行うという機能を持っているというような定義を与えて、その定義を持ったものに対して研

究していくという学問ですけれども、ある意味対照的なのですよね。

これは、本来生命だと思っていることに対し、既知のものとしようというふうな目的を持って研究する学問であると。ある意味こちらは、生命の機械化というのが見えてくる学問であるのに対して、こちら側は、今度は機械の生命化という方向に向かっている対照的な学問であるというふうに見えるわけです。

人工知能というのは、これからいろいろ発展してくると思うのですが、7番目ですけれども、現在の段階においては、人間の保護下において発展していくわけです。したがって、ここに書いてあります、エリート的人工知能というのは歓迎されるでしょう。だけれども、実際に人工知能としてつくれないのが、こっちのほうだと思います。

『道』という映画がありますけれども、主人公が転落人生を歩んで、最愛の人を失ったときに自分の人生を振り返って、その過ちに気づくという映画ですけれども、そういう人工知能というのは、多分つくれないのだらうなと思う。それがつくれるようになってくるのは、もし人間を人工知能が超えるということがあるのでしたら、超えた後に、一度はやっぱり人工知能は転落人生を経験してみる必要があるのではないのかと。つまり、このレジュメにある、一旦、カーツワイルの指数関数を逆向きにたどって、ゼロに戻すべきだというふうに思いたいものです。

最後に、シンギュラリティは本当にあらわれるかどうか、これはまだ2045年というのに達していないのでわからない。正確に、厳密に言えば、わからないわけなのですが、こういった議論をネタにして、実は人間とは何かとか、知能とは何か、生命とは何かというある意味古来からある、我々自身に対する哲学をもう一回想起する機会になるのではないかと、一種のスフィンクスの問いかけの一つではないかというふうに考えたいと申し上げて、講演を終わりたいと思います。

ありがとうございます。(拍手)

○司会 森先生、ありがとうございました。

10分休憩をとることになっておりますので、今、3時10分ですが、10

分間休憩をとりまして、15時20分にパネルディスカッションを開始したいと思います。

(休憩)

講演資料

2016 年度人文学会「A I がヒトを超えるとき」

国際日本文化研究センター
文化資料研究企画室 准教授
森 洋久

平成 28 年 11 月 12 日

$$1 \quad \lim_{\xi \rightarrow +\infty} e^{\xi} = +\infty$$

レイ・カーツァイル

- 「2045 年にコンピュータの能力が人類の能力を超える」
- (技術的) 特異点、シンギュラリティ(Technological Singularity)

なぜこのような予測に至ったのか。その背景には、「ムーアの法則」がある(図 1)。

wikipedia 訳

「部品あたりのコストが最小になるような複雑さは、毎年およそ 2 倍の割合で増大してきた。短期的には、この増加率が上昇しないまでも、現状を維持することは確実である。より長期的には、増加率はやや不確実であるとはいえ、少なくとも今後 10 年間ほぼ一定の率を保てないと信ずべき理由は無い。すなわち、1975 年までには、最小コストで得られる集積回路の部品数は 65,000 に達するであろう。私は、それほどにも大規模な回路が 1 個のウェハー上に構築できるようになると信じている。」 ”Cramming more components onto integrated circuits”, Electronics Magazine 19 April 1965[1]

最近、どうも CPU やメモリの容量は頭打ちになっているようにも思うが…。カーツァイルは、ムーアの法則は必ずしも工業製品だけの話ではないという。生物進化や、技術革新にも同様の指数法則があるという。ところが生物進化および人間の技術革新の進化係数と、コンピュータの技術革新の係数を比較すると、後者の方が大きいという。つまり、いつか、前者は後者に抜かれる。計算すると、その交差点がだいたい 2045 年に位置する。

2 外挿のトリック

人間の感覚を外挿すると思ってもよらぬ結論に達する。

地球は平面か、丸いか。

我々の生活圏では地面は平面と考えても、さほど問題ではない。この感覚をリニアに外挿していくと、地球全体は平面である、という古来からある結論に達する。ガリレオの大議論にもあったようにそれは間違っていた。

物理学的な問題には多い。ニュートン力学と相対性理論も同じ関係にある。

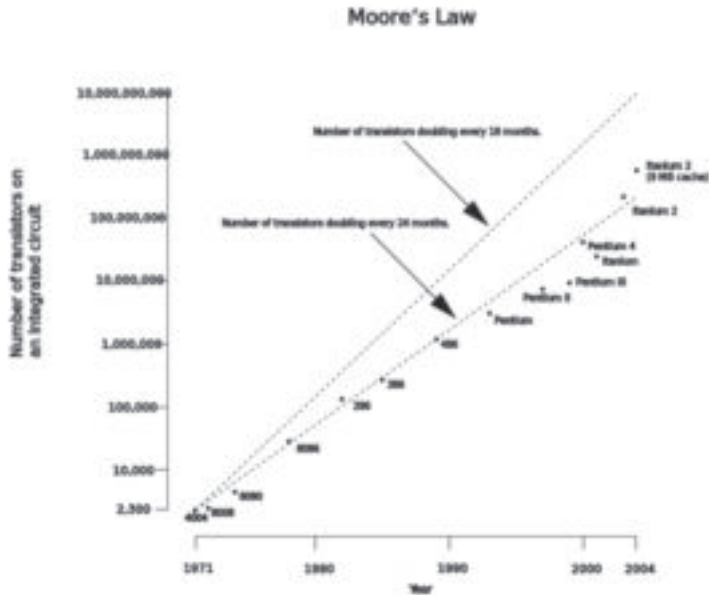


図 1: Moore's law

将棋の神様 小林秀雄

将士の接戦を見ていると、素人にはないその迫力と面白さに圧倒される。もし、このまま将士がどんどん腕を磨き強くなっていったら、さらに、面白さが増すのではないかと考えた作者が、ある工学系の教授に質問をする。究極の将士、将棋の神様同士が対決をしたら、さぞかし、肉薄した対局が展開されるのではないか。教授の答えは思いもよらぬものだった。

究極の飛行機

トランプ氏がついに大統領になった。最近世界中をまたにかけ、ビジネスを展開する彼はトランプ・フォース・ワンというプライベートジェットに載っている。プライベートジェットにより世界は近く、小さくなる。速度競争の果て、たとえば、世界中どここの都市でも 30 分～1 時間で行けるような飛行機が開発されたらどうだろうか。それは宇宙第一速度に達している可能性が高く、地球に戻ってこれなくなりますね。

バブル期のグラフ

バブル期の地価の高騰は「青天井」と言われ、様々なプレゼンで示される開発の可能性は、右肩上がり。みなさんそれを信じたんですね。

バタフライ効果

- 気象学者のエドワード・ローレンツ 1972 年の講演 “Predictability: Does the Flap of a Butterfly's Wings in Brazil Set Off a Tornado in Texas?”

- カオス現象の、初期鋭敏性や予測不能性に関わる論考。

偶然と必然

ジャック・モノー

3 進化論から見た2045年問題

レイ・カーツァイルは、シンギュラリティーという目を引く楔を打ち込んだことで注目された。シンギュラリティー以降、コンピュータが地球を埋め尽くし、宇宙へも進出していくという世界観は、現時点では奇異と思う人が多いであろう。

レイ・カーツァイルの議論は、おそらく2045年から少々あまでの彼の予測は当たっている。しかしそれ以降はどうだろうか。つまり、人間である彼が予測する世界像、宇宙像は、彼が言う通り、皮肉にも、おそらくコンピュータにあつさり裏切られるだろう。

ムーアの法則に代表されるような産業技術理論の外挿である。産業理論は往々にして、バブル期の右肩上がりグラフのように線形外挿論法に終始しがちである。つまり、この理論への偏向が強すぎるのではないか。

これに対する対極として進化論的観点を提示せざるを得ない。一つのシンボリックな理論がジャックモノーの偶然と必然モデルである。必然フェーズにおいては、(外挿可能な)線形的に進化するが、ごくごく稀に、偶然によって、それが方向転換をする。レイ・カーツァイルが見ている世界は、必然フェーズでしかないのではないか(図2)。

偶然フェーズはごくごく稀なので、2045年という短いスパンでは、まず起こるとは思えない。しかしこの話を太陽系、宇宙へと拡大してくるならば、その範囲に必然的に偶然が紛れ込み、全体は予測不能とんある。

レイ・カーツァイルはそういった進化論的予測不能性も見越して、全体はそれでも線形であると主張している。が、それは、進化論的過程が、部分が線形であると同時に全体も近似的に線形になるという自己同相形を持ついわゆるフラクタル構造をしているからだ。だが、台風を、様々な地形や気象変化で起こる多数の渦の集合としての渦であるとするならば、一つ一つの小さな渦、たとえば、家の近くで起こった渦から位置地域を飲み込む全体の渦が予測できようか。逆に台風という全体の渦から、地域で起こる小さな渦を予想できようか。つまり、フラクタルの予測不能性はそこにある。進化も同じである。

4 ネアンデルタール人とクロマニオン人と現代人とコンピュータ

レイ・カーツァイル的外挿を進化論で展開するならば、こうなる。現代人とコンピュータ、および2045年以降の未来のコンピュータの3者の関係をネアンデルタール人とクロマニオン人、現代人に置き換えることができる(図3)。

こうしてみると、レイ・カーツァイルの言っていることが、2万年前にすでに起こっているともいえる。その結末が我々現代人だ。かつて、現代人は、ネアンデルタール人の子孫ではないと言われていた。しかし、近年現代人の遺伝子に5~10%ほどネアンデルタール人の遺伝子があることが発表された¹。

- 我々現代人を、クロマニオン人とは呼ばないように、未来のコンピュータはコンピュータではない。(しかしいま代わる言葉がないので、未来のコンピュータと呼び続けることにしよう。)
- 現代人はネアンデルタール人の混血があることが認められるように未来のコンピュータは現代人の混血でもある。

¹ “Special Feature: The Neandertal Genome”, Science (アメリカ科学振興協会), (5 2010) 2010年8月12日閲覧。

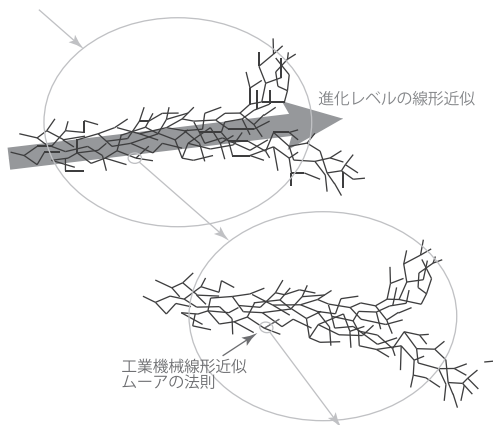


図 2: 進化のフラクタル

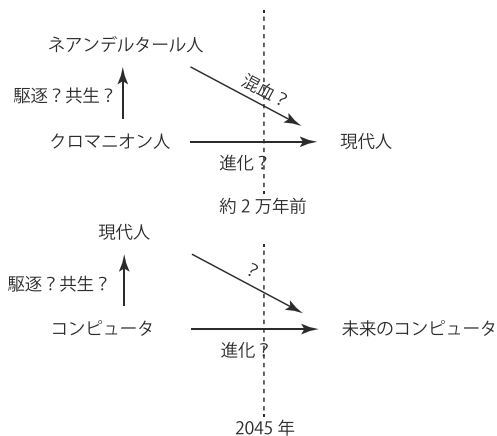


図 3: 進化論による一外挿

事態はそれほど単純ではない。

クロマニオン人やネアンデルタール人が現代人がなにをしているか予測しただろうか。現代人というたかだか5万年ほどの歴史しかない種は今から5万年後はよくも悪くも存在しない。コンピュータも同じだ。絶滅しているか、進化していた場合はいまから予測不能な形態になっていることは明らかで、それをネアンデルタール人たる現代人がとやかく言っても始まらない。

混血という意味でいうならば、バイオ系デバイスをつかったコンピュータや、コンピュータを組み込んだ人間も現れ、コンピュータ対人間という二項対立そのものが成り立たなくなってくるだろう。

いまのインターネット自身がすでに、人間というバイオデバイスを組み込んだコンピュータであるとも言える。

5 生態系から見た2045年問題

人間対コンピュータという図式に目を奪われがちであるが、実際には、進化は人間以外の動植物も連関して進行するものである。その中で、人間対コンピュータという図式はそのごく一部にしかすぎない。

生物学者というのは、とすると、自分の研究対象への思い入れを強く持ちすぎるものだが、おそらく、その共通する意見の一つが、食物連鎖の頂点にある人間は絶滅しても、他の生物への影響は極めて低いというものだ。

その一方で、例えば、タコ・イカ学者は、タコやイカは食物網の中間に位置する捕食者であり、それより下の生物と上の生物を結びつけることによって、生態系全体の食物連鎖を結びつけるハブの役割をしている。ということは、タコやイカが絶滅すると、地球全体の食物連鎖に大きな影響を与えることになると同時に、様々な生物の増減のフィードバックとなって、生態系の安定化に寄与していることだろう。

コンピュータや人間が共進化的に進化し、宇宙へ飛び出していったとしても、生態系の主役は依然として、タコ・イカであり、単に生態系から人間が消えただけのこと、とらえることも可能。カーツァイルがいうように人間とコンピュータが神になったと思いついて入っているのは人間だけで、(いやカーツァイルだけかも)つまり、ロマンチストの豚(やなせたかし)となってしまうかもしれない。

6 生命とは「理解できないものである。」

西垣生命論

西垣通は絶対的に生命とコンピュータを分ける立場であるが、それは、西垣通が情報科学系のエンジニアとしての過去があるから、彼にとってコンピュータは「既知」であるからだ。

対象定義と関係性定義

「生命」を我々の外部にある対象として正しく定義することはできない。「生命」はあるモノをさすものではなく、「私」と「対象」の関係として成り立つものである。

「未知」＝「生命」

人が「理解できない対象」として対象を見ることができたとき、そこに生命が宿る。Siriを一般の誰かは理解することのできないものとしてみている人もいよう。その人にとってSiriは生命である。西垣にとっては生命ではない。

民俗学、文化論への連続性

シャーマニズムや宗教において、山岳信仰や妖怪の対象となり得るものは、それは「理解できないもの」として生命化されるのと同じである。

人工知能 <=> 分子生物学

人工知能が人間にとって理解不能となったとき、人工知能に生命が宿る。一方で、分子生物学において生命を定義し研究する行為は、「生命」を「機械」として捉える第一歩であるといえよう。つまり人工知能研究と生化学研究はベクトルの真逆の研究である。

$$7 \quad \lim_{\xi \rightarrow +\infty} e^{\xi} = 0$$

- 東大を目指していた人工知能²が進路を変えたそう。文科省の要請で、一浪をしたくないならば、そうせざるを得ない。
- 受験を終えてもまだ受験をやりたい、変わった学者さんがいるもんだなあ。私は受験が嫌いだ。いまでも嫌いだ。

²<http://21robot.org/>

- だが、多少浪人すればいつかは入れるようになるだろう。だが、転落人生の人工知能を作ることはできない。輝かしい成果を出す人工知能には開発費が出ても、そのような人工知能には開発費が出ないからだ。今後 2045 年に向けて、進んでいく社会の構造がそうになっている。
- だが、人間の世界は、過去も未来も、エリートの道もあれば転落人生もある。すべてが揃って人間の社会なのだ。ここが、人間を育てる世界と、人工知能を育てる世界の大きな違いである。であるから、思いやりをもったロボットとは、所詮は真似事であると言われても仕方ない。
- 2045 年以降、人間の理解不能な対象となつてしまい、一度は転落人生を経験した人工知能が現れるとしたら、そこで初めて、人工知能にも、思いやりや、慈愛がわかるようになるだろう。

8 さいごに

シンギュラティイーが本当におこるのかどうか、それが 2045 年かわからないが、ある時点で、レイ・カーツァイルの主張に類似したことは起こりうるだろうと考えるが、事態は彼が言うように単純ではない。そのため、その 2045 年という状況が具体的にどういう状況なのか、いまの我々には想像し難いのである。

だが、一方で、いま人工知能を考えることは、過去において人類が連綿と重ねてきた哲学的議論、前回お話ししたクオリアの問題や哲学的ゾンビ、今回の生命や知能といった問題を改めて、人工知能という視点から考えるよい機会である。いわば、スフィンクスの謎かけなのだ。