

人間と人工知能(AI)の共働

産業技術総合研究所 人工知能研究センター
脳型人工知能研究チームチーム長 中田 秀基

はじめに

中田と申します。宜しくお願い致します。

今日は人間と人工知能がどのようにして共働していくのかについてお話を進めていきたいと思います。

初めに産業技術総合研究所の紹介ですが、旧通産省の工業技術院の中に10～20の研究所があり、2001年の行政改革に伴い独立法人化する際に他の電総研、機械研、生命研などを一つに集めたものが産総研です。筑波に本拠地があり3,000人くらいの研究者がいる組織で日本でも最大の研究機関の一つです。よく聞く名前ではロボット研、地質調査所、があります。私の居る人工知能研究センターは経産省からNEDOを通して予算を頂いて、一昨年の5月に設立された研究センターです。正規の研究員は少ないのですが、ポスト・ドクターとか外部の研究員を併せて300人強います。

今日お話しする内容は、最終的には「雇用と人工知能がどのように関わってくるか」が焦点です。

まず今どんな状況かを整理したいと思います。そもそも人工知能は何なのか、今何ができ、何ができないのか、将来どうなるのか。この辺りを整理して話を進めたいと思います。

●人工知能とは何か

初めに人工知能は何か、から話を進めます。人工知能の定義は“人間と同じ知能を持つ機械である”。では知能があるとの定義は何なのか。非常に難しいと思います。人工知能の研究のメインストリームの一つにチェスの研究がありました。チェスのプレーヤを人工知能で実現するという事です。チェスを打てるのが、知能がある事だと思っていました。皆さんはチェスや将棋や囲碁が強いコンピュータを持ってきても、これらに知能があるとは思われません。ある意味では人工知能の研究は逃げ道みたいなものであり、“ここまでできました”と言っても、“それは人工知能ではない”と言われ、どんどんハードルが高くなって行きます。“機械にできることは知的ではない”と言われると原理的に人工知能は実現不可能になってしまいます。このように非常に曖昧なことを相手に我々は仕事をしています。そうはいっても仕事が進まないのです、知能というものをある程度条件をつけて考えたいと思います。

【人工知能の定義】

ここに4つ定義を上げました。色々な定義があると思いますので、これがDefinitionというつもりはありませんが、一つの考え方だと思って下さい。

「自律性」勝手に動くことです。つまり詳細な指示を与えないで何か物を取ってきて、と言って動いてくれるのが、自律性があると言い、知能の一つの条件です。

「意味性」データの意味に応じて動作させる。人の顔が映っている写真がある場合、写真のこのPixelが黒だったら何かする。というのは余り人工知能らしくなく、“顔の中にある目を見つけて下さい”というのは人工知能らしいと思います。これが意味性です。

「技巧性」例えばオセロのプログラムで、相手がここに打ったら次はここに打つというようなツリーを広げるような探索は可能で、仮に全部探索する、つまり計算パワーに任せて探索する事はあまり知的ではなく、巧みに解決方法を見つけることが技巧性です。

「適応性」これは重要で、ある条件に対して解を見つけることは余り難しくなく、条件が変わっても同じ方向に向かって何回も失敗を繰り返しながら解を見つけることは人工知能らしいと思います。

このあたりが全部満たされれば、より人工知能らしくなると思います。

話は変わり人工知能とロボットがどのような関係であるか整理したいと思います。鉄腕アトムはロボットであり人工知能です。又鉄人28号はロボットですが、人工知能では無い。アトムは自律的なロボットで、おしゃべりをして自分で考えて善悪の判断もします。鉄人28号はリモコンで動くだけ

のロボットです。形こそ人型ロボットですが、建設機械と大差ないと言えます。

ここに HAL9000 があります。映画「2001年宇宙の旅」に出てきます。宇宙船そのものが肉体のコンピュータで自意識まで持っているかわかりませんが、知的な応答ができる人工知能です。人工知能は知性が問題であって肉体はあまり関係ないという事です。

【物語の中の人工知能】

映画「アシモフのロボットシリーズ」でロボット3原則が出てきます。①ロボットは人を傷つけてはいけない、②命令には服従する、③自分が壊されるのを防がなければならない。これらのルールを決めておいて、その制約の中でミステリーを解くという事がこのシリーズです。「ブレードランナー」も面白い映画で、これもアンドロイドです。「マトリックス」は1999年の映画でこれもある種の人工知能です。「2001年宇宙の旅」も人工知能を扱った映画です。これらに共通するのは、人間を裏切る事です。ロボットは人間が作った機械であるが、人間に対して犯意の有無に関わらず結果的に悪い事をします。これらに共通していることはフランケンシュタイン・コンプレックスであろうと思います。これは何かを作ったものに対して、裏切られてしまうことへの恐怖ですが、非常にプリミティブな我々の中にある感情を人工知能が揺さぶるので、このような扱われ方をよく物語の中ではされているのだと思います。何が言いたいかというと、雇用が奪われるのではないか、人工知能に裏切られて変なことが起きるのではないか。それは理性的な側面もあるが、やはり感情的な側面が大きいのではないか、ということのを頭に置いておいた方がよいと思います。

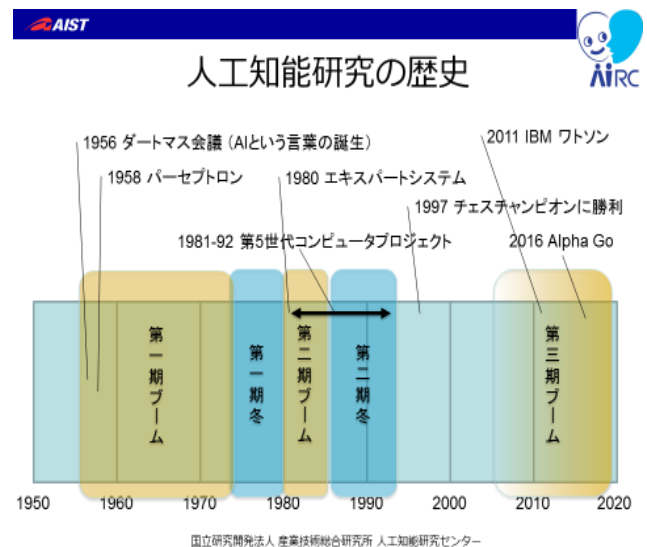
【人工知能研究の歴史】

話は変わり「人工知能研究の歴史」を振り返ってみたいと思います。人工知能の歴史は大変古く、コンピュータの歴史と変わらないくらいです。コンピュータは第二次世界大戦の時1940年代ころに砲弾の軌道を計算するために作られ1956年にダートマス会議で「AI」という言葉が作られました。そこから長い第1期ブームがあり、冬の時代があり1980年代に第2期のブームがあり、また冬があって今のブームに繋がっています。1981年から92年にかけて通産省のリードで富士通、日立、三菱電機、東芝、沖電気などから人とお金を出させて第5世代コンピュータ・プロジェクトが実行に移されました。これはうまく行くのかと思いましたが大失敗でした。そのようなこともあって第2期の冬はかなり深刻だったようです。

この第5世代コンピュータの失敗が尾を引いて第3期ブームが大分遅れてしまった感じがします。この時は経産省がメーカーに物凄く迷惑を掛けました。人や物を出させたが殆ど何も成果が無く終わってしまい、それでメーカーの経営者にAIとか機械学習に対してアレルギーが残ってしまいました。海外では2000年位から機械学習に対する研究が始まったのですが、日本は今になって慌ててキャッチアップしようとしており、この後遺症が意外に長く続いた感じがします。

ここで第1期2期3期ブームを整理すると、

「第1期ブーム」は、知能はなんなのかよく解らないけれど人間がすでに知能を持っているから計算機上に実現するのは簡単に違いない、と非常にプリミティブな感情で動いていました。これは歩行ロボットに関しても同じで50年60年代のSFを読むと、いとも簡単に歩き、話し、感情を持つロボットはたくさん出て来ますが、実際にロボットが歩けるようになるのは1995年位になります。その頃は歩行に対する難しさは全然わかってなく、作ってみたが全然うまく行きませんでした。それから繰返し作り続けて、歩行に関しては1995年にできるようになると、その後誰が作っても動くようになりました。人工知能に関しては残念ながらまだそのようにはなっていません。第1期ブームは要するになんの成果も得ることなく終了しました。



「第2期ブーム」はエキスパートシステムを作りました。エキスパートシステムは限定的な言い方ですが、医者が患者を診断する場合、患者に対して“熱がありますか、痛いところがありますか、それはどこですか”などの質問を通して専門家としての知識、推論を重ねて答を求めます。その場合、知識が十分であれば、その事に相応した答に到達できます。

このように、専門家の知識をシステムに取り込み、このことから、何かを推論するのがエキスパートシステムなのですが、実際には、うまくいく場合とうまくいかない場合とがあります。

例えば、人間だと誰にもできるのに、機械だとむずかしいようなケースがあります。

因みに、写真に映っている人の顔を見て、顔だと認識する事は非常に難しいのですが、その表情が怒っているのか、笑っているのかは、人間は誰しも、容易に判断できます。しかし、この事を機械にあてはめた場合、非常にむずかしい判断になります。このような場合、人間に認識の中身、その仕組みを問うても、説明できないと思います。実際どのような判断をして、人は笑っているかは皆さんでもわからないのではないのでしょうか。結果として、人間はできてはいるのですが、その仕組みを、説明できない領域がまだ多くあることが判明し第二期ブームは終了しました。

「第3期ブーム」はディープラーニングというテクノロジーが2000年代の終わりから使えるようになったことで発生しました。それによってパターン認識、つまり笑っているか、怒っているか認識できるようになったのです。このパターン認識が人並み以上にできるようになったことで工業的な応用も広く知的な意味でのインパクトも大きく今盛り上がっています。ディープラーニングの性能が向上したのは、一つは計算機の能力が著しく向上したことによります。たとえばアップルウォッチのCPUパワーは1990年代初頭のスパコンより早く指数関数的に性能が向上しました。1980年ころにはできなかったことができるようになりました。ディープラーニングは膨大なデータが必要で1980年ころに自然言語処理の研究をする場合は新聞社に行って新聞記事を貰ってくるしかなかなく、当時電子化された日本語の情報はなかったのですが、今は幾らでもありWebで何テラバイトでも集めることができます。この2つの要素によって技術的にブレークスルーが可能になりました。

[ディープラーニングとは]

ディープラーニングを簡単に説明すると、Deep「深い」Learning「学習」つまり「深層学習」という言い方をしますが、要するにニューラルネットワークが深いという意味です。ニューラルネットワークとは大変古いシステムで1958年にできたパーセプトロンが一層のニューラルネットワークになります。ニューラルニューロンとは脳の中にある神経細胞です。脳の中の計算回路がこの様なものと当時解って、それを真似して計算機機構を組み立てて、一定の成功を取めたものがパーセプトロンです。

計算機の構造は単純です。神経細胞の場合の一つの神経細胞から足が出てその先に別の細胞に繋がっています。一つの細胞に対して入力がたくさん繋がっています。それぞれの入力に対して重みを加え掛け足して関数を掛けたものを出力するという仕組みになっています。計算機素子としては簡単に電気回路で作っても簡単にできます。これをたくさん繋げると色々なことが出来ます。これがパーセプトロンで1958年には見つかっています。それで色々なことができるのではないかと盛り上がったのが第1次AIブームの一つの潮流だったのですが、最後にマービン・ミンスキーという『心の社会』を書いて大変有名になった学者が“パーセプトロンではできないものがある”と言ったので第1次AIブームが終わりました。ところが“XORが出来ないのは少し工夫すれば逃れられる”ということが解りその後復活します。「誤差逆伝搬法」という今も使われている技術が見つかって色々なことが出来ると思ったのですが、この当時はデータ量や計算パワーもないので続かなかったようです。その後暫くニューラルネットの研究をしていると笑われた時

Deep Learning

- Deep = 多層のニューラルネットワーク
 - 能力は高いが学習が困難
 - 特に誤差逆伝播学習は層が増えると困難
 - 局所収束 + 学習の遅さ
- ブレークスルー
 - 1) 問題に適した構造を予め組み込む
 - 畳み込みネットワーク
 - 2) 層ごとの教師無し事前学習
 - Restricted Boltzmann Machine
 - Stacked Auto Encoder

Deep learning の構造の一例:
2層教師なし機械学習器を組み重ねたもの

入力

国立研究開発法人 産業界技術総合研究所 人工知能研究センター

代が 20 年位続きました。

このころ「ニューロ&ファジー洗濯機」があり TV CM になるほど流行ったのですがその後尻すぼみになりました。その後は別の研究が進んでいたが 2006 年になってニューラルネットの研究を続けていた Geoffrey. Hinton が技術的なブレークスルーを発見して大流行しました。

「ディープラーニング」は何かというと、基本的には入力として信号がニューロンのたくさん並んでいるニューラルネットワークに入り、そして次の層に行きます。電気的にはたくさん重ねるほど色々なことが出来ると原理的には解っていたのですが、技術的限界があり上手く実現することが出来なかったのです。「誤差逆伝搬法」は基本的には上がってきたデータが合っているか間違っているかを伝えていくことで、個々の重みを修正して学習していく理屈ですが、それが深くなると教える信号が消えてしまい上手く学習できませんでした。これが第 2 期 AI ブームの限界だったのですが、それをいくつかの方法で回避できるようになり、今やありとあらゆる所にディープラーニングが使われて大成功を収めています。

「質疑応答」

Q: iPod の様な革新的技術は外国から来ているようだが、日本で開発されているものはあるのですか？日本人はハードウェアには強いがシステム、ソフト面が弱いのは日本人に向いてないのか、教育がずれているのかうまく育ってないのはなぜか？

A: AI に限らず日本の産業界は元気がありません。AI に関していえば、第 5 世代コンピュータ・プロジェクトの失敗の後遺症で他のジャンルより増して出遅れている感があります。トップカンファレンスと言われる最高の会議に行っても日本からの発表が殆どありません。

ソフトに関していえば、日本にも元気で若い企業があります。PFI、PFN は AI でも元気で世界的に徐々に認められつつあります。よって日本人が本質的に向いてない事はないと思いますが、日本の会社組織がソフトウェアに向いてない事はあると思います。つまり雇用形態があまりにも硬直しています。プログラマーの生産性は他の産業と全然違い、並みのプログラマーと本当に良いプログラマーと生産性が 10 倍ではきかないのです。他の産業では人によって生産性が 10 倍違う事はあまりないと思います。生産性が 10 倍違うのに 10 倍の給料で報いる事は日本ではできない。だからそのような人は日本では育たないし、居たとしても海外に行ってしまうという事になります。

我々もその問題に直面しており、AI センターで海外の優秀な研究者を雇おうと思っても 2,000 万円出せば雇えるのに日本では 2,000 万円出せないのです。理由はそのようなシステムになってないからとしか言いようがありません。我々は公的機関なので高い報酬は出せないのはしょうがないですが、例えば日立、ソニー、東芝なら出せるかという、私企業なので理論的には出せる、もしくは出さなければいけないのに、日本の企業では、そのつもりはありません。この分野が遅れるのはこのような理由からだと思います。日本人が向いてないのではなく、日本の会社組織が向いてないかもしれないということです。

Q: ディープラーニングで起動しているコンピュータのプログラムはどのような言語で記述されているのか。又バグに対して必ずデバッグを行うが、どのように行われるか。又私の持っている囲碁のソフトは動かなくなりパスしてしましますが回避法はありますか。

A: 何で記述されているか分かりませんが、一番下の部分は「CUDA」という GPGPU という特殊な計算機用の C 言語の一種で書かれています。又バグがあるか無いかは別の話ですが、それはいまだ全然解決していません。デバッグの手法とかデバッガー は物凄く発達していますが、本質的にバグが無いという事を証明して行っているわけではありませんので現状ではあまり進んでいません。現に将棋のプログラムでは最近棋士がなっても損が無い状況で、ならないという事によってコンピュータが混乱して、負けを宣言して棋士が勝ったことが有りましたので実際その様なことはあり得ます。

Q: ニューロンの Layer があり、下の Layer から上がってゆき一番上の Layer で間違えを発見した

場合、どの Layer にフィードバックして重み付けを変更するのはそのように行うか。

A：基本的に上から落としてゆくしかありません。つまりここで違いがあった場合この違いを作るにはこの下がこれだけずれていけば良いでしょう、みたいなことを考えます。これをこれだけずらすには、ここをこれだけずらせば良いだろう、のように上から戻ってきます。これがどんどん小さくなって無くなって落とす量が無くなった場合、ここから下が学習できなくなり、あまり深く学習できなくなるというような現象がおこります。要はコントロールしてなく構造で決まります。

●人工知能は何ができて何ができないのか

できる事

次は今何ができて何ができてないかを話します。できている内容はもうかなり話しましたが、できるようになった事は「パターン認識」で、できない事は真の意味での「理解、自意識」の辺りは全く手つかずのまま残っています。始めにエポックメイキングだった事件について話します。

[IBM DeepBlue] 1997年にIBM「DeepBlue」という計算機が、チェスのチャンピオンに勝っています。このカスパロフというチャンピオンは大変強い人で、将棋界の羽生さんに相当するような人で、チェスの世界の Champion of Champion でその人に勝ったから当時非常に話題になりました。

[IBM Watson] 2011年IBMの「Watson」というシステムが「Jeopardy」というクイズ番組で、人間のチャンピオンに勝ちました。これはアメリカの非常に人気ある番組で社会的にはインパクトがありました。クイズは、“これは日本で一番高い山です。さてこれとは何でしょう？”のような質問に対して“富士山”と答えるようなクイズ番組です。つまり Fact をどこからか引っ張ってくるというような原理です。これを例えば Wikipedia とか各種辞書とかのデータを全て機械に取り込んで、この質問文を解析して、それらしい回答を全部探し出してきて一番それらしい回答を探し出し、又はそれらしい回答が無い場合はスルーするような事も含めて判断し人間のチャンピオンに勝ちました。

ちなみに「IBM Watson」はブランド名として、このシステムに関係ない物も含めて「Watson ブランド」で商売をしています。

[AlphaGo] 次は囲碁の話になりますが、「AlphaGo」が2016年に韓国の棋士イ・セドルに勝ちました。これは研究者にとって非常にインパクトがありました。チェスの次は将棋だろうと言われ、事実上昨年まで人間が負けている状況でした。将棋では今もかなり強い棋士にコンピュータが連勝しているはずですが、但し、羽生さんが負けない限り人間が負けたことにはならないかもしれませんが、そのような状況です。なぜインパクトがあったか、というと、チェスの次は将棋だろう、2～3年の内に負けるだろう。ただ囲碁に関しては問題の難しさが全然違うので、あと10年くらいかかるだろうというのが昨年初頭までの業界全体のコンセンサスだったのですが、Googleの「AlphaGo」でイ・セドルに勝ってしまったので非常に大きなインパクトがあったのです。これはチェスが負けた時とは構図が違います。チェスに勝つプログラム、「DeepBlue」ですが、これを作るスタッフの一人がチェスの名人「グランドマスター」でした。つまりチェスの強い人が作ったコンピュータがチェスの強い人に勝ったということです。チェスの場合は人間の持っている知識を何らかの形で記述して、それに応じてプレイしたプログラムが人間に勝っただけです。「AlphaGo」はそういうことではなく本質的には機械同士が対戦していることになります。「AlphaGo」を2台並べて対戦し、勝ち負けを繰り返すことにより棋譜が溜まり、勝ち負けの結果情報も溜まります。これをコンピュータの中で無限に繰り返して人間が教えなくても学習できます。これを「強化学習」と言いますが、これによって急速に強くなりました。

始めは人間が教えたプログラムがありそこから始めるが、その後は機械同士が対戦を繰り返して強くなります。こうなると基本の棋譜の中には全く無かったような突飛な手で打つ事もあり、解説者が解説できなくなり、しばらく経ってから“これはここで効いていたのか”という感じで後から分かることになります。よく人工知能と人間の差は Creativity（創造性）みたいなものと言われますが、これに関していうと、機械のほうが創造性があるのではないか、と言われる部分もあります。われわれ

は、囲碁という問題を人間が解いたとっていました。つまり、“勝つにはこのような定石がありこのように読み切れればよい”とある程度解っていて実践できる強い人がいるという構図だと思っていたのですが、実は囲碁を人間は理解してなかったということがここにきてわかってしまったということです。このような意味でショッキングでありエポックメイキングな出来事でした。

【自動翻訳】皆さん、日本語・英語翻訳は使ったことが有りますか？これは随分前から来ていましたが、日本語に関しては使い物になりませんでした。フランス語・英語間の翻訳はかなり出来ています。日本語を入れると変な英語が出てきて使い物にならなかったのですが、昨年12月からパーフェクトではありませんが急に良くなりました。以前は単語単位で訳していましたが、今は文章全体を見て翻訳するので大分洗練された文書が出来るようになりました。これもディープラーニングで行っていると言われていました。

【画像認識】この図を見た時に“ここに車がある”というのが答えですが、業界では長く画像認識を競うコンテストがありました。画像のセットを用意して“車があります”、“鳥がいます”、“飛行機が飛んでいます”などのタグを付けます。この120万枚の画像を学習させて15万枚の画像に対してどれ位当たるか、というゲームです。そして当たる確率が高い方が良いとされます。認識精度は2010年、2011年だと誤り率が20～30%位で推移してきましたが、ディープラーニングの技術が導入されて急に良くなります。2015年の段階で人間の識別率5%ですが、人間より当たるようになりました。よって画像認識に対して問題は解かれたのでこのコンテスト自身が今年で最後になりました。

【特徴抽出細胞の学習】これはGoogleが出した面白い論文です。沢山絵だけを見せると特定の画像に反応する細胞ができました。つまり人の顔の概念を自動的に獲得したと同じです。人の顔や猫の顔に対応して事前に教えなくても自然に反応するということが話題になりました。

【音声認識】音声認識に対しても大変進歩しています。「Apple Siri」、「Windows Cortana」等の音声認識エンジンがあるので、相当識別精度が高くなっています。このあたりも全てニューラルネットワークで処理しています。

【文の感情分析】ある文章を読んで、これがポジティブなのかネガティブなのか。これは商業的にもすごく重要で、例えば「Amazon Reviews」で、これが肯定的か否定的か、Twitterで流れてくる自社製品の反応が肯定的か否定的かはマーケティング的に重要な情報です。例えば皮肉の様な文章を書いた場合経験的に表示データが充分あれば識別できると言われています。

【画像と文の対応づけ】これは画像に対して文を割当てます。つまり画像に対して文章を生成することです。この市場の画像を入力すると、“屋外の市場で買い物をしている”というような、それらしい文章が出てきます。これもかなり凄い技術です。

【強化学習との融合】インベーダーゲームやブロック崩しなどのゲームを解かせることも簡単そうですが、ゲームのやり方を教えないで、コンピュータには画面と取れた点数だけを与えてプレイできるようにする。つまり点数と文字だけを読める子供が試行錯誤しながらうまくプレイできるようになる、と同様なことができるようになりました。

【画像生成】次は画像認識だけでなく画像生成もできるようになりました。たとえば寝室の画像をたくさん憶え込ませると寝室の画像をいくらでも生成できるようになります。しかも寝室には違いないけれど今迄にない新しい画像を生成することが出来るようになりました。これも大変面白い応用です。

できてない事

何ができないか？ですが、「理解」ができません。絵を見て文章が出てくれば理解していると思われませんが本質的に全然違います。これはネットワークに情報を入れると長い数字の列が生成されます。それをネットワークに入れると文章が出てくるわけです。その場合長い数字が理解していることになるわけですが、その列に対して、“この写真はだれが居ますか？売っているものはバナナですか？”と質問しても全然答えられません。非常に限定的なことしか理解していない状況です。

理解の一つの定義はその絵に対する質問が全て答えられるということになりますが、現在そこには到達しておりません。又“そのコーヒーどこにあった？”との質問はそのコーヒーがどこにあったか

を知りたいのではなく“僕もコーヒーを飲みたい”という意味が裏にあるわけです。このようなことは全く理解できていません。一見会話できているようですが、単にある特定のパターンに対して特定の答え方をしているに過ぎず、何かを理解して会話しているとは程遠いのです。そこまで到達する道筋が全然見えてない状況です。

理解することの一つの例として「中国語の部屋」があります。部屋の中に中国語が解らないが辞書を持った人がいます。外から中国語の文字列を送ると特定の変換を行って戻す作業を、辞書を使って行います。

これは外部から見ると、何か質問すると、それらしい答えが出てくるので、“この人は中国語を理解しているといえるか？”については、この人と辞書のペアは理解しているといえますが一般的には理解してないと思います。インプットとアウトプットだけを見て理解しているか判断するのは不可能ですが、できるような方向に向かって進む必要があると思いますが、どのようにして進むかについては解っていません。

もう一つは「意識」です。これは哲学的な話で、“意識とは何ですか？ 意識を持っていることを外から判断できますか？”ということになりますが、今はまだ問題にはなっていませんが、前項の「理解」が到達した時点で「意識」を持っているかどうかの問題になると思われます。現状ではなにも解っていません。

[質疑応答]

Q：日本に「京」というコンピュータがあるがこれはAIですか？

A：違います。「京」の上で人工知能を動かすことはできます。「京」はハードウェアでAIはソフトウェアなので「京」が人工知能になることは可能ですが、今の処、その様な使い方はしていません。

Q：ニューロンへの入力は電気刺激ですか？又AIの具体的な形はPCとかスマホの中の配電盤の様な形ですか？

A：人間の脳の中では電気刺激や化学的な刺激になります。又コンピュータの中では電気刺激で表現されます。配電盤で考えている場合はそれが具体的な電線のネットワークだとすると、そうではありません。計算機の中では固定した電線のネットワークがあるので、その上でソフトウェア的に表現するという形になります。

Q：囲碁でAIが10手先を読んだとき負けると解っても負けを宣言しないで、次の手までどうにか生き延びる対処、つまり自暴自棄の問題をどのように解決するのか？

A：自動運転を例にすれば、このまま進むと人を轢き、ブレーキを踏んでもハンドルを切っても無駄な場合、ブレーキを踏まないAIは考えられないと思います。このような反応をした時にこのようなことが起こるという可能性の高さと、それが起きた時の影響の大きさを掛け合わせて、結果が最小になるように持ってゆく。例えば人を轢くのと猫を轢くのでは結果がまったく変わってきます。制御の構造は同じでも挙動は違うべきです。そのあたりの重み付けは人間の匙加減ですが、そのあたりを踏まえて自暴自棄にならない様にAIを設計する必要があります。

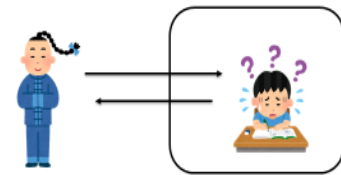
Q：先程の市場の写真でこれがバナナのたたき売りをしているとの認識は出来るのか？

A：それは教えることはできると思います。たたき売りの画像とそうでない画像を沢山準備すれば識別はできるが、それがたたき売りを理解したことにはなりません。要はパターンを認識しているだけで、全く新しいたたき売りの写真を見た時、人間なら認識できてもAIは認識できないと思います。

●これから何がおこるか？

● 中国語の部屋

- 哲学者ジョン・サールによる思考実験
- マニュアルに従って、見知らぬ記号を処理
- 外からは中国語を理解しているようにみえる



国立研究開発法人 産業技術総合研究所 人工知能研究センター

AI がこれから進歩した時に何が起こるか、についてはいろいろ意見が出ています。「楽観論」として AI が発達すると、次の世代に行くと Post-Human が出てきて現行の人類ではなくなる。という人がいます。「悲観論」としては AI が人類を駆逐して人間はコンピュータのおこぼれで生きて行くという人もいます。

まず「楽観論」ですが、Google に在籍しているレイ・カーツワイルが提唱している「Singularity」数学用語では「特異点」で、何かが連続ではなく変化するという意味があります。

「人類史のイベント」つまり人類史上起こった主要な出来事をプロットすると、きれいな直線になりますがこれは対数軸で、近年になればなるほど大事なイベントと次のイベントの間が徐々に短くなっています。つまり技術は一直線に伸びるのではなく、急激に立ち上がるという主張です。

「収穫加速の法則」は、ものすごい勢いで物事が進化するので今後ある時点を超えると凄い事が起きるとい主張です。具体的には人工知能の賢さの度合いが、ある一点を超えると AI が AI 自体を改良するので加速すると、AI が科学のあらゆる問題を解いてゆき Nano

Technology が発展して AI によりあらゆる問題が神のように解決する。最終的には人間の脳を Nano Technology Device によって解析すると人間の意識がどのようにになっているかも解るであろう。その情報を吸い上げてコンピュータに入力すれば、自分の体を捨てて、コンピュータの中に移住することが出来る。2045 年まで生きていて充分なお金があれば永遠にコンピュータの中に生きられるという主張です。これを真面目に主張することは凄いことだと思います。

反対に「悲観論」は言っていることはほとんど変わらず、“ある瞬間を以ってコンピュータが制御不能な進歩を遂げる。その結果生まれた知生体は人間とは全く価値観が違った知生体であろう。しかも圧倒的な知力を持っているので、そうなった時に人間はどうなるでしょうか？”この手の話は欧米では流行っているが、日本ではあまり聞かれませんが、シンギュラリティがこんなアグレッシブな話だと解釈されてないと思います。

昨日面白い本を見つけました。『そろそろ人工知能の真実を話そう』という題名で、要するにこれは宗教的な話で“「グノーシス主義」があり善悪二元論では魂は善で肉体は悪であり、悪である肉体の中に魂が捕らわれていて、いつか肉体を離れて善なる世界に旅出つ”という教えがあります。具体的に宗教と結びついているのは「マニ教」ですが、キリスト教の世界では異端とされつつも何となく一般に共有されているような世界観があり、それから出てきているのではないかと。このように解釈すると自然で、意識のアップロードは魂が解放されるということに他ならないわけです。

悲観的なほうは黙示録的な世界観で、“獣の素性を持つものがやってきて世界の終わりが来る”との歴史観に近いように感じます。歴史観に基づくもので、そんなに根拠がないのではと思います。特に「収穫加速の法則」は、これがどこまでも続くのだという話ですが、確かにプロットされている所までは事実ですが、これがどこまでも続くという保証はなく、特に物理的な限界が必ずあるわけで、それについては全く言及していません。だから与太話として聞けばよいと思います。したがって「シンギュラリティ」がきて凄く良いことや悪いことが起こることはないのではないかと思います。

「起こりそうもないこと」として例えば自意識を持った人工知能ができて悪を持って人類を滅ぼすようなことは我々が生きている間はあまり心配しないで良いのではないかと思います。

「もっと心配した方がよいこと」として、社会のインフラ管理を先進的な計算機システムが担う場合、今後自律性が強化されてより賢くなるから計算機が最良の状態で動く範囲が広がっていきます。それがバグや何かの拍子に二つのシステムが相互作用を起こして暴走することはすごく考えられます。



シンギュラリティ



- レイ・カーツワイルらが提唱
- Singularity - 技術的特異点
 - 技術が指数関数的に発展
 - 人工知能がある段階に達すると、人工知能がより優れた人工知能を開発するようになり、人類の制御を離れる
 - 2045年頃??

国立研究開発法人 産業技術総合研究所 人工知能研究センター

Q: いまの AI は知識の自己増殖を繰り返して賢くなるとの話ですが、囲碁の例で Deep Learning が最強の人間に勝てるなら、あと 10 年もすればとても人間は AI に太刀打ちできない状況になるのでゲームソフトの面白みが無くなるのではないかと？

A: 囲碁のソフトを今買っても既に勝てない状況です。でもそれがつまらないかと言えば、考え方の問題で、つまらなくないと言えます。それは AI が手を抜くからです。こちらのレベルに合わせて AI が囲碁を打つのでゲームが成立します。オセロは既に 20 年前にそのような状況になっています。また囲碁もすでに絶対に勝てない状況になっています。また 1997 年にチェスはすでに AI に負けているが、だからと言ってチェスのプレーヤーが減ってはいません。もう一つの考え方で、ブルドーザと力比べをして負けても何とも思いません。それは人間の知性は全方位で一点勝負の AI と勝負をして負けてもなんてことない。そのような話だと思います。

Q: AI の進化によって、例えば 1 平米で 1 億人くらい食べさせられる米を作るような、全体的な人類の生活が豊かになる方向にむかうスピードや可能性はいかがですか？

A: エネルギー保存はどうにもならず 1 平米に対して 1 億人食べさせることは物理的に絶対に無理です。エネルギーの投入量の限界があるので、このような物理的限界は絶対に越えられません。雇用との関係はこの後話します。

●雇用と人工知能

AI の発展に対して雇用は当然影響します。「技術的失業」とはタイピスト、電話交換手など今では完全に消えています。例えば「Amazon Go」ではレジの無いスーパーで、店に入って品物を持って店を出るが会計は自動的に行われているのでレジは必要なくなります。次はアマゾン配送センターで、現在は労働集約的に沢山の人間で箱詰め作業を行っているが、これを全て自動化しようと進めています。するとここでも雇用は失われます。

「雇用の未来」について オックスフォード大学のオズボーン教授がコンピュータで米国の職業の性質をパラメータで調べ、コンピュータに置き換えられるかどうか調べたレポートですが、半分の職業が近い将来無くなるだろうという非常にショッキングな結果がでました。特に注意するのは、レジとか箱詰め作業は無くなるだろうと皆さんも思いますが、そうではないような知的と思われる職業も含まれているのでショッキングな結果でした。

「なくなりそうな職種」このリストでは、アメリカの 702 の職種の中でどれもが 99% なくなると言われています。

例えば銀行の融資担当で、お金を貸して良いか判断するのは典型的に知的な職業に属しますがなくなるだろう。又保険の審査も税金の申告補助も知的な労働のはずですが、これらも全てなくなると言われています。これらはある種のパターン・マッチングなので計算機の方が得意で一網打尽に無くなってしまおうと言われていています。

「残りそうな職種」では、リクレーション療法士、作業療法士、技師装具士、ソーシャルワーカーなどは人間と対峙して多様な状況に対してそれぞれ適応する動作が必要になります。又現場監督者も、現場によって状況が違うのでその場その場で最適な作業を選ぶ必要があるのでこのような職業は残るだろうと言われています。

AIST 残りそうな職種

- リクレーション療法士
- 機械、設置、修理の現場監督者
- 緊急事態管理ディレクタ
- メンタルヘルス薬物乱用ソーシャルワーカー
- 聴覚訓練士
- 作業療法士
- 義肢装具士
- 医療ソーシャルワーカー
- 口腔外科
- 消防現場監督者

国立研究開発法人 産業技術総合研究所 人工知能研究センター

AIST なくなりそうな職種

- 銀行の融資担当
- 写真現像、現像機操作
- 税金申告補助
- 貨物取扱人
- 時計修理工
- 保険審査担当
- 数学技師
- 縫い子 (手縫い)
- 不動産調査
- 電話販売員

国立研究開発法人 産業技術総合研究所 人工知能研究センター

人工智能がどうこうというよりも IT 化によって色々な商売が無くなっているとの話ですが、ゴールドマンサックスの株式トレーダーは、これも比較的知能労働と言われていますが、2000年には600人いたが今では2人になったといわれています。計算機システムの維持に200人雇用したので、差し引き400人が首になったという話です。

例えば弁護士はどうかというと、弁護士は依頼者の話を聞いたり、整理したり法的な組み立てをしたり、賠償額を決定する仕事ですが、話を聞いたり、整理をしたりすることは非常に人間的な部分なので、なかなか置き換えられないと思うが、他の部分はある種のパターン・マッチングです。特に定型なもの、過払い金訴訟とか交通事故の損害賠償額の決定は前例から引っ張って来ているだけですから、比較的簡単にできるので、弁護士の仕事も徐々になくなるだろうと言われています。

では医者はどうか、診療という行為は心理的部分が相当あるので置き換えは難しいのですが、たとえば画像診断では、心電図を見ておかしいと判断したり、肺の写真を見てガンではないかと判断することはすでに置き換わっています。「IBM Watson」が膨大な論文を全て読み込み、その情報を使って非常に特殊な白血病を診断した事例も話題になりました。これはどちらかということと医者を補助する形になりますが、かなり医療の領域にも入っています。

●では何をすべきか。

さきほど「技術的失業」でタイピストの話をしました。これはもう珍しい事ではなく、産業構造は移り変わると言われています。第1次(農業)、第2次(工業)、第3次産業(サービス業)があり、第1次から第3次に向かって労働人口がシフトすることは学生のころ習ったと思いますが、米国の農業人口の推移は1840年には70%の就業割合が今では2%に減っています。工業人口は一度増えて又下がっています。サービス業は単調に増加して今は80%近くまで増えています。これは農業従事者の70%位が他の業種に移行していることを意味しています。この様に産業構造の変化による雇用の変化は特に珍しい話ではないと思います。

1800年初頭の第一次産業革命の時に「ラッドライト運動」がありました。この当時は家内製紙工業で家庭の中に織機を入れて機織りをしていたが、蒸気機関を使った巨大な機織り機が導入され、大量に失職したことがありました。これにより自暴自棄になった労働者が機械の破壊活動を行い何人もが死刑にされました。一時的な騒乱は確かにありましたがわずかな期間で収束しました。これは産業構造が変わるだけで雇用そのものが失われるわけではなく別の所に雇用が発生するわけです。又経済が発展すると需要が拡大するので新たな雇用が発生します。だから失職したからといってそんなに困らなかったのが今までです。

では AI による雇用の変化も今までと同じように吸収できるのかを考えた場合、これはできないと思います。違いはスピードと量です。「スピード」に関してですが、文科省の懇談会資料には今の子どもたちの65%は大学卒業時に今は存在していない職業に就くと言われているように、ものすごい勢いで職業が減っていき新しい職業が増えています。

次は「量」です。今迄は、雇用は別のところで創出されてトータルでは吸収できたけれど、今回は違うだろうと思います。そもそも社会としての富の生産量が充分大きい場合は働く必要があるのかという議論に戻ります。もう一つ考える必要があるのは「貧富の差の拡大」です。横軸が収入で低収入の肉体労働に、一番多い中収入の事務労働、高収入で少数の頭脳労働の構造だとすると、今回の AI によって職を奪われるのは事務労働に集中します。この人たちが雇用の移転先として頭脳労働は難しいので肉体労働に向かいます。結果としてごく少数の頭脳労働者と多数のコンピュータ化が難しい肉体労働者に分けられるという問題が残ります。

私はこの辺り問題については素人なので本当なのかわからないが、今週専門家に聞く機会があり、大体同じ結論だったのでそんなに間違っていないはず。一つ考える必要があるのは「教育システム」です。一人の人間がある職業から別の職業に移る時に頭脳労働間での移動とか事務労働から頭脳労働に移るのは教育が非常に重要です。だから今までの「教育システム」は、子供の時から長期間の教育のあと就労したらその後は何も教育しないというのが一般的ですが、これを就労の間に教育を挟ん

で常に Up-to-Date な職業形態に適応できるような労働力を確保することが必要です。

これだけでは十分でなくもう一つ考える必要があるのは、そもそも働く必要があるのかを真面目に考えることです。ここで出てくるのが「ワークシェアリング」です。オランダでは非常に導入が進んでいると聞きますが、個人の労働時間を週休3日4日で1日3時間のように減らして行き雇用を担保しようという考えです。では収入はどうするのか？ とてもこの労働時間では暮らしていけませんので、「ベーシックインカム」で補うべきだというのが駒澤大学の井上先生の考えです。

「ベーシックインカム」は最近話題になっており、負の所得税ともいわれますが、全ての人間に対して収入にかかわらず定額のお金を支給する代わりに、生活保護とか年金を中止するシステムです。

フィンランドでは実際に導入が検討されていて一部の地域だけですが月 12 万円位のお金を支払ってやってみようという話になっています。

スイスでは導入が国民投票で否決されました。だから日本では話題になってないが、「ベーシックインカム」の導入はそんなに実現性のない話ではありません。では財源はどうするのかとの疑問には、消費税を上げる、富裕層、法人からの税金などで賄うとの議論が出ています。又通貨発行益で賄うという議論を駒澤大学の井上先生がしています。

『限界費用ゼロ社会』という本があり、通常物を作る時作れば作るほど費用は掛かりますが、ソフトウェアなどは一度作れば作る量とコストは関係なくなります。つまりコピーするだけです。色々なものがこのようになると“資本主義のベースである価格決定は需要と供給で決まるので、供給がフラットになると価格が決まらなくなり資本主義の根幹が崩れて今後は別の社会が来る”と言っている人もいます。

「シェアリングエコノミー」で Uber, Airbnb などのコミュニティベースの社会主義の方向に社会は行くべきだという人もいます。結局問題は、“雇用はなんのために働いているのか？”“お金が潤沢にあるのであれば働く必要が無いのでは？”というところまで立ち返って社会的な構造を見直すところまで我々は来ているのではないかと思います。

[質疑応答]

Q：日経新聞で AI が基本的人権を侵すリスクがあると言われていますが、いかがでしょうか？

A：非常に大きな問題だと思います。基本的に今までは何かを公開しても公開した情報に対してアクセスが難しく問題にならなかったです。公開している事と実際に見つかってしまう事の間には大きなギャップがあったわけですが、それが AI を含めて Big Data 解析でギャップが果てしなく小さくなってしまいました。法的対策は有効かどうか、非常に微妙で色々な規制というのはバランスなので技術によってバランスが崩れているのが現状です。それに対して規制で元に戻すのは非常に難しいと思います。となると一つはリテラシー、何かをどこかに書くことは世界の全員が読む事を前提に書くという事に尽きると思います。就職活動をしている人が皆見えていると思って書けという事です。これが良い社会かは別の議論ですが、この方法しか解は無いです。つまり友達は読めるけれど会社の人は読めないという事は不可能です。昔の少人数の村の中で暮らすような気持ちでいるという事です。それは技術的な解決ではなく社会的に解決する事だと思います。

Q：10年後に仕事が無くなる。とネットで話題になっていますが、仕事は生産性の高い AI に任せて我々が行う仕事は無くなるほうがよいと思うが、ただし適切な政策が必要になるわけで、我々は政府にそのような政策を取らせる意識を持つ必要があると思いますが？

- 政治的な対策が必要
 - 職種間の移動、新たに発生する職種への適応を容易にするための職業訓練
 - ベーシック・インカム、ワークシェアリング
 - 新たな社会形態 - コモンズ
- 何のために働くのか

A: 全くおっしゃる通りです。ベーシックインカムが導入されなければ地獄です。仕事が無いわけですから貧富の差が広がる一方です。最後に残る職業は多種多様な仕事を対面で行うという非常に高度な仕事であるコンビニの店員ではないかと思っています。ですからベーシックインカムの導入の議論は今後深まっていくと思いますが、世界的に見て一番反応が悪いのは高齢者の方です。

つまりベーシックインカムは年金を統合して行う様になります。又年金の額はベーシックインカムの額よりは大体下がりますので年金受給者は大抵反対します。それをどのように考えるかは非常に難しい問題です。人間より生産性が高い AI が代って働いても全員の生活が豊かにならないのは富の再分配が上手く行ってないので政策を変えるしかありません。

Q: 先生の働いている人工知能研究センターは日本の政府のどこの管轄ですか？

A: 人工知能研究センターの管轄は経産省傘下になります。

中田秀基（なかだひでもと）先生のプロフィール

現職 独立行政法人産業技術総合研究所人工知能研究センター
脳型人工知能研究チームチーム長

経歴

1995 年 東京大学情報工学博士課程修了博士（工学）
同年 工業技術院電子技術総合研究所入所
1999 年 カリフォルニア大学サンディエゴ校客員研究員
2001 年 産業技術総合研究所に改組
2004 年 ウィスコンシン大学マジソン校客員研究員
2001 年-2007 東京工業大学客員助教授
2008 年 ENS リヨン客員研究員
2011 年- 筑波大学連携大学院教授併任
2013 年 7 月- 2014 年 6 月 経済産業省商務情報政策局情報処理振興課出向

著書

- ・『トコトンやさしい人工知能の本(今日からモノ知りシリーズ)』（一部著），日刊工業新聞社，2016 年
- ・『すっきりわかる Google App Engine for Java クラウドプログラミング』ソフトバンククリエイティブ，2010 年
- ・『雲の世界の向こうをつかむクラウドの技術』（一部著），アスキー・メディアワークス，2009 年
- ・『岩波講座インターネット 4 ネットワークアプリケーション』（一部著），岩波書店，2003 年
- ・他

翻訳

- ・『Python ではじめる機械学習』，オライリー・ジャパン，2017 年
- ・『ZooKeeper による分散システム管理』，オライリー・ジャパン，2014 年
- ・『Java サーブレットプログラミング』，オライリー・ジャパン，2002 年
- ・他

監訳

- ・『データ分析によるネットワークセキュリティ』，オライリー・ジャパン，2016 年
- ・『Cython -C との融合による Python の高速化』，オライリー・ジャパン，2015 年
- ・『Head First C -頭とからだで覚える C の基本』，オライリー・ジャパン，2013 年
- ・『デバッグの理論と実践なぜプログラムはうまく動かないのか』，オライリー・ジャパン，2012 年 ・他