

レクチャーシリーズ 「人工知能とは」 [第7回]

人工知能とは？ (7)

What's AI? (7)

松原 仁
Hitoshi Matsubara
公立ほこだて未来大学
Future University of Hakodate.
matsubar@fun.ac.jp

Keywords: artificial intelligence, Astro Boy, artificial general intelligence.

問い：人工知能とは何か。

答え：究極には人間と区別がつかない人工的な知能のこと。

人間と区別がつかないというのは狭い意味でのチューリングテストに合格することではない。普通に付き合っている関係で人間でないとは気が付かない人工的な存在（ロボット）ができることを意味している。例えば転校生がクラスに来て3か月あるいは半年ほど経過して先生やクラスメートは皆それが人間だと信じてきたけれど、実はそれが人工物（ロボット）であることがカミングアウトされてみんなが驚くという状況を想定している。著者はこれを「風の又三郎」テストと称している。食事や排泄などの生理的な部分まで完璧に真似することは大変でなおかつ意味があまりないと思われるので、生理的な部分はごまかせる程度でよしとしたい。ちなみに人工的な存在を実現するものはコンピュータとロボットには限定されていないものとする。

著者の表現で言えば「鉄腕アトム」が人工知能のシンボルである。手塚治虫の鉄腕アトムは見掛けでロボットとわかる形状をしているが、浦沢直樹がリメイクした鉄腕アトムは人間と区別がつかない。浦沢直樹のリメイク版鉄腕アトムが人工知能に相当する（浦沢直樹の鉄腕アトムをご覧になっていない方は彼の「Pluto」の第1巻（小学館，2004）をぜひご覧いただきたい）。

人工的な知能は物理的な実体でなくコンピュータの中の存在でもよいという立場もあると思う。著者は人間が五感を含めて相手と応対して人間か人工物か区別できないという点にこだわりたい。チューリングテストなどはインタフェースのチャンネルが細過ぎる。また人間とは全く異なる人工の知能も存在の可能性はあると思うが、その知能と人間とはともにコミュニケーションできないと思うので除外する。

問い：人工知能研究とは何か。

答え：人工知能を実現しようという試みを通じて知能を構成的に理解すること。

これは標準的な回答だと著者は思っている。中島秀之氏の言う（説明原理ではなく）動作原理を追求することである。風の又三郎あるいは鉄腕アトムをつくる過程を通じて人間の知能はどうすればできるのか（人間の知能はどのようにしてできてきたのか）を、ほぼ同じものをつくることのできる程度に理解することを目指しているのが人工知能の研究である。

人工知能を実現することが研究の直接の目的ではないことには意味がある。人工知能は簡単には実現できない。チェスや将棋で名人に勝ってもクイズで世界チャンピオンに勝ってもそれは知能のごく一部の個別の問題で人間に追いつき追い越したに過ぎず、人間の知能がもつ汎用性は実現できていない。後で述べるように知能で一番重要なのは汎用性である。汎用性を有する人工物はまだしばらくは実現できないと思う。これまでの人工知能の研究は失敗を繰り返した歴史であり、これからしばらくの間も失敗を繰り返すと予想される。偉大な失敗（こうやっても人工知能は実現しないとわかること）こそが人工知能の研究であり続ける。

問い：intelligence amplifierの研究は人工知能の研究か。

答え：情報処理の研究ではあるが、人工知能の研究ではない。

intelligence amplifierをつくる過程を通じて人工知能の実現（人間を助ける補助ではなく自律的に働く知能の実現）を目指すのであれば人工知能の研究といえるが、役に立つintelligence amplifierをつくること自体が目的であるならばそれは工学的に意義はあるが人工知能の研究ではないと思う。著者の考える人工知能の研究には役に立つ道具をつくるという目的は含まれない。それは情報処理の研究であろう（人工知能と情報処理のどちらが上で下だというつもりは全くないが、役に立つ道具をつくることまで人工知能研究に含めるとあまりに範囲が広がり過ぎるという認識である）。人工知能の研究をしている過程で副産物として役に立つ道具ができることは十分にあり得る。

問い：脳科学の研究が進めば人工知能研究は不要になるのではないか。

答え：いいえ。

最近の脳科学の進歩は目覚ましく多くの知見が得られている。人工知能はそれらの知見を生かして進めていくことが期待されるが、それは人工知能が今後脳科学にとって代わられるのとは異なる。脳科学はあくまで分析的に知能を理解することが目的（の一部）であって、人工知能のように構成的に知能を理解することとは異なる。分析的に理解することと構成的に理解することの違いは、前者はいわば十分条件を求めているのであり、後者は必要十分条件を求めているということである。今後さらに脳科学が進歩していったとしても人工知能研究の意義はなくならないと思う。人工知能研究の意義がなくなるとすれば、それは人工知能が実現されたときである。

ついでに言えば、一部の人工知能の研究がタコソボ化して面白くなくなってきたのは、数学あるいは物理学に対する過度な依存の影響があると思う。人工知能が知能という未定義の概念を目標としたあやしいものである。人工知能をまっとうなものにするためのつつい既存の権威にすがってしまいがちなのである。そうすると一見完成度は上がったように見えるかもしれないが、人工知能の実現という大目標からは離れてしまいがちである。定式化したりグラフを書いたりするのはあくまで手段に過ぎないのであって、それらが目的化してしまってはならない。

問い：知能とは何か。

答え：未知の状況に対して（死なない程度に）適切に対応する能力のこと。

進化論によれば、人間は種の保存のために知能という能力を獲得してきたと考えられる。その人間が地球上で（少なくとも今現在は）最も高度な知能をもって生物の頂点に君臨できているのは、人間がもっている未知の状況に対する対応能力のおかげである。そのおかげで幾多の困難を乗り越えて種として保存され、個としても生き延びてきた。チェスや将棋で未知の局面で次の良い指し手を求めるのも同様であるが、ゲームでは可能な合法手が離散的で数も限定的である。人生においては無限の連続的な選択肢の中から適切なものを選ばなくてはならないのではるかに難しい。もっともチェスや将棋は 100 点ないしは 100 点に近い手を指すことが求められるが、人生は死なない程度に 60 点以上の点を取り続けることが求められる。

知能の具体的な中身についてはこの連載でこれまでに述べられてきたのでここでは触れないこととする。

問い：知能にとって本質は何か。

答え：どのような状況に対してもそれなりに対応できる汎用性。

未知の状況とはあらかじめ予想がつかないということである。そのすべてにうまく対応するのは不可能であるけれども、多くの場合にある程度（繰り返しになるが死なない程度に）対応する能力が知能であるので、知能の最も重要な要素は汎用性である。溝口理一郎氏および堀浩一氏が言及している「分散認知」、「環境とのインタラクション」、「symbol grounding（記号接地）」、「身体性やインタラクション」などは著者としては汎用性という本質に迫るための最近の人工知能研究の道具立てあるいは切り口とみなしている。例えば、実際にロボットをつくるほうが人工知能は実現しやすいと信じているが、シミュレーション技術がさらに進歩すれば、コンピュータの中に人工知能を実現することも理屈としては不可能ではないと考える。

浅田 稔氏がこの連載の中で IJCAI に日本の参加者が非常に少ないことを嘆いていたが、著者も最近 IJCAI に参加していない。それは著者にとって参加する意義が薄れてしまっているからである。正直に言って、そもそも論文を投稿しようという気にもなくなってしまう。聞いていて面白い発表、あるいは読んでいて面白い論文が極端に少なくなってしまった。これは著者が年を取って感受性が弱くなったという側面も否定できないが、人工知能の研究領域が細分され過ぎてしまったほかの領域の研究者には意味が理解できない発表が増えてしまったことに主な原因があると思う（著者の周りの若い人工知能研究者もみんな IJCAI は面白くなくなってきたと言っているので著者だけの感想ではないだろう）。人工知能の初期の頃はあやしいけれど夢のある発表が IJCAI でもよくあったのであるが、人工知能という研究領域が成熟していくに従って細分化されて洗練されていって完成度は高いかもしれないが夢が感じられず何が面白いかわからない発表が増えていった（悪く言えばタコソボ化していった）。それで IJCAI など全般的な会議を敬遠するようになってしまったのである。この質問の文脈で言えば IJCAI が知能の本質を追及する学会ではなくなったとみなされていることだと思う。知能の研究者は機械学習やエージェントなど、それぞれの研究領域の国際会議に参加して発表する傾向が顕著である。

浅田 稔氏が指摘したように、人工的な知能を実現するうえで人工知能とロボティクスの連携は非常に重要である。著者も人工知能研究を（非公式に将棋のプログラムを書いていたものの）ロボットのビジョンの研究から始めたので、一部の人工知能研究者がパターン認識とロボティクスへの意識が低いことは残念に感じる。意識が低い人達はいわば「きれいごと」に終始しているのである。知能は少数の公理で説明できるようなものではない。情

報は常にノイズを含んでいる。人工知能研究を目指す人には一度はパターン認識を扱ってロボットに触ってみることを強く勧めたい。知能とはきれいごとではすまないことを実感してから先に進んでほしいと願っている。

日本の人工知能学会の全国大会は全般的な会議ではあるが、おかげさまで数多くの参加者を得ている。同じ全般的な会議で IJCAI と異なるのはなぜかを考えることには意味があるだろう。日本の人工知能研究がガラパゴス化して世界の潮流から乖離しているという見方もあるかもしれない。確かに日本での流行がアメリカやヨーロッパの流行と異なってきている。それは従来の欧米追従からするとむしろ望ましいことだと思う。ガラパゴス化したままで終わるのか、日本発の研究が世界を変えることができるのか、それこそ我々の今後にかかっている。ガラパゴス化といえば最近では欧米のほうにそれが目立つような気がする。例えば非単調論理の一部の研究は欧米でとても盛んであったが、知能の人工的な実現にはほとんど何も貢献しない(例えば、狭い意味でのフレーム問題が解けたという研究があったが、知能とは無関係である)。あれらは人工知能の研究ではなく数学の研究である。それなのに完成度が高いためか IJCAI や AI Journal にその関係の論文がいくつも出ているので聞く気も読む気もなくなってしまふ(採録率があまりに厳しくなってしまったことによる悪影響であろう)。さいわい今の人工知能全国大会のオーガナイズドセッションはたとえあやしくて完成度が低くても(もちろんまっとうで完成度が高いものもあるが)人工知能の未来を感じさせるものが多い。ぜひこの傾向が続いてほしいと思っている。

問：フレーム問題は人工知能に解けるのか。

答：人間が解いているとすれば解ける。人間が解いていないとすれば解かなくても人工知能は実現できるはずである。

コンピュータに個別の問題、例えばコンピュータ将棋をさせているときにはフレーム問題は生じない(いわばプログラマがフレーム問題を回避させているのである)が、汎用性をもたせようとするときと直面する。我々は McCarthy & Hayes らの狭い意味(記述の爆発)と区別して広い意味で一般化フレーム問題と名付けたが、一般化フレーム問題は直面してしまうと人間にも解けない。フレーム問題は人工知能すなわちコンピュータやロボットだけの問題で人間にはフレーム問題は存在しないという批判を心理学者や哲学者から受けるが、著者としては彼らが人間には存在しないという言い方を好むのであればそれはそれで構わない。しかし知能を構成的に実現しようとするとき一般化フレーム問題をどこかで回避しなくてはならないという意味で、一般化フレーム問題の存在が明らかになったのは人工知能研究の知能に対する大

きな成果なのだと思っている。哲学や心理学はずっと人間あるいは動物の知能だけを対象としてきた(進化的にうまくいくようになった知能だけを見てきた)ので、うまく働かない知能に目が届かなかったのだと思う。著者が25年ほど前に一般化フレーム問題を提唱してきたときからフレーム問題にまつわる状況はほとんど変化がない。

ちなみに子供のロボットをうまくつくって実世界で育てれば一般化フレーム問題をほとんどの場合に回避できる(あるいはフレーム問題が存在しない)ロボットに成長できると考えている。いまはまだ子供のロボットがうまくつけれないが、将来は可能と楽観している。

問：記号接地問題は人工知能に解けるのか。

答：解ける。

記号接地問題(symbol grounding problem)はフレーム問題と並んで人工知能の難問として取り上げられることが多い。今はリングという記号とリングの実体が人間のようにコンピュータは接地できていないのはそのとおりであり、そのためにリングに対してコンピュータは人間のような適切な対応ができない。

今のコンピュータに記号接地問題が解けないのはコンピュータがリングを「体験」できていないからである。(一般化)フレーム問題のように子供のロボットをつくらせて(人間の子供のように)リングの「体験」を積み重ねればそのロボットはそのロボットなりに(おそらく人間とはかなり異なった形で)リングという記号とリングという実態を接地できるはずである。

問：コンピュータ将棋のような個別の問題を扱って汎用性につながるのか。

答：いきなり汎用性を得るのは難しいので研究の方法論として個別の問題をつぶしているつもりである。

Minsky は個別の問題ばかり解こうとしている傾向を一貫して批判している。人工知能のパイオニアの研究者はみんな一度はコンピュータチェスを研究しているが、Minsky だけはコンピュータチェスなどやっつけてもだめだと言って手を付けなかった。著者が最も尊敬している人工知能研究者は(ずっと相変わらずに)Minsky である。Minsky は人工知能研究で最も難しくても意味があるのは5歳ぐらいの子供の遊びを再現できるシステムをつくることだと言いつづけている。

5歳ぐらいの子供の遊びを再現するのが最も難しくても意味があるというのには賛成である。それが知能の汎用性という本質に関わっているからである。しかし Minsky と著者はそれを実現するための方法論が(僭越ながら)異なるのである。最も難しい目標を一気に達成

するのは当然ながら非常に難しい。Minsky も多くの優れたアイデアを提供しているが、最も難しい目標を達成できてはいない。これからもそう簡単には Minsky 以外の研究者も達成できないと思われる。

異なる方法論というのが個別の問題に取り組んで一つ一つつぶしていくという地道なものである。それをできる人間はそれなりの知能をもっていると思える個別の問題をコンピュータで解いていく。チェスの世界チャンピオンに勝つという問題がそうであり（1997年に達成した）、クイズ番組のジョパディのチャンピオンに勝つという問題がそうであり（2011年に達成した）、将棋の名人に勝つという問題がそうであり（あと数年で事実上達成される）、東大入試の合格点を取る（2021年を目指している）という問題がそうである。その過程を通じて知能について何らかの知見を積み重ねていくことで汎用性という知能の本質に迫ることができるのではないかと期待しているのである。また、コンピュータ将棋がそうなるかどうかはわからないが、個別の研究から人工知能全体のブレークスルーが生じる可能性がある。具体的な例題を解くという方法論はそう捨てたものではないと思う。この期待がうまくいくかどうかは今の時点ではわからない。人工知能が実現できたときになって初めてわかることである。

問い：コンピュータは心をもてるか。

答え：もてる。

問い：コンピュータは意識をもてるか。

答え：もてる。

これらの問いに対する著者の回答は長尾 真氏のものに近い。哲学者や心理学者の中には「コンピュータが心をもっている」ことと「コンピュータがあたかも心をもっているように見える」ことを区別したがる人達がいるが、これらは決して区別できないという立場を取る。あえて心や意識を定義するとすれば、堀 浩一氏が述べているように下位の要素の間の相互関係の総体であると思う。

よくいわれるように人間同士でも他人が心をもっていること、意識をもっていることを証明することはできない。他人も（自分と同じような）心や意識をもっているとみなしてやり取りをすることが自分にとって便利なのでそうしているに過ぎないと思う。そのことと同様にこのロボットに心がある、あるいは意識があるとみなしてやり取りすることが人間にとって便利ということになれば、それはそのロボットが心あるいは意識をもっていることである。

ロボットがある程度以上に汎用性をもって複雑な挙動を示すようになれば、そのロボットを理解したり行動を予測したりするのに心や意識の存在を仮定したほうが便利になるはずである。そういうロボットは実現すると確信している（今はできていないが、将来にわたってでき

ないという理由が存在しない）。

問い：コンピュータは創造性をもてるか。

答え：もてる。

この連載で堀 浩一氏が書いているように、特に「創造性」という特殊な能力が人間（だけ）に備わっているのではない。「創造性」は人間の知能の働きをある側面から見た概念に過ぎない。創造性が新しいものを発想することであるならば、ずっと前からコンピュータは創造性を有している。問題なのは発想した新しいもののほとんどが的外れで使いものにならなかったということである。

例えばコンピュータ将棋を取り上げてみよう。コンピュータ将棋はもっぱらプロ棋士の棋譜から機械学習によって評価関数をつくっている。教師データがプロ棋士の過去の棋譜ということは、できた評価関数は過去を反映したものに過ぎない。しかし最近のコンピュータ将棋は未知の局面（学習データにはなかった局面）でプロ棋士が高く評価する新手を「創造」している。例えば第2回電王戦でコンピュータ将棋の GPS 将棋が三浦八段相手にある局面で初めて指した「8 四銀」はその後プロ棋士の間での定跡となった。あるいは 2013 年の名人戦で羽生三冠相手に森内名人が指した「3 七銀打ち」はコンピュータ将棋の Bonanza が指した手を森内名人が偶然知っていわば真似をしたものである。コンピュータ将棋は明らかに新手を創造しているのである。

パズルは制約条件が厳しいのでそれだけコンピュータにとって新しい作品を創造しやすい。数独（ナンバープレース）の新しい問題はコンピュータがかなりの数を創作している。著者は以前詰将棋の創作の研究に従事していたが、一定水準以上の作品ができるようになっていく。

パズル以外の領域でも今後はコンピュータによる創造が少しずつ実現していくものと思われる。創造性というのは決して神秘的な能力ではないのである。

問い：コンピュータにショートショートを自動創作させることができればコンピュータが創造性をもったことになるのか。

答え：個別の問題をつぶすという作戦の一つであるが、一般の人に対して象徴的な良い例になると期待している。

著者はずっと 30 年以上手掛けてきたコンピュータ将棋がその目標（名人に勝つこと）達成が時間の問題になったので、それに代わる研究テーマとしてコンピュータにショートショートを自動創作させるというプロジェクトを開始した。ショートショートに厳密な定義はないが原稿用紙 20 枚以内（8000 字以内）程度といわれている。ショートショートの第一人者は星 新一で彼は生涯

に1000作以上のショートショートを書いている。著作権者の協力が得られて彼の作品の電子データを扱うこともできるようになったので、それを教師データとして星新一のようなショートショートをコンピュータに創作させることを目指している。

コンピュータにとってパズルなどに比べて小説の創作は難しい。制約条件が緩いので候補作品が絞り込みにくいのである。しかしある水準のショートショートを創作できるものと期待している。その過程で人間がどのようにショートショートを創作しているかの知見も得られるとうれしい。

コンピュータにショートショートがつくれたからといって人間のもつ汎用の創造性に比べると個別の問題を解いているに過ぎない。しかし小説の創作は人間にとっても特別な能力とみなされているので、コンピュータにそれができるといえることになれば人工知能の象徴としてかなりのインパクトがあると考えている。

問い：人工知能は実現できるか。

答え：できる。

技術的にできない理由が存在しない。著者の生きているうちには実現できなくてもそう遠くない将来に実現できると確信している。だからこそ人工知能の研究をしているのである。

ゲームの例が多くて恐縮であるが、コンピュータチェスの研究が始まってしばらくの間は絶対に永久にコンピュータが名人に勝てないといわれていた。人工知能批判の哲学者の Dreyfus が言っていたのは有名である（彼は「コンピュータは自分にも勝てない」と言いすぎてしまっただけで公開対局で当時の弱いコンピュータチェスに負けて大恥をかいた）が、当時はそう言う人が多数派であった。1997年に世界チャンピオンの Kasparov がコンピュータチェスの Deep Blue に負けると彼らは一転して「チェスは人工的なゲームでルールが明確なのでコンピュータが勝つのは当然だ。チェスは単純な問題だったのだ。チェスで人間に勝ってもコンピュータが知能をもったことにはならない」と言っている。チェスで勝ったからといってコンピュータが人間並みの知能をもったことにはならないのはそのとおりであるが、発言がぶれるのはみっともない。

将棋も同様である。30年前のコンピュータ将棋はとても弱かったので多くの人が永久にプロ棋士には勝てないと言っていた。最近強くなってきてプロ棋士と良い勝負をするようになるとプロ棋士側は大慌てである。時期を逸しないためにはこの数年の間に名人とコンピュータ将棋の対戦を実現するしかない（それを過ぎるとコンピュータが圧勝して勝負にならなくなる）のであるが、プロ棋士側の大慌てのために時期を逸してしまうことを危惧している。

これまで到底できないといわれていた個別の問題が解決されている。最終目標である（汎用の）人工知能の実現もできないと考える理由が存在しない。

あることが人工知能でできるようになるとそれはたいしたことではなかったと言いだすことを「人工知能効果」と呼ぶそうである。これは人間しかできないと思われていたこと（例えば道具を使うこと）をほかの動物ができるとわかったときに見られる効果（そのことは人間にとってたいしたことではなかったと言いだすこと）をコンピュータにも適用したものである。人間の尊厳を保ちたいという意識から来るものであろう。確かに将棋はルールが限定されていて実世界からはほど遠いので将棋に強いことなどたいしたことはないという言明を最近よく聞くようになってきた。コンピュータに対するいわば「負け惜しみ」である。人工知能の発展によって一般の知能に対する考え方が変わってきたのである。当然ながら、たとえコンピュータ将棋に負けるようになったとしても人間の将棋の強い人すなわちプロ棋士の価値は変わらない。四則演算がコンピュータのほうが速くて正確でも人間の価値が変わらないのと同じである。

問い：人工知能研究は今後どういう方向に進むべきか。

答え：個別な研究はある程度進んできたのでそろそろいったんは汎用性を追求すべきである。人工汎用知能 (artificial general intelligence) は人工知能研究がもともと目指していたものだと思う。

すでに述べたように、最終的には汎用な人工知能を実現することを目指しながらいまは多くの研究が個別の問題を対象としている。それは方法論として妥当と考えているが、ずっと個別の問題にこだわっていると汎用性という本質を見失ってしまう危険がある。分散と統合とすれば個別の問題は分散に相当するので、ときどきそれらの統合を試みて汎用性を検討するのがよいと思う。

とはいえ、今統合してもすぐに汎用の人工知能が実現できるとは思っていない。統合の試みはまず失敗する。それでもときどき統合を試みてその時点で何ができていて何ができていないかを問うことが重要である。

最近になって人工汎用知能というものが取り沙汰されるようになった。これは今の人工知能研究が個別の問題ばかりを対象にしている現状に対して汎用性を目指す試みとみなすことができる。人工知能の研究も最初の頃は一般的な枠組みを志向するものが多かった。Newell & Simon の GPS (General Problem Solver) や Minsky のフレーム理論などが代表的である。知能の本質は汎用性にあるので一般的な枠組みを志向するのは当然であるが、いきなり一般的な枠組みを構築しても解くことが期待されている個別の問題は解けないので、その後の人工知能研究は直接個別の問題の解決を目指す方向に進んだのである。もともとの原点への回帰をしようとしている

のが人工汎用知能だと思う。

問い：人工知能の研究が進み過ぎると問題が生じるか。

答え：問題が生じる可能性はあるので人工知能研究者には技術的な指針を示す責任がある。

西田豊明氏と堀 浩一氏が言及しているように、人工知能が人間を凌駕することがそろそろ視野に入ってきている。凌駕すると人間には人工知能を制御することが難しくなることが考えられる。Kurzweilは凌駕することを技術的特異点 (technical singularity) と呼び、技術的特異点に来る時期を 2045 年と予想している (2045 年問題と呼ばれている)。

それが 2045 年なのかどうかはともかくとして、それほど遠くない将来に凌駕する時期が来る (「超知能」ができる) と考えられる。人工知能の実現を妨げる理由がないのと同様に、技術的特異点を訪れることを否定する理由は存在しない。そのときの「超知能」が著者の言う「人工知能の実現」に相当するのかは現時点ではわからない (直観としては異なるものとなるような気がしている)。ともあれ「超知能」と人間がどう付き合っていくか今のうちからよく考えておく必要があると思う。

すでに指摘されているように、「超知能」は複雑系 (人間の脳よりも複雑) なのでその挙動を正確に予測したり制御したりするのは不可能である。「超知能」は人間をそれまで以上に幸福にしてくれるかもしれないが、不幸のどん底にたたき落とすかもしれない (そういう能力を有しているのである)。最新の科学技術のほとんどがそうであるように悪い目的に使われるとひどいことになる。人工知能も我々研究者が望もうと望まざるとにかかわらずそういう段階に来たことをよく認識する必要がある。

西欧の SF では「超知能」は必ずといってよいほど人間に対して敵対する。それはこれまで人間だけがほかと

区別された特別な存在とされてきたことに対する裏返しだと思う。ボスとして君臨してきたので、ボスの座をほかに奪われると大変なのである。日本では人間とほかの存在は連続的であって明確に区別されたものではない。おおげさに言えば「超知能」の時代に人間と「超知能」が共存するのに必要なのはこの日本的な考え方である。我々人工知能研究者は (このままいけば) 遠くない将来に技術的特異点を訪れる可能性が高いことを世の中に伝えて社会的に「超知能」をつくるべきなのか、つくるとすればどのようにするのが望ましいか (堀 浩一氏が述べているように完全には制御できないにしてもある程度制御しやすい構造にすることは十分に可能である) という議論を始めないといけない。人工知能学会誌でもこのテーマの特集が近いうちに組まれるということなのでそれを議論のきっかけにできればと期待している。

謝 辞

本稿に対して編集委員長の松尾豊氏に貴重なコメントをいただいた。深く感謝する。

2013 年 12 月 12 日 受理

著 者 紹 介



松原 仁 (正会員)

1986 年東京大学大学院工学系研究科情報工学専攻博士課程修了。工学博士。同年、通産省工技院電子技術総合研究所 (現 産業技術総合研究所) 入所。2000 年公立はこだて未来大学教授。人工知能、ゲーム情報学、エンタテインメントコンピューティング、観光情報学などに興味をもつ。著書に「将棋とコンピュータ」(共立出版、1994)、「鉄腕アトムは実現できるか」(河出書房新社、1999)、「コンピュータ将棋の進歩 1-6」(編著、共立出版、1996～2012)、「先を読む頭脳」(共著、新潮文庫、2009) など。人工知能学会副会長、NPO 観光情報学会会長、情報処理学会理事ほか。