

特集 ■ 第 15 回大会

特別講演 2

ロボカップから生命体都市ロボシティコアへ



浅田 稔
Asada Minoru

大阪大学

川原: 特別講演 2 は大阪大学大学院工学研究科知能・機能創成工学専攻の浅田先生にお願いしました。

大変お忙しい中、今回の日本バーチャルリアリティ学会で「ロボカップから生命体都市ロボシティコアへ」という題でご講演を引き受けていただきました。

それでは浅田先生お願いします。

浅田: 大阪大学の浅田です、こんにちは。バーチャルリアリティ学会ですが、実は僕は会員ではないのです。けれど会員の先生方は良く存じていまして、プロの前で話すのでちょっと恥ずかしいんですが、学生さんが多いということで易しめで話しますが中身は高等だと思っています。

これはバーチャルリアリティ学会初代会長の館先生の言葉で非常にインプレッシブです。「バーチャルとは本質のあるいは効果として現実であり原物であることである」ということで、これはまさしく人とロボット、特に人を真似た人工物としてのロボットは、何を本質として捉えて設計すべきかという点で非常に意味のある言葉だと思います。

我々は人工物設計と言ってるけれど、じゃあ人工物とはなんぞやというと、台湾のアーティストの言葉を借りれば、「人工物 = 芸術 + 自己 + 真実」であるということです。これは何の話かということ「art + i + fact」ということで、我々の人工物とはロボットのことでロボットとはアートであり自分を表現するものであり、しかもリアリティという事実を持っているという意味ではこの英語は正しいと思います。

今日は 1 時間のトークでスライド 124 枚お見せできるとは思わないのですが、自分たちが持っているロボットに対する思いをいろんな側面から説明できればと思います。タイトルを「ロボカップからロボシティコアへ」と

したのですが、ロボカップはイントロで、ロボシティコアはエピソードで、ロボットを通じて人間の心を理解するというところを中心にお話しようと思ってます。

実は僕はバーチャルリアリティ学会には一回も出たことがないのですが、プログラムを見るとかなり近いことをやられているなというイメージがありまして、人間を知るための道具としてロボットを捉えたときに、人間の構造や人間のいろんな側面を知らなきゃいけないという意味で、かなり共有する部分があると思います。

IT ですが、最初は R の中に I があるじゃないかと言っていたのですが、I の人たちがそれを許さなくて IRT、さらに最近では ICRT 「インフォメーション・コミュニケーション・ロボットテクノロジー」と呼んでます。(スライドを示し)上はロボットらしいロボットで、下はロボットらしくないロボットで左からナノマシン(難波プロトニックナノマシンプロジェクト)で最終的にはドラッグデリバリーシステムを目指していて、ロボットが体の中に入ってくる。

その隣がダ・ヴィンチという手術システム、次が顎の矯正システム(早稲田大学)、知的交通システム、宇宙ということで実はロボットらしくないロボットがいろんなところで働いていて、ロボットらしいロボットが実はあまり働いていないようで、癒しなどの何らかの価値があるかもしれない。

IT と RT は別物ではなくて、クラウドロボティクスという形で情報ネットワークとリアルな世界がシームレスにつながりながらいろんなものを作っていくだろうということで、ロボシティコアというものを私は提案しており、それは、今日の話の一つです。

昨日、ATR の萩田所長に別のところで講演していたのですが、インビジブルという話がありました。既

にロボットもコンシャス・アンコンシャス、それにバーチャル型というインビジブルな部分、環境に埋め込まれた部分と連携していかなければならず、Wi-Fiがキーテクノロジーになってくると言われています。

ロボットを何のために作るかという話なのですが、我々のプロジェクトでは基本的に人間というものをどう理解するかという第3の方法として提案しています。第一は、脳科学や神経科学などの自然科学的アプローチであり、第二は、哲学、心理学、社会学などの人文科学的アプローチです。何の役に立つかと言われたら、最終的には役に立つと思っていますが、すぐさまの機能を実現するという意味でのロボットにはなっていないので、そこは言い訳しています。

いずれにしても、従来はロボットというと「センサーで情報を取ってきて」、「判断して」、「動かさない」という三つの要素の集合と言ってたんですが、そんなことは言われてられなくて、実は皮膚が大事だよとかいろんな技術が入ってくる、環境が入ってくるという意味では人間理解から始まるもっとも集約度の高いサイエンステクノロジーだろうと思っています。そういった中で今のロボットはまだ満足のものになっていない、人々の期待する漫画やSFのロボットとは違ってまだまだギャップがあるということでそのギャップを埋めるためにどうしたらいいんだろうということで始めたのがロボカップです。

アイデアとしては1993年くらい、もう17、8年前に構想し、第1回大会が1997年です。ロボカップで何が良かったかという、標準問題を決めてみんなを進めるといって特に複数のロボットが協調するということは人間と人間のコミュニケーションの問題、人間と人工物のコミュニケーションの問題、インタフェースの問題も入ってきてます。困難なテーマにチャレンジすることで次の世代の産業を作っていこうということです。更にもかく面白いことをやろうと思ったわけです。一番大切な事は皆が面白いと思うことです。

皆が面白いと思うという、非常に根源的な情動的な部分にどう訴えていくかということは、非常に大きな課題だと思っています。センセーショナルなターゲットとしては人間に勝つよと言ってるんですが、これがなかなか難しい。それで97年にやったときにはたったの40チームだったのが今は400から450チームくらいで、時間と空間の限界で2000年半ばで飽和しているんですね。10年くらいで右肩上がりです。これだけ増えてきて、素晴らしく大成功なんですけど、運営はむちゃくちゃ大変です。そ

して喜ばしいことは参加者の半分は小学校、中学校、高校までの生徒たちです。子供たちが凄く夢を持って参加しています。つたない英語でもそれを使いながらコミュニケーションをするということが広がれば、全世界的に紛争もなくなるんじゃないかと思っています。

ロボカップの技術躍進は非常に速くて、当初は全然動いていなかったのに1年2年経つとすぐさま技術躍進。チームが増えるとレベルが下がったんですけど、逆の現象が起きて、どんどん新しいことにチャレンジしています。小型リーグでは天井からリモートセンシングしてるんですが、中型では自律分散で各ロボットが視覚・運動系を全部持って赤いボールを追っています。今までは色情報に頼っていたんですが、最近ではボールも通常のサッカーボールに替えて、形とかいろんな情報を使ってやりましょうとチャレンジしています。あとは全方位移動・全方位視覚ということで人間とかなり違った視覚・運動系になってるんで、これがゆくゆくはヒューマノイドになるだろうということです。

アイボリーグはある意味一番人気でして、小型から中型、中型からアイボと段々と人間に近づいています。アイボは鼻のところにカメラがあって非常に視覚が狭いので、倒れこみヘディングシュートとかいろんなチームがいろんな技術を開発しています。

ここでのポイントは、ロボカップは公開競技でいろんな人たちが入ってきて、そのことによってまず技術の公開、交流がなされるのでアームレース的に技術が上がってくる。

例えば2000年のアイボリーグでは、UNSWというオーストラリアのニューサウスウェルス大学がシュートの手法を開発しまして、翌年にはすべてのチームが使うということになったわけです。もちろん技術を盗むことはOK、お互いに盗み合いながら閉じるのではなくオープンにすることで技術を挙げていこうという一つの例です。

2005年大阪大会では遺伝的アルゴリズムを使ってハードウェアに最適な歩行法を見つけて、全チームが同じ歩き方をするというのもありました。2002年からヒューマノイドリーグが始まりまして、アシモに来てもらってデモンストレーションをしてもらいました。背中のランドセルが非常に大きいんですが、8年前なので今ではずっと小さくなっていて、ちょっとずつ確実にマイナーチェンジをしています。このデザインをどう評価するか色々議論があるんですが、これも一つのレファレンスにはなるかなというところなんです。

ホンダのアシモには PK のデモンストレーションをしてもらいました。8 年前の大会にうちもロボットを出しました。Senchans という富士通の HOAP を改造したもののなんですが、当時は PK 戦しかなくて有線も許されていませんでした。シュートは蹴ることよりも蹴ったあとに立っているということの方が難しいんです。キーパーの時には色を識別して右か左か判断して飛び込むんですが、彼は右肩を脱臼しまして、彼は痛くないんですけど私は痛みを感じまして。なぜかというと彼はベースが 550 万するんですよ、改造費が 100 万くらいで全部で 650 万位するんですよ。2004 年から石黒浩阪大教授が影の監督をするチーム大阪が参加しまして、非常に強いんですよ。当時はキーパーが寝転んでゴールを塞ぐということが許されていたんです。頭に全方位カメラがついてまして、狭い隙間を狙ってゴールするということが出来てまして、これ以来、連続 5 年ベストヒューマノイドに選ばれています。

それから 3 年前に初めて人間のチームとロボットのチームを戦わせました。ロボットチームは中型の優勝チームで、人間側は国際委員会の理事メンバーです。ロボットチームは人間と対戦するということを想定してなかったもので、単純に相手チームメンバーを障害物とみなしています。ロボットは自動プログラムで人間は全然介してません。手がないのでスローインが出来ないので代わりにキックインするんですが、誰が蹴るかもロボットチームが勝手に決めて動いています。2-0 でロボットチームが負けたんですが、いろんな研究課題が出てきました。ロボットチームは、それなりに動いて人間チームを苦しめました。

まだまだヒューマノイドは進んでいないのですが、大学やその他のチームも参加し、年々技術が上がっています。ただ、いわゆる人工知能とロボットの融合だと言っているんですが、まだまだ人間側が明示的に行動を指定するプログラムが多いです。

私自身は最初は強化学習を使っているような研究をやっていたんですが、もう少し人間の方を知りたいと思ひまして、そちらの方をやっています。「心の理論」は霊長類学者のプレマックとウッドラフが 1978 年に言いはじめまして、チンパンジーの生活を見ながらあたかもチンパンジーが心を持っているかのごとく行動を示すということで心の理論が始まりました。「心の理論」とは自分や他人の心の状態を推測できる能力であるということです。これが人間の心理学の対象を広げているんですが、最近では自閉症や障害心理学でよく使われています。

それで心はどこにあるかということ、脳にあると多くの方が思われています。感覚や運動の末梢系に関しては脳部位の局在性が比較的良好に研究され理解が進んでいるけれど、認知などの高次系は未だよく分かっていない。最近のイメージング技術により関連する部位が見えつつあるが、それがどのようにして獲得されたかに関してはまだまだ多くの謎が残されている。

それで我々は実は赤ちゃんが発達の手がかりになりそうだと考えています。この間、赤ちゃん学会の副理事長に推されまして、なんであんなに赤ちゃんのって良く言われるんですが、私の答えは簡単にロボットを賢くしたい、それじゃあ人間はどうやって賢くなったのかということで赤ちゃんというわけです。

赤ちゃん学会は異種格闘技戦という感じで、医学、心理学、脳科学、工学などの異分野が交わり、非常に面白いです。以前は赤ちゃんは静的で受動的、純粋無垢で生まれてくるなんて言われていたんですが、そんなことはなくて胎児の時期から、学習を始めている非常に動的かつ能動的なエージェントとの見方になってきています。発達をどういうふうにつかえるかということですが、いろんなアスペクトがあって、もちろん環境との相互作用がありますし、先程も述べましたが受動的ではなく能動的に探索を行っているらしいということがいろんな事実から分かっています。それから神経修飾や価値、神経可塑性などで脳の中での学習の素地があり、ステップバイステップで進んでいる。

特に我々が注目しているのが拘束の集合で、これは俗に身体性という言葉として使っています。身体性という言葉について議論し始めるときりがありませんが、ここでは物理的な要素の組み合わせでできる物理的な拘束としています。もう一つが社会的相互作用で、いわゆる赤ちゃんとお母さんとの社会的相互作用が重要だと考えています。これも社会性をどう定義するかいろいろ議論があるんですが、ここではお母さんと赤ちゃんの相互作用の関係として捉えています。

(スライドを示し) これは神経科学の教科書から取ってきたんですが、日本ではこういった教科書を見かけないんですよ。アメリカに行くとおよそ 5 万人の神経科学者がいるらしく、教育にちゃんと燃える人がいます。この本なんですが凄く良い本で、全編イラストがついていて、凄く分かりやすい。左のイラストは受精から始まって 18, 20, 22, 24 日の様子でたった 2 日おきに平たい胚から脳神経系を作っていく様子を表しています。これは奇跡としか言い様がない。この時期は遺

伝子的な設計図があって脳神経系を作っていくんだらうと考えられていて、25、6週になると大人の脳と同じような構造が出来つつあると言われています。ただしコネクションに関しては完全ではないと言われています。我々は生まれたあとのことを考えがちなのですが、実は胎児の時点で既にいろんなことが起きている。7週から26週の間色々な動きをしている、例えばあくびとか吸いつき、これは12、3週から始まっていてかなり早い時点で運動が始まっているわけです。それから感覚に関しましては、触覚が7週あたりから、視覚に関しても20週あたりから光に反応するらしいと言われています。そうすると既に体内の中でいろんなことが起きている（胎児の動きの映像を映しながら）。これは14週の胎児の動きなんですけど、お母さんのお腹の中であぐましている。これは探索していると考えていただいて構わないと思います。これは26週ぐらいですが特に注目してもらいたいのは顔を触っています、これは手を探索の手段として顔をたくさん触っています。それから36週で生まれる直前になるともう目が開きつつあります。マウスオープニングリフレクションというんですが、手を持ってくると口が開きます。

生まれてから1年の間にもいろんな行動が出てきます。自分の手をじっと見るとか、抱いた人の顔をいじるとか、いろんなものを落とす等のいろいろな行動が出てきます。我々ロボティクスの言葉で言い換えると、自分の手を見るのは手の順・逆モデルを学習している。つまり関節角度を指定して手がどういう最終姿勢にあるか、もしくはその最終姿勢に持ってくるには、どういう関節角度を指定しなければならないとか、いわゆるフォワードキネマティクスだとかインバースキネマティクス。それから抱いた人の顔をいじるのは顔の視触覚情報の統合だとか、特に大事なものはいろんなものを落とすことで3次元物体認識の学習、因果性・永続性、物体の動力学モデルの学習をすることです。

正直に言いますと、我々ロボット学者はこれだけのことを1年の間に自律的にやってくれるロボットを作ること出来ません、ギブアップです。逆にいうと赤ちゃんは自律的に、もちろんお母さんとの相互作用の中でこういったことを学習していくわけです。

「じゃあどうやってこういったことが可能になるか」というミステリーはいろんな学問に関わるのです。生物学や認知心理学という自然科学の分野と我々人工物を作っていく工学との相互作用の過程の中で、人間と似たようなもの、本質的に等価であるバーチャルなもの

をやって作っていくかで悩むわけで、それによって人のあらたな理解が可能になるのではと考えており、それが新しいサイエンスに繋がらないかと期待しています。特にその中で認知発達ロボティクスと謳っているのは、従来のロボティクスでは設計者がいろいろ考えてロボットを動かしているのだけど、そうではなく、なるべくロボットに考えてもらおう、そのために何を埋め込むべきかちゃんと考えようというアプローチです。

構成的アプローチと呼んでいるんですが、作ることによる理解ということでATRの川人氏は「作るほどに理解するんだ」と言われていますが、我々は「作ってみたいと分からない」と言っています。実際に作ってみたいと分からないことが出てきて、よくよく考えれば当たり前なんだけど、作ってみたいと分からないことがあります。従来の科学的手法では、神の視点でアナリシス主体で進めており、従来の工学的手法も境界条件を明示して評価関数を最大か最小にして解決してるんですが、人間を始めとする生物の環境では問題がオープンですから、評価関数が動的に変化する。だから神の視点から状況に埋め込まれた主体の視点へ持っていくことで突破口が開けるのではないかと、その基本問題としてコミュニケーションをどう科学するかということが挙がってくると思いますが、これにアプローチしています。

いろんな課題を扱っているんですが、基本的には我々人というものの存在をどう形で捉えるかということで問題を他の分野の方々とシェア出来るのではと思っています。具体的には設計論の観点から、ロボットにどのような能力を埋め込むかという内部構造設計の問題、それから環境、特に他者とのどのような相互作用がコミュニケーション能力の創発を導くかという環境設計の問題、さらに、どのような時間的発展の構造を持つかという問題に答えなければならない。特にその場合には生物のアナロジーの相同性を議論するためには脳科学、神経科学、それから行動学レベルの検証が必要です。

関連する分野は、計算脳科学と発達心理学で、前者では神経科学を基盤とした計算モデルの構築があり、後者は観察を中心にした行動の検証があるわけです。ところがこの二つの表現のレベルが完全にずれているわけです。我々は、それをロボットを使って埋めるだけではなく、新しい科学を作っていきたいと思っているわけです。

「やわらかな遺伝子」という本は、nature v.s. nurture という二元論ではなくて、そういったものが複雑に絡み合っていて我々人間が出来ていると語っているんですが、この本の話と我々是一緒のところもあるんですが、設計の

観点から言うと、便宜上分けています。

我々のプロジェクトは四つに分かれていて、「身体的共創知能」が人工筋肉を使ってダイナミックな動きを生成をしています。従来のロボットは電動モータを使っていて、制御的には安定するんですが、生物とはかなり違うわけです。アシモは走っていますけど、あれはかなり無理して走ってます。現場に行くとギアの軋む大きい音がするらしいです。電動モータとは違う人工筋を使って人間の身体を再現して、認知がどういふふうに進化するかということをやっています。

「対人的共創知能」では対人のシミュレーション・開発を行っています。

「社会的共創知能」はアンドロイドを用いたコミュニケーション実験、「共創知能機構」は認知神経科学ということで、イメージング研究やウィリアムズ症候群と自閉症の症状を踏まえたモデリングをやっています。

心理学では発達心理が関係して、脳科学では計算脳科学、ロボティクスでは認知発達ロボティクスと言っているんですが、これらを合わせて共創知能と呼んでいます。それで、当初は個体ベースの認知発達と個体間の相互作用を分けてやってきました。前者は脳科学・神経科学が主で、後者が行動観察になっているんですが、本来シームレスなんですけど、表象が全然違うのでそこをどう埋めるか、埋めるだけじゃなくどうまくやっていくかということを考えています。方法論としては計算モデルに基づくコンピュータシミュレーションや人間やロボットを使った実エージェントの実験、片や人間を知る新たな手段ということでロボットを使って今までできなかった実験などを行っています。

図 1 はちょっと古いんですが、プロジェクトの全体像です。個体発生は系統発生を繰り返すということで、脳

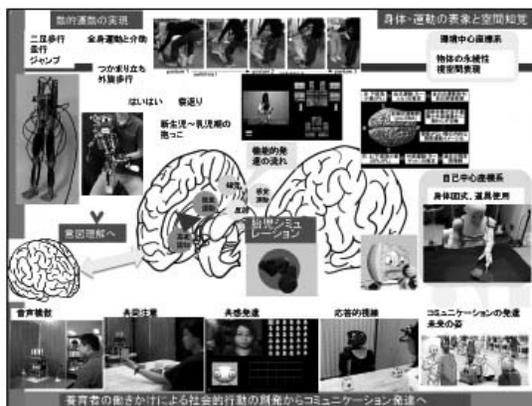


図 1 プロジェクトの全体像

の進化の流れを個体の脳の発達となぞらえて、脳の後ろから前へ発達すると考え、それぞれの段階のモデルを作りました。(身体性共創知能グループの空気圧人工筋を使ったロボットの動画を見ながら) 人間の筋肉はいくつかの種類があるんです。いわゆる通常の拮抗する筋肉では単に力を出すだけなんですけど、バイセプターと呼ばれる 2 関節筋などの関節をまたぐ筋肉があって、この種の筋肉は力を出すだけではなくて、他の筋肉との調整をやるんです。最初、ジャンピングロボットはバイセプターを使わずにやろうとしたんですが、全然安定せず、バイセプターをつけたらジャンプが安定しました。

これは生理学的には定性的に言われていたんですが、我々のグループでは、構成的に示したことになると思います。最初は、細田グループは非常に単純な拮抗する筋肉だけをつけて歩かせていたんですが、とても人工的に見えました。ところが、大腿筋やハムストリングといった本来の筋肉をつけていったら、むちゃくちゃ艶めかしくなりました。やっぱり生体の持つ構造を学ぶことで、いろいろなことが出来るようになると感じました。

最近是人肌に近づけるようにもしてるんですけど、シリコンの選択によって艶めかささも変わってきます。

ところで分かりやすいチャレンジとして、新生児模倣を考えます。これは Meltzoff and Moore が 1977 年にサイエンスに出した論文の中で、生後 2 時間以内で舌を出したら舌を出す、口を開けたら開けるということで新生児は生まれながらにして模倣能力を持っていると主張しています。これは非常にコントラバーシャルで、だいたい認知科学者はみんなこれを試しています。僕の個人的統計では、40% 真似してくれる。60% は偶然であるとか、赤ちゃんが真似する前に大人が真似してるだとか、赤ちゃんが生まれ持って模倣能力を持っていると主張することは簡単なんですけど、それじゃあどうやってそれが出来るのかということを考えないといけない。恵まれたものと言ったらそこで思考停止。なるべくそうじゃなくて学習する可能性はないか、学習できる可能性があるとしたらその根拠は何かということを探りたいわけです。

それで生得しかないと言われたら悔しいので何か学習できる場所はないかといういろいろ実験してみました。二つほど示します。胎児は顔を触るということをしてしていますから、探索運動で手を動かせるわけです。体性感覚が未発達なので顔の触覚分布が分からない、つまり顔の触覚はあるんだけどそれがどういう配置か知らないという仮説をおきます。視覚は未発達なんですけど、光のたまかな方向は分かるかと仮定する。これは

実験がありまして、光を当てると嫌がって顔を背けます。眼球運動は出来ないんだけど首をふるというのをします、なのである程度の光の方向はわかるんじゃないかなろうかという仮説をおきます。

こういった仮説を使ってシミュレーションをしました。

顔の上に21×21の触覚センサーを配置して、手で探索して触覚の分布のパターンを調べてみようとしたわけです。この21×21のパターンを最初ランダムに置いて触って自己組織化していきます。アイデアは非常に簡単で、自分の顔を触ったときに例えば左下を触ってから右上に行ったとします。その時にちょっとだけ目を開けて視覚的イメージでこっち行ったというのが分かったとすると、それを自己組織化し、さらに触覚特徴によって右目、左目、鼻、口などの大まかなパターンが出てくるわけです。ということは胎児は自分の顔を直接観測できないんだけど、触覚パターンで顔のイメージが分かるのではなからうか。

最近、新生児の触覚野と視覚野の間にある種の結合の存在を仮定し、胎児が母体内で自分の顔を触る時の指と顔のダブルタッチのみによって、顔の視覚的イメージを形成するモデルを共創知能グループから提案しています。そうすると実は胎児は見えないんですけど、自分の顔を学習できる可能性があるということが一つです。

もう一つは、対人的共創知能グループの胎児脳の発達シミュレーションです。非常に簡単な左右脳モデルを構築し、全身198個の筋肉、骨格、筋紡錘、腱、脊椎回路、延髄回路と皮質モデルからなる筋骨格・神経系を含む非常に単純な赤ちゃんモデルを作りました。基本的には非常に簡単な回路でカオティックなパターンが出ることを事前に調べておいて、今度はそれを実際の赤ちゃんのデータに基づくモデルに適用しました。19個の円筒と球で身体を近似していて、198の筋肉それぞれに神経振動子を与え、それらを相互作用させてやりましょうということです。神経振動子の間で相互作用するのではなく、神経振動子が筋肉、筋骨格、それから胎内環境を通じて相互作用します。それらの神経振動子の相互パターンによって行動が出てくる。

これによって出てきたことが二つあります。まず一つは体性感覚マップ、つまり最初は自分の身体の分布が分からないんだけど、ヘブ学習と自己組織化をやっていくことで、例えば最初は足にすべてマップされていたものが、体性感覚として腕、首ということが分かってくる。胎内の流体の中で動いてるってことでやっていったときに、例えば脚を伸ばしたときに頸が曲がるだとか、そう

いった環境との相互作用を含めて体性感覚野と運動感覚野のマップを学習すると、先程のマップが出てくるということです。もう一つはそのマップが出来ることで行動系が変わってくるということです。

それで、これを外に出してみました。胎内環境では重力がない状況でしたが、外に出してくると重力の影響を受けますから、生誕直後と2ヵ月後で動きに変化が見られます。早産児・未熟児という24週くらいで外に出た赤ちゃんはICUでケアされて普通の子供と同じように健康に成長するんですが、完全には証明されていないんですが、小学生から中学生くらいで幾何学的推論などいろんな能力が劣っていると言われてます。そうすると24週くらいで胎内環境でないところから早く重力環境下に出されることで、どういう変化が起きて最終的に、発達障害があるとしてどんな影響があるか研究し始めています。

ヒトという種の特徴として二足歩行、道具使用、言語使用の三つがあります。

ロボット研究として二足歩行、うんそうだね。道具使用、これはロボットの研究だね。言語使用についてはみんなロボットじゃないと言うんですが、我々はロボットの研究だと思ってます。特に言語がどんなふうに発達するのかいろいろアプローチしているんですが、我々はハードウェアということで身体性を考慮しなくちゃいけないので肉声ってことでやってきています。

一番最初にアタックしたことはなにかというと、ちゃんと「あいうえお」が言えるロボットを作ろうということです。その時に赤ちゃんは聞くことと声に出すこととの関係について知識がないというふうに想定しています。胎児はお母さんの声を聞いているので先に聴覚テンプレートができていると言われており、生誕直後はお母さんの声には反応するけど、悲しいかなお父さんの声には反応しません。そして未成熟なのでお母さんの声をそのまま真似できない。設計論的にはどういうハードウェアを用意すれば良いか、母親はどうすればいいのか、という二つのクエスチョンに対して仮説を立ててみました。

実際のお母さんと赤ちゃんを見ていると、互いに真似し合いあっているということがあります。お母さんは赤ちゃんの真似をすることで赤ちゃんが喜ぶ、赤ちゃんはお母さんの真似をすることでお母さんが喜ぶので赤ちゃんも喜ぶわけです。赤ちゃんの発声機構は、声道理論で近似できるということで、シリコンの筒を変形させることで音を出すんですが、それと学生さんの中で環境を再構成しています。もう少し言うと、このシリコンの筒を

5 個のモータを使って変形させると、音源に対して出てくる音が変わります。実際の人工声帯と同じ手法です。発声にはこれを使います。耳の方は何を使うかというフォルマントを使いました。フォルマントとは音声スペクトルのピーク周波数の低い方から第 1 フォルマント、第 2 フォルマントとしています。この第 1 フォルマントと第 2 フォルマントは母音の識別に有効と言われていて、霊長類や鳥も利用しているらしい。

これは日本人女性のフォルマントを分析したものののですが、汚い塊がなにかというと、これがロボットのフォルマントなのです。音域的には全然違うわけです。それじゃあどうやってコミュニケーションをとるかという、我々は最初の仮説として養育者のオウム返しを立てました。クーイングと呼ばれる 4 ヶ月から 5 ヶ月くらいに赤ちゃんはランダムに発声します。その時に養育者はそれを聞いて「あ」と聞こえたら「あ」と返してやる、「い」と聞こえたら「い」と返してやりませう。そうするとロボットの方は何をするかという脳の方は聴覚層と構音層から出来ていて、例えばランダムに発声させて、たまたまお母さんが「あ」と聞こえたと思って「あ」と返したとすると、フォルマントが出てくるのでフォルマントのベクトルと構音のベクトルをへブ学習させます。

これは皆さんご存知のミラーニューロンシステムで、お猿さんが何か物を持ち上げたりしたときに発火するニューロンが他者の同じ行動を観察しても発火するというものです。それが何を意味するかというと、自分の行為の実行と相手の行為の視覚認識が結びれているということ、不可分であるということ。我々はその意味で聴覚層でフォルマントのベクトルを認識するという、それから構音ベクトルで発声するということを対応学習させることで、知覚と行動が不可分であるということ、密につながっていることを示しました。

つまりなぜ我々が「あ」ということを「あ」と理解できるかという、自分が「あ」と言えるからであるということ。つまり理解できる本質がなにかということ、それを再現する能力があるからこそ出来るんだってこと。例えば踊りのお師匠さんが舞い散る雪を取るよと言ったところで素人は分からない、ある程度スキルを獲得していくとその意味が分かってくる。だから理解の本質はまさに自らが身体をもって再現できる能力を持つことによって相手を理解できる。子供は大人の会話を聞いて全部を理解出来ないけれど、子供の持っている語彙をもって大人の会話を構造化する。だから

正確さは欠けるかもしれないけど、そんなに間違っていない、それが理解の本質である。というわけであろうといった形で学習をやらせたわけです。

(発声している映像を流して) どうですか、「あいうえお」って聞こえましたか。(別の動画を流して) これは「いやいや」ではなくて「いあいあ」と言っています。音だけではなくてちゃんと唇のかたちも見てみよう、視覚情報も使って学習させると 3 倍速くなりました。

このとき人間側に二つのバイアス効果があります。人間がインタラクションしているときにどんな反応をしているんだと聞かれたので、どういう現象が起きているか二つの仮説を立てました。

一つは聴いた音声を自身の母音音声と知覚しようとする知覚マグネット効果、これはどういったことかと言いますと、例えば 6 ヶ月までの赤ちゃんは「あ」から「い」までの連続母音をそのように識別します。6 ヶ月を過ぎると母語に影響されて、いくら「あ」から「い」の連続音を聞かせても「あ」か「い」にしか聞こえないわけです、これが知覚マグネット効果です。我々はこのに加えて構音の方にもバイアスがあるだろうということで知覚運動マグネット効果と呼んでるんですが、なかなか証明することが出来ない。もう一つが実は大事で、相互模倣により相手の発声音は自身の発声音に近いと期待するオートミラリングバイアス。

この二つのバイアスがあることによって相互模倣が起きているんじゃないだろうか。

図 2 は先程述べました知覚マグネット効果です。入力に対して知覚の方は図のような特性を示します。

6 ヶ月までは直線的なんです、6 ヶ月を過ぎると母語にチューニングされて「あ」か「い」にしか聞こえない、これが知覚マグネット効果です。

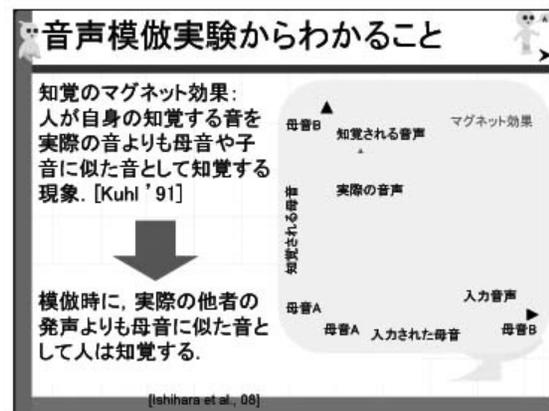


図 2 音声知覚のマグネット効果

実際に実験やってみたんです。うちの学生さん5人くらいに、コンピュータの方で「あいうえお」のフォルマットの周りに適当に分散させた音を出して、母音ではなく音を真似して下さいと言いました。そうしたら何が起きたかという、音を真似するから同じようにばらまかれるかなって思ったんですが、なんと自分の母音に引き込んでいました。つまり大人は聞いた音に対して自分の母音のバイアスがかかって、無意識のうちに自分の母音で応答しようとしています。もっと極端に言うといくら散らばっていても、そこに近い母音だと無意識に知覚しているらしい、これが知覚マグネット効果です。

それでもう一つがオートミラーリングです。

これに関しても、心理実験を行いました。コンピュータが音を出します。その時に一つ目のグループは統制群・期待なしで「音声の真似をして下さい。音声はランダムに合成されます。」と言います。もう一つのグループは、実験群・期待ありで、「音声の真似をして下さい。相手も真似してくることがあります。」と言います。

この二つのグループで、基本的にロボットは同じ音を出します。この実験をやってみて、被験者の発声の変化の幅を調べました。

模倣されることを期待していた群（実験群）ほど、被験者自身の直前の発声に似るように相手が模倣したと錯覚するので、発声音の変化が少なくなります。期待なしのグループは真似るともなんとも言われていないので、聞いた音をそのまま返すだけなんです。だから連続音の発声の差が、実験群よりも大きいわけです。統計的な差が出ています。

オートミラーリングバイアスがあることがわかりました。それでこれを使ってモデル化すると、ちゃんとマグネットバイアスとミラーリングバイアスがあると相互模倣して適したところに行くわけです。ところがマグネット効果だけだと収束はするんだけど、正しいところには行かない、オートミラーリングバイアスだけだと方向は正しいところには行くんだけど収束はしない。両方ないとばらばらだけど、バイアスがあるとある程度収束することがわかってます。

図3は機械とのインタラクションということで、いわゆる Vo-Cal, Vivid Oral Communication through Acquiring Language というわざとらしい名前なんです。 (交代数え上げインタラクションの動画を観ながら) これ正しくオートミラーリングなんです、皆さん次に何が来るか予測してるから聞こえるわけで、単音で聞いたら何言ってるか分からないんです。これは構音そのものの構



図3 より円滑なコミュニケーションを目指した Vo-Cal(グッドデザイン賞受賞) *口絵にカラー版掲載

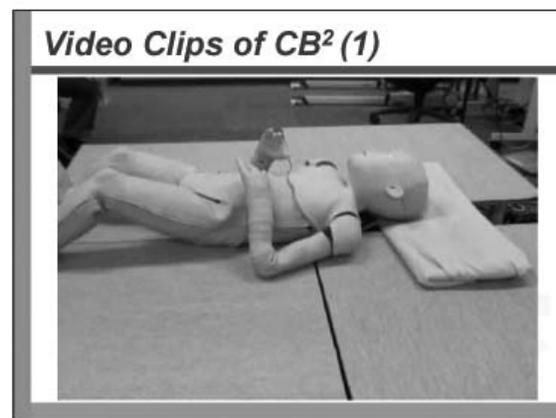


図4 CB2(Child robot with Biomimetic Body)の行動の様子

造を見せるとどういった影響があるのか、機械特有のインタラクションがどうあるべきかが研究対象です。顔の話もあるんですがまだ作ってなくてですね、一応グッドデザイン賞は頂きました。

図4は我々が開発したCB2というチャイルドロボットで、身長1.3メートル、体重33キロ、視覚と聴覚、あとはシリコンに覆われていて全身に198個の触覚センサーを持っているのでどこを触られても分かります。局部だけよじれるのでセンサーは付いていません。簡単な「あー、あー」位の発声が出ています、それで音を聞かせると反応する、いわゆる反射的な行動を埋め込んだロボットです。光が好きだという想定をさせておいて、光の方向を与えて光の方向を向きます。光を横のところを持ってくると寝返りをうちます、この映像を小児科医に見せたらこれは自閉症の子供の寝返りの仕方だと言われました。健康児は片足ずつ寝返りをうつんだけど、何故か両脚一緒になっちゃいました。それで最初にやった

実験はですね、実はこのロボット自分では立てないんですよ。なので支援するんですが、成功失敗が出てきます。

自由度が 56 個あってそのうち 52 個が空気圧で、このとき 52 個を制御することは制御的に難しく空気圧では位置制御が非常に難しい。これは非常に難しい実験に見えるんですが、簡単に出来てしまう。実際の赤ちゃんとお母さんにならって初期位置と最終位置と中間位置を決めてスイッチングと PID 制御を使ってやっています。けどほとんど位置制御できないわけです。それで人間が助けるわけですが、僕自身は 2 回目で成功したんですがちゃんと 1 回目で成功する人もいます。成功するときにどんな形になっているのか、ロボットの動きと人間の動きの時空間相関を取ります。取ってみると成功失敗の違いがはっきり分かるわけです。

例えば成功例だとすると、最終的に人間とロボットの動きが同期してくるんですけど、その同期具合が最初のずれからだんだん合ってくる。高い相関の領域がつながってるんですが、失敗するときにはこれが離れている、つまり時空間相関が取れないわけです。物理的には 52 個のパラメータが有るんですが実は 1 個のグローバルパラメータで識別することが出来ます。

もう一つは学習する話なんですけど、人間の主観的判断（よしあし）のデータに基づいて、ロボットが行動改善学習するという方法です。それによってロボットのパフォーマンスをあげたという話です。一つ象徴的なんですけど 52 個のパラメータがあったんですが、主成分分析したら 2 個に落ちました。

それで時間がなくなってしまってロボシティコアの話が出来なくなってしまったんですが、ロボシティコアではですね、こういったロボットのテクノロジーを一般の人にどういうふうに観てもらおうかということで研究と教育、産業化を同時に進めているわけです。JR 大阪駅の北のほうに貨物操作場跡があって、その 24 ヘクタールの土地に新しい街を作ろうというので、第 1 期 7 ヘクタールの土地で、A ブロック、B ブロック、C ブロックがあるんですが、A ブロックは商店街、B ブロックにロボシティコアというロボットの実証実験をする街を作るようにしています。既に着工されていて、大阪駅もかなり新しくなっていて、ここにロボットの街を作って、一言でいうとマイノリティ・リポートのような、環境が人々に対してちゃんと作用するような環境のサイドとロボットのサイドの両方含めた形で街を作っていこうと提案しています。

ご清聴ありがとうございました。

河原：どうもありがとうございました、多少時間をオーバーしてるんですが折角の機会ですのでフロアの方から 1 件だけでもお受けしたいんですが。

質問者：浅田さんたちのやっていることを聞いて人間に近いロボットを作ろうとしているように思えたんですが、それでいいのでしょうか。

浅田：初代会長の館先生のお言葉ならば、生命と宇宙と同じでどんなに近づけても近づき得ないと思いますが、だからこそ近づきたいということです。近づき得ないからこそ差が生まれてくる、当然です。その差に本質がある、つまりその差があつてなお共有できるものは何かってことを見定めるって言うことが、ここで言っている本質です。効果としては現実であり、実物であることにロボットとしての意味があるのではないかと考えています。

その時に何が本質かという議論をしなきゃいけない、それは我々の間でも喧々諤々の議論をしていて、うちのグループリーダーはみんなえらいので僕の言うことは聞かないんですが、いつも言っているのは例えば発声ということを考えてときに、何故スピーカーを使わずにシリコンチューブを使うのか。我々は身体というものにケアしたくて、身体があるからこそいろいろな制約、例えば実は今日お話ししてない中にへブ学習をして「あいうえお」のそれぞれのクラスタを作るんですが、人間ってわがままなんで「あ」にもたくさんあるんですね。

そうするとロボット側のいろんな「あ」が出来てしまう。それがカテゴリとしてまとまらないんですね。それで、へブ学習に疲労度合いを減らす制約を加えると、ちょっと絞られるわけです。そうするとカテゴリとして簡単になってくるんです。そうすると、その時には身体を持っているという物理的な拘束が効いてくるだろうと思ってます。

その時の本質は何かって言われると生体、発声する物理的実体そのものがエネルギーを持ってやっているし、それが疲労してくるしということが入ってくると思います。実際はまだ肺を作っていないんですね。ロボットに呼吸する意味があるんかかって話があるんですが、ターンテキング、単語の切れ目であるとか会話の切れ目を自然にだそうとすると、やっぱりロボットは肺を持たないと、肺を持って発声しないといけない。今のところ空気を流しっぱなしにすると 10 分ぐらい「あー」って言い続けられるんですけど、人間じゃそんなことあり得ないですよ

ね。そうすると呼吸系を作るってナンセンスやないかそんなもん意味有るんかって言われますけども、そこにちゃんとこだわりたいと思います。

その時には呼吸によって単語セグメンテーション、会話セグメンテーション、ターンテーキングをどうつくっていくか。だから本質は何かは、所々で微妙に違うんですが、今はそこを探索してる状態で、それをまとめるときに何が残るかって言われるとあんまり自信がないんですね。でいろんな価値観、多様性があると思っていて、自分で自分の首を絞めるんですけど、人間としての本来の意味が人間以外の種と共有する部分と、異なる部分をどのように捉えるかという課題。これも社会性の問題も含めて考えなきゃいけないところがあってですね、連続性と非連続性、これは生物でよく言われると思うんですが、その課題には立ち向かっていかないといけないだろうと思っています。

質問者：さっきの川崎先生に対する質問に対する川崎先生の答えにも関係するんですけど、人間にはいい部分もあるけれど悪い部分もある。ロボットにいいところを取り込むのはもちろん、害があるのは問題ですが、人間らしい悪さがあると思うんですが、こういったものも取り込もうと考えているのかお聞かせください。

浅田：研究者としては当然悪いところも含めないと出てこない。ロボカップでフェイントをかけるロボットをつくるって話をして、それもトリッキーなんですけど、それが他者を騙すってことになっているかは非常に疑問なんですけど、ただもとはそういうことがあるんじゃないかな、例えば根源的にはあるんじゃないかなと思っています。

ただ社会性に至るほど賢くなってないんで、2歳の赤ちゃんでもものすごいことをいっぱいやるんですよ。それに手一杯になっていて、そこから言葉、言語、社会性

ということにどのようにつながっていくか、そのところで人らしさ、その人らしさに、いいも悪いも含めて人らしさですけど、そうじゃないといいことだけ含めてってことは僕はあり得ないと思っていて、けどそこは、人間の知恵で、メタな意味でそれをどう構造化しておさえるかが課題です。弱気ですが、おさえられない部分もあると思う。

当然ロボットによる犯罪がそのうち出てくると思うんですけど、それは単にロボットをコントロールしてる人が悪いんだってことになるのかどうかも含めてですね。色々問題があると。

話しだすときりがないので、以上で。ありがとうございます。

【略歴】

浅田 稔 (ASADA Minoru)

大阪大学大学院 工学研究科 教授

1982年大阪大学大学院基礎工学研究科後期課程修了。1995年大阪大学工学部教授、1997年から現職。知能ロボットの研究に従事。情報処理学会研究賞(1989)、IEEE/RSJ IROS'92 Best Paper Award(1992)、日本ロボット学会論文賞(1996, 2009)、人工知能学会研究奨励賞(1997)、日本機械学会ロボティクス・メカトロニクス部門貢献賞(1999)、文部科学大臣賞・科学技術普及啓発功績者賞(2001)、日本機械学会ロボティクス・メカトロニクス部門賞：学術業績賞(2001)、人工知能学会研究会優秀賞(2004)、科学技術政策研究所 科学技術への顕著な貢献 in 2006(ナイス ステップな研究者)「イノベーション部門」(2006)、(財)大川情報通信基金大川出版賞(2007)、2008年度グッドデザイン賞、博士(工学)、日本機械学会(フェロー)、日本ロボット学会(元理事)、IEEE(Fellow)等の会員。