パネル討論

ロボットにおける心とヒューマンインターフェース†

The Mind and Human Interface of Robot

菅野重樹*¹ 原文雄*² 佐藤知正*⁴ 大熊道雄*⁴Sigeki Sugano*¹, Fumio Hara*², Tomomasa Sato*³, Michio Okuma*⁴佐野明人*⁵ 中谷和夫*⁶ 広瀬茂男*⁷ 加藤一郎*⁸ 吉田和哉*⁹Akihito Sano*⁵, Kazuo Nakaya*⁶, Sigeo Hirose*⁷, Ichiro Kato*⁸ and Kazuya Yoshida*⁹

司会(菅野重樹:早稲田大学): ロボット学会では1992年の7月より、早稲田大学の加藤一郎先生を委員長にお願いし、ロボットにおける情動のメカニズム研究専門委員会を発足させて、ヒューマン・インタフェース、あるいはロボットの知的制御、などの分野に興味をお持ちの先生方、また、心理学、認知心理学といった分野からロボットの制御にご興味ある先生方にご参加いただき、ロボットにおける心、そしてヒューマンインタフェースといったテーマでディスカッションを続けております。

この特別セッションは上述の委員会の第6回研究会としても位置づけられます。今日は3人の先生方にパネラーをお願いしています。まず原先生お願い致します。

1. ロボットの感情と性格

原文雄(東京理科大学): 人間同士のことも考えると、機械と人間との間のコミュニケーションは、単に命令だけの情報のやりとりだけでなく、いわゆる心理的な情報のやりとりも必要ではないかと考えています。人間同士ですと相手がどんな気持ちになっているかをわりと察します。それからじゃあどうしてあげようか、ということも分かる。

従来の機械は、必ずしもこういう機能を持っていませんので人間が合わせていた。そうすると、ワープロでも時々皆さん、かんしゃくを起こしてキーを叩く。

そこで、機械の方が機械のパートナーになっている人間の気持ちとか感情とか、最終的には心と思うんですが、そういうものを理解する必要があるんじゃないか。私のところでやっている顔の表情認識、表情と心とは必ずしも一致しないかもしれませんが、少なくとも外に出てくる情報か

ら、目とか声とか、表情とか、身振りとか、当然言葉の調子も入ると思いますが、そういうものを認識する。そうしたときに、次はどんなふうなレスポンスをそのハードが伝えるか、そういうものを徹底するような一つのデジジョンメイキングのプロセスが必要なのか、そういう決めたものをどういう形で人に表現するかという、この3つができると、もう少し僕らも楽しめるような機械になるんじゃないか。

ここではその2番目の特に感情を人工的に作る方法論にどんなものがあるかをご紹介しますと思います。

心理学の知見をベースにしながら、それを工学的に表現するヒドラ・ネットの一つにハーモニー・セオリーというものがあります。それとジネティカルボリズムを使うとそれらしいものができる。

ハーモニー・セオリーというのは、結論的には2層のネットワークになっていて、例えば下層がその入力の実現特徴みたいな話で、それをいくつか結びつけると一つ意味が出てくる、そういうネットワーク。これをいくつもつなげて、一つのネットワークという全体のつながりの具合、調和度を定義します。いちばん調和がとれているようなつながりで、入ってきた入力が必要なことを意味しているか解釈をするようなネットワークです。

堀哲郎先生が『脳と環境』という本を共立出版から出しておられますが、そこに感情の機能論的なモデルが書いてある。自己保存の原理が感情の評価に非常に役立っている、そういう見方もあるかもしれません。

根本においては、生命体は自分が生き延びるために外からの刺激がプラスかマイナスか、そういう評価をして、それに対するリアクションを出すためのエモーションが出て、それから身体コールが出てくる。

最終的には、人間の顔を考えてますから、目、口、眉、その他いろいろあると思いますけど、私のところでは、基本的には三つで顔の表情で表そうとしていますので、それにどういうふうに結びつくか。

*本文は紙面の都合で約1/3に圧縮されている。全文をご希望の方は事務局へお問い合わせ下さい。

原稿受付 1994年1月19日

*¹早稲田大学理工学部(司会)

*²東京理科大学工学部(パネラー)

*³東京大学先端科学技術研究センター(パネラー)

*⁴横浜国立大学工学部(パネラー)

*⁵電気通信大学機械制御工学科(討論者)

*⁶東京大学文学部(討論者)

*⁷東京工業大学工学部(討論者)

*⁸早稲田大学理工学部(討論者)

*⁹東京工業大学工学部(討論者)

†1993年11月12日開催の第11回学術講演会より

ネットワークの結びつき方は、基本心理の顔の表情とか感情の問題を研究して、どれとどれが結びつくかというのは経験的にも結びつけられると思いますが、ここではそうではなく、人工的にもっともらしいネットワークの構造を決めていくというやり方をします。

外からいくつかの刺激を与えて、ネットワークのつながりを何種類かランダムに選んで、その一つ一つについていわゆるハーモニー・セオリーの考え方でそのネットワークのつながり方自身に対して、外から入った刺激がどういう感情の意味を持っているか、感情になっているかというのを出してそれが顔の表出に全部一致しているかどうかという評価をします。

あるネットワークのつながりが、例えば怒りに対応するのが出て来ても、目は笑ったようになってたり口は泣いたようになってたりしてしまいますので、GAを使ってその中でいいものを選びながら何回も繰り返す。最終的には全体が感情ネットワークのモデルみたいになります。

それから、そういうふうな調和理論とGAリターンで、ネットワークの構造化ができるか、ということをやってみました。心の理論なしに、一つのネットワークとしてある刺激、例えば相手の怒った顔を見たとき、自分が怒るという感情を出すというようなことを考える場合、こういうネットワークで自然に作ることもできます。

こういう人工の感情を出すときに、例えばロボットが笑った顔、怒った顔を見たときに、次にどういう表情を出すかは、ロボットの人格、性格を決めないと思えない。そうすると、いつも優しい顔をするのか、いつも怒った顔をするのかというのは、たぶん心に関連して来て、その問題を進めていくと、やはり心がどういうものかということまで行かざるをえない。

それからもう一つは、人工の心を持つ人工システムができて——これは広い意味の機械でいいと思うんですけど——そういうものと人間と一緒に仕事をする場面を考えてますから、新しい心理学がここから発生するんじゃないか。

今までの心理学は人間個人の問題、人間と人間の間でのコミュニケーションの問題、それをやっていた。従来の機械は心を持っていませんから、ある意味で人間と機械のコミュニケーションは非常に少なかった。

そう考えてくると、これまで人間と人工のシステムとの関連を扱う心理学は必要なかったけれども、機械がますます知能化して人工心理をインプリメントされると、新しい心理学が出て来ることが期待される。そういう意味で、新しいロボットから新しい学問が出てくるんじゃないかと思っています。

司会：ロボットの性格をどうすればいいのか、私も必ずそれは問題として持ち上がると思います。続いて佐藤先生、お願い致します。

2. 意識・意志・意図と無意識

佐藤知正（東京大学）：心を脳の活動の我々の世界への表象である、ある意味でのパフォーマンスであると考え、そういった観点から、ロボットの心と人の心についてお話しします。

二つトピックスがあって、一つはロボットの心についてです。「知・情・意」ということが言われます。AIにおいて確かに「知」「情」については研究されているのですが、「意」は意外とやられていない。その「意」の情報処理について考えていることをまず話題として述べます。もう一つのトピックスは、人の心についてです。

人間の行動をロボットが観察して、その人の意図を理解して支援をする、そういう話をしたいと思います。まずロボットの心の「意」についてですが、具体的に例えば意識、意志、意図が今のロボットに欠けている。

モデルベースアプローチのある目標つまりミッションが与えられると、作業計画をロボットが作り出す。次に、その計画された作業を実施するとどうなるか予測を立てる。さらに、作業を実施してみて、それがうまくいっているかどうかを予測と照らし合わせて評価する。それがうまくいってなければ、また新たな計画をするという、そういうループを描くことがロボットがロボストに作業を実施する上で大事なことです。「意識」はまさにこの評価の機能であって、今自分が何をしているのかということをやちゃんと認識する機能のことをいいます。

産業ロボットは手を伸ばしてボルト締めをしているようにはたからは見えるけれども、そこにボルトがなくても手を伸ばします。そういった意味ではロボットが何をしているかを知る機能、つまり意識の機能がない。これに対して意識をロボットが持つようになると、自分の思うようになかなか作業がいかないなあ、とロボットがノイローゼになるんじゃないか。ロボットがノイローゼになるぐらいの研究が進歩すると、意識の研究は完成でありロボストなロボットが実現する、と思っています。

「意志」とはここでは多少のことがあっても目標を追求し、作業を遂行する働きのことをいいます。これはまさにフィードバックの機能です。今のロボットでフィードバックがうまくいっているのは軌道レベル。要するにどこの位置に腕の先を持っていかねばいけなかないかということに関しては、非常にうまくフィードバックがかかるけれども、もう少し上のシンボリックなものが入って来たところでは、ちゃんとできていないんじゃないか。そういった意味では今のロボットは「意志」の機能というか、とにかく作業を遂行しようという気迫がない(笑)。そういう機能があったらいいと思う。

「意図」、これはロボットの意図ではなく、人間の意図。

人が何をしているのかというのをロボットが読み取る機能をいいます。人間が何か作業目的を持ってやっていると、その目的をちゃんとロボットが理解してくれると、それは作業教示になる。それから、例えばうちわであおぐというような行動を見ていて、部屋の温度を下げてくれるようになると、それは人間の生理まで考えた人間支援ということになる。

こういうように人間行動の意図の読み取りがうまくできるようになると、心が通っていると人間が感じるロボットができるんじゃないか。そういうロボットがあってもいい。

以上述べたような「意」の処理の基本は行動理解ということです。ロボットが「人間は何をしているのか」と人間の行動を理解してもいいし、ロボットがロボット自身の行動を理解してもいいし、他のロボットの行動を理解してもいい。

ロボットがロボット自身の行動を理解していると、意識機能が実現されるだろう。エラーを発見し、それを修正することで、意志をもって何か作業を遂行するといった機能が実現し、これが一つの要の技術になる。意図という観点から言うと、ロボットが人間を見ていてその行動を理解するという技術。またロボットがたくさんいるときには、ロボットたちが協調するために、他のロボットが何をしているのかを理解する必要がある。そういう意味では行動理解というのは行動の認識機能ですが、わざわざ「行動理解」という言葉で言いたい理由は、いわゆる音声理解を考えるからです。音声理解は1次元の情報理解です。それから画像理解、これは2次元情報の理解ですね。それに対して行動理解というのは、そういったものをすべて含んだもっと多次元の、難しい情報処理領域じゃないかと思うわけです。

ロボットの「意」の研究が必要なことを述べましたが、これは人間がロボットを心が通う機械であると感じられるところまででいい。いいかえるとロボットが心を持つことは必要ないと感じております。しかし僕自身が老人になったとき、心が通うロボットがいて相手になってくれると、気味悪いなあという感じと、そんなロボットがいてもいいという気がします。

もう一つのトピックス、「人の心」。「意」の話にもかかわるけれども、今度は無意識も含めた話をします。人の心が読めるロボットが本当に心が読めたかどうかについてはかなり疑問がある。それにしても、心が読めなくても、意思疎通が図れるロボットぐらいは欲しい、できるんじゃないか。そのベースになる考え方は、脳の働きである心の動きは、何らかの形で世界に向かって発せられるはずである。そうすると、何か適切なセンサーを用意すれば、心の動きはある程度検出可能なんではないかということです。

そういうセンシングをちゃんとやることによって、ある

程度心が通っている機械だと人間に感じさせることぐらいはできるんじゃないか。そのときに意思疎通のチャネル、いちばん人間らしい強力なチャネルはもちろん言葉です。それから行動というチャネルも大事。これは自動車の運転を言葉でしろと言われると非常に困ることからも分かります。

人間からロボットへの意思疎通の手段として行動を利用するときに、これまでの人たちは、意識的な行動だけを利用してるんじゃないか。人間の無意識的に出る行動、動作、反応、そういうものを利用すると面白いことができるんじゃないかと考えている。

これまでのインターフェースというのは、例えば、マウスとかキーボードとかジョイスティックとか、これらは人間の意識的な行動を自動的に入力するデバイスに過ぎないから、能力の限界がある。無意識的なところまで含めて利用しようとする、ロボットが能動的に情報を取りに行かなきゃいけない。それには視覚が有力な手段を提供する。場合によっては、環境変化装置なんてのがあって、「この人本当に怒ってるかどうか」ということを、例えばちょっと部屋の色を変えてみたときに、その人がどういう反応を示すのかというので判断する。また場合によっては相手をひっぱたいてみて、能動的に状況を知るというようなこともあっていい。相手の情報を取りにいく能動性があると、人間の心の動きがとらえられるのではないか。

このような考えから、微細作業において無意識的な行動により粗動と微動の自動切り換えをするという研究をやっています。顕微鏡下で実施する微細作業では、微動と粗動の切り替えスイッチがついている。この切り替えは結構わずらわしく、作業する人を悩ませる。人間は作業を実施したいのであって、切り替えスイッチを操作するのは余分な負担となるから、それを無意識の人間の動作で切り替えられるのではないかと考えたのです。

研究の具体的内容についてはスペースが限られているので省略しますが、意識的な動作のみではなく、人間の無意識的に出る行動を理解すれば、よりスムーズな意思疎通ができるんじゃないかというわけです。

別の例でいえば補聴器がちょっと聞き取りにくかったりすると人間は、エッ？ とこうやりますね。そういうような動作をもし認識できるなら、補聴器の感度が自動的に変えられるんじゃないか。患者が病室で寝ていてのどが渇いて、ゴクンとしながら冷蔵庫の方に目をやると、「ジュースを取ってほしいのか」と聞いてくれる。

人間というのはコミュニケーションするときに、相手がどういう状況にあるのかということに、自分が持つほとんどの情報処理能力を費やしている。ロボットは何もしていない。人間がどう考えているのか、全くおかまいなくやっている。そういう意味で、今のロボットが使いやすいロボットであるはずがない。もっともっとロボットは人間の意

図1 人を見るロボットの広がり

図を理解するためにコンピューティング・パワーを割くべきではないか。

ここからは夢のお話になるんですが、こういう働きを見せるためにデモルームを作ろうじゃないかと考えています。ロボットルームと呼ぶんですが、人間が寝ている部屋でも病院でもいい、とにかく部屋にセンサーとかアクチュエータがついていて、人間の状況を一生懸命見てくれている。

例えばテレビ欄のある部分を指で指してこれを録画したいと言うと、ロボットたちが連携を取りながらこのことを理解をしてビデオをセットしてくれる。例えば「あれ取って」と言う、あるいは取ってくれというような顔をするのと取ってくれる、というような支援ができる。そういう部屋(ルーム)を作ろうというのです。人間がロボットに見られてというのは嫌だという拒否反応を示す方がきつとあるので、これは病室だと言ってるわけですけども、病人あるいは人の手助けが必要になった人間にとっては、こういう支援はありがたいんじゃないか。

こういうふうに、人を見る機能に重点を置いたロボットは非常に大事で、図1に示すようにいろいろスペクトルがある。例えば、人間の中にいて人間を見るという保健ロボット。小動物型で人間を見るというのもいい。それから今ご紹介したような部屋の中がそういう環境になっているというのもありますし、それから街灯にそういうのがくっついていて、歩行者が飛び出しそうになったときに、近づいて

くる運転者を支援する。今はやりの言葉で言うと、新社会資本になる。もしこれができ上がるとたくさんのものが必要になるので、これが非常に大きな産業になるに違いない。

結論は、ロボットの心に関しては「意」の情報処理が大事だ。それから人の心に関しては、ロボットによる人の意図の読みとりがこれから先非常に大事になるんじゃないか、と考えているということです。

司会：ロボットにノイローゼが生じるかという話から

具体的なインタフェースのお話までご意見をいただきました。それでは大熊先生にお話ししたいと思っています。

3. 心の構造と心らしさ

大熊道雄(横浜国立大学)：私がやってますのは、生物のシステム的な特性を機械システムの要素技術に適用しようということです。

心の構造から応用して、有限エマージェント・プロセスというものを開発しました。

DNAの些細な差から、エクспRESSION(発現)して、それでタンパクの分子の形状の変化が順に積み重ねられ、それを含んでいるホスト、つまり生物の特性が変わってくる。その個体が、環境というコンテキストによって選択され、適応しなければサバイブできないで、死という生物にとって非常に重大な意味が出てくる。そういう下位情報から上位の構造物に転換するものを有限エマージェント・プロセスと呼んでいます。

人間が情報領域からの情報を意味領域に解釈するときには、コンテキストが必要で、それは心から来ている。このコンテキストは、情報が入るにつれてまた変化していく。これを機械にあてはめると、情報領域から入ってきたデータを解釈して、構造物を作る。こちらは物理領域である。従来のものと、プログラムで全部手順を教えますが、これを構造物の方から場の情報として入れることによって、コンテキストとして使います。数列情報を解釈して、構造物を作るということと、構造物からコンテキストを入れてやるということとの組み合わせによって、非常に簡単な方法で多様な構造物ができる。

今度はロボットの心に関する考え方ですが、心を定義するのは難しい、ただ、定義というか、規準ということで示されているものは有る。ベイトソンは『精神と自然』の中で六つあげている。①相互に反応する部分ないし構成要素の集合、②精神の各部分で起こる相互反応を活性化するのは差異または変化である。③エネルギーの流れを切り替える決定機構がある。④循環的な決定の連鎖を必要とする。⑤メッセージを記号化し、部分を知覚して全体を推測できる。⑥論理階型の切り替えができる。例えば仮説を作るとか、遊びをするとか。そういうものがあると言っています。

困難だとはいえ、こういう研究をすることは工学的に役立ちそうである。社会的な必要性が感じられる。

硬直したマニュアル人間は、マニュアルに書いてないことはやらない。ちゃんと規定されてない問題、アンダー・スペシファイド・プロブレムといますが、これを、解こうとしないような人々の代わりに、気働きのいい心のあるロボットで置き換えたい。

心の定義が難しければ心らしさを感じさせる機能を調べてもいいんじゃないかと、やってみました。

以下の(1)(2)(3)…は、先程のベイトソンの①②③…に対応しています。まず、

(1) 相互に反応する部分、引き込みがあるんじゃないか。心はノンリニアですから、エントレインメント(entrainment)、あるいは反発が起こるんじゃないか。ロボットもその引き込みが起こるものを作ると、案外よく言うことを聞くかもしれない。しかしあんまり引き込みばかりじゃ退屈してしまうから、反発があるものもあっていいんじゃないか。

(2) 動的になったときに受け手に抵抗感を生じる(中村雄二郎)、動的機能です。だから、ハートはハードの向うにあるハートを見過してしまう。

(3) 決定機構ということで、評価主体として、美的統一感に準拠するというのが、ベイトソンとかパースが言っていることです。こういう準拠するものをどう設定するか。中村雄二郎先生は自明性とか共通感覚と言われてます。整合性だけでなく、自由裁量の余地があって、それをどうするのかということもある。

(4) 循環的決定の連鎖はただの循環ではなくて、循環しながらヘリカルに向上していくのである。密教で瞑想をやる時、「身口意(しんくい)」の一巡毎にテーマを切り替えていくのもそうでしょうし、Hofstadter が言ってる「エマージェント・フェノメナ」。これは、上位レベルと下位レベルがあって、下位レベルで整合性のあるものを取り上げて、結論を上位レベルで判断して、繰り返すというもの。この循環をストレンジ・ループと呼んでいる。我々の脳の中で起きている例として、アイディアとか希望とか、意識とか自由意思とか、こういうものを処理しているらしいと彼は主張している。

有限エマージェント・プロセスは、このストレンジ・ループを切って、部分だけを取っておくことで、有限という名前を付けている。それからオリジナリティーとか、先程もありました無意識の過程、こういうものがある。

(5) 記号化。それには推測して補完する部分がある。いろいろあってゲシュタルトと言われます。帰納的情報処理。それからバルナビリティ(vulnerability)を見通すということ。先程の生物が持っているハードウェアの機能にはライアビリティとアダプタビリティしかないんですが、人間はシステムのいろいろなレベルにおけるバルナビリティ、すなわち攻撃を受けやすい弱点を推定することができる。これが心の機能だと思われる。

(6) 心は論理階型を切り替える。仮説とか演繹とか。最後に、死を自分のこととしてとらえる。これはロボットでは死をどう考えるかという問題になる。

さらに提案なんです。正統的なアプローチ——これは菅野先生とか皆さんがここでやりになっているんですが——に加えて、機能面からアプローチして、副産物で、何

かメリットを出していくというのも手近にあるのではないかな、ということです。

司会：中村雄二郎先生のお名前も出て、別の視点からのお話でした。

4. パーソナルロボットと性格

司会：「ロボットの性格を考えた場合には、感情とか心をロボットの中で考えなければならぬんじゃないか」と、原先生が最初に問題提起をされたと思うんですが、佐藤先生からは「ロボットはある意味で見せ掛けの心でいい」というお話がありました。

では具体的にロボットの性格をどうするのか、佐藤先生、大熊先生のご意見をいただきたいと思います。

佐藤：とにかくロボットは人間がどういうふうなことを考えていて、どんなことをやりたがっているのかというのを、一生懸命計算してほしい。そういった意味でロボットの性格を定義づけると、律儀にきちんと人の意図を把握してくれる、全部ちゃんと話を聴いてくれる、というような性格でしょうか。

大熊：例えばお金を払って、切符を自動販売機で買っていると、「毎度ありがとうございます」。それしか言わないんですね。それから初めに「いらっしゃいませ」って言うんですけれども「えー、今日はいいお天気ですねえっ」とか、あるいは「お疲れのようですね」とか、一言違うこと言ってくれたらいいと思うんです。それもあまりに使われると「お疲れ」っていうのもまた退屈してしまう。そこらへんの論理階型、それを一階層上がった機能を持ったものが欲しいと思うんです。

原：自動販売機がそういうことをする必要はない。私自身がロボットにイメージしているのはどちらかというとパーソナルな機械、自分に付き合ってくれる機械です。機械の方も相手をする僕の性格をある程度読み取ってそれなりに、日によっては違ったレスポンスをする。要するに、性格をつけるなら、大勢の人に使われる機械か、それともパーソナルな機械か。機械である以上それがどういうふうに使われるか、誰に使われるか、そういうところをベースにしてどういう機械がいいのか、考えなくてはいけないと思っています。

佐野明人(電気通信大学)：「パーソナル」という言葉が出ますが、例えば、お年寄りが仏壇に向かって話し掛けるといのはよくあることですが、例えば原先生が開発しておられるような感情ロボットが、仏壇の扉をあけると出て来て、そのロボットが自分の伴侶の、いわゆる心とか感情までも模擬していると。それが本当に残された者にとって慰めになるのだろうか。

そのとき感情ロボットとコミュニケーションしているのではなく、それをインターフェースにしてあの世の自分の

伴侶と心を通わせてるだけで、そのロボットと心を通わせてはいないのではないか。

原：答えを持っていません。それは先程最後に言ったように新しい心理学の分野だと思う。僕はそういう機械を作る技術としてどういうものがあるかという議論をしているんであって、それをしたときに人間が本当に幸福になるかならないかというのは別のことだと思う。それから2番目。あの世とコミュニケーションしてるのか、僕には分からない。相手とコミュニケーションするとき、また犬とかとコミュニケーションしているとき、また子供が人形とコミュニケーションしているとき、人形としているのか、人形の方に何かイメージを持ってしているのか。心理学の中谷先生に一言いただいた方がいいと思います。

中谷和夫（東京大学）：心が分からない人ほど心理学をやりたいという傾向があったりします。ロボットの心というのも現在そういうところがある。心が分からないというのは確かに正直であり、しかも正確な認識じゃないでしょうか。

人の心というのはなかなか分からないですね。分かったとしてもその心が何であるかが分からないのではないか。一人の心がどれだけ読めるか、それを心理学的な実験でやるにはどういうふうにデータを集め、どんなふうな処理をしたらいいのか、初級心理学実験の演習の中で、単純な実験デザインで検討したことがあります。

実験者になった学生はトランプのカードをランダムにめくって、それが赤か黒か見て、半々で本当を言ったり嘘を言ったりします。そのときに向こう側にいる人は実験者の顔色を見ていて、今この人は嘘を言ってるか本当を言ってるか、分かるかどうか。何回当たったか、それを符号検定するという、そんな課題を出します。そして信号検出理論を使って、一体被検者の方はどれだけ疑い深いかに信じやすいか、というふうなその人の受信者操作特性と、それから、その嘘を言ったりしてるときの嘘が顔色にどれだけ出るかという発信者の特徴との両方を測るというように実験が使われる。情動反応の理論に、瞳孔の大きさを見てるとその人が嘘を言ってるか本当を言っているかよく分かるというのがある。自律神経系の活動が瞳孔反射としてよく出るからです。もちろん日が照ったり陰ったりすると瞳孔が大きくなったり小さくなったりしますが、それ以上に心の作用というのが瞳孔の大きさに出るので、その瞳孔の大きさの変化を人間は人と対話をしてるときにいつも無意識のうちに読んでコミュニケーションの制御をやっていると。

実際にやってみると、その“瞳孔反射説”は嘘らしい(笑)。ヨーロッパ人は瞳孔がよく見え、瞳孔の変化は確かに手掛かりになってるらしい。しかし日本人の目は黒いばかりで、変化は分からないんです。瞳孔よりも口の周りの筋肉なんかの動きが手掛かりになっている。心が読まれ

やすい人と読まれ難い人といっていますが、全体の統計をとってみると、結局分からない。人の心を人間は50%も読めない。それを機械に読ませることによって、もっと読めるようにさせることができるかどうか。全くできないとは思わないんですが、人間と同じ方法では人間を超えることは大変難しいと思う。

5. 利他主義と犬の心

広瀬茂男（東京工業大学）：ロボットの性格をどうすべきかというお話にもどりますが、今この世の中には、心が存在するのは生物しかない。ということは、生きていうことと心とは不可分の関係にある。ひょっとすると生物の進化の過程において、利己主義じゃない、利他主義を持っていた心もあったかもしれない。でもそれは存在し続けられないから今はない。だから今存在している生物の心というのはすべて、利己主義と不可分につながっている。

我々が知っている心というのはこのようなものしかないから、それが心だと言わなければならないかもしれないが、しかしそれが心の本質であるとは私にはどうしても思えない。ロボットの性格を我々が知っている利己主義とながった心にしなきゃいけないという必然性はないのではないか。人間という生物を前提とした学問である心理学から離れ、全く新しいロボットの心理学を作らなければならない理由がここにあると思う。

自己保存でなく他己保存、つまりロボットが2台あったら相手の状態の最適性のみ考えて、それをいつもよくするような行動を発生する心もあり得るわけで、それをどのように発生していくかを探っていくことによって、人間とロボットが存在するときに、人がよくなることだけを考えて、そのことに喜怒哀楽を示すような好ましい心を生み出すことができるのではないか。人間を知るというのももちろん重要だけでも、それ以外の新しい種類の心の発生というものこそ探っていくべきと思う。

加藤一郎（早稲田大学）：佐藤先生の人心が読めるロボット、私もそこから出発しております。長寿社会の中で、人の支援をしてくれるロボットというテーマを設定したときに、そのロボットが機能するためには、相手の人間の心のある程度忖度できなければ機能できないだろう、ということから出発したわけです。

もう一つ、「工学的人間学」という言葉を私は使って来ておりますが、心の問題、これは古くは哲学から、心理学、生理学、最近では認知科学もこういう分野を扱っていますが、工学の分野からもこの問題に対するアプローチができるのではないかと問題提起です。そこで私は生物の歴史をたどってみることによって心が分かるのではないかと考えました。

最初に生物が登場したのは、化学進化の段階を受けて、

原核生物としてですね。その原核生物の代表例はバクテリアです。先日この委員会で相沢先生から、バクテリアにはセンサー、モーター、それから一種のメモリ装置さえあるというお話をうかがいました。単細胞生物バクテリアに、それらがすでにすべてある。これを我々はこれまで「走性」と呼んでいます。この走性の中にすべての原点があるんじゃないか。私は今そういう仮説を立てております。

走性は本能行動の原型である。その走性の中には遺伝的に伝えられていることだけでなく、バクテリアの段階でも、新しい環境に対応して行動を変えていくということが分かってきた。これは好みをかえているということだろう。特に教えたわけでもないけれども新しい状況に適応している。これは学習の始まりといえるだろう。新しく適応していくということは、情動行動の原点がここにあると考えてよいだろう。

それからもう一つ。いろいろなセンサーがあって情報をいろいろ採り入れる。そしてまた一種のコントローラー、情報処理システムがプリミティブな形で存在する。それは意識の原点ではないか。そういったことを総合すると、バクテリアに心の出発点がある。心は走性から始まったと私は考える。

走性は一種の自己保存。生物すべて自己保存からスタートしている。35億年前に登場した「原核動物としてのバクテリア」には自己保存があるが、自己保存といっても代謝による自己増殖である。

それからまたかなりの年月がたって、真核生物が出てくる。先程ロボットの死というお話が出ましたが、生物に初めてここで性が登場し、死が登場する。性が始まったところで、利他的な行動も始まった。

それから現在まで来て、人間は進化の一つの頂点にある。その進化の頂点は2つある。もう一つは昆虫。今昆虫のことは置いて人間、ここで考えるのは人間の心。人間には進化の過程の中で情報媒体としてホルモンと電気信号が加わる。電気信号はいわゆる神経系。それがまた、かなりの年月を経て脳幹・脊髄レベルの脳ができる。それから大脳辺縁系ができて、さらに新皮質ができるという進化をしてきた。我々が「心」というのは通常はこのレベル、精神作用としての心です。

こういう長い年月の中で、人間はバクテリアが持っていた心を持って新しい精神作用を獲得したというのではなく全部積み上げて来ている。我々の行動にはバクテリアと同じようなものもあるし、あるいはその途中のいろんな生物のレベルの心すべてがある。

ここで利他的なロボットが欲しいという広瀬先生のお話に戻ると、私もそう思う。SFにあるように、ロボットが心を持ち天下を取って人間を滅すというのではたまったものでない。我々の手助けをしてもらおうというのがそもそ

もロボットの発想だから、そうでなくては困る。

自己保存から始まる以外の心もありうるかなと私も思わないわけではない。生命と心が裏表の関係のものかどうかはまだ結論を出さない方がいいと思うのはそういうことで、違う形の心がありうるかもしれない。

それを見つけるにつけても、人間から、あるいは生物から出発するアプローチの中から、それが見つかっていくんじゃないか。人間にまさに従順な心を持つ動物に犬がいる。生物から学んで心が実現できるとしたとき、そういう犬のような心ができないかなと思う。

6. 秘書ロボットと心

吉田和哉(東京工業大学)：昨日SFのロボットの討論会で類似のテーマを扱いました。その討論の中で私は理想の秘書ロボットはどうあるべきかということを題材として提示しました。そこで私が問い掛けたかったことは、まさに原先生が最初にいわれたことでした。人間と機械が共同で作業するときには機械はどうあるべきか。機械と心をもし考えたとすれば、その関係はどうあるべきか。特に機械と人とのパーソナルな関係を考えるときに、両者のコミュニケーションはどうあるべきか、もしそれを「心」と呼ぶならば、その「心」はどうあるべきかという問題です。

私自身の個人的な意見は、心を「知・情・意」と分けたとき、「知」は知能あるいは知識で、これは必要である。ただし、「情」と「意」はロボットには要らない。「情」の部分はロボットの性格はいかにあるべきかという先ほどの命題にからむが、ロボットの性格は全くないか、あるいはあってもニュートラルであるべきで、変に性格はないほうがいい。「意」もその機械自身が勝手に意図を持つと使いやすい機械じゃない。ただ、全否定するわけではなく、その「意」の部分に関しては佐藤先生の立場、人間の意図を理解してくれる機械が理想と考える。機械自身の意図はいらないけれど、人間の意図をくみとってくれる機械であってほしい。

加藤：昨日秘書の話、たいへん面白くうかがいましたが、秘書を使った経験のほとんどない方の議論だった(笑)。例えばいろんな装置が増えればいいと言うのではホーム・オートメーションやOAです。機械がどんどん増えてもそれは秘書ではない。横小路先生でしたか以心伝心とおっしゃいました。私の秘書は私の目の動き、ちょっとした言葉からさっと分かってくれる。そのかわり私、何かと指摘されることも多い。それでこそ秘書ですね。操縦しているつもりが逆に操縦されているというお話もありましたが、まさにそうで、それが秘書との関係なんですね。自動機械ではだめです。それはせいぜい現在の産業ロボットのレベルです。うかがっていてそう感じました。

それから「知・情・意」というのは、私自身の仮説では

切り離せない。さっき意識を私なりに定義しましたが、その「意識」があって、それから先天的に持っている「自己保存」、これが感情のいちばん根源になっている。無意識のときは何の判断もできませんから意識のある状態で、自己保存を基本評価基準として、我々はすべてのことを判断している。それが「知」の構造です。

我々、知能ロボットとか人工知能と言ってますが「意・情」がないと先程言った精神的な意味での知能になりえない、と仮説を立ててアプローチしているのです。

それから意図を理解できればいいというお話。私は機械が人間の意図を理解できるということは、その機械は心を

持っている状態にあると考える。人間どうしてもあるいは人間と他の動物の間でも、他者に対して心を感じるのは、我々がそういう判定基準をもってやっていることです。その基準も人それぞれ微妙に違うかもしれないけれども、大局的に共通の基準を持っていてそれで済んでいる。そういうふうに思います。

司会: この問題、まさにこれからこの学会を舞台にして議論を深めていくべき重要なテーマだと思います。専門委員会としては、さまざまな先生方とさらにディスカッションをする場を、これからも提供したいと思います。ご参加ありがとうございました。