

ロボットの社会的発達と「心の理論」の獲得

小嶋 秀樹

通信総合研究所 けいはんな情報通信融合研究センター
〒 619-0289 京都府相楽郡精華町光台 2-2-2
xkozima@crl.go.jp

いまロボットに必要とされているのは、人間の社会的活動にコミットする能力である。社会的活動は、他者との効率的な協調・競争を実現するものであり、他者の行動を予測・制御する能力によって成り立つ。本論文では、この他者行動の予測・制御という視点から、他者の心的状態にアクセスするための「心の理論」を、とくにその発達に焦点をあててモデル化する。この社会的発達モデルは、原初的な共同注意と模倣を出発点とし、他者の行動を間接的に疑似体験することによって社会的な行動パターンを蓄積していく。また、遅延模倣やランダム反応といった自発的行動を、他者（とくに養育者）からのフィードバックに応じて意味づけしていく。このような他者とのインタラクションをとおして、他者が物理的・社会的環境についてもっているモデルを探索し、さらに社会のなかで共有されたプロトコル（ジェスチャ・言語など）を探索していく。この探索活動は、マクロ的には社会的発達であり、ミクロ的にはコミュニケーション（社会的コミットメント）でもある。

Epigenesis of social intelligence and “theory of mind”

Hideki Kozima

Keihanna Human Info-Communication Research Center, CRL
2-2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0289, Japan
xkozima@crl.go.jp

Commitment to the social activity of human beings is an important requirement for symbiotic robots. Human social activity, which is for and maintained by efficient cooperation and competition with others, is based on the human ability to predict and control others' behavior. In terms of the prediction and control, this paper describes an epigenetic model of “theory of mind” as a device for accessing others' mental states. In this model, infants/robots start their social development from a primordial form of joint attention and physical imitation, then they accumulate patterns of socially meaningful behavior through indirectly experiencing others' behavior. Infants/robots also acquire the social meaning of their spontaneous behavior, such as deferred imitation and operant responses, according to the feedback given by their caregivers. Through this interaction with others, infants/robots explore the models that others have of the physical and social environment; they also explore the communication protocol, like that of gesture and language, shared in the community.

1 はじめに

近い将来、介護ロボットや家事ロボットが家庭に入り、案内ロボットや秘書ロボットが職場に入ってくるだろう。ロボットを物理的・社会的存在 (physical/social being) として人間の活動に組み入れるには、ロボットを人間と同じ物理的・社会的環境に人間と同じように埋め込まなければならない。

ロボットを人間と同じ物理的環境に埋め込むには、ロボットに人間と同じような感覚運動モダリティを与えればよい。ロボットは人間と同じように物理的環境を捉え、物理的環境に働きかけることが可能になる。ゆえに、無声映画やダンスのように、行動そのものがコミュニケーション媒体となる。これは人工知能研究で「身体性 (embodiment)」と呼ばれている要請である。最近のヒューマノイド研究 [1] にみられるように、この分野の進歩はめざましい。

ロボットを人間と同じ社会的環境に埋め込むのも同じであろう。ロボットが人間と同じように社会的環境を捉え、社会的環境に働きかけることができればよい。このためには、まず人間と同じようなコミュニケーション能力 (言語や文化的慣習を含む) をもつこと、そしてこの能力をつかって人間の社会的活動にコミットすることが必要である。ここでいう「コミット」とは、その場にふさわしい役割をもって社会的活動に参加することであり、これによって個々のコミュニケーション行動が意味をもつようになる。これは人工知能研究で「状況性 (situatedness)」と呼ばれている要請にちかい。しかし、Eliza 以来、この分野に質的な進歩はほとんどない。

どうすれば身体をもった人工システムを人間の社会的活動のなかに埋め込むことができるのだろうか。本論文では、発達とその障害 (とくに自閉症) を手がかりに、状況に埋め込まれた身体 (situated embodiment) [2] にアプローチする。

新生児でも、泣き・ほほえみ・人への注目などによって、他者の行動をある程度コントロールすることができる。ただ泣くことによって養育者から食べ物を手に入れたり抱っこしてもらうことから始まり、指さしなどによって欲しいものを選択的に手に入れるようになる。やがて言葉を使うようになり、より複雑な社会的インタラクションに自らを巻き込んでいく。この発達プロセスは、まるで作り込まれたプログラムであるかのように効率的であると同時に、養育者を含めた社会的環境がなければ成り立た

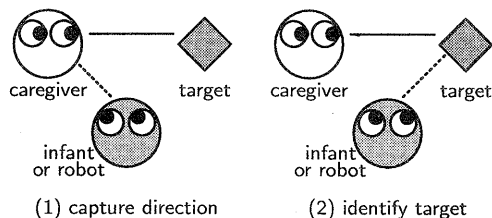


図1. 共同注意をつくりあげるプロセス。

ず、また社会的環境が変れば発達の方向も (個別言語の獲得のように) 変ってくる。[3]

一方、自閉症児/者は自らを社会的インタラクションに巻き込もうとしないことが多い [4,5]。もちろん、欲しいものがあれば養育者の手をつかんで取らせようとし、VIQ の高い自閉症児であれば質問に対して文法的にも意味的にも正確に答えるだろう。しかし、相手の意図を読むことが困難であり、語用論レベルでのインタラクションに失敗することが多い。たとえば、「時計もってる?」「うん」のように、相手の意図に気づくことなく、字義どおりの意味で回答してしまう [4]。このような自閉症児/者のもつコミュニケーション障害は、現在のヒューマン・インタフェースやヒューマノイドにもそのまま当てはまるだろう。

おそらく自閉症児/者は、自らを社会的インタラクションに巻き込ませるための「フック」を先天的 (もしくは妊娠・周産期など発達初期) に失っていたのではないだろうか。ならば、その「フック」をヒューマノイドに与えれば、そのヒューマノイドは養育者 (人間) とのインタラクションをとおして、人間社会のなかで共有されているコミュニケーション・プロトコルを探索することができるだろう。本論文では、この「フック」として、原初的な共同注意と模倣を想定し、これら「フック」からどのように社会的な行動が構成されるのかを論じたい。

2 共同注意と間接経験

共同注意 (joint attention) とは、他者 (養育者) が注意を向けている対象に自分も注意を向ける活動である。図1に示すように、共同注意を実現するには、(1) 養育者の注意方向を視線・顔の向き・指さしなどから読み取り、(2) 養育者が実際に注意を向けている対象を同定する (切り出す) ことが必要であ

る。厳密に言えば、さらに (3) 両者が注意を共有していることを相互に了解することも必要であるが、ここでは (1) と (2) だけを満足するものを「共同注意」と呼び、さらに (3) を満足するものを「相互共同注意」と呼んで区別する。

2.1 共同注意 — その発達と障害

健常児では、一般に生後6ヶ月前後から原初的な共同注意 — 養育者が左右どちらかに向いたとき、その方向の一番近い対象を捉える — が可能であり、12ヶ月では視線の角度に応じて正しい対象を捉え、18ヶ月では自分の背後にある対象を振り返って捉えることができる [5]。また、生後2ヶ月でも養育者が向いた方向に首を向けることが観察されており、もし目立つ色・動き・明確な輪郭などをもつ対象がそこにあれば共同注意が実現されるだろう。また、ヒト以外の霊長類にも共同注意がみられる。たとえば、チンパンジーやオラウータンなどの大型類人猿は、他者の指さし・顔の向き・視線などを追跡することができる [6]。

一方、自閉症児/者は共同注意に障害をもつことが多く、共同注意の欠陥は現在の診断基準 [7] にも含まれている。しかしながら、実験室環境で対面した人物の注意対象を同定するタスクをやらせると、正確にその対象を同定できることが多い [5]。このことは、自閉症児/者の視知覚系は正常であるが、他者の注意を読もうとする動機がないことを示唆している。また、自閉症児/者は他者とのアイコンタクトを避ける傾向が強いと言われる。これは、健常児/者が他者の目に注意を向け視線を読もうとするのに比べて、自閉症児/者にはそのような選好がない（または弱い）ためかもしれない。

以上から、健常児はほぼ生まれながらに他者の注意に敏感に反応するが、その一方、自閉症児は他者の注意に気がつくことなく発達の初期段階を送ると考えられる。もちろん、自閉症児も他者の視線・指さしなどが何らかの対象・事象と随伴することを学習できるだろうが、これを生得的な制約なしに達成するにはかなりの時間がかかると思われる。

2.2 間接経験

共同注意の役割とは何だろうか。それは他者の行動を観察するためのデバイスであろう。図2に示すように、養育者が対象と何らかのインタラクションをしているとする。共同注意によって、乳児（もし

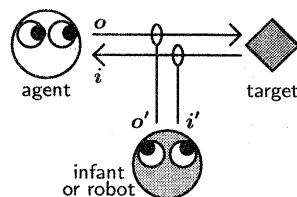


図2. 他者の感覚入力と運動出力の観察.

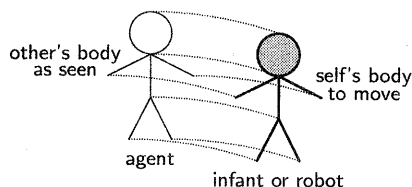


図3. 自他同形性マッピング.

くはロボット) は、この養育者がどのような感覚入力 i を対象から受け取っているのかを推定することができる。同じ対象に同時に注意を向けるのだから — 実際には視点取り (perspective taking) が必要である — 両者の感覚入力 i と i' はほぼ同じものになる。また、養育者の注意方向を読み取ろうとするとき、乳児は養育者がどのような運動出力 o を対象に対して働きかけているかを観察することになる。つまり、共同注意によって、他者の感覚入力と運動出力を観察することができるわけである。

ここで注意が必要なのは、乳児が観察した他者の運動出力 o は、あくまで遠感的 (おもに視覚的) なイメージにすぎないことである。観察した他者行動を模倣可能にするには、この遠感的イメージ o' を乳児自らの筋肉運動イメージ $\mu(o')$ に変換するマッピングが必要となる。残念ながら、この遠感的な他者から運動的な自己への同形性マッピング $o' \mapsto \mu(o')$ を実現するメカニズムはよくわかっていない。おそらく、新生児にみられる表情模倣 [9] や共鳴動作のような原初模倣 (すなわち視覚・運動的な反射) を土台として、養育者との模倣あそびや自分の身体の一部を見ること (たとえばハンドリガード) などをとおして構成されるのだろう。

他者の感覚入力 i を共同注意によって i' として取り込み、また他者の運動出力 o を遠感的イメージ o' として捉え、それを運動イメージ $\mu(o')$ として取り込む (図4参照)。これが他者行動 $\langle i, o \rangle$ を

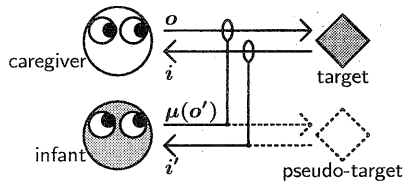


図4. 間接経験から模倣へ。

$\langle i', \mu(o') \rangle$ として疑似体験すること，すなわち間接経験 (indirect experience) である。(ここでは行動を感覚入力から運動出力への対応関係として捉えている。)

2.3 模倣と学習

間接経験した他者行動を再生すれば模倣 (imitation) となる。図4に示すように，共同注意によって得られた感覚入力 i' にあわせて，間接経験した養育者の行動 $\langle i', \mu(o') \rangle$ を参照し，対応する運動出力 $\mu(o')$ を実行すればよい。このとき乳児が働きかける対象は，養育者が働きかけている対象そのものでもよいし，同等の別のものでもよい。さらに，働きかける対象が異なる形状・機能をもつ場合や対象そのものが存在しない場合もありうるが，これらはより高次の発達段階で「ふり (pretence)」や「ごっこ遊び (make-believe play)」として出現する。

健常児の場合，身体運動の模倣は6ヶ月前後から始まり，9ヶ月ごろから対象への操作を模倣するようになる。一方，自閉症児は模倣にも障害をもつことが多い。興味深いことに，健常児では身体模倣のあとで音声模倣が出現するが，自閉症児では音声模倣が早い時期から出現することが多い。

間接経験は，他者がどのような状況にどのように対処するのかを学習するためのデータを（正例だけでなく負例も）提供する。この学習によって，他者がある対象（たとえば電話器）を前にしたとき，どのような行動をとるのか（たとえばダイヤルする）を予測することができる。また，他者が何らかの動作（たとえばバイバイと手を振ること）をしているとき，その動作がどのような対象（たとえば視野外にいる人物）からの感覚情報に向けられたものなのかを推定することができる。また，他者に何らかの行動をさせたければ，それに対応する感覚入力（または対象）を提示すればよい。

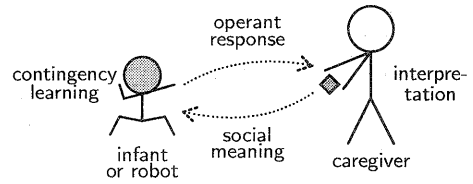


図5. 養育者による社会的意味づけ。

3 探索活動としての発達

間接経験によって，外因的に誘発される行動のパターンを学習することができる。しかしこれだけでは目の前にキャンディがあっても，誰かがキャンディを食べるのを模倣せずには，それを手に入れることができない。では，自発的にこのキャンディを食べ，その心地よい甘さを知るにはどうしたらよいだろうか。すなわち，自発的に環境（対象）に働きかけ，報酬を得るにはどうすればよいだろうか。おそらく，間接経験によって得た運動出力（およびその変形版）やランダム運動（オペラント反応）を実行し，その結果として返ってきた感覚情報を（たとえば快・不快として）評価し，それを強化子として学習を進めるという戦略であろう。

3.1 養育者の役割

養育者も探索活動の対象（上の例ではキャンディ）になりえる。図5に示したように，乳児（もしくはロボット）が既存の運動レパトりの変形版やランダム運動を提示すると，それを受けた養育者はその社会的意味を解釈し，それを乳児に返してやる。社会的に適切な提示であれば報酬（注目・笑い声など）が得られ，社会的に不適切な提示であれば罰（無視・怒り声など）を受けるだろう。社会的に無意味な提示であれば，養育者はそれを無視するだろう。

さらにこの探索活動によって「意図」というものが発生する。たとえば，乳児が何の意図もなしに腕をある方向に伸したとする。それを見た養育者は，乳児の腕の先にあるおもちゃを取り，乳児に渡してやる。結果として乳児は願ってもない報酬を得ることになる。この行動 — 腕伸しを原因としおもちゃを結果とする — は強化される。何度も試行を繰り返すなかで，おもちゃを腕伸しの方向の合致したときの行動が選択的に強化され，ついには欲しいおもちゃを手に入れる行動が形成される。つまり，原因から結果をつくりだす行動が，手段によって目的を

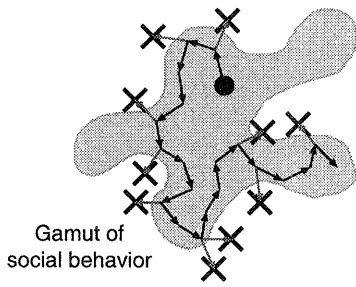


図6. 社会的行動の探索.

達成する行動として機能するようになる。意図的行動とは、欲求（オモチャが欲しい）つまり目的に対応する手段（腕伸し）の行使である。

3.2 社会的行動の探索

間接経験（模倣）によって行動パターンの「タネ」を取り込み、ときに全くランダムな行動を「タネ」として、養育者からのフィードバックに応じて社会的に意味のある（報酬をたくさん得られる）行動をシェイピングしていく。たとえば、図6に示した可能な行動からなる空間のなかで、「タネ」にランダムなゆらぎを与えていき、その各ステップで養育者をとおしてその行動を評価し、社会的に有意な行動クラスを探索することが考えられる。もっとも、この行動クラスは明確な輪郭をもつ2値的なものではなく、むしろファジー集合的な性質をもつかもしれない。

このような社会的行動の探索によって、社会のなかで共有されているコミュニケーション・プロトコル（ジェスチャ・言語など）を獲得することができる。あるレベルのコミュニケーション能力を獲得できれば、それを利用してより複雑な探索活動が可能となる。このメカニズムが人間のコミュニケーション能力を特異的に高度なものにしたのだろう。

4 ミクロ探索 — 心を読むこと

いままでコミュニケーション能力がどのように発達するのかを、共同注意・間接経験による社会的行動の探索としてモデル化してきた。この探索活動によって獲得されるものは、社会の構成員の平均的な行動パターンである。その一方、個々の探索活動は、特定の個人の特定の状況での行動を対象として

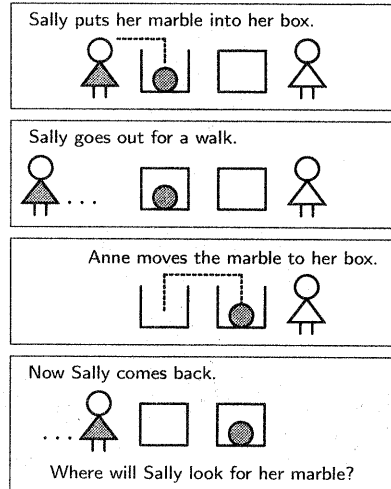


図7. 誤信念課題「サリーとアンのテスト」.

いる。そこで前者をマクロ探索、後者をミクロ探索と呼ぶことにする。

このミクロ探索とはどのようなものだろうか。個々のミクロ探索は、すでに述べたとおり、コミュニケーション行動そのものである。個々のコミュニケーション行動がもつマクロ平均からの「ゆらぎ」は、おそらくランダムなものではなく、ある程度の時間スパンをもったバイアスとなるだろう。たとえば「行動Bを実行すると、個体の行動パターンは...のように変化する」といったバイアスである。このようなバイアスを「心の(部分)状態」とみなせないだろうか。楽しいときと楽しくないときでは異なるバイアスが生じる。何かを知ってからと知るまえでは異なるバイアスが生じる。このようなバイアスがいくつも重なりあって「心の状態」がつくられていると考えられる。

このバイアスはどのようにして認識されるのだろうか。おそらく、つきあう相手ごとの行動を間接経験によってフォローし、その「心の状態」をシミュレートしていると思われる。健常児の場合4才前後になってはじめて「誤信念課題」— 他者が自分とは異なる信念を持ちうることへの理解についてのテスト(図7)— にパスする。おそらく、複数のスペースに分かれた長期記憶を管理する能力が備わったとき、このシミュレーションも可能になるのだろう。バイアスを認識すること、すなわち「心を読むこと」は、相手の行動履歴をシミュレートすること

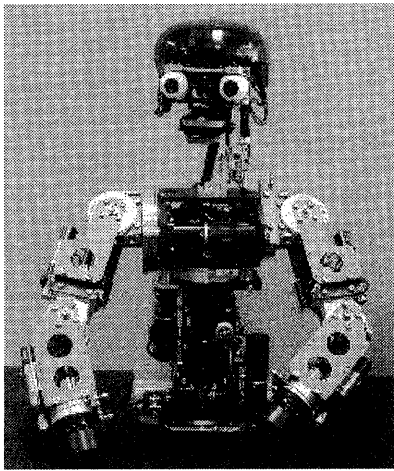


図8. 赤ちゃんロボット Infantoid.

であり、そのためには間接経験（あるいは共感）の能力と複数にスペースに分かれた記憶（そして記憶管理能力）が必要である。

5 おわりに

本論文では、原初的な共同注意と模倣を「フック」として、養育者や他の人間との社会的インタラクションをとおして社会的能力（とくにコミュニケーション能力）を獲得していく発達モデルについて述べた。このモデルは、健常児の発達と自閉症児の発達障害からの示唆を受けつつ、コミュニケーション能力を機械的に獲得できるアルゴリズムを模索したものである。まだ多くの矛盾点・欠陥点があるが、発達・進化・社会についての研究との融合をとおして、改善していきたい。

また著者は、この発達モデルを機械的に実証するために、Infantoid というヒューマノイド [10] を開発している。Infantoid は、図7に示すように、3歳児の上半身とほぼ同じサイズをもち、ハンドを除いて23の自由度をもつ。とくに共同注意に関係する眼球は、人間のものに近い形状をもたせている。また、唇を複雑に動かして、いろいろな表情をつくることができる。この Infantoid に、原初的な共同注意や模倣の能力を工学的に与え、それを使って人間の養育者とインタラクションさせることによって、どのような情報が社会的環境から獲得できるのかを調べていく予定である。

6 参考文献

- [1] Rodney A. Brooks, *et al.* (1998), The cog project: building a humanoid robot, In C. L. Nehaniv (ed.), *Computation for Metaphors, Analogy and Agents*, Lecture Notes in Artificial Intelligence, Vol. 1562, Springer-Verlag.
- [2] Jordan Zlatev (1999), The epigenesis of meaning in human beings, and possibly in robots, *Lund University Cognitive Studies* 79, Lund University.
- [3] Hideki Kozima (1999), Attention-sharing, behavior-sharing, and acquisition of language, *International Workshop on Ecology of Language Acquisition*, University of Amsterdam.
- [4] Uta Frith (1989), *Autism: Explaining the Enigma*, Blackwell.
- [5] Simon Baron-Cohen (1995), *Mindblindness: An Essay on Autism and Theory of Mind*, MIT Press.
- [5] George Butterworth and Nicolas Jarrett (1991), What minds have in common in space, *British Journal of Developmental Psychology*, Vol. 9, pp. 55-72.
- [6] Shoji Itakura (1996), An exploratory study of gaze-monitoring in nonhuman primates, *Japanese Psychological Research*, Vol. 38, pp. 174-180.
- [7] *The ICD-10 Classification of Mental and Behavioral Disorders* (1993), World Health Organization.
- [8] Andrew Meltzoff and Alison Gopnik (1993), The role of imitation in understanding persons and developing a theory of mind, in Simon Baron-Cohen, *et al.* (eds), *Understanding Other Minds*, Oxford University Press.
- [9] Andrew Meltzoff and M. Keith Moore (1983) Newborn infants imitate adult facial gestures, *Child Development*, Vol. 54, pp. 702-709.
- [10] Hideki Kozima (2000), Infantoid: an experimental tool for developmental psycho-robotics, *International Workshop on Developmental Study*, Tokyo.