解　説

発達ロボティクスからみた模倣とコミュニケーションのなりたち

小　嶋　秀　樹*

要旨

「他者の意図（目的を達成するための行為プラン）を再現すること」としての模倣は，ヒトという生物種を特徴づける問題解決・学習ストラテジーである．模倣能力をもったロボットを実現できれば，言語や慣習などさまざまな社会的スキルを，観察と実践をとおして獲得させることができる．では，ヒトはどのようにして目に見えない他者の意図を目に見える身体動作から読みとるのだろうか．ロボットによる他者意図の読みとりを可能にするには，どのような機能が必要だろうか．本論文では，ヒト乳幼児のコミュニケーション発達を手がかりとして，模倣を成り立たせている認知メカニズムについて考察してみたい．

1．はじめに

あるパーティ会場で，空のグラスを手に持った人があなたに近づいてくる．この人は，あなたにグラスを手渡したいのだろうか，それともあなたにビールを注いでほしいのだろうか．いや，グラスを凶器にあたなに襲いかかるのかもしれない．目に見える動作は同じでも，その裏にある意図がわからなければ，この人の行動を予測・説明することはできない．目に見えない意図を読みとり，それに応じて自分の行動を調整していくことで，他者との社会的インタラクションが実現される．これが広い意味でのコミュニケーションであろう．

このことは「模倣」についても当てはまる．他者の身体動作をそのまま自らの身体によって再現すること（mimicking）や，他者が生じさせた結果（報酬の獲得など）を目的として取りこみ，それを試行錯誤（行為レパートリの組合せ）をとおして再現すること（emulation）だけでは不十分だ．さまざまなスキルを効率よく獲得するには，目に見える身体動作と結果の裏にある意図（intention）――つまり「目的を達成するための行為プラン」――を再現すること（imitation）が必要になる．この意図レベルの模倣学習によって，個体による発明・発見がすみやかに集団内に拡散・定着，世代を越えて蓄積され，ヒト特有の高度な文化が形成されたのだろう[1]．そして，このような文化発展の副産物として，ヒトは意図レベルの社会的インタラクションを営むようになったのだろう．

ならば意図レベルの模倣能力をもったロボットであれば，ヒトとの社会的インタラクションが可能であり，その実践をとおして言語や慣習といったヒトの文化を学んでいくことで，ひとりの社会的存在になれるのではないだろうか．言い換えれば，ヒト乳児の社会的発達をロボットによって再現することができるのではないだろうか．本論文では，このような「発達ロボティクス（epigenetic robotics）」[2]の考えから，意図レベルの模倣を成り立たせている認知メカニズムについて考察してみたい．次章では，模倣研究における重要課題のひとつである身体マッピングについて考察し，第３章では，目に見える身体動作と目に見えない意図をつなげるアイデアとして「目的論的スタンス」[3]を紹介する．第４章では，これら考えをふまえて，ロボットによる意図理解・模倣への発達的アプローチについて展望する．

2．身体マッピング

意図レベルや目的レベルの模倣のまえに，まず身体動作の模倣を実現すべきだという態度は，工学的な立場からは自然にみえる．しかし，視覚イメージとして捉えた他者の身体動作と自分の運動イメージ（運動指令あるいは自己受容感覚の系列）とを対応づけることは，これらモダリティ間の異質性ゆえに難しい問題となる（図１）．この「身体マッ

2004 年 11 月 18 日受付
* 情報通信研究機構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 3-5
Tel : 0774-98-6892　Fax : 0774-98-6963
E-mail : xkozima@nict.go.jp
キーワード：認知発達，社会的学習，意図理解，目的論的スタンス
身体マッピング，社会的インタラクション

ピング」が生得的であると考える研究[4]もあるが，最近の発達心理学研究や脳科学研究からは，むしろ目的レベルの模倣が先にあり，それを使った他者とのインタラクションをとおして身体マッピングが構成されていくという考えが広がりつつある．

2.1 モダリティ未分化な知覚による動作模倣

身体マッピングがなくてもある種の身体動作を模倣することができる．たとえば，生れて間もない乳児が，至近距離で対面した他者の口や舌の動作を模倣できること（新生児模倣）[5]や，２カ月の乳児が，対面している養育者の顔方向変化に呼応して，自分の視線や顔の向きをシフトさせること[6]などが知られている．また，形の異なるオシャブリをしゃぶっただけで視覚的にそれらを区別できること[4]や，刺激のリズムあるいは強さを視覚と聴覚の間で対応づけられること[7]なども興味深い．これらは，おそらく生得的なメカニズムによって実現されているのだろうが，特定の感覚モダリティ間（たとえば視覚→運動）に特化した反射経路の集まりとして実現されているのではなく，むしろ皮質下におけるモダリティ未分化（amodal）な知覚を反映しているという仮説[8]が提案されている．

モダリティ未分化な知覚とは，皮質上に各モダリティごとの情報処理システムが発達する以前（「２カ月革命」まで[9]～生後４カ月頃まで[8]）の知覚スタイルである．さまざまなモダリティの感覚イメージ（そして運動イメージ）が，リズム・テクスチャ・エネルギー・方向といった次元によって皮質下で統一的に知覚される．リズムやテクスチャは刺激の時間的・空間的な強弱パターンであり，前者は粗いもの（手拍子など），後者は細かいもの（音色・手触りなど）をカバーする．また，エネルギーは刺激の大きさや速さであり，方向は自分からみた刺激の位置や移動方向を意味する．これら次元によって外界からの刺激を捉え，情動的な価値づけを加えることで，意味のある身体動作――定位（刺激に向きあうこと）・リーチング（刺激に手を伸ばすこと）・回避（刺激を避けること）など――を引きだすことができる．このような知覚スタイルによって，顔動作模倣など，ヒト乳児のもつコンピテンス（環境に応じていく能力）の多くを説明できるだろう．

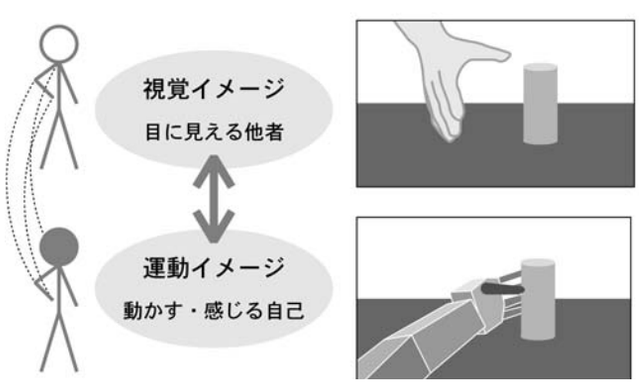


図１　自他間の身体マッピング（対応づけ）

2.2 自分の手を見ること

モダリティ未分化な知覚によって，原初的な動作模倣をある程度実現できるだろうが，道具使用のようなより複雑な動作を模倣するには別のメカニズムが必要だろう．その手がかりとして最近の脳科学研究で注目されているのが，サル（マカク）の前運動野に発見されたミラーニューロン（mirror neuron）[10]である．ミラーニューロンは，他者が特定の動作（とくに手の動作）を見たときと，その動作を自分で行なったときの両方で発火する．丸い物をつかむ・細長い物をにぎる・小さな物をつまむなど，対象のもつアフォーダンス（対象との自然な関り方）ごとに異なるミラーニューロンが存在する．このようなミラーシステムがヒトにも（それもブローカ言語野のあたりに）存在することが示唆されている．

このミラーシステムは生得的なものだろうか．ヒトの生活世界につぎつぎと新しい人工物（道具）が導入されてきたことを考慮すると，経験をとおしてミラーシステムが構成されていくと考えるほうが自然であろう．その可能なルートのひとつとして，「自分の手を見ること」による学習[11]が提案されている．その要点は，ある対象に関わるときの，(a) 自分の手動作の視覚イメージと，(b) そのときの運動イメージ（運動指令あるいは自己受容感覚の系列）の相関を学習していくというものだ．この相関学習は，さまざま対象のアフォーダンスを探索していくなかで実行できる．そして，他者の手動作を見たときに，同じような手動作をエンコードしているミラーニューロンが発火し，対応する運動イメージが喚起される．この運動イメージを実行すれば動作模倣が実現する．

2.3 機能的マッピング

上述のミラーシステムの学習方法にも問題点は多い．たとえば，自分から見える自分の手と他者の手は，単純に比較可能ではない．とくに対面状況での他者の手は，自分の手ではありえないような角度で見えるだろう．手以外の身体部位については，さらに難しくなるだろう．また，先天的に腕を失っている場合はどうだろうか．そのような乳児でも，他者の手動作を足や口を使って「模倣」できるだろう．

このような問題をクリアするために，機能的ミラーリング（functional mirroring）[12]という学習方法が提案されている．これは，動作によって対象にもたらされた効果（目に見える結果）を軸として，他者の動作と自分の動作を対応づけていくという手法である．たとえば皿の上に置かれた握り寿司を口元にもってくるという動作の場合，手でつかもうが，箸でつかもうが，結果として寿司が口元にくればよい．このような等価性を基準として，目に見える他者の動作と，実際に実行する自分の動作の間の対応を学習して

いく．言い換れば，動作そのものを対応づけるのではなく，同じ機能をもった動作のクラスとして，視覚イメージとしての他者動作と運動イメージとしての自己動作を対応づけていく．一見，自由度が大きすぎて学習困難にもみえるが，対象のもつアフォーダンスは同じような身体をもつ他者にも同じように知覚され，適用可能な動作の範囲は十分小さくなるため，自他間のマッピングは学習可能になるだろう．

3．目的論的スタンス

前章でみたように，身体マッピングは，意図レベルや目的レベルの模倣に先立って獲得されるのではなく，さまざまな対象への身体動作を他者とやりとりするなかで構成されていくのだろう．しかし，身体動作と意図の間にはまだ大きな隔たりがある．ここでは，これらをつなげるアイデアとして Gergely（ゲルゲイ）[3] が提案する「目的論的スタンス（teleological stance）」を解説する．

3.1 合理性マッピング

ヒト幼児が他者の身体動作を模倣するとき，同じ結果を得るために「合理性（rationality）」という基準で自分の身体動作を選択すること[13] が実験的に明らかにされている．この実験では，幼児が対面したモデルの動作を真似るように指示される．その動作とは，自分の耳（左右どちらか）を自分の手（左右どちらか）で触るという単純なものだ．結果として，幼児は触るべき側の耳を正しく再現できたが，耳とは反対側の手を使う条件のとき，幼児は正しい耳を同側の手を使って触ることが多かった．この結果から，モデルと同じ結果（正しい耳に触ること）を最も合理的・効率的に実現するように，自らの身体動作を選択していったと解釈できる．

また，モデルが両手で両耳を触る場合（平行・交差）も実験された．モデルが交差条件を提示した場合，上述の合理性原理に従えば，幼児は平行条件で応じることが予測される．しかし実際には，幼児は交差条件で応じることが多かった．これは，腕を交差させるというモデルの動作が大きな情報量（顕在性）をもつため，腕を敢えて交差させること自体がモデルの目的であると認識したためであろう．

3.2 目的論的な行為解釈と模倣

12 カ月の乳児が他者の動作を「目的を達成するための合理的な手段」として解釈できることが一連の心理実験によって明らかにされている．これを可能にするのが「目的論的スタンス」[3] と呼ばれる行為解釈システムである．図２に示すように，目に見える他者の行為〈動作 A，結果 R，状況 S〉を，〈手段 M，目的 G，制約 C〉として，すなわち「制約 C のもとで目的 G を達成するための最も合理的な手段 M の実行」として解釈する．そしてこの解釈のもとで，行為の３つ組のうち１つが未知の場合，それを予測することができるようになる．

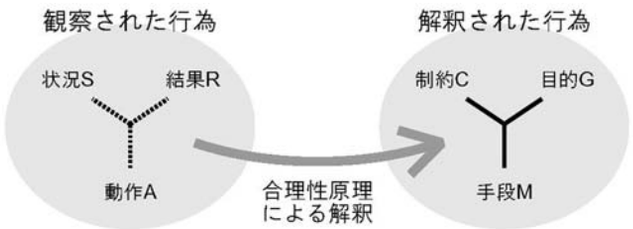


図２　合理性原理による行為解釈[3]

このことを実証した心理実験[3] を紹介しよう．図３の左側のアニメーションを 12 カ月児に何度も提示する．たとえば実験 A のアニメーションは，白玉が壁を飛び越えて黒玉と出会うというものである．このようなアニメーションに十分に馴れさせたところで，中央あるいは右側のアニメーションに切り換える．もし乳児にとってそのアニメーションが「期待はずれ」であればより長く凝視するだろうし，「期待どおり」であれば凝視時間は増えないだろう．結果は，３つの実験いずれも，中央のアニメーションをより長く凝視し，右側のアニメーションには凝視時間の変化をみせなかった．この意味を実験 A について説明しよう．左側のアニメーションを何度も見た乳児は，白玉のふるまいを〈手段＝ジャンプ，目的＝黒玉と会う，制約＝壁〉という合理的な行為として解釈する．つぎに見るアニメーションには壁がない．中央のものには既知手段＝ジャンプがそのまま残されているが，右側のものには新奇手段＝直進が現われる．そして乳児は，それが合理的な手段かどうかという観点から，前者を「期待はずれ」，後者を「期待どおり」と捉えた．

最後に，このような目的論的スタンスと模倣の関係を明らかにした実験[14] を紹介したい．14 カ月児に２つのモデル動作のいずれかを見せる．どちらも毛布をかぶったモデルがテーブルに置かれた照明箱（上から押すと点灯する）を

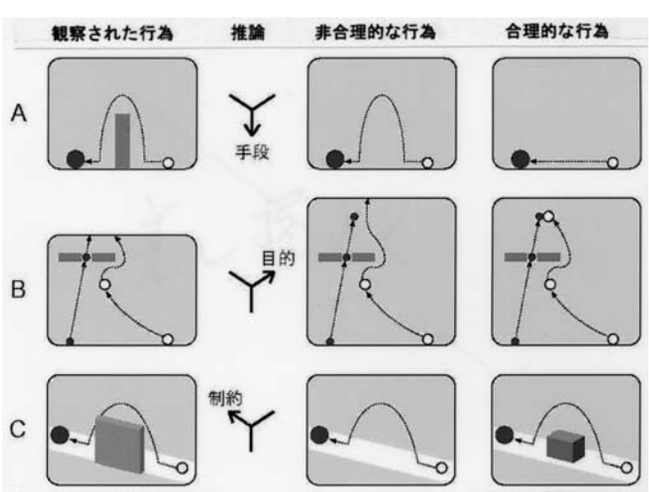


図３　目的論的スタンスによる目的・手段・制約の推論[3]

額で押すという動作だが，乳児群Aには両手とも自由に使えるモデルによる動作を，乳児群Bには両手とも使えないモデルによる動作を提示する．1週間後，両手が自由な乳児の前に，同じ照明箱を20秒間だけ提示した．その結果，両群ともすべての乳児が手を使って目的（照明箱の点灯）を達成したが，乳児群Aの69%が額を使ったモデル動作も再現したのに対して，乳児群Bでは21%にとどまった．乳児群Bは，両手が使えない状況での合理的なモデル動作を観察しており，それを両手が使える状況に合せて合理的に再現しようしたのだろう．

4．ロボットによる意図理解・模倣

これまで見てきたように，身体マッピングは模倣に先立って存在するものではなく，むしろ他者（とくに養育者）とのインタラクションをとおして構成されるものである．乳児は，さまざまな生得的コンピテンスを駆使して，他者との原初的なインタラクション――行為のやりとり――をつくりだし，すこしずつ他者の行為を（よく慣れた状況を共有するなかで）予測できるようになっていく．そして12カ月ごろまでに，他者の行為を「目的を達成するための合理的な手段」と解釈できるようになる．このような目的論的スタンスによるインタラクションをとおして身体マッピングが段階的に構成されていき，さまざまな動作単位として，あるいはその時系列・空間配置として，他者の身体動作を捉えられるようになっていく．こうして獲得された目的論的スタンスと身体マッピングによって，意図理解，つまり「目的を達成するための行為プラン」として他者の行為を捉えることへの道がひらけていく．

しかし大きな謎が残されている．そもそもなぜ乳児は他者の行為を模倣するのだろうか．なぜヒトは，自分と他者を重ねあわせ，目的や身体動作，そして意図を，共有しようとするのだろうか．たとえロボットに模倣能力を実装できたとしても，この「共有への動機づけ」を組み込むことができなければ，社会的スキルや文化的知識の実践学習には結びつかない．

この謎に答えを与えてくれるのは，やはり乳児のコミュニケーション発達だろう．母子間のコミュニケーションは，自他未分化な情動共有からはじまり，すこしずつ互いの知覚・動作・意図を，理解・調整・共有していく．このことを，生後1年間のコミュニケーション発達[15]からみてみたい．

- 0～3カ月：発声・表情をともなったアイコンタクト（見つめあい）が確立する．養育者が乳児の反応パターンを予測することで，視線・表情・声のやりとりにリズミックな時間的構造（ターンテイキングなど）が形成されていく．
- 3～9カ月：養育者が乳児の欲求や快不快などを積極的に読みとり，それに応答してやることで，情動を共有したインタラクションが形成されていく．乳児も，そのような養育者からの関わりを予測できるようになっていく．



図4　ロボットとヒト養育者の注意・情動のやりとり[15]

- 9～12カ月：乳児は，情動だけでなく，注意を養育者と共有できるようになる．注意対象を共同化しつつ，対象についての情動を相互に参照していく．そして，対象にもたらされた変化とそれに随伴した情動――ポジティブ・ネガティブ――から，達成された（あるいは未達成の）目的に気づくようになる．

このように，まず情動の共有があり，つぎに互いの注意対象を参照しあうようになる．そして，さまざまな行為について，目に見える結果と自他間で共有された情動変化を経験するなかで，「目的」というヒトの行動を方向づける何かを捉えていく．

いまのところ，ロボットの模倣能力は，ここで考察してきたような「目的」あるいは「意図」レベルには達していない．身体マッピングの獲得や目的の共有に養育者とのインタラクションが不可欠であるならば，ロボットでそれらを実現するにもヒト養育者とのインタラクションをとおした発達が必要だろう．これがヒトの社会的知能への発達的アプローチ[2,15]である．これを実現するための要素技術として，物理的環境とのインタラクション機能だけでなく，たとえば図4のようなヒトとのアイコンタクトや共同注意の機能，表情や声をとおして情動を認識・表出する機能などが必要とされている．この分野のさらなる発展に期待したい．

参考文献

1) Tomasello, M.: The Cultural Origins of Human Cognition, Harvard University Press, (1999).
2) 小嶋秀樹，高田明：社会的相互行為への発達的アプローチ，人工知能学会誌，16(6)，812-818，(2001).
3) Gergely, G.: What should a robot learn from an infant? Mechanisms of action interpretation and observational learning in infancy, Connection Science, 15(4), 191-209, (2003).
4) Meltzoff, A. N. and Borton, R. W.: Intermodal matching by human neonates, Nature, 282(5737), 403-404, (1979).
5) Meltzoff, A. N. and Moore, M. K.: Imitation of facial and manual gestures by human neonates, Science, 198(4312), 75-78, (1977).

6) Scaife, M. and Bruner, J.: The capacity for joint visual attention in infant, Nature, 253(5489), 265-266, (1975).
7) Spelke, E. S.: The development of intermodal perception, In: Cohen, L. B. and Salapatek, P. (Eds.): Handbook of Infant Perception, Academic Press, (1987).
8) Baron-Cohen, S.: Is there a normal phase of synaesthesia in development?, Psyche, 2(27), (1996).
9) Rochat, P.: The Infant's World, Harvard University Press, (2001).
10) Rizzolatti, G. and Arbib, M. A.: Language within our grasp, Trends in Neuroscience, 21(5), 188-194, (1998).
11) Oztop, E. and Arbib, M. A.: Schema design and implementation of the grasp-related mirror neuron system, Biological Cybernetics, 87(2), 116-140, (2002).
12) Kozima, H. et al.: Emergence of imitation mediated by objects, In: Prince, C. et al. (Eds.): The Proceedings of the Second International Workshop on Epigenetic Robotics, 59-61, (2002).
13) Bekkering, H. et al: Imitation is goal-directed, Quaterly Journal of Experimental Psychology, 53A, 153-164, (2000).
14) Gergely, G. et al.: Rational imitation in preverbal infants, Nature, 415(6873), 755, (2002).
15) 小嶋秀樹：赤ちゃんロボットからみたコミュニケーションのなりたち，発達，24(95)，52-60，(2003).



小嶋　秀樹（こじま　ひでき）
1994 年電気通信大学大学院情報工学専攻博士課程修了．工学博士．同年通信総合研究所入所．1998 ～ 1999 年：MIT 人工知能研究所客員研究員．2000 年通信総合研究所けいはんな情報通信融合研究センター主任研究員．現在に至る．ヒトのもつコミュニケーション能力の成り立ちをロボットを道具として解明することに興味をもつ．人工知能学会，情報処理学会，日本発達心理学会等各会員．