

再考：HAL 設計論

Rethinking of Building HAL 9000

浅田 稔
Minoru Asada

大阪大学大学院工学研究科
Graduate School of Engineering, Osaka University.

Keywords: vision, SGP, frame problem, mind, language, cognitive development.

1. はじめに

今回の原稿を書くにあたり、原作(訳本[クラーク 93])を引っ張り出し、映画のビデオを見直した。人の記憶は曖昧で、また、読んだり見たりする立場や目的が異なるので、かなり異なる印象を受けた。多くの方がHALは究極の人工知能との認識があり、通常のマシンの能力をはるかに超えた存在と頭ごなしに信じきってしまった感がある。もちろんその能力はかなりのものであるが、宇宙船という閉じた空間でのヒトとのコミュニケーションは、原作を読み直したり、映画を見直す前に比べ、限られた印象を受ける。もちろん、それでもなお、その知性の高さ、意志、意図の表出、他者の意図理解は、現状の人工システムの能力をはるかに超えているが、自然界の知性のあり方は、動物行動学などでも知られているように、多種多様で、自然淘汰の結果(経過?)とはいえ、目を見張るものがある*1。ヒトの知性を自然界のトップに持ってくるのは、ヒトの傲慢だとしても、ヒトの知性、もしくは知的行動が他の動物のそれらと異なる点は多くある。二足歩行、道具使用、言語使用はその最たるものである。これらと並行して、あまりに未熟で生まれるがために、誕生後、比較的長い間、親の過保護のもとで多くを学び、発達する能力も多いように見える(ただし、どちらが原因で結果かは分からないが)、人工物としてのHALが、自然界の知性のあり方をフォローする必然性はまったくなく、人工物としての知性のあり方を追求すればいい。ただし、そのあり方を理解する(もしくは設計しようとする)ヒト側にとって、自身の知性のあり方と比較対照するのは必然であろう。

筆者はことあるごとに、HALの存在を否定してきた

[瀬名 00]。その理由は、一見、徹底した機能モジュールの実装により、あたかも超越した存在として宇宙船に君臨可能な人工システムHALが、いかにしてヒトの行動を理解し、意図をくみ取る能力を獲得できたかに疑問があるからである。これは、「心の理論」の問題[Premack 78]や、意識、意図の発生とも絡む。本稿では、筆者が漠然と抱いているHALの知能(知的能力)と、ヒトの知性のあり方を比較しながら、人工物としての知能の設計論を論じてみたい。

2. 徹底した機能モジュール

宇宙船ディスカバリー号の航行システム、人工冬眠の生体維持システム、その他宇宙飛行におけるあらゆる管理機構を制御しているように見えるHALは、その意味で徹底した機能モジュールを実装していると考えられる。特に、宇宙探査プロジェクトなどでは、出来立てはやはやの技術ではなく、熟成され、完全に動作が保証されたものが利用されているのが現状で、その意味では正しいかもしれない。境界条件が明確で最適化が容易な具体的な機能実装技術の延長線上に「人の心を読む」という機能まで可能かという自然な疑問がわいてくる。

2.1 HALは何を見ているか?

映画に出てくるHALの視覚は、撮像管タイプのカラーテレビカメラを連想させ、魚眼レンズがついていて広範囲でものを見ていることをうかがわせる(HALから見たとおぼしき映像からそのように判断)。映画で見ている限り、最初は視線は固定のように見えたが、リップリーディング時には望遠レンズを通した能動カメラ(二人の唇の動きを追っていた)で視線を動かしていた。ここで、HALが有していると思われる視覚情報処理機能を列挙してみよう。

- (1) 人の発見、追跡
- (2) 人の同定
- (3) リップリーディング
- (4) 人の行動理解

*1 例えば、ガとコウモリの餌と捕食者の関係では、ガは発見されればコウモリに比べ飛翔速度が遅いので捕らえられてしまうが、そのように察知した瞬間、ガは自身の飛行システムにノイズを与え、それまでの飛翔から自然落下にスイッチングすることでコウモリから逃れることができると言われている [McFarland 93]。

- (5) 人の表情の同定
- (6) (人の意図理解)

もちろん船内は HAL のボディ内部なので、その構造や機構は百も承知で、その上で人が存在可能な場所もおおよそわかったとし、しかも船員も既知（データベース入力済み）の仮定のもとで、以上の機能をシミュレートしてみよう。画像処理技術は、近年大いに発達した分野であるが、その多くは、画像変換、エッジ抽出、オプティカルフローの抽出など初期処理の高速ルーチンが廉価で手に入ったことである。照明条件一定のもとで、動くものの発見は、環境固定なので背景差分などを使えば、(1) はそんなに難しくなく、(2) は、DB があるし、最初に初期位置を与え、あとは追跡することで困難でないかもしれない。要は、限られた環境内で限られた人の発見と同定は、少ない資源で可能である。また、原作では、「正面には魚眼レンズがあった。レンズは船内の要所要所に配置され、HAL にとってのビジュアル入力装置となっている」ので多数の視点からの情報を統合しシステムの頑強性を保持しているのかもしれない。まさしく分散協調視覚システムである*2。

問題は (3) 以降である。ポッド中でのボーマン船長とプール副船長の会話を視覚を用いて人の唇の動きを読む（リップリーディング）ことは、至難の業であろう。しかも正面を向いていない。映画のなかでは、二人の会話中の唇の動きを横から、魚眼ではなく望遠レンズで観測したかのごとき HAL からの映像が映し出されていた。リップリーディングの現状は、話者特定でしかも正面の高分解能画像を用い、語彙が限られた範囲であろう。とても、通常の会話を唇の動きだけで読むのは困難と考えられる。これを機能モジュールで実現しようとする、ボーマン船長とプール副船長のさまざまなありとあらゆる発話映像を事前に獲得し、リップリーディングの訓練をしなければならないが、出発前にやっているとは思えない。それとも HAL には、音声と視覚情報から得られる唇の動きの対応が、あらゆる人間に対して完全にわかっているのであろうか？ これとてたやすいことではない。

さらには、人の表情、行動、意図を理解するに至っては、とても視覚情報処理に閉じた機能モジュールでは実現できそうにない。まさしくフレーム問題である。つまり、何を見ているかは、漠然と視覚情報を処理しているのではなく、目的に応じた注意の焦点が作用し、絶え間なく作動する意図の働きと密に結合しなければならない。HAL が観測しているとおぼしき映像を見た観客自身が、そこに「注意の焦点」を作用させ、そこ

から HAL の意図をくみ取ろうとした時点で、HAL にもそのような能力があるかのごとく錯覚していないか？ ハルは、一体何を見ているか？

2・2 HAL は何を聞き何を発話するか？

人工システムの聴覚・発話は、音声認識・音声合成を基本としていると考えられる。近年の音声認識技術は発達し、最近では話者特定ではあるがパソコンレベルで十分認識可能な範囲も広まりつつある。また、スピーカを使えば、PCM 録音した自然な音声を再生可能である。文庫本からは、以下が述べられていた。「プールとボーマンはあたかも人間に対するように話せるし、HAL もまた、数週間の短い電子的な幼年期に学んだ非の打ちどころのない慣用英語で答えるのである」。

原作や映画から察すると、

- (1) 船内のシステムの状況に対する質問応答。

宇宙船のあらゆる面の制御を遂行しているので、これらのシステムの動作状態に対する応答は、限られた語彙で十分なので可能であろう。

- (2) ボーマン船長とプール副船長との日常会話。

ELIZA を拡張すれば、この程度の会話は可能であろうか？ それでも、なおかつ知的と思える会話を実現しているので、かなり強力な事例 DB を有しているかもしれない。

- (3) 他者の意図を読むより、だますための発話を実現しているように見える。発話能力というよりヒトの行動理解、意図理解能力に関連する。実は、この点が HAL の一番大きな能力と考えられる。

3. ヒトの行動、意図を理解すること

HAL に対して抱く一番大きな疑念は、彼（男性の声の持ち主なので）が、あたかもヒトの行動を理解し、予測し、だまそうとしているように見える点である。これを機能モジュールとして捉えれば、ヒトのさまざまな意図の種類、行動の種類をすべて列挙し、それぞれに対して、HAL の意図に応じた行動を生成することに対応する。これは、いわゆるフレーム問題そのもので、列挙のしようがない。この問題の解決は、互いに関連する二つの観点から、困難と思える。

3・1 記号接地問題？

最初の問題は、記号接地問題（Symbol Grounding Problem）[Harmad 90] に関連する。SGP は、AI の分野の基本問題であるが、シンボルを獲得するエージェント側からすれば、記号創発問題（Symbol Emergence Problem）である。自然言語理解の問題は、最近の学会誌の AI マップ [郡司 00] にあったように、シンボルの意味をどのように獲得するかにあると思われる。

*2 <http://vision.kuee.kyoto-u.ac.jp/CDVPRJ/index-jp.html> 参照。

「健は千香を自分の車に乗せた。」と「健は千香を自分の三輪車に乗せた。」で、前者は「健」を指し、後者は「千香」を指すことの推論が非常に困難であることの指摘があった。これは、当然の話で、最初に「シンボルありき」とするからである。最初にあるのは、物理的実体とそれとの相互作用であり、そのような物理的意味を内包しない限り、問題は無限後退する。つまり、問題は、物理的実体を伴ったアイコンック、インデクシャルの連合（これは動物でも可能）から、シンボル学習に至るメカニズムを解明しない限り、解決されない [Deacon 98] (邦訳 [Deacon 99])。つまり、シンボルはそれ自身で作用するのではなく、アイコンック、インデクシャルの連合を下位の構造として持って、初めてシンボルと呼ぶに値する。まさしく言語のミッシングリンクを解く鍵である。

アイコンック、インデクシャルの連合は、感覚・運動などの末梢系と前頭前野を結ぶ領野に重きがあり大脳後頭部の連合野が重要な役割を果たす。これに対し、シンボル操作は、それらのバックアップのもとに推論、計画する前頭前野と小脳が働くと考えられている。これらの働きを際立たせる症例が、自閉症とウィリアムズ症候群だと Deacon [Deacon 98] は指摘している。自閉症では、前頭前野と小脳が萎縮し、その代わり大脳後頭部の連合野が肥大している傾向がある。これが、人との社会的交流を嫌い、自身のスケジュールを厳守し、数学パズルなどが得意などの傾向を生み出す要因と考えられている。それに対し、ウィリアムズ症候群では、逆に大脳後頭部の連合野が萎縮し、語の意味を実世界に結びつける能力が劣っている。ハイパーソシアルな傾向にあり、終始笑みを浮かべ、饒舌である。それゆえ、幼年期は天才児かと間違われるが、成長するにつれ、言葉の使い方が奇妙で意味が通じないことが多く、周りがそのことに気づくそうである。これは、先のアイコンック/インデクシャルの連合が切れた状態で、発する言葉に意味がない。

アイロニカルであるが、ウィリアムズ症候群が現状のAIに対応し、自閉症が現状のロボティクスに対応しているように見える。では、人工システムがシンボルを獲得する手だては何であろうか？ 少なくとも、アイコンック/インデクシャルレベルの連合に関する実体験が近道ではなかろうか？ そのためには、物理的実体と相互作用するための身体が不可欠である。中島氏も指摘しているように [中島 01]、このとき不完全性、不確実性、資源（まさに身体！）制約が知的行動発現の源と考えられる。理想的には、実環境での体験がエッセンシャルであるが、コンピュータシミュレーションが本質的な情報を維持している限り、実環境での若干の補正で十分な場合もある [浅田 99]。多感覚多自由度ロボットを対象とした場合は、何が本質的な情報が未知の場合、シミュレーションそのものが現実的で

なくなる場合もあり得る。HALは、どのような体験をしてきたのだろうか？ 「HALもまた、数週間の短い電子的な幼年期に学んだ非の打ちどころのない慣用英語で答えるのである」ということが可能かどうか？

3・2 身体性再考：理解すること

身体を持つことが知能の発現に重要であることは、現状では大いに認識され、筆者もその要件を示した [Asada 99a]。身体性の本質は、物理的実体として限られた処理能力の範囲で、主体の「多様な感覚情報」、 「多自由度の運動能力」に基づき、タスクや環境の複雑さに適応するために学習およびその経時的発展を可能にすることである。この意味では、時間的发展、すなわち発達の結果が、知的行動を生み出すことを期待するが、保証はない。つまり内部構造をいくら作り込んでも、良い方向に変化（発達）すべき環境が与えられないかぎり、知能が発現するとはいえない。この意味で外部環境の設計は重要である。このことは、身体性の本質が陽にはロボットの内部構造に関する要件を述べているように見えながら、同時に提示されるべき外部環境の条件を陰に示していることに注意しなければならない。特に、知的行動をヒトのレベルまで求めるなら、ヒト以外の動物にも可能な連合学習から、ヒト特有のシンボル生成/利用の記号学習、すなわち言語獲得に至る過程（言語創発）が、ロボットの内部構造と外部環境の多様かつ制約的相互作用のなかに見いだされなければならない。

さらに他者の行動を理解すること、その行動の意図を理解することについて考えよう。最近、サル運動前野のF5と呼ばれる領域にミラーニューロンと呼ばれるニューロンが発見された。これは、サルが自分で「ものをつまむ」、「紙をちぎる」などに該当する行為を遂行するときに、それぞれに対応するミラーニューロンが発火する。さらに、同じニューロンがサルが観測する他者が同じ行為をしたときも発火する。Rizzolatti and Arbib [Rizzolatti 98] は、同様の部位にヒトの場合、発話に深く関連するブロッカ野があることから、行動の模倣の能力や行動の理解能力がコミュニケーションスキルの発達に役だっており、最終的には、マインドリーディングにつながるのではと推測している。すなわち、模倣を契機として、

- (1) 観測による模倣学習により行動レパートリーを作成する。
- (2) 他者の行為を観測したときに、自己の行動レパートリーと照合することで、他者の行動が何かを知る。
- (3) すなわち、自身に再生する能力があること（レパートリーを持っていること）により、他者の行動が理解できる。これが理解の本質。
- (4) 他者の行動が理解できることにより、他者の行

動の意図がわかり、他者の行動予測が可能となる。というストーリーが描ける。ヒトがヒトの行為を理解するとは、他者が自身と同じ身体構造をもっている仮定のもとに成り立ち、これは、コミュニケーションの本質でもある。ヒトがHALと会話できるのは、ヒト側が、勝手にHALの全知全能性を信じるがゆえに、HALが自分の意図を理解できると錯覚しているのではなかろうか？ HALの身体が宇宙船全体であり、HALがそのような身体しか持たないということは、ヒトの行動を理解できないのではないかと考えてしまう。HALがヒトと同じ身体構造を持ち、ヒトとの相互作用を実験することが必要条件と筆者は考えている。この意味で「認知発達ロボティクス」[Asada]の重要性を問いたい。すなわち、環境との相互作用を通じて学習・発達可能な内部構造の設計とともに、同じ重要度で埋め込まれた能力がうまく開花するような環境設計が大事である。環境設計の中に他者の存在があり、子供の言語獲得に見られるお母さんの役割が重要である。このように考えると、HALのお母さん役は誰だったか？

4. おわりに

ヒトの知能発現に関連する知見をもとに、HALの身体構造に起因する知的行動発現の困難さを列挙し、HALの設計が難しいことを説いた。先にも述べたように、HALの知能設計はヒトのそれをフォローする必然性はない。そもそもヒトの知能設計ですら明らかでないので、人工システムでの実現がヒトの知性モデルになる可能性すらある。本稿では、ネガティブな意見を述べたが、これもそれを乗り越える人工システムの出現を期待するからである

◇ 参考文献 ◇

[Asada] Minoru Asada, Karl F. Mac Dorman, Hiroshi Ishiguro, and Yasuo Kuniyoshi: Cognitive developmental robotics as a new paradigm for the design of humanoid robots. In *Proc. of*

the 1st IEEE/RSJ International Conference on Humanoid Robots.

- [Asada 99a] Minoru Asada, Eiji Uchibe, and Koh Hosoda: Cooperative behavior acquisition for mobile robots in dynamically changing real worlds via vision-based reinforcement learning and development, *Artificial Intelligence*, Vol. 110, pp. 275-292, 1999.
- [浅田 99b] 浅田 稔: ロボカップによる認知ロボティクス, 日本ロボット学会誌, Vol. 17, No. 1, pp. 34-37, 1999.
- [Deacon 98] Terrence W. Deacon: *The Symbolic Species: The co-evolution of language and the brain*, W. W. Norton & Company, New York, London, 1998.
- [Deacon 99] T. W. Deacon 著, 金子隆芳 (訳): ヒトはいかにして人となったか—言語と脳の共進化—, 新曜社, 1999.
- [郡司 00] 郡司, 西田, 梶, 田中: AI マップ「言語理解—SHRDLUの先にあるもの—」へのコメントと回答, 人工知能学会誌, Vol. 15, No. 6, pp. 1035-1042, 2000.
- [Harnad 90] S. Harnad: The symbol grounding problem, *Physica D*, Vol. 42, pp. 335-346, 1990.
- [クラーク 93] アーサー・C.クラーク 著, 伊藤典夫 訳: 2001年宇宙の旅—決定版—, 早川書房, 1993.
- [McFarland 93] D. McFarland and T. Bösser: *Intelligent Behavior in Animals and Robots*. MIT Press, 1993.
- [中島 01] 中島秀之: HALの謀反, 人工知能学会誌, Vol. 16, No. 1, pp. 82-85, 2001.
- [Premack 78] D. Premack and G. Woodruff: Does the chimpanzee have a theory of mind?, *Behavioral and Brain Sciences*, pp. 515-526, 1978.
- [Rizzolatti 98] G. Rizzolatti and M. A. Arbib: Language within our grasp. *Trends Neuroscience*, Vol. 21, pp. 188-194, 1998.
- [瀬名 00] 瀬名秀明: ロボット開発最前線をゆく—鉄腕アトムをつくれ—, 文芸春秋, pp. 354-385, 2000.

2000年11月21日 受理

著者紹介



浅田 稔 (正会員)

1982年大阪大学大学院基礎工学研究科後期課程修了。同年、大阪大学基礎工学部助手。1989年大阪大学工学部助教授。1995年同教授。1997年大阪大学大学院工学研究科知能・機能創成工学専攻教授となり現在に至る。この間、1986年から1年間米国メリーランド大学客員研究員。知能ロボットの研究に従事。1989年情報処理学会研究賞、1992年IEEE/RSJ IROS '92 Best Paper Award。1996年日本ロボット学会論文賞受賞。博士(工学)。日本ロボット学会、電子情報通信学会、情報処理学会、日本機械学会、計測自動制御学会、システム制御情報学会、IEEE R&A, CS, SMC Societiesなどの各会員。