

ヒューマノイドはクオリアを夢見るか？

茂木健一郎

〒141-0022 東京都品川区東五反田3-14-13 高輪ミュージズビル
ソニーコンピュータサイエンス研究所

要旨

近年の脳科学における「システム論的転回」により、ヒューマノイド・ロボットなどの構成論的アプローチによる仮説検証を必要性が増している。クオリアと志向性という二つの心的表象の要素を通して、環境にダイナミックに適応する脳のシステム論的な性質を把握できる。ボディ・イメージや言語などを支える脳のシステム論的な機構の理解が検証が、今後の重要な課題である。

Do Humanoids dream of Qualia?

Ken Mogi

Sony Computer Science Laboratories
Takanawa Muse Bldg. 3-14-13, Higashigotanda,
Shinagawa-ku, Tokyo, 141-0022 Japan
kenmogi@csl.sony.co.jp

The "systematic turn" in brain science in recent years has made it necessary to test a hypothesis about brain function through constructive approaches such as humanoid robotics. I present a paradigm for understanding the systems level properties of the brain through the dynamic interaction of neural mechanisms underlying qualia and intentionality. Binocular rivalry, language, and body images are the discussed as specific instances of this paradigm.

1、脳科学のシステム論的転回

脳科学は、現在、「システム論的転回」とでも言うべき転換期を迎えている。近年、脳のニューロンの活動は、感覚や運動の特定のモダリティの中の情報処理だけでは理解できないということを示すデータが相次いで報告されている。視覚の中核である第一次視覚野は、触覚による空間把握の際にも必要不可欠な活動を担っていることが明らかにされた

(Zangaladze et al. 1999)。自分で自分をくすぐることはできないが、これは、小脳からの自己運動の情報が、皮膚感覚のシグナルをキャンセルするからだということが示された (Wolpert et al. 1998)。一部では分子生物学におけるDNAの発見に比肩すると評価される「ミラーニューロン」の発見は、従来の意味での感覚と運動の両方にまたがる、さらには「自己」や「他者」といったより大きな枠組みの情報処理に関わる神経回路網の存在を明らかにしている (Gallse & Goldman 1998)。

これらの発見がなくても、そもそも脳の情報処理メカニズムを、システム論的にとらえる必要があるということは、自明のことであった。しかし、HubelとWieselの第一次視覚野ニューロンの反応特性（いわゆる反応選択性）の発見 (Hubel & Wiese 1962) 以来、脳科学は、どちらかと言えば、各モダリティ内部の情報処理を分離して扱うことにより発展してきた。このようなやり方ではうまく行かないという認識が、次第に広がりつつある。

脳機能をシステム論的に理解する上で、所謂構成論的なアプローチは極めて重要である。工学においては、あるシステムを理解したということは、そのシステムを実際に組んでみて動くことを確認することである。千億個のニューロンが数千のシナプスを通して結合しあう、天文学的な複雑さを持った脳を人工的に再構成するのは容易ではない。しか

し、脳機能の理解においてシステム論的な視点が不可欠であることが明らかになりつつある今、ヒューマノイド・ロボットなどのシステムを実際に組んでみて、脳の情報処理に関するモデルを検証すること (浅田1999、銅谷1999) は、必要不可欠なステップであるといえることができるだろう。

2、クオリアと志向性の機能

赤い色の質感 (redness of red) などのクオリア (qualia) は、心脳問題の核心にある (茂木1997)。すなわち、物質的過程である脳の中のニューロンの活動から、いかにしてクオリアが生まれるのかが、心と脳を結び付ける際の「困難な問題」 (hard problem) だとされるのである (Chalmers 1996)。

クオリアと並ぶ、心的表象の重要な要素が、「志向性」 (intentionality) である。志向性とは、簡単に言えば、私たちの心が、何かに向き合っている状態である。志向性は、Franz Brentano (1874-1895)によって我々の心の持つ固有の性質とされた。クオリアと志向性という現象論的な性質を通して心的表象を解析するのが、現代的な文脈での心脳問題の標準モデルである (信原1999)。

クオリアも志向性も、脳の中のニューロンの活動によって生じていることは疑いない (Barlow 1972)。本質的な問題は、いかにしてニューロンの活動の時空間パターンに特定のクオリアが重生起 (supervene) するかということである (Mogi 1999b)。

我々の認識は、クオリアと志向性のマッチングが生じることによって成立している (茂木1999a)。このうち、クオリアは、後頭部にある第一次視覚野を中心とする神経回路網によって生まれ、一方、志向性は、高次視覚野から第一次視覚野へ向かう神経回路網によって生じる。クオリアと志向性は、主観的には異なる感覚を伴って表象される (図1)。例えば、カニツアの三角形などの錯

視図形の場合、錯視によって生じる形の知覚は、錯視を引き起こすクオリア（カニッツアの三角形においては、3つの黒いパックマン図形）にくらべると、より「抽象的な」感覚を伴う志向性として成立している。通常の三角形の認識においては、三角形を構成するクオリアと志向性の間に整合的な関係が生じるが、所謂錯視図形の場合には、クオリアと志向性が整合しないのである。

錯視図形と所謂正常な視覚認識の間には厳密な境界はなく、どちらも、私たちの脳が自らの内部の世界モデルに基づいて、能動的に外界のイメージを再構築するプロセスを反映していると考えべきである（下條 1999）。

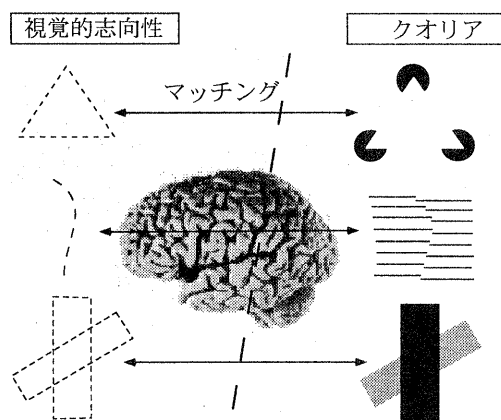


図1 クオリアと志向性

3、クオリアと志向性のダイナミックなマッチング

私たちは、普段、両眼を通して入力する環境からの情報に基づいて、安定した外界の立体的イメージを構築している。安定した両眼

視が成立するのは、右目と左目からの入力の間に関連があるからである。縦縞と横縞のように、二つの目から入る刺激の間の相関が低く、両者を整合的な立体像に融合できない時は、視野の中に右目が勝つ領域と左目が勝つ領域（眼優位性（ocular dominance）の分布）が入り交じり、時間とともに変化する、いわゆる両眼視野闘争（binocular rivalry）と呼ばれる現象が起こる。Low pass filterのcut off frequencyを調節して左目、右目からの信号を通し、両眼からの入力の中の相関を調節すると、安定した両眼視と両眼視野闘争の間の相転位を見ることができる（Tamori & Mogi 1999）。

両眼視野闘争は、末端の感覚器からのボトムアップの神経回路網によって生み出されるクオリアと、視覚的志向性の相互作用の結果生じる現象である。神経生理学的な知見によって、クオリアを中心的に生み出している第一次視覚野のニューロンの活動レベル自体は、両眼視野闘争における見えの変化と連動せず、連動するのは、ニューロンの発火の同期であると考えられている（Fries et al. 1997）。そして、眼優位性の変化に伴うニューロンの活動の変化は、クオリアを直接コードしていない高次の視覚野のニューロンに見られることが分かっている（Lumer et al. 1998）。これらの知見は、クオリアを生み出す第一次視覚野を中心とするニューロン活動の同期が高次視覚野における志向性を生み出し、クオリアと志向性のマッチングを通して初めて視覚的アウェアネス（Crick and Koch）に上る視覚認識が生じるというモデルと整合性がある。すなわち、クオリアと志向性のマッチングがとれることが、我々が「見る」ための必要条件であり、両眼視野闘争において観察されるのは、このマッチングのダイナミックな変化であるという仮説が立てられるのである。

両眼視野闘争においては、外界の変化にダイナミックに適応する視覚系の能力が観察さ

れる。田谷と茂木は、左目と右目にそれぞれ異なる均一の色を背景に動く円を呈示した場合、両眼視野闘争の結果生じる視覚的アウェアネスの中では、動いている円が両方とも常に見えることを見出した (Taya and Mogi 2000a, 2000b)。このような見えを生じさせるためには、右目、左目で動いている円の周囲の領域でそれぞれ右目、左目のイメージが優先されるように、眼優位性の空間分布がダイナミックに調節されなければならない。

実際、このことは、動く円の近傍の方が、その円が載っている背景の色 (ipsilateral color) が見える確率が高くなることで確認できる (図2)。この「近傍効果」は、静止した円では低減する。

動いている円は、環境において顕著な (salient) 情報である。両眼視野闘争の結果生じる視覚的アウェアネスは、いわば「心のデスクトップ」であり、その後の注意や運動といったモジュールの活動は、視覚的アウェアネスに呈示された視覚特徴を基に組み立てられる。脳は、クオリアと志向性のダイナミックな相互作用を通して、重要な情報を高率的に「心のデスクトップ」に呈示する、動的適応性 (dynamical adaptability) を持っているのである。

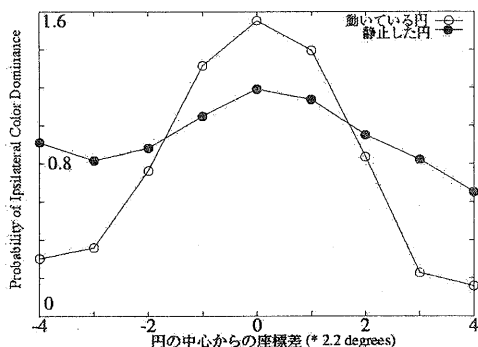


図2 円が載っている背景色を見る確率

4、言語

人間の認知プロセスにおいて、クオリアと志向性の相互作用のダイナミックスが最もリッチに表れるのが言語である (茂木 1999b)。

言語理解のプロセスは、音声的及び視覚的クオリアに、言語のコンテキスト、意味を担う志向性が張り付けられるプロセスである。視覚的クオリアに誘導されて作られた志向性が、音声的クオリアの認識に影響を与えるプロセスも見い出されている (McGurk & MacDonald 1976)。

言語認識のプロセスは、脳全体に広がるクオリアと志向性の相互作用のシステムの一つの表れであると考えられる (図3)。

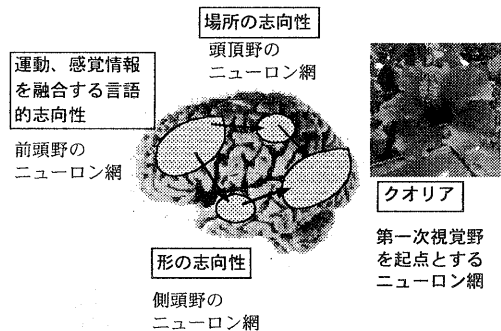


図3 クオリアと志向性のネットワーク

5、ボディ・イメージ

ヒューマノイド・ロボットを作る際に、最も重要なテーマの一つは、ボディ・イメージの神経コードをどのように理解し、実装するかということである。

新宿駅の雑踏の中を歩く時、私たちは他人にぶつからないようにうまく避けて歩くこと

ができる。このようなことができるのも、私たちの脳が、手や足などの身体の部分が空間的にどこにあるかを、リアルタイムにアップデートしているからである。

熊手を使って餌をとる訓練を受けた猿の体性感覚野においては、今迄手のボディ・イメージをコードしていたニューロンが、熊手のイメージもコードするようになる (Iriki et al. 1996)。しかし、この変化は、熊手を使わないようになると、急速に元に戻ってしまう。このように、ボディ・イメージをコードしているニューロンは、身体と環境の相互作用の変化を反映して、ダイナミックにその活動を变化させる能力を持つ。

幻肢 (phantom limb) のような、よりドラマティックなボディ・イメージの改変の例 (Ramachandran & Ramachandran 2000) を見ると、ボディ・イメージが、固定されたデカルト座標のどの部分集合を身体が占めているかというような形で表現されているのではなく、むしろ、感覚と運動が複雑に融合した、一種の「エージェント」の集まりから、デカルト座標に類似の空間イメージが逆に出てくると考えるべきだと思われる。

私は、ボディ・イメージの神経コードを理解する鍵は、両眼視野闘争におけるクオリアと志向性のダイナミックな相互作用と共通点を持つ、体性感覚野における運動の情報と感覚情報のダイナミックな融合における、体性感覚のクオリアとそれを組織化する志向性の相互作用にあると考えている。視覚認識において、視野という空間的構造をつくり出しているのは志向性であり、その秩序化に基づいてクオリアが配置される。似たようなことが、ボディ・イメージの構成においては、よりダイナミックに起こっていると考えられるのである。

6、ヒューマノイド はクオリアを夢見るか？

ヒューマノイドは、文字どおり人間そっくりの存在である。だとすれば、私たち人間が、いかに「人間そっくり」ということを定義するかによって、ヒューマノイド・ロボットの意味合いは変わってくる。サッカーのような、認知プロセスとしても行動のフレキシビリティとしてもタスクをテスト・ベッドとして、「人間そっくり」を定義するのも一つの手である (浅田1999)。一方、人間の本質を考える時に、喜びや哀しみといった感情を含む心的表象は無視できない。しかし、このアプローチを真っ当に追求しようとするれば、心が脳にいかにして宿るかという深刻な心脳問題に直面せざるを得ない。心的表象は、あくまでも第一人称の視点からのみアクセス可能である。我々にできることは、両眼視野闘争に表れているような視覚におけるクオリアと志向性の相互作用を通しての動的適応性の理解と実装を通して、あるいはボディ・イメージなどの、感覚と運動の双方にまたがる志向性のダイナミックスの理解と実装を通して、脳の構成論的理解と検証を少しずつ進めていくことだろう。

どうやらロボットが心を持っているらしいということになれば、様々な倫理的な問題が浮上してくる。しかし、その時には、我々は他の人間に対して適用している倫理規範を適用することを選択するしかないだろう。そもそも、他人の心の存在は、推定されるものであって、直接確かめられるものではないのだから。それに、もしヒューマノイドがクオリアを夢見ることが原理的に可能だとしても、それが実現するのはかなり先の話になりそうである。

References

- Barlow. H. B. (1972) Single units, sensation: a neuron doctrine for perceptual psychology? *Perception* 1, 371-394.

- Chalmers, D. (1996) *The Conscious Mind* Oxford University Press
- Jay Schulkin (2000) Theory of mind and mirroring neurons. *Trends in Cognitive Sciences*, 4, 252-254
- Crick, F. and Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature* 375: 121-123.
- Fries P, Roelfsema PR, Engel AK, Konig P, Singer W (1997) Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. *Proc Natl Acad Sci U S A* 94(23):12699-7
- Gallese V. & Goldman A. (1998) Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences* 2, 493-500
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746-748.
- Hubel D.H., Wiesel T.N. (1962) Receptive fields, binocular interaction and functional architecture of cat's visual cortex. *J Physiol Lond* 160:106-154.
- Iriki A, Tanaka M, Iwamura Y (1996) Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport* 1996 Oct 2;7(14):2325-30
- Lumer, E. D., Friston, K. J., & Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science*, 280, 1930-1934.
- Mogi, K. (1999a) Response Selectivity, Neuron Doctrine, and Mach's Principle. in *Understanding Representation in the Cognitive Sciences*. Plenum Press. 127-134.
- Mogi, K. (1999b) Supervenience and qualia. Comment on Gold, I. & Stoljar, D. A neuron doctrine in the philosophy of neuroscience. *Behavioural and Brain Sciences*. 22(5).p84-85
- Mogi, K & Barlow, H B. (1999) The source of variability in neural responses from MT. *J.Physiol* 515, pp. 101-102
- Ramachandran, V. S. & Rogers-Ramachandran, D. (2000) Phantom Limbs and Neural Plasticity. *Arch Neurol*. 57:317-320
- Taya, F. and Mogi, K. (2000a) What is salient in binocular rivalry? *Consciousness and Cognition* 9, pp.53.
- Taya, F. and Mogi, K. (2000b) Dynamical Adaptability of the visual system in binocular rivalry. Submitted.
- Tamori, Y. & Mogi, K. (1999) Maximisation of information in binocular rivalry. *Perception*. 28 supplement 36.
- Wolpert, D. M., Frith, C. D. & Blakemore, S.-J. (1998) Central cancellation of self-produced tickle sensation. *Nature Neuroscience* Vol. 1 pp 635- 640.
- Zangaladze, A., Epstein, C.M., Grafton, S.T. & Sathian, K. (1999) Involvement of visual cortex in tactile discrimination of orientation. *Nature* 401, 587 - 590.
- 浅田稔 (1999) 「ロボカップによる認知ロボティックス」 *日本ロボット学会誌* Vol.17 pp.34~37.
- 銅谷賢治 (1999) 「脳科学とロボット」 *日本ロボット学会誌* Vol.17 pp.7~10.
- 信原幸弘 (1999) 「心の現代哲学」 勁草書房
- 下條信輔 (1999) 〈意識〉とは何だろうか 脳の来歴、知覚の錯誤 講談社
- 茂木健一郎(1997) *脳とクオリア* 日経サイエンス社
- 茂木健一郎(1999a) *心が脳を感じる時* 講談社
- 茂木健一郎 (1999b) 「言語の物理的基盤 — 表象の精密科学へ向けて」 「言語」 (大修館書店) Vol.28 No.12 p.49-57