

# 身体性による知能の発現

## Emergence of Intelligence Based on Physical Embodiment

浅田 稔, 大阪大学大学院研究科

Minoru Asada

Graduate School of Engineering, Osaka University

### 1 はじめに

知的人工システムの構築に「身体」が必要であることは、最近、人工知能や計算機科学の研究者に認知されつつある。「身体」を持つことによる外環境との相互作用が可能になり、これが「知能」の発現につながるのではという期待である。一方、多くのロボット研究者にとって、ロボットの身体の必要性は、当たり前すぎて議論にならないことをよく経験する。実際の環境の中でロボットを作動させるのだから、身体は必要であるというわけだ。しかしながら、ロボット研究者側から、なぜ、どのような意味で身体が知能の発現に関わり得るのかについては、これまであまり深い議論があまりなされてこなかったように思う。

本稿では、身体性が知能の発現に、いかに本質的であるかを説明することを試みる。筆者の研究目標は、身体を有する物理的実体(ロボット)が環境、特に他者の存在によって、行動の複雑さを露呈する基本的内部構造の構築であり、結果として現れる行動の「観察者」の立場からの形容が「創発的」であったり、「知的」であったりすると考えられる [1]。このことは、「創発」や「知能」を直接的かつ明示的に構築する事の矛盾や困難さを表している。知的行動を発現するためには、システムと物理的環境(複雑で、不確実だが、統一のとれた自然の拘束集合を含む)とが「意味のある」相互作用をするために物理的身体が必要である。

身体性の本質は、主体の「多様な感覚情報」、「多自由度の運動能力」を、物理的実体として限られた処理能力の範囲で、タスクを遂行するために「抽象化」することであり、その結果として「知的」振舞いが可能になると考えられる。環境構造の複雑さ、特に「他者の存在」が、抽象化のレベル向上を必然的に伴うことを、視覚を例に説明を試みる。以下では、まず、身体性の意味を説明し、身体と環境との相互作用から生まれる内部記述の複雑さについて議論する。そして、工学的手法の一つとして、状態ベクトルの推定によ

り環境内のエージェントの複雑さを規定し、それにより強化学習を通じて、結果として協調行動を獲得する例を説明する。この例から、今後の課題を述べ、まとめる。

### 2 身体性の意味

身体性の意味は、以下のようにまとめられる。

1. 環境の多様な構造や自身の状態を知覚できる能力と環境に働きかける多様な運動能力は不可分であり、密に結合していること。
2. 限られた資源(メモリなど)や処理能力の範囲でタスク(究極のタスクは、存続すること)を遂行するために、知覚・運動空間を抽象化すること。
3. 抽象化は、身体への基本埋め込みと経験(環境との相互作用)に依存するので、抽象結果は、身体を有する主体中心の主観的表現であり、その評価は行動結果でなされる。
4. 実世界では、エージェント間、エージェントと環境の間の相互作用は、並列、非同期であり、充分複雑である。エージェントは与えられたタスクを遂行するために増加する環境の複雑さに対応しなければならない。

身体性の意味をより明確にし、重要さを強調するためには、ロボティクス研究の新たな側面で、この重要さを発揮できる具体的なタスク領域が必要である。すなわち、チェスが人工知能の標準問題としての役割を果たしたように、実世界における知的行動の様々な側面を捉えていると考えられる新たな標準問題が必要である。

そのようなタスク領域として、ロボットによるサッカー競技、ロボカップ [2] をとりあげる。ロボカップは、広範な最新技術が集積され、試される標準問題を提供することにより、人工知能とロボティクスのあらたな研究課題を生み出す試みである。

筆者の研究グループは、既に、(1,2,3)に関連した手法を発表している。それらは、ボールをゴールにシュートするタスクに対して、自分の知覚空間(ここでは視覚)と運動空間を経験を通じて抽象化し、状態・行動空間を形成している。一つは、オフラインの学習であり [3]、もう一方はオンライン学習である [4]。

さらに、敵や味方などの他のエージェントが存在する環境では、相互作用の複雑さが一挙に増す。学習者が視覚情報を通して環境を観測するとき、敵や味方などの他のエージェントも環境の一部として観測され、区別がない。そこで、環境の複雑さを定義することにより識別の可能性がどうか探る。

### 3 他者の存在と「知能」

知能の本質に関する議論として、社会性、他者の存在、模倣などがキーワードとして挙げられている [5, 6]。特に大澤は、自己を否定し得る可能性のあるものやことを「他者性」とし、未来性も含めており、時間との関連性を印象づける。離人症の例 [7] から、自己の存在を時間と直接関連づける木村 [8] の説との関連性もあると考える。更に模倣に関連づけて、言語の獲得と身体運動、特に母親との接触との関連など正高 [9] の実験などは、霊長類を「知的」と形容させている「2足歩行」、「言語」、「道具の利用」[10] との関連も含め面白い。

知能のレベルが、環境との相互作用の複雑性にあると考えると、視覚を基にした環境との相互作用をベースとする我々の研究でも、他者の存在が、複雑さを増すであろうことや、自己の規定などに関する議論が可能と考えられる。

1. **自己の身体**：運動機能を有する主体が、運動指令と視覚情報との直接的な相関がとれる場合(例えば、他の多くの人たちと手のひらを重ねあわせた状況を観測していて、自分の手を発見するとき、自分の手をちょっと動かすだけで分かる。)、すなわち、運動指令を發した自己の運動結果が視覚で確認され、大澤の言を借りれば、運動指令を出した自己を裏切らない範囲が自己と規定し得る範囲であろう。
2. **静止環境**：自己の運動との直接的な相関により、切り出し可能(視線を変えると、静止環境に対するオブティカルフローは、逆方向に流れたりする。)。切り出せれば制御可能。全ての運動は相対的なので、「静止環境」と「自己の身体」との区別は容易ではないが、重力方向などによる他の参照座標系との関係から、識別可能となる。
3. **受動エージェント**：自己や他者の運動の帰結により、静

止環境内で運動したり、停止したりするもの。自らは状態変化を起こさないもの。ロボカップでは、ボールが相当する。状態空間の自律的構成 [11] では、静止環境(ゴール)と受動主体(ボール)を同時に含む空間なので、静止環境のみを主体とするゴール探索より、複雑度が高いと考える。

4. **他の能動エージェント**：直接的には、自己の運動との単純な相関を持たない。大澤の意味では、自己を否定する可能性のあるもの。視覚情報からの相関では、単純な相関を持たないものとして、最初、排他的に扱われるが、自己の運動との高度な相関(協調、競合、干渉)により、規定される対象。複雑度は一挙に増す。行動選択の複雑さは一挙に増す。

アフォーダンス [12] は、環境から直接的にロボットのロボスタな行動を引き出す上での基本原理として魅力的であり、熟練、熟達といった、個々のエージェントの腕の磨き上げの説明にも有効に見える。但し、設計原理から観た時には、他のエージェントの存在による環境の複雑さにどう対応するかまでは教えてくれない。もちろん、不変項を抽出する過程で、抽象化を必然的に伴うので、他者という概念がでることも期待できる。状態空間分割でも、分割された状態がラベルを表し、行動とペアになり世界の分節化ルールを築けば、そのルール(通信規約になりうるか?)に基づいて、コミュニケーションが成り立たないかと考える。但し、これまでの単純な作業環境は、能動的な他者の存在しない環境での世界の分節化であり、他者が存在した場合の分節化はより複雑で高度になる。

以上の議論を具体化すべく研究を行っているが、以下では、エージェントの識別及び行動学習に対する具体的手法として、状態空間の推定に基づく他者の発見とその対応に対する手法について説明する。

### 4 複数ロボット環境下での行動学習

ロボットが自律的に目的行動を獲得する手法として、強化学習法が注目されている。しかし、マルチエージェント環境では、学習者の行動が、必ずしも自分自身の観測と1対1には対応しないため、通常の強化学習をそのまま適用することは困難である。学習を成功させるためには、学習者は他者の行動を自分自身の観測と行動を通して予測できる必要がある。そこで、学習者の観測と行動を通して、学習者と他者の行動の関係を局所予測モデルとして推定し、その結果をもとに強化学習をおこなうことを考える。これを簡単な1対1のサッカーゲームに適用し、その有効性を確認する。詳細は、直接文献 [13] を当たって頂くとして、以

下に簡単に概要を述べる。

### 4.1 アーキテクチャ

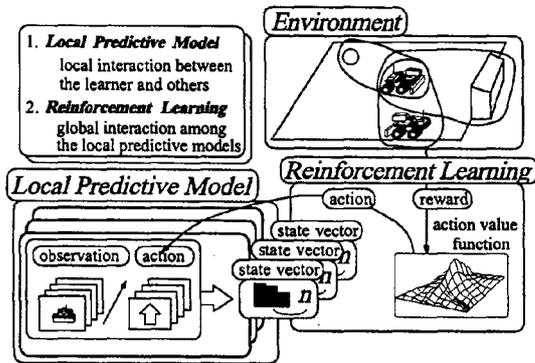


図 1: 提案するアーキテクチャ

図 1 は各ロボットに与えられる行動獲得のためのアーキテクチャである。はじめに、学習者はセンサ情報だけでなく、学習者自身の行動のシーケンスも考慮して局所予測モデルを構築する。ここでのポイントは、観測のみによって対象のモデルを推定するのではなく、自分の行動との関わりを通じて、対象の動きを予測することである。この結果、自分との関わり(味方(協調), 敵(競争), 審判(干渉?))が判別可能となる。次に推定された状態ベクトルをもとに、協調行動獲得のための学習を開始する。

### 4.2 局所予測モデル

局所予測モデルは、多入力(行動)多出力(観測)の関係を記述する必要がある。状態表現の方法として、システム同定の一つである正準変量解析(Canonical Variate Analysis) [14]を用いて、局所予測モデルを構築する(詳細は文献参照)。CVAは離散時間で線形の状態空間モデル

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t), \end{aligned} \quad (1)$$

を用いる。ここで  $\mathbf{u}(t) \in \mathbb{R}^m$  と  $\mathbf{y}(t) \in \mathbb{R}^q$  はそれぞれロボットの行動ベクトルと観測ベクトルであり、 $\mathbf{x}(t) \in \mathbb{R}^n$  は状態ベクトルである。また、 $\mathbf{A} \in \mathbb{R}^{n \times n}$ ,  $\mathbf{B} \in \mathbb{R}^{n \times m}$ ,  $\mathbf{C} \in \mathbb{R}^{q \times n}$ ,  $\mathbf{D} \in \mathbb{R}^{q \times m}$  はパラメータ行列である。学習者は観測と行動のシーケンス  $\{\mathbf{y}, \mathbf{u}\}$  から状態ベクトルを次数を含めて推定しなければならない。状態ベクトル  $\mathbf{x}$  は過去の観測と行動のシーケンスの線形和

$$\mathbf{x}(t) = [\mathbf{I}_n \ \mathbf{0}] \mathbf{U}\mathbf{p}(t), \quad (2)$$

表 1: The estimated dimension

observer	target	$l$	$n$	$\log  \mathbf{R} $	$AIC$
computer simulation					
shooter	ball	2	4	0.23	138
	goal	1	2	-0.01	121
	passer	3	6	1.22	210
passer	ball	2	4	0.78	142
	shooter	3	5	0.85	198
real experiments					
shooter	ball	4	4	1.88	284
	goal	1	3	-1.73	-817
	passer	5	4	3.43	329
passer	ball	4	4	1.36	173
	shooter	5	4	2.17	284

によって状態を表現する。ここで、

$$\mathbf{p}(t) = [\mathbf{u}(t-1) \cdots \mathbf{u}(t-l) \ \mathbf{y}(t-1) \cdots \mathbf{y}(t-l)]^T,$$

であり、 $\mathbf{U} \in \mathbb{R}^{(m+q) \times l(m+q)}$  はCVAによって計算される行列であり、 $l$ は考慮する履歴長さである。また  $n$  は状態ベクトルの次数であり、情報量規準によって決定する。

### 4.3 タスク

提案手法を、2台のロボットが存在する環境下での、簡単なサッカーゲームに適用した(図2(a)参照)。各ロボットはTVカメラを一つ搭載し、そこから得られる画像情報から環境の状況を観測する。モータコマンドとして、各ロボットは2自由度を持つ。そこで、ロボットへの制御入力  $\mathbf{u}$  は2次元ベクトル

$$\mathbf{u}^T = [v \ \phi], \quad v, \phi \in \{-1, 0, 1\},$$

として表現する。ここで、 $v$ は台車の移動速度であり、 $\phi$ はステアリングの角度である。また、各ロボットが観測できる画像特徴量(観測ベクトル)は、例えば、領域の位置(重心や4隅の座標)、大きさ(面積)などであり、冗長に与えれば、十分である。

### 4.4 実験結果

最初にシューターとパスナーは、ボール、ゴール、そして互いの局所予測モデルを、計算機のシミュレーション上で構築する。次に、シューターを静止させた状況下で、パッ

サーは行動の学習を実機で開始する。パスサーの学習が終了した時点で、パスサーの行動政策を固定し、シューターの学習を開始する。パスサーは、ボールをシューターにパスしたときに報酬 1 を受け取り、シューターはボールをゴールにシュートしたときに、報酬 1 を受け取る。さらに、ロボット間で衝突が発生した場合、 $-0.3$  の報酬が与えられる。

表 1 は計算機シミュレーションおよび実環境での推定された状態ベクトルである。ここで、 $R$  は局所予測モデルの誤差の共分散行列であり、 $AIC$  は赤池の情報量である。これらから、

1. ゴールの次の状況を予測するための履歴長さ  $l = 1$  であり、静止環境であることをしめす。
2. ボールの履歴長さは  $l > 1$  であり、シミュレーションでは速度的なもの、また実機ではノイズなどの影響により加速度的なものを含んでいると考えられる。パスサーとシューターで次数は同じであるが、中身は異なり、交換してもうまく動作しない。このことは、ハードウェアは同じであるにも関わらず、タスクと経験により異なった内部表現を構築しており、表現の主観性が窺える。
3. 能動主体もノイズなどの影響により、互いに履歴長は実機のほうが大きい。また、明らかにボールより複雑な相関を持っていると考えられる。

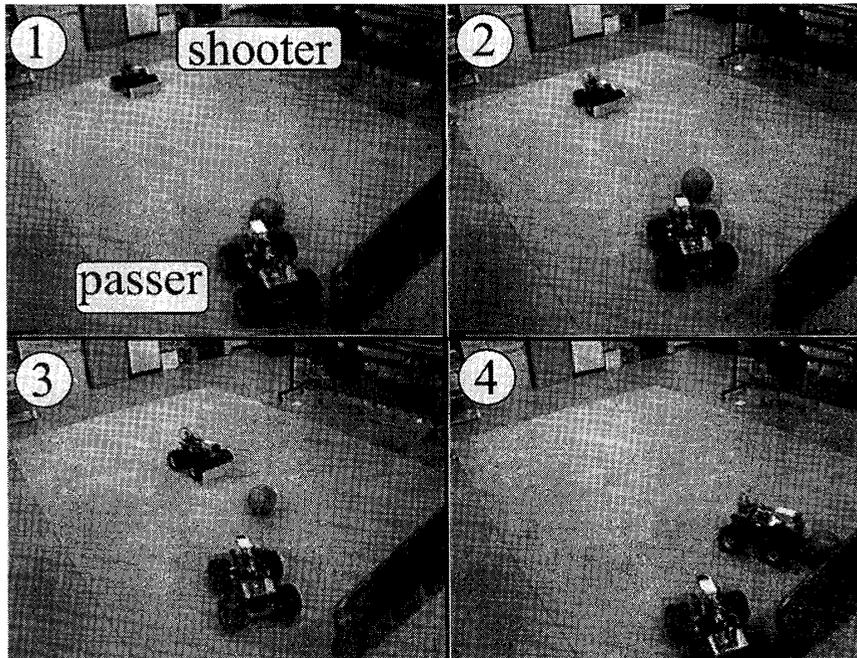
図 2 に獲得された行動の例を示す。まず、パスサーがボールをシューターに向かってボールを蹴り、シューターはボールをゴールにシュートする。パスサーはボールを蹴った後は、シューターとの衝突を回避するための行動をしていることがわかる。

## 5 おわりに

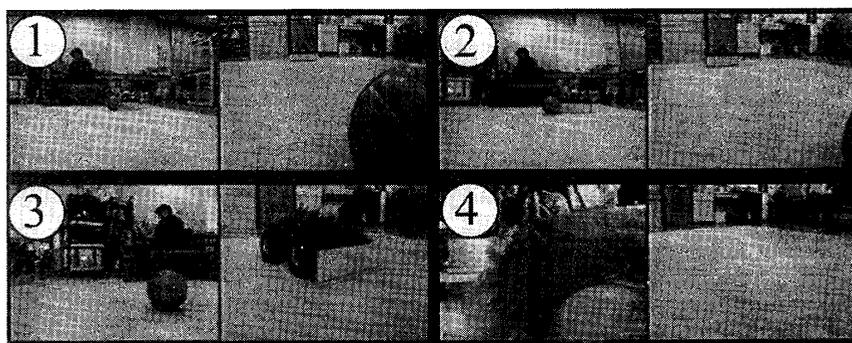
欺瞞が自己とか自意識と関連することが、動物行動学 [10] などでも論議されているので、学習、発達過程で「フェイント」できるサッカーロボットができないものかと考えている。このためには、相手の複雑な行動を十分予測可能な状態ベクトルが推定されただけでなく、行為系列の分節化によるより複雑な内部表現 (より高度な抽象化) が必要であり、今後の研究課題である。

## 参考文献

- [1] D. McFarland and T. Bösser. *Intelligent Behavior in Animals and Robots*. MIT Press, 1993.
- [2] H. Kitano, M. Asada, Y. Kuniyoshi, I. Noda, E. Osawa, and H. Matsuura. "robocup: A challenge problem of ai". *AI Magazine*, 18:73-85, 1997.
- [3] M. Asada, S. Noda, and K. Hosoda. Action-based sensor space categorization for robot learning. In *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems 1996 (IROS '96)*, pages 1502-1509, 1996.
- [4] Y. Takahashi, M. Asada, and K. Hosoda. Reasonable performance in less learning time by real robot based on incremental state space segmentation. In *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems 1996 (IROS'96)*, pages 1518-1524, 1996.
- [5] 大澤. 「知能の社会性」. 日本ロボット学会誌, 14(4):15-19, 1996.
- [6] 國吉. 「実世界エージェントにおける注意と視点-情報の分節・統合・共有-」. 人工知能学会誌, 10(4):507-514, 1995.
- [7] 中島. 「手足を持った人工知能」. 日本ロボット学会誌, 14(4):12-14, 1996.
- [8] 木村敏. 「時間と自己」. 中公新書, 1982.
- [9] 正高. 「身体運動は言語獲得にどのような役割を果たすか」. 日本ロボット学会誌, 14(4):31-34, 1996.
- [10] リーキー著 and 馬場悠男訳. 「ヒトはいつから人間になったか」. 草思社, 1996.
- [11] M. Asada, S. Noda, and K. Hosoda. Non-physical intervention in robot learning based on lfe method. In *Proc. of Machine Learning Conferen Workshop on Learning from Examples vs. Programming by Demonstration*, pages 25-31, 1995.
- [12] 佐々木. 「オーガニゼーション・オブザブト」. 日本ロボット学会誌, 14(4):26-30, 1996.
- [13] E. Uchibe, M. Asada, and K. Hosoda. "state space construction for behavior acquisition in multi agent environments with vision and action". In *Proc. of ICCV 98*, pages 870-875, 1998.
- [14] W. E. Larimore. Canonical variate analysis in identification, filtering, and adaptive control. In *Proc. 29th IEEE Conference on Decision and Control*, pages 596-604, Honolulu, Hawaii, December 1990.



(a) 鳥瞰図



(b) ロボットからの映像

図 2: 獲得された行動