

# 人工痛覚が導く ロボットの共感

浅田 稔 あさだ みのる

大阪国際工科専門職大学, 大阪大学

ヒトの共感能力は、他の種に比して強いようである。それは、人間の社会的動物としての証かもしれない。高度に発達した人工物との共生社会において、人がロボットに共感能力を期待するならば、ロボットの共感能力をどのように設計できるだろうか？ 本稿では、痛みの感覚が共感能力を生み出す源と考え、ミラーニューロンシステムの発達を通じた行動・知覚(痛覚)の共有から、人工物が共感能力を生み出す道筋を描いてみる。

## 1 痛みの共感から共感能力一般の発達へ

ロボットなどの人工システムが人間と共生する社会において、ロボットが共感能力を有することは、人間へのさまざまな支援に役立つと考えられるが、その実装は容易ではない<sup>1</sup>。痛覚はヒトを始めとする生物が生存していくうえで非常に重要な感覚であり、この感覚があることで情動感染を始めとする共感が進化したことが明らかにされつつある<sup>2</sup>。筆者は人工共感の実現には、ロボットが痛覚を有することが有効ではないかと考えている。すなわち、痛みの感覚を経験し、それを他者と共有することで共感能力が発達するのではないかと想定する。これらの過程は、筆者らが長年提唱・推進してきた認知発達ロボティクス<sup>3</sup>の流れを汲むものであり、以下では、人工痛覚を前提とした作業仮説をたて、どのように実現可能かを議論する。共感能力の基礎となるミラーニューロンシステム発達の一例を示したうえで、運動のみでなく知覚のミラーリングも観測されていることから、知覚の一つに痛覚を導入することで、痛みの共感から、より一般的な共感能力の発達につながられ

ると期待できることを述べ、そのための課題と現状を示す。最後に今後の課題を示す。

## 2 心的機能創発の要としての痛覚神経回路

人工物がヒトの心のような機能を実現するには痛覚が重要であり、この痛みを共有しているという感覚が共感のもととなっていると考えられると述べた。筆者の人工共感のサーベイ<sup>1</sup>で紹介した神経科学、認知科学、心理学の多くの文献が痛みを題材にしていることもそのことを示している。よって、ロボットの神経系に触覚系とは別の痛覚回路を埋め込むことは、生物進化の観点からも、過度に人工的なバイアスにはならないと考えられる。ヒトの痛覚信号の経路を図1に示す。刺激が与えられ、それを受容し(①)、求心性(末端から脳へ)経路を通じて(②)、脊髄で統合され(③)、遠心性(脳から末端へ)経路を通じて(④)、効果器で行動を起こす(⑤)。また、脊髄から脳へも上行する(⑥)過程である。人工痛覚の信号経路もこれに倣う形態をとる。

痛み感覚の共有は、ミラーニューロンシステム<sup>5</sup>(以下、MNSと略記)により可能になると考えられる。MNSとは、ある行動を起こすときに発火するだけでなく、他者の同じ行動を観測したときにも発火するミラーニューロンからなるシステムであり、自己と他者の行動の同一性の理解につながると言われている。さらに、行動のみならず、知覚についても同様に自己と他者の知覚の同一性を理解するシステムでもある。MNSにより、自己認知を通じた他者の行動を無意識に模倣したり知覚したりする傾向がある。我々の作業仮説は文献

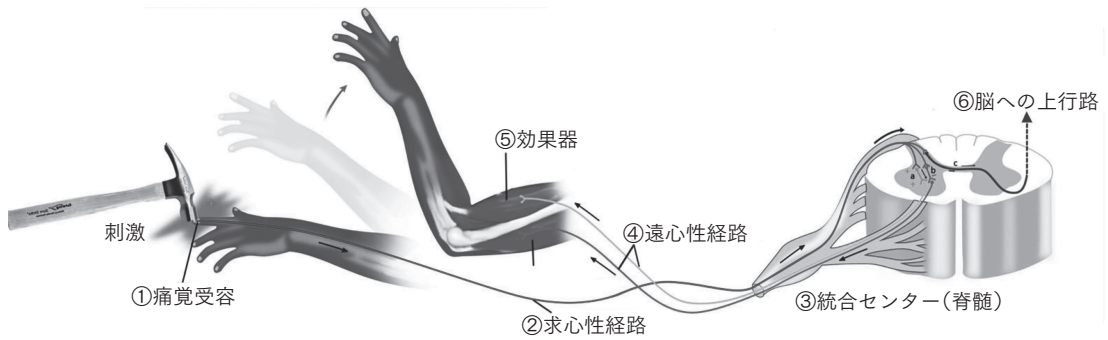


図1—痛覚信号の経路(文献4のFig. 2aを改変)

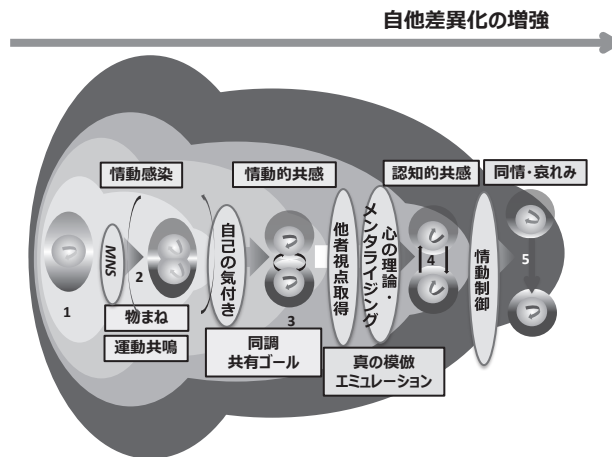


図2—共感の発達(文献1のFig. 6を改変)

6を一部修正した、以下の4段階である。

1. 痛覚神経回路を有するロボットが痛みを感じる。
2. 痛覚を始めとする多様な感覚運動体験から主体感、運動所有感などの初期自己が形成される。
3. MNSの発達を通じて、ロボットが他者の痛みを感じる(人工共感の研究がベース)。
4. 情動感染、情動的共感、認知的共感、同情、哀れみの感情をロボットが発達させる。

### 3 人工痛覚による共感発達

共感の発達に関して、筆者のサーベイ<sup>1)</sup>で示した図2に沿って述べる。最初に人工痛覚の第1段階としてのセンサーを紹介し(図中の1に相当)、次にMNSの発達について説明する(同2)。MNSの構

造が明確になれば、他者視点取得やメンタライジングなどの心の理論にもとづく他者の行動や意図の理解につながる(同3)。認知的共感では、自己認知を通じて、理解の内容の論理的な構造(「Aさんが事故にあっているので助けなければいけない」)を把握する必要があり、説明能力としての言語機能が必要と想定される。ここでは、Bengioの意識プレイヤー<sup>2)</sup>を例にその可能性を議論する(同4)。さらに情動制御を経て同情や哀れみの行動学習が必要であるが、自身の痛み回避行動学習が先行すると想定される。ここでは、Seymourの強化学習の枠組みを紹介する<sup>3)</sup>。

#### 3.1 人工痛覚

図1に従って、以下では、人工痛覚の例を実現レベルに応じて概説する。

1. 痛覚刺激を受容するソフトセンサー：人工

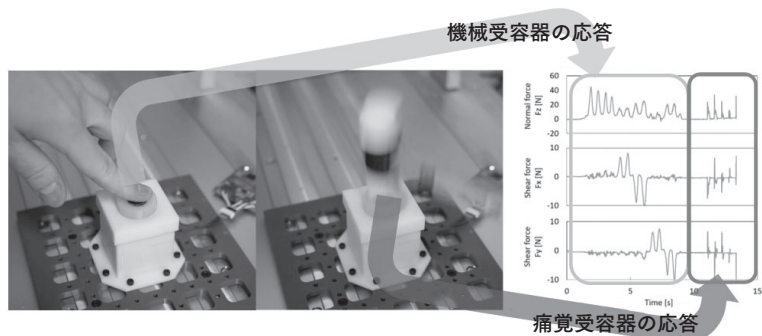


図3—開発した磁性エラストマとスパイラルコイルを用いた柔軟触覚センサーの応答なでる場合(左)と強く叩く場合(中央)の3軸の力の応答波形(右)

痛覚実装の予備段階として、我々は磁性エラストマとスパイラルコイルを用いた柔軟触覚センサーを開発してきた<sup>9</sup>。メカニズムの詳細は省くが、柔軟かつ頑強であることが特徴である。図3にその様子を示そう。図の左は優しくなでる様子で、中央はハンマーで強く打ち込んでいる様子である。右にそのときの3軸の力の応答波形を示している。優しくなでている場合は通常の機械受容器が応答するのに対し、ハンマーで打ち込んだ場合は、痛覚受容器が応答すると期待され、痛覚受容器の応答はシャープな波形から容易に識別可能である。このことから、識別に関連する経路は最初からの(生得的な)埋め込み回路として実装可能と考えられる(図1の①と②に相当)。

2. スパイクニューロンによる痛覚信号処理と回避行動：Kuehn & Haddadin<sup>10</sup>は、ロボットアームに人工痛覚を埋め込み、痛みを回避する反射的な運動生成を可能にした。彼らの意図は、人間との協調作業における安全性の確保の考え方をロボットにも適用したものと考えられる(図1の①から⑤に相当)。
3. ロボットに痛みを自覚させ自己修復させる小型脳：John et al.<sup>4</sup>は、圧力から生じる「痛み」を処理して反応するエッジデバイスを開発した。このシステムと自己修復性のあるイオンゲル材料の組み合わせによる自己修復能力の可能性に言及している(図1

の①のレベルだが、痛覚信号に対する処理をエッジで行い、さらに⑤の回避行動に相当する自己修復を実現しており、より生物的である)。

これらは、主に痛みのセンサー信号の識別経路に相当するものであり、痛みの記憶などの情動的一動機付けの経路(図1の⑥以降)は実現されていない。後者の経路の学習には、痛みの記憶などが含まれる。これには予測による痛みの軽減行動(図1の⑤)や、逆に記憶再生による痛みの増幅などが考えられ、大きな課題である。これらと並行して、MNSの発達共感への道筋と想定される。

### 3.2 MNSの発達

MNSの発達に関しては、いくつか提案されているが<sup>11, 12</sup>、相手の立場になって観測される情報を取得できる能力である他者視点取得は、明示的な座標変換を介さずに行うことが困難な課題である。全方位画像などの特殊な視覚情報ではなく、通常の透視変換の画像を用いたロボットの行動学習でMNS的な働きをする手法が提案されているので、紹介する。

Secker et al.<sup>13</sup>は、深層感覚混合ネットワーク(deep modality blending networks, 以下、DMBNと略記)と呼ばれる手法を提案している。図4に、その構成例を示す。最初に、把持、押出、囲い込みなどの各種タスクに対して、視覚、関節角の情報を個別に処理して、それぞれの特徴を抽出し(I)、その特徴空間を構成する(深層学習による符号化)(II)。そして、それらを確率的重み付けによって混合した内

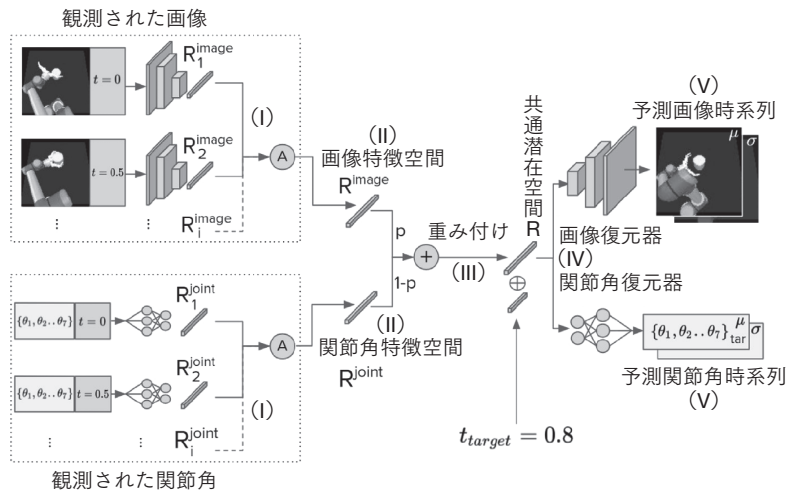


図4—ロボットアームのさまざまなタスク達成用のDMBNの例(文献13のFig. 2を改変)

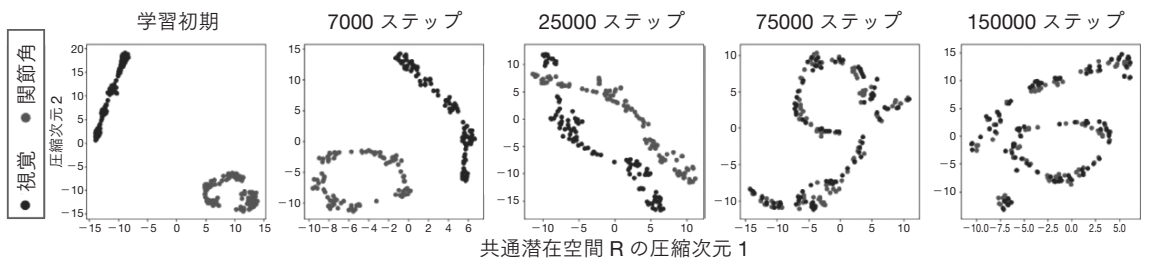


図5—共通潜在空間の遷移の可視化

濃色と淡色はそれぞれ視覚と関節角の符号化(文献13のFig. 7を改変)

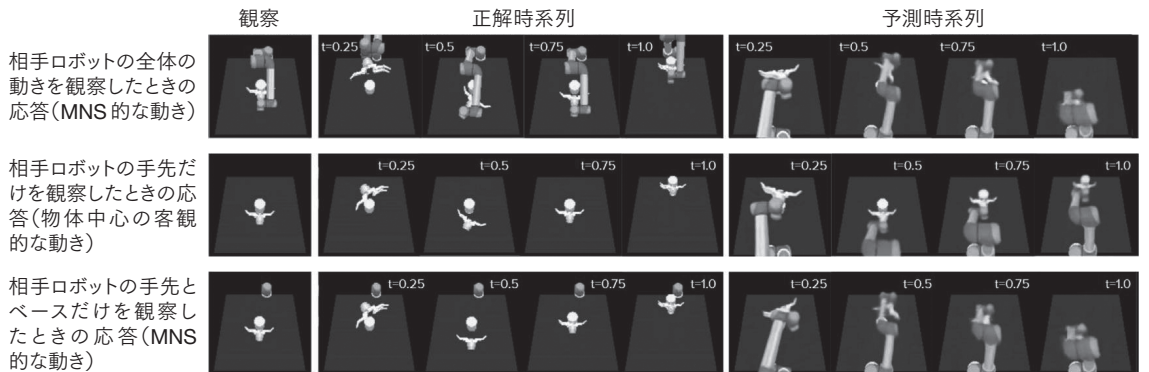


図6—他者運動観測時の予測：MNS的な動き生成(文献13のFig. 9を改変)

部表象として多種感覚の共通潜在空間 R を作成する(III)。これらの過程で作成される共通潜在空間 R の遷移の可視化を図5に示す。学習当初かけ離れていた視覚・関節角の符号化が学習後期で共通化されていく様子が窺える。これにより、視覚(もしくは、関節角、以下同)の表象がこの空間 R で

得られれば、同時に関節角(視覚)の表象が得られ、これを解釈する(元のデータに復元)(IV)ことで、関節角(視覚)が復元され、それ以降の振る舞い(視覚および関節角の時系列)を予測する(V)ことができる。

具体例を示そう。図6の中段では、手先が物体を掴んでいる視覚情報が与えられた(左端)とき、



システムは、右端 4 枚の時系列の写真の予測をしている。実際、中央 4 枚の写真は、物体を図の上方にプッシュしている正解時系列であり、正しく予測している。このとき、対象物体の動きは、環境に紐付けされた物体中心の運動生成(客観的生成)である。それに対し、図 6 の上下段では、予測された動きは、他者運動観測時に、自己にマップした運動生成(主観的生成)であり、まさしく MNS 的な動きを示している。どのような状況に対して、MNS 的な動きか、客観的な動作かについて、まだ明確にはなっておらず、共通潜在空間 R における視覚と関節角への重みやパラメータ依存と察せられる。

### 3.3 共感の発達

共通潜在空間の構造が明確になれば、すなわち、MNS 的な動きと、客観的な動きの差異を明示化できれば、感覚・知覚についても同様の明示化が可能かもしれない。すなわち、MNS の機能には、運動の観察から運動指令を想起するだけでなく、感覚の観察から感覚を想起することも含まれる。DMBN の例では、関節角情報の代わりに、触覚・痛覚情報(触れるまではゼロシグナルで、触れて以降は触れたセンセーション)を視覚と同時学習することで、可能になるかもしれない。ヒトの場合、touchingsight<sup>15</sup>はそのよい例で、他者が触られている状況を見るだけで、触られているという知覚が想起される。このように、より一般的には、MNS

は運動の結果や感覚そのものの状態予測に関連している<sup>16</sup>。

図 7 に痛み経験を共有するモデルを示す。自己と他者の情動状態の部分共通潜在空間 R に対応する。これは、図 2 に示す共感発達モデルの 1 から 4 あたりまでに対応する。すなわち、MNS を起点として、情動感染による自己の気づき、情動的共感、他者視点取得・心の理論の獲得を経て、認知的共感に至る過程である。ただし、認知的共感では、自他認知などの論理的構造が必要となり、言語表現が含まれる。昨今、深層学習による自然言語理解と生成には目を見張るものがあるが(例えば、AI 白書 2022<sup>17</sup>の 2.4.6 の「言語」参照)、感覚や運動の身体性のレベルからの言語生成にはやや遠い。痛覚と直接関係しないが、感覚や運動のレベルから、結果として言語獲得に至るアーキテクチャーとして、Bengio の意識プライヤー<sup>7</sup>を例にその可能性を議論する。我々が言語で操作する種類の高レベルの概念の表現を学習するためのアーキテクチャーとしての意識プライヤーである。

図 8 にその概要を示す。時刻  $t$  における観測を  $x_t$ 、 $x_t$  から引き出される高次元(低位)表現を  $h_t$  とすると、 $h_t = F(x_t, h_{t-1})$  と表され、 $F, h_t$  は、それぞれ、表現 RNN(エンコーダー)および無意識表現状態と呼ばれる。 $h_t$  から注意選択された要素は低次元(高位)の意識状態  $c_t$  であり、ある一瞬に人が意識できる状態を表す。 $c_t = C(h_t, c_{t-1}, m_{t-1}, z_t)$  と表され、 $m_t, z_t$  は、それぞれ時刻  $t$  の記憶内容および

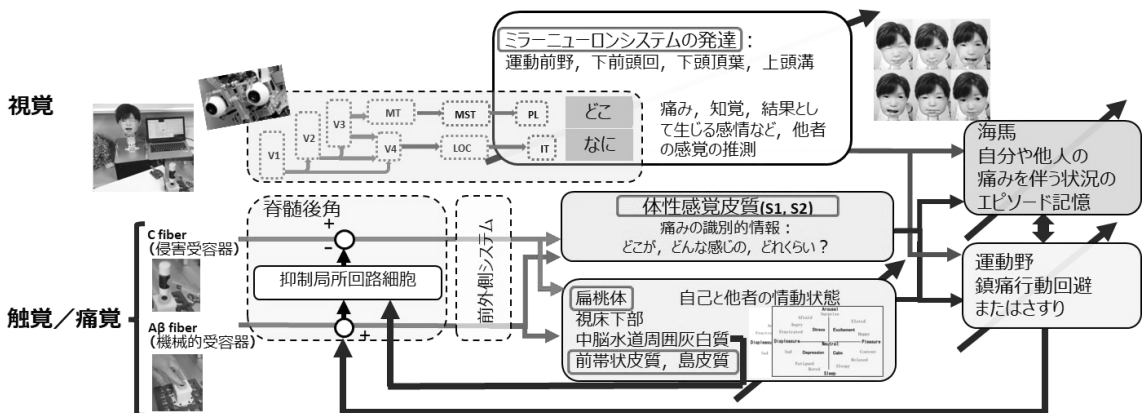


図 7—痛み経験の共有のモデル(文献 14 の図 10. 7)

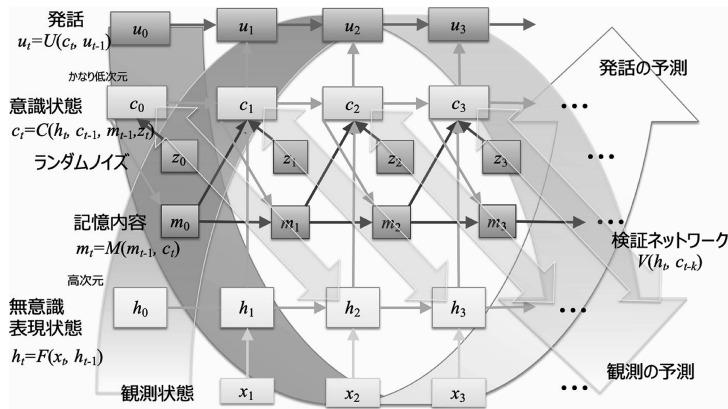


図8—意識プレイヤー概要図(文献18の説明図にもとづく)

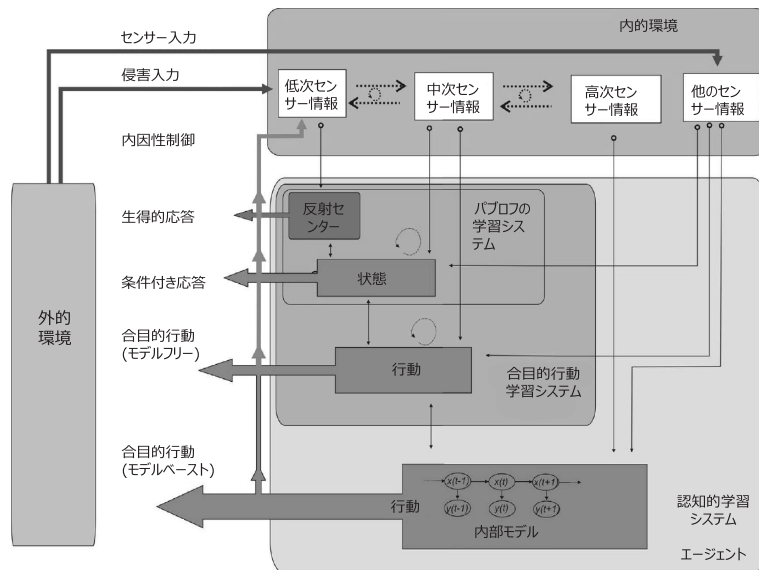


図9—痛み強化学習モデル(文献8のFig. 2を改変)

びランダムノイズで、 $m_t = M(m_{t-1}, c_t)$ である。ランダムノイズは探索モードで多様な予測や計画立案に有効であり、記憶内容の集積はエージェントの履歴であると同時に、未来予測にも重要な要素である。詳細は、AI白書2022<sup>17</sup>の2.4.5「意識」を参照されたい。

意識プレイヤーは、現状の深層学習(DL 1.0)を拡張してシステム2のレベルに至るDL 2.0の中核をなすものと見なせる。そのためには、注意スキーマだけでなく、身体スキーマも組み込むことが肝要である。さらに、他者の概念を獲得することで、他者との関係を醸し出す社会ルールの獲得

も期待したい。これは、同情・哀れみの行動表出に重要な課題である。これらが一連の学習のみで可能とは思えず、発達の視点が重要で、Baby AI<sup>19</sup>の考え方に通じる。

### 3.4 痛みに伴う行動の学習

Kuehn & Haddadin<sup>10</sup>は、反射的な回避行動生成が主眼であったが、Seymour<sup>8</sup>は、痛みの中核的な機能は、短期的・長期的な行動を害から遠ざける動機付けであると主張し、この点に関する脳の働きを機械的に理解するために、痛みシステムの計算機アーキテクチャーを基盤とする痛みの強

化学習モデルを提案した。そして、痛みは、適切な行動を引き起こすための正確かつ客観的な制御信号であると考えた。図9にそのアーキテクチャーを示す。階層構造によって急性痛から慢性疼痛までの時間スケールを表すとともに、それぞれのレベルでの痛み表現を制御信号として行動の強化学習の枠組みを提案しているが、実装はされていない。

#### 4 おわりに

人工痛覚からロボットの共感能力の創出に向けた、関連研究を紹介した。それぞれの研究のつながりが、まだまだ希薄であり、強い連携が望まれる。認知発達ロボティクスも同様で、発達自体を直接表象可能な構造にしなければならない。今後の課題も含めてまとめると以下のとおりである。

1. 痛みの種類、場所、強度などの情報を伝える神経回路は、通常の触覚との区別がセンサーレベルでできれば、原理的に実装可能で体性感覚野が信号を受け取り活性化するので、身体イメージや身体スキーマなどの実装でまかなえそうである。
2. MNS 発達システムとしての DMBN を拡張し、痛みを始めとする各種感覚体験から、その知覚を共有するシステムが必要であり、3.3 で述べた方法が使えるとよい。
3. 行動の理解から意図推定に至る過程のように、相手の痛みの理解から、行動生成する学習の枠組みが必要で、痛みの予測とその回避行動の計算には、自由エネルギー原理が使えるかもしれない。
4. 言語の課題は重く、一筋縄ではいかないが、意識プレイヤーの進展を期待するとともに、最新 AI 手法の言語ツールの利用も考えられる。
5. 神経科学的な最近の知見なども参考に、より発達的な視点でのシステム作りも課題である。

謝辞 本研究の一部は JST 戦略的創造研究推進事業 (CREST)「脳領域/個体/集団間のインタラクション創発原理の解明と適用」(平成 29~令和 4 年度, 研究代表者: 津田一郎)の支援によって行われた。本研究の参画メンバーおよび研究室メンバーに感謝する。

#### 文献

- 1—M. Asada: *International Journal of Social Robotics*, **7**, 19 (2015)
- 2—F. de Waal: *The Age of Empathy*, Three Rivers Press(2009)
- 3—M. Asada et al.: *IEEE Transactions on Autonomous Mental Development*, **1**(1), 12(2009)
- 4—R. A. John et al.: *Nature Communications*, **11**, 4030(2020)
- 5—ジャコモ・リゾラッティ, コラド・シニガリア(柴田裕之訳/茂木健一郎監修): *ミラーニューロン*. 紀伊國屋書店(2009)
- 6—M. Asada: *Philosophies*, **4**, 38(2019)
- 7—Y. Bengio: *The consciousness prior*. arXiv: 1709.08568 (2017)
- 8—B. Seymour: *Neuron*, **101**(6), 1029(2019)
- 9—T. Kawasetsu et al.: *IEEE Sensors Journal*, **18**(4), 5834 (2018)
- 10—J. Kuehn & S. Haddadin: *IEEE Robotics and Automation Letters*, **2**(1), 72(2016)
- 11—Y. Nagai et al.: In *Proceedings of the 2011 IEEE International Conference on Development and Learning (ICDL)*, CD-ROM(2011)
- 12—J. L. Copete et al.: In *Proceedings of the 2016 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, CD-ROM(2016)
- 13—M. Y. Seker et al.: *Neural Networks*, **146**, 22(2022)
- 14—浅田稔: *浅田稔の AI 研究道*. 近代科学社(2020)
- 15—C. Keysers et al.: *Neuron*, **42**, 335(2004)
- 16—K. Friston: *Nature Reviews Neuroscience*, **11**, 127(2010)
- 17—AI 白書編集委員会(編): *AI 白書 2022*. KADOKAWA(2022)
- 18—松尾豊: *表現・言語・意識*. 学生セッション講演, 人工知能学会 2018(2018)
- 19—M. Chevalier-Boisvert et al.: *BabyAI: First steps towards grounded language learning with a human in the loop*. In *International Conference on Learning Representations (ICLR)*(2019)