

OS-18

感情とAI

Emotion and AI

日永田 智絵

大阪大学

Chie Hieida

Osaka University.

hieida@otri.osaka-u.ac.jp, <https://www.hieida.com/>

堀井 隆斗

大阪大学, 東京大学国際高等研究所ニューロインテリジェンス国際研究機構

Takato Horii

Osaka University. / International Research Center for Neurointelligence, The University of Tokyo.

takato@sys.es.osaka-u.ac.jp, <https://takatohorii.github.io/>

長井 隆行

大阪大学, 電気通信大学人工知能先端研究センター

Takayuki Nagai

Osaka University. / Artificial Intelligence eXploration Research Center, The University of Electro-Communications.

nagai@sys.es.osaka-u.ac.jp, <http://www.rlg.sys.es.osaka-u.ac.jp/>

Keywords: emotion, body, interoception, predictive coding, artificial intelligence, machine learning.

1. はじめに

画像認識分野での進展を皮切りに深層学習が目覚ましい成果を上げたことで、人工知能（AI）が人の知能を超えるのではないかという考えや期待が広がっている。このようなAI研究において、感情は知能とは異なる機能や存在として、人や生物と人工物を分かつものとして捉えられてきた。一方で認知科学や神経科学の分野では、感情は知能（いわゆる論理的な情報処理）と背反するものではなく、その一部であると捉えられており、感情に関わる知能では、身体的・社会的な情報を扱うことが重要だと考えられている。このような知能に関する考え方の中で、感情を理解できず感情をもたないAIが真に知能を獲得し、そして人の知能を超えることがあり得のだろうか？

本オーガナイズドセッション（OS）「感情とAI」は、こうした著者らの問題意識のもと、2019年の人工知能学会全国大会で開催した新設OSである。招待講演1件、一般講演8件で構成され、一般講演では発表件数上限を超える多数の原稿が投稿された。聴講者は100人を超え、立ち見も出るほどであり、会場に入れない人も多数いた（図1）。このことは、AI研究者の感情研究に対する興味や期待を端的に表している。

本OSは、AI研究における感情へのアプローチを通して、感情の本質やメカニズムについての議論を活性化することを目的の中心に据えた。そして、心理学や神経科学、認知科学など異なる分野における感情に関する知



図1 OSの様子

見を取り込み、構成的アプローチによる感情研究を促進することで、これまで曖昧なものだと思われてきた感情の本質を明らかにするための足掛かりをつくりたい。

初開催となる本年は、名古屋大学大学院情報学研究科心理・認知科学専攻の大平英樹先生を招待講演者としてお招きし、「内受容感覚の予測符号化と感情の創発」と題してご講演いただいた。またOS当日は、会場内の情報共有とOSを（時間的にも空間的にも）超えた議論の実施を目的とし、Sli.doを導入した。セッション中の限られた時間だけでなく、Web上でも活発な議論が進んだことから、多くの参加者に興味をもっていただけるOSになったと考えている。

本稿では、まずAI分野におけるこれまでの感情研究について簡単に述べる。また一般講演について言及した後、大平先生の講演内容を概説する。最後に感情とAIに関する議論を展開し、まとめと今後の展望を述べる。

2. 人工知能分野における感情研究

AIにおけるこれまでの感情研究は、ロボットやエージェントによる感情の表出 [Breazeal 02, Kishi 13] や、Affective Computing [Picard 00] に代表される人の感情(表情)認識が主であった。つまり感情それ自体に迫るのではなく、コミュニケーションツールとしての側面を捉えたものであるといえる。実際、感情表出の研究では、人工物は感情と呼べるものをもたず、あくまでも行動として感情があるように振る舞うための設計論が議論の中心である。また感情認識研究では、認識精度に限界があったり、認識すべき感情カテゴリーの文化差 [Gendron 14] なども問題となる。これらは、感情を行動出力の側面のみで捉えたアプローチの限界を表している。

感情がもつ一部の機能だけではなく、感情そのものを人工物に与えようとした研究も存在する。例えば尾形らは、人の内分泌系における複数のホルモンの挙動を参考に、ロボットの内分泌系を考えた。そして、自己保存の観点からホルモンの挙動をモデル化し、それらの相互作用によって変動するロボットの内部状態を感情として捉えた [Ogata 00]。こうした、センサ値や他者との相互作用により変化する感情をもとに、ロボットの情緒的な行動生成を試みている。また Lim らは、ガウス混合モデルを用いて複数の感覚情報に共通する感情的特徴量を表現可能な感情モデルを構築している [Lim 15]。この感情モデルは外部刺激から自身の感情を変容させるとともに、ミラーシステムとして他者の感情を推定することも可能である。実験では、人-ロボット相互作用場面において自己と他者の感情状態を考慮した音声感情コミュニケーションを実現した。共感コミュニケーション、つまりは自分や他者を考慮した感情の本質を捉えるために、浅田は共感発達ロボティクスの枠組みを提唱している [Asada 15]。

著者らは、感情の発生や発達・分化に注目した研究を進めている。堀井らは、視覚・聴覚・触覚の感覚様式に着目し、それらの感覚情報を統合することで感情発達を再現する計算論モデルを提案している [Hori 18]。また日永田らは、ダマシオの概念モデルをもとに感情モデルを提案、深層学習を用いて実装したうえで母子インタラクションの中での感情分化を検証している [Hieida 18]。これらの研究の共通点は身体に着目していることであり、ダマシオが提唱している「情動は身体反応であり、その認知が感情である」[Damasio 03] といった考えに基づいている。感情を議論するうえで身体が重要なことは、感情の機能的側面を考えれば明らかである。感情は、自身の身体の状態と深く関わっている。体の外側の世界に関するセンサ情報が処理され、外受容感覚を生み、世界に関する概念や知識をつくり出すのだとすれば、自身の体に関する情報が内受容感覚を生み、身体に関する概

念や知識をつくりだす。そしてこれらが、情動や感情と深く関わっている。

残念ながら、現状の感情生成を目指した研究では、身体反応や内受容の設定に問題がある。また、「どのような条件を満たせばそれが感情といえるのか?」という究極的な問いに正確に答えるのは難しい。しかし重要なことは、心理学や神経科学、認知科学などさまざまな分野における知見を取り入れるとともに、成果をフィードバックすることである。この相互のやり取りにより本質を見据えることが、今後の感情研究において有用である。また、相互に曖昧な部分を残したまま仮説を共有し研究を進めていくアプローチも有用かもしれない。

後述する大平英樹先生の招待講演は、異なる分野の知見を取り入れるという点で非常に優れている。こうした考えを基盤に、AI研究が構成的なアプローチを推し進めることで、感情研究の発展につながると思う。

3. 一般講演

本 OS では、8 件の一般講演が採択された。投稿された論文は発表上限数を超過していたため、情報開示されたタイトルとアブストラクトに基づいて選定した。研究分野をなるべく広くカバーするように選定しており、論文の質は考慮されていないことに留意されたい。

まず著者の一人である日永田は、「ソマティック・マーカー仮説に基づく行動選択」と題して発表した [日永田 19]。また、共著者の堀井は、「制限ボルツマンマシンによる自由エネルギー最小化に基づく能動的知覚」と題して発表した [堀井 19]。これらの発表は、本 OS の主眼でもある身体と感情の密接な関係について、計算モデルを用いて議論することを目指しており、次章で述べる招待講演の内容とも関連している。特に堀井は、その中心的役割を担う自由エネルギー原理について議論している。

続いて静岡大学の佐久間一輝氏は、「深層学習の可視化による神経科学的知見の抽出」と題して発表をした [佐久間 19]。本研究は、感情に神経科学的にアプローチしており、興味深い知見を得ている。

大阪大学の内田貴久氏は、「共感を目的とした対話におけるユーザの選好に対する概念獲得手法に関する検討」と題して発表した [内田 19]。また、奈良女子大学の塚本麻衣氏は、「論理モデルに基づく感情生起と感情推測を行うエージェント」と題して発表した [塚本 19]。これらの研究は、対話や物語といった文章情報より、ルールベースで感情推論や感情生起を実現しており興味深い。また、広島市立大学の見尾和哉氏は、「LSTM を用いた本心でない発話の自動検出」と題して発表し [見尾 19]、東京農工大学の熊谷和実氏は、「経験に基づく個人の感情推定モデル構築方法およびロボット行動生成への利用」と題して発表した [熊谷 19]。両者の研究ではピッチや韻律などの音声特徴量や表情などの顔特徴量といっ

た情報から、相手の感情を推定している。さらに熊谷氏は、感情推定結果に基づいてロボットの行動を生成しており、人とロボットの相互作用を実現する手法の提案につなげている。これらの研究は **Affective Computing** の考えに近く、実社会応用や人とのインタラクションを視野に入れている。

玉川大学の宮田真宏氏は、「感情の価値計算システム仮説にもとづく人の推論システムの提案」と題して発表した [宮田 19]。本研究は推論に主眼においたものであり、複合的な機能に注目する重要な視点である。

本稿ではそれぞれの研究の概要のみを示しているため、詳細は各論文を参照されたい。

4. 大平英樹先生による招待講演

本章では、大平先生の講演「内受容感覚の予測符号化と感情の創発」を概観する。この講演は、これまでの感情研究の難しさから、近年脳の基本原則として注目されている自由エネルギー原理（特に予測符号化）の紹介と計算モデルの提案、心理学実験における生理反応や脳活動計測の結果から感情を明らかにしようとする試みなど、多岐にわたる内容であった。

4.1 感情研究の難しさ

我々は感情について強い実感をもっており、直感的に簡単に理解できると考えてしまう。例えば、「怒りという感情が存在する」、「誰でも同じような主観的な怒りを体験する」、「怒りは主観的な自己報告を記述することで研究できる」といった具合である。これらの考え方から導き出されるのは、自己報告を「教師信号」として脳活動や生理的活動を測定すれば、「怒り」を機械学習で推定できるのではないかという考えである。しかしながらこの考えは、誰ももっている人間の「心」についての信念や思い込みである素朴心理学を十分に排除できていない。思い込みを排除し感情をどのように捉えるべきかは、大きな問題である。

4.2 感情の理論：基本情動説 vs. 心理学的構成主義

基本情動 (basic emotion) 説とは、Paul Ekman が提唱した、進化の過程で形成された少数の生得的な基本情動が存在し、この基本情動には固有の神経生理学的基盤が存在するという考え方である [Ekman 92]。この説にはいくつかの反証がある。例えば、表情認識研究のメタ分析では、表情による情動認識は、文化差が大きいことが知られている [Gendron 14, Gendron 18b]。また、神経画像研究のメタ分析では、個々の知覚、認知、感情で特異的に活動する脳部位は存在しないといわれており、ほとんどの精神機能は脳の大規模ネットワークにより実現されていると考えられる [Lindquist 12]。また、Lisa Feldman-Barrett は、表情、脳活動、生理的反応のいず

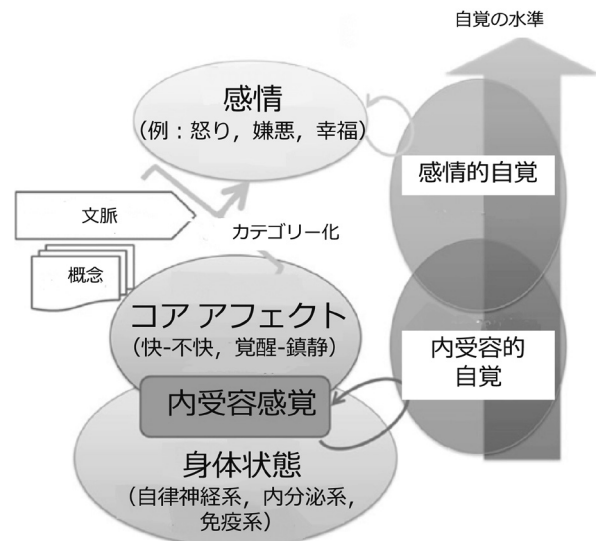


図2 感情の心理学的構成主義（大平先生の資料をもとに再編集）

れにも、私達が経験するさまざまな感情を1対1で指し示すパターンは存在しないと主張しており、このことを、感情の「指紋」は存在しないと表現している [Feldman Barrett 17]。これらの反証と呼応して、近年の感情科学では、感情は身体内部の知覚である内受容感覚からコアアフェクトという基礎的状态が形成される過程と、コアアフェクトを言語により解釈するカテゴリー化の過程から創発されるとする心理学的構成主義 (psychological constructivism) に根差した感情のモデル [Moriguchi 13] が提案されている (図2)。コアアフェクトは、脳の生成モデルによる予測に基づいた身体信号のベイズの推論により形成されると考えられている。

4.3 内受容感覚、コアアフェクト、意思決定

上述のコアアフェクトの形成は、予測的符号化 (Predictive coding) の考え方に基づいている [Friston 10]。予測的符号化では、知覚は単なるボトムアップ的処理ではなく、内的モデルによる予測と実際の入力信号との比較により生起すると考える。そして脳は、内的モデルの更新か身体状態の変更のいずれか、もしくはそれらの組合せによって予測誤差を最小化するように行動を制御する (図3)。知覚経験はベイズ的につくられる、つまり、入力された感覚信号と内的モデルにより生成された予測 (事前分布) との差から、ベイズの定理に基づいて事前分布が事後分布に更新される一連の過程が意識されたものである。そう考えると、内受容感覚は予測と現実の信号により創発されるバーチャルなシミュレーションであるといえる [Farb 15]。予測的処理の神経基盤の研究 [Seth 16] は、島皮質 (insula) がハブとなっていることを示しており、さらに島皮質は感情において重要な役割を果たしていることから、上述の議論の妥当性が示唆される。これまで説明してきたように、内受容感覚 (身体の知覚) が、感情の創発にそれほど重要なのは

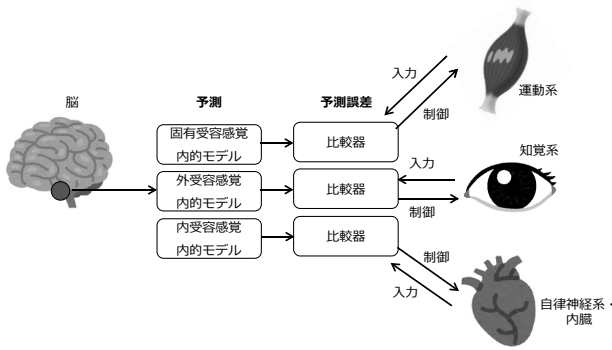


図3 予測符号化 (大平先生の資料をもとに再編集)

なぜであろうか？ William James の以下の主張が、このことをよく説明している。「身体の変化は喚起的事実の近くに直接引き続いて起こり、そしてそれらの身体変化が起こるときに生じるその同じ変化の感覚こそが情動なのである」[James 84].

報酬, 学習, 意思決定, 感情をシームレスに統合した例として, Keramati らの研究がある [Keramati 14]. この研究では, 報酬予測誤差により身体状態のセットポイントを変更するシステムを提案している. また Joffily らは, 予測誤差 (自由エネルギー) の変化のパターンと感情経験のつながりを議論している [Joffily 13]. 大平先生は, 予測的符号化のハブ領域と意思決定の関係を, MRS (Magnetic Resonance Spectroscopy) を用いて検証している [Ohira 14].

4.4 内受容感覚の計算論モデル

大平先生は, Stephan らの提案をもとに内受容感覚の計算モデルを提案している (図4) [大平 18, Stephan 16]. このモデルでは, 生理的状態の内的モデルが予測を出力し, 生理的状態の神経表象との予測誤差により内的モデルを更新する. 生理的状態は自由エネルギー原理の枠組みの一つである能動的推論に基づいて制御される. またこの計算モデルを用いて, 不安障害やアレキシサイミア様の身体反応をシミュレーションしている.

4.5 カテゴリー化：感情創発における概念と文脈の意義

例えば, 駅のホームで吐いている人を目撃したとき, 我々は予測誤差が拡大し驚きを知覚するかもしれない. そして, 心拍数が上昇すれば怒り, 心拍数が抑制されれば嫌悪を知覚するであろう. このように, 人の感情がカテゴリー化されているとしたとき, 概念 (例えば, 怒りや嫌悪) を高次の予測生成モデルの出力と考える. 概念を適用することによって, 内受容感覚モデルの精度が高まり, 分化し, 安定した生理的反応が生じる. また, 言語ラベルは予測誤差最小化のために有効である [Gendron 18a]. 文脈の影響としては, ヨークトパラダイムを用いた実験で, 同じ報酬であっても文脈により生理的状態に変化が出ることが知られている. 生体反応の

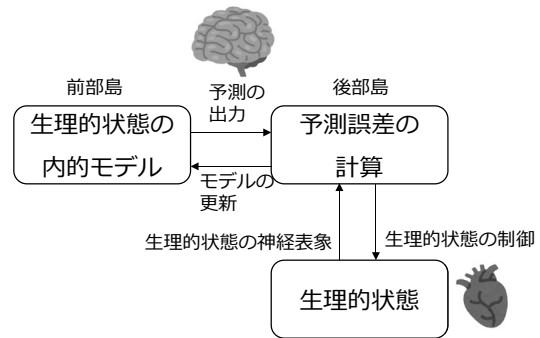


図4 内受容感覚の計算論モデル (大平先生の資料をもとに再編集)

時間特性も見逃すことのできない重要な点である. 脳と身体の生物学的システムは, 異なる時定数で, しかも相互に影響し合いつつ反応が進行する. すなわち, 大きなぜい弱性 (vulnerability) をもつことになる. これが, 感情の本質かもしれない.

4.6 結語：感情をどう理解すべきか

我々は感情について強い実感をもっており, 直感的に簡単に理解できると考えてしまう. しかしその一方で, 我々はブラックホールを直感で理解しようとはしないし, 簡単に理解できるとも思わない. 感情も研究対象として同じであり, いったん素朴心理学を徹底的に廃して, 定義と数理化を追求して研究すべきかもしれない.

5. 議 論

本章では, 本 OS を通して得たものを振り返りつつ, 今後の方向性などについて議論する.

5.1 身体と感情と人工知能

大平先生の講演は, 感情が身体反応や内受容感覚と密接に関わっていることを明確に示すものであった. そして身体反応や内受容感覚は, 予測符号化でそのメカニズムの大枠を説明可能である. こうした感情に対する視座は, 異なる時定数で相互作用し, 非常に複雑でぜい弱な身体システムと, そのシステムを安定化するための知能のメカニズムがその背景にあることに気付かせてくれる. 生物にはこうした身体という制約があり, その制約下で行動を決定しなくてはならない. 一方, 人工物は死ぬことはなく, 一度バッテリーが切れても充電すれば再び動くことができてしまう. では, AI に身体は必要だろうか? この問いに対する正しい答えは今のところないが, 著者らは必要であると考え. それは我々作り手が人工物に制約を与えることにもつながるが, この制約こそが感情の根源であり, 人の感情メカニズムを再現し知るといふ文脈では不可欠である. 次に来る問いは, 身体をどのように実現すべきかである. 近年は, シミュレーション技術の進展も目覚ましく, 多くの AI 研究が

シミュレーションによって支えられているのも事実である。ただし、シミュレーションでは再現できないことも多く、実ロボットを使った研究の重要性の一つはそこにある。また、実際につくことで実応用にも生かせる可能性があるのも実ロボットの利点である。しかしその一方で、ハードウェアを用いるコストは莫大であり、どのようなハードウェアであれば人間の感情メカニズムを再現できるかは未知である。そうした中、近年のロボティクス分野では、ソフトロボティクスが急速に発展しており、身体的に柔らかいロボットの制御困難性の中に、感情の一端を見いだせる可能性がある。

5.2 自由エネルギー原理に対する批判

自由エネルギー原理は必ずしも万能な原理ではなく、批判もある。その一つが「暗室問題」である。もし生き物にとって予測誤差=自由エネルギーの最小化だけが目指されるのであれば、予測誤差が生じない、すべての感覚が遮断される暗室に閉じこもっているのが、最も適応的なはずだという疑問である。これに関して、大平先生は下記のように述べている。1) 現実の環境は多次元で不確実であり、否応なく常に予測誤差が生じるので、その対処のために行動が必要になる。2) 生き物にはエネルギーが必要であり、それは常に消費されるので、獲得するための行動が必要である。この暗室状態に最も近いのが、薬物、アルコールなどへの依存症だと思われる。依存症になると依存対象への渴望が極大化する一方で、その他の物事は報酬として機能しなくなる。こうした依存症では、もし上記の2)が満たされるのであれば、薬物などの依存対象への予測誤差だけが問題になり、それを最小化するために喜んで暗室に引きこもるだろう。実際そうした患者は、古来多く存在している。その意味で「暗室問題」は、自由エネルギー原理の批判というより、むしろ支持するものになっている。

5.3 感情研究の方向性

感情を知るためには、本来、人を調べるべきであるという考えに対して、我々AI研究者はAIやロボットという新しいツールを使った方向性を強く打ち出すべきである。そして今我々がすべきことは、まさにつくることである。例えば、自由エネルギー原理は概念的な原理であり、それをどう実装するかは明らかではない。そして実装し動かすことで、人間では不可能な実験や内部状態の可視化が可能となる。それ以前に、道具としてのAIが、身体のせい弱性を乗り越えることができるかは、非常に興味深い点である。実際、著者らはこれまで、外受容や固有受容を対象にカテゴリー化(概念化)や言語獲得の問題を、AIやロボットという道具を使って研究してきた[長井 12]。そこで目にしたのは、完全に人間とは同じでないにせよ、ロボットが概念を獲得し、言語を理解できる可能性である。これを内受容に拡張することは、

ある意味必然であり、究極的には、感情を人工的に作り出せる可能性はゼロではない。そのためにも、感情と呼ぶための要件は何か? という問いに、向き合わなければならない。

AI研究では、何かをつくるのが先に来てしまい、その本質を検証するところが弱くなりがちである。本OSでの知見を受け、それをシステムとして作り、実際のコミュニケーションにおける検証や、シミュレーション、システム内部の可視化などを通して、より深くメカニズムの検証を行っていくことが重要である。

5.4 応用としての感情研究

最後に、応用という視点で感情研究の方向性を考えてみたい。ここでは再び、「AIに感情は必要か?」という問題に行き当たる。その答えは、応用先に依存するであろう。例えば決まった表情を認識するための識別器が、感情をもたなければならない理由は見当たらないし、大量のデータによって学習されたニューラルネットワークは、人間を超えた精度や速度で表情を見分けるであろう。しかしこれが、他者の心的状態を認識・理解するという応用になれば話は別である。人は他者の心的状態を、自分を通して認識し理解することができる。これは、共感や社会的な高次感情にも関係している。情動伝染やミラーシステムなどの存在を考えれば、人に共感するAIやロボットには、感情が必要かもしれない。

SFでは、人の意識をサイバー空間にアップロードするような話も存在する。人が身体を捨てるというのは夢のある話ではあるが、身体を捨てたときに果たしてその存在は人と呼べるものなのかは疑問である。つくれるかどうかや新しさだけではなく、最終的にどのようなものを求めるべきなのかを、我々は考えていかなければならない。感情のメカニズムを明らかにすることは、そうした人工物と人との関わりや、倫理的な問題を考えるうえでもさまざまな示唆を与えてくれるはずである。

6. まとめと今後の展望

感情に関する研究は少なくない。しかし、その本質になかなかとり着けないのは、素朴心理学を廃せないことと、各分野の感情研究がバラバラになされていることに起因しているのかもしれない。また感情の研究者はそれぞれ、感情についてのこだわりをもっており、そのこだわりが他を受け入れにくい傾向にあるように思える。一度、それぞれのこだわりをぶつけ、徹底的に感情の本質について語る必要があり、その輪を広げていくことが、感情の解明につながると考える。本OSは、感情の本質について語る場として新設したものであるが、感情というワードにこだわり過ぎず広く議論できることが、最終的には感情の本質にたどり着く道筋をつくるのかもしれない。そしてAIは、感情メカニズムの解明と感情認識

など社会的な応用という二つの側面で重要な役割を果たすはずである。

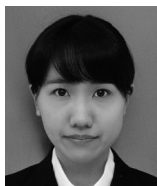
謝辞

本 OS の開催にあたり JST CREST (JPMJCR15E3) の助成を受けた。また、多大なるご協力、ご議論をいただいた大平英樹先生に感謝する。

◇ 参考文献 ◇

- [Asada 15] Asada, M.: Development of artificial empathy, *Neuroscience Research*, Vol. 90, pp. 41-50 (2015)
- [Breazeal 02] Breazeal, C.: *Designing Sociable Robots*, The MIT Press (2002)
- [Damasio 03] Damasio, A. R.: *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*, Houghton Mifflin Harcourt (2003)
- [Ekman 92] Ekman, P.: An argument for basic emotions, *Cognition & Emotion*, Vol. 6, No. 3-4, pp. 169-200 (1992)
- [Farb 15] Farb, N., Daubenmier, J., Price, C. J., Gard, T., Kerr, C., Dunn, B. D., Klein, A. C., Paulus, M. P. and Mehling, W. E.: Interoception, contemplative practice, and health, *Frontiers in Psychology*, Vol. 6, p. 763 (2015)
- [Feldman Barrett 17] Feldman Barrett, L.: *How Emotions are Made: The Secret Life of the Brain*, Boston & New York: Houghton Mifflin Harcourt (2017)
- [Friston 10] Friston, K.: The free-energy principle: A unified brain theory?, *Nature Reviews Neuroscience*, Vol. 11, No. 2, p. 127 (2010)
- [Gendron 14] Gendron, M., Roberson, D., Vyver, van der J. M. and Barrett, L. F.: Perceptions of emotion from facial expressions are not culturally universal: Evidence from a remote culture, *Emotion*, Vol. 14, No. 2, p. 251 (2014)
- [Gendron 18a] Gendron, M. and Barrett, L. F.: Emotion perception as conceptual synchrony, *Emotion Review*, Vol. 10, No. 2, pp. 101-110 (2018)
- [Gendron 18b] Gendron, M., Crivelli, C. and Barrett, L. F.: Universality reconsidered: Diversity in making meaning of facial expressions, *Current Directions in Psychological Science*, Vol. 27, No. 4, pp. 211-219 (2018)
- [Hieida 18] Hieida, C., Horii, T. and Nagai, T.: Emotion Differentiation based on Decision-Making in Emotion Model, *2018 27th IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN)*, pp. 659-665, IEEE (2018)
- [日永田 19] 日永田智絵, 堀井隆斗, 長井隆行: ソマティック・マーカー仮説に基づく行動選択, 2019 年度人工知能学会全国大会 (第 33 回) 論文集, pp. 3G3OS18a02 (2019)
- [Horii 18] Horii, T., Nagai, Y. and Asada, M.: Modeling development of multimodal emotion perception guided by tactile dominance and perceptual improvement, *IEEE Trans. on Cognitive and Developmental Systems*, Vol. 10, No. 3, pp. 762-775 (2018)
- [堀井 19] 堀井隆斗, 長井隆行: 制限ボルツマンマシンによる自由エネルギー最小化に基づく能動的知覚, 2019 年度人工知能学会全国大会 (第 33 回) 論文集, p. 3G3OS18a03 (2019)
- [James 84] James, W.: What is an emotion?, *Mind*, Vol. 9, No. 34, pp. 188-205 (1884)
- [Joffily 13] Joffily, M. and Coricelli, G.: Emotional valence and the free-energy principle, *PLoS Computational Biology*, Vol. 9, No. 6, p. e1003094 (2013)
- [Keramati 14] Keramati, M. and Gutkin, B.: Homeostatic reinforcement learning for integrating reward collection and physiological stability, *Elife*, Vol. 3, p. e04811 (2014)
- [Kishi 13] Kishi, T., Kojima, T., Endo, N., Destephe, M., Otani, T., Jamone, L., Kryczka, P., Trovato, G., Hashimoto, K., Cosentino, S. and Takanishi, A.: Impression survey of the emotion expression humanoid robot with mental model based dynamic emotions, *2013 IEEE Int. Conf. on Robotics and Automation*, pp. 1663-1668 (2013)
- [熊谷 19] 熊谷和実, 水内郁夫: 経験に基づく個人の感情推定モデル構築方法及びロボット行動生成への利用, 2019 年度人工知能学会全国大会 (第 33 回) 論文集, p. 3G4OS18b03 (2019)
- [Lim 15] Lim, A. and Okuno, H. G.: A recipe for empathy, *Int. Journal of Social Robotics*, Vol. 7, No. 1, pp. 35-49 (2015)
- [Lindquist 12] Lindquist, K. A. and Barrett, L. F.: A functional architecture of the human brain: Emerging insights from the science of emotion, *Trends in Cognitive Sciences*, Vol. 16, No. 11, pp. 533-540 (2012)
- [見尾 19] 見尾和哉, 石野亜耶, 目良和也, 竹澤寿幸: LSTM を用いた本心でない発話の自動検出, 2019 年度人工知能学会全国大会 (第 33 回) 論文集, p. 3G4OS18b01 (2019)
- [宮田 19] 宮田真宏, 大森隆司: 感情の価値計算システム仮説にもとづく人の推論システムの提案, 2019 年度人工知能学会全国大会 (第 33 回) 論文集, p. 3G4OS18b04 (2019)
- [Moriguchi 13] Moriguchi, Y. and Komaki, G.: Neuroimaging studies of alexithymia: Physical, affective, and social perspectives, *BioPsychoSocial Medicine*, Vol. 7, No. 1, p. 8 (2013)
- [長井 12] 長井隆行, 中村友昭: マルチモーダルカテゴリゼーション: 経験を通して概念を形成し言葉の意味を理解するロボットの実現に向けて (特集: 記号創発ロボティクス), 人工知能学会誌, Vol. 27, No. 6, pp. 555-562 (2012)
- [Ogata 00] Ogata, T. and Sugano, S.: Emotional communication between humans and the autonomous robot Wamoeba-2 (Waseda amoeba) which has the emotion model, *JSME Int. Journal Series C Mechanical Systems, Machine Elements and Manufacturing*, Vol. 43, No. 3, pp. 568-574 (2000)
- [Ohira 14] Ohira, H., Ichikawa, N., Kimura, K., Fukuyama, S., Shinoda, J. and Yamada, J.: Neural and sympathetic activity associated with exploration in decision-making: Further evidence for involvement of insula, *Frontiers in Behavioral Neuroscience*, Vol. 8, p. 381 (2014)
- [大平 18] 大平英樹: 内受容感覚の予測的符号化: 福島論文へのコメント (特集: 叡智), 心理学評論 = *Japanese Psychological Review*, Vol. 61, No. 3, pp. 322-329 (2018)
- [Picard 00] Picard, R. W.: *Affective Computing*, MIT Press (2000)
- [佐久間 19] 佐久間一輝, 森田純哉, 野村太輝, 平山高嗣, 榎堀優, 間瀬健二: 深層学習の可視化による神経科学的知見の抽出, 2019 年度人工知能学会全国大会 (第 33 回) 論文集, p. 3G3OS18a04 (2019)
- [Seth 16] Seth, A. K. and Friston, K. J.: Active interoceptive inference and the emotional brain, *Philosophical Trans. of the Royal Society B: Biological Sciences*, Vol. 371, No. 1708, p. 20160007 (2016)
- [Stephan 16] Stephan, K. E., Manjaly, Z. M., Mathys, C. D., Weber, L. A., Paliwal, S., Gard, T., Tittgemeyer, M., Fleming, S. M., Haker, H., Seth, A. K., et al.: Allostatic self-efficacy: A metacognitive theory of dyshomeostasis-induced fatigue and depression, *Frontiers in Human Neuroscience*, Vol. 10, p. 550 (2016)
- [塚本 19] 塚本麻衣: 論理モデルに基づく感情生起と感情推測を行うエージェント, 2019 年度人工知能学会全国大会 (第 33 回) 論文集, p. 3G4OS18b05 (2019)
- [内田 19] 内田貴久, 港隆史, 中村泰, 吉川雄一郎, 石黒浩: 共感を目的とした対話におけるユーザの選好に対する概念獲得手法に関する検討, 2019 年度人工知能学会全国大会 (第 33 回) 論文集, p. 3G4OS18b02 (2019)

著者紹介



日永田 智絵 (正会員)

2014年電気通信大学情報理工学部知能機械工学科卒業。2016年同大学院情報システム学研究科情報メディアシステム学専攻修士課程修了，修士（工学）。同年，日本学術振興会特別研究員（DC1），2019年電気通信大学大学院情報理工学研究科機械知能システム学専攻博士後期課程単位取得退学。同年，大阪大学先導的学際研究機構付属共生知能システム研究センター特任研究員，現在に至る。HRI，情動，感情モデルの研究に従事。



堀井 隆斗 (正会員)

2013年大阪大学大学院工学研究科博士前期課程修了。2017年同研究科博士後期課程単位取得退学。博士（工学）。電気通信大学大学院情報理工学研究科特任研究員，同特任助教を経て現在，大阪大学大学院基礎工学研究科助教。東京大学国際高等研究所ニューロインテリジェンス国際研究機構客員研究員を兼務。感情の発達原理に興味をもち，計算モデルによってその過程を明らかにすることを目的とする。



長井 隆行 (正会員)

1993年慶應義塾大学理工学部電気工学科卒業。1995年同大学院理工学研究科電気工学専攻前期博士課程修了。1997年同後期博士課程修了。博士（工学）。1998年電気通信大学助手，2003年カリフォルニア大学サンディエゴ校客員研究員，2015年電気通信大学教授を経て，2018年より大阪大学大学院基礎工学研究科教授。電気通信大学人工知能先端研究センター特任教授，玉川大学脳科学研究所特別研究員，産業技術総合研究所人工知能研究センター客員研究員を兼務。IROS Best Paper Award Finalist，Advanced Robotics Best Paper Award，本学会論文賞など多数受賞。知能ロボティクス，認知発達ロボティクス，ロボット学習に関する研究に従事。