

[暮らしの中で活躍する AI とロボット]

## ②なぜロボットを使うの？



今井倫太 | 慶應義塾大学

### AIにとってロボットって何？

人工知能ブームを発端に人とコミュニケーションするロボットも脚光を浴びている。レストランの店先で順番待ちの整理券を配ったり、ホテルのフロントで受付をしたりと実際に働いている(?)姿も見受けられるようになってきている。普段から研究を行っている立場から現状のブームを見ていると、人とロボットのコミュニケーションの発展に皆が期待を寄せているのを感じる。しかしながら、一歩引いて考えると「ロボットって本当に必要なの?」という疑念が湧いてくる。筆者も長年の研究人生の中で何度もさまざまな場でその必要性を問われてきた。物理的な体があるので場所を取るし、動く部品がたくさんあるので故障する可能性があるし、コンピュータや携帯端末に比べるとコストがかかる。導入にあたっての大きなデメリットがあるにもかかわらず、人とコミュニケーションするロボットが行っているタスクは、道案内や情報検索、商品の購入、整理券の発券などコンピュータや携帯端末で事足りることばかりである。違いがあるとすれば置いてあると目立つので客引き効果があることかもしれない点が挙げられる。本当にほかにメリットはないのか? その疑問について本稿は解説したい。

最初に筆者の考えを表明しておく。「ロボットはAIにとって本当に必要である」ということになる。ただしここでいうロボットとは機械の体を持ついわゆる「ロボット」に限定するのではなく、コンピュータグラフィックで描かれるエージェン

トや、スピーカを利用した音声エージェントといった自律的に判断して動き、言葉や時にはジェスチャなどを用いて人とコミュニケーションできる人工物全般を指している。人工知能や情報通信分野の技術が応用されることで、自動運転を始めとしてさまざまな機械や機器が自律的に判断して動き、我々の生活を支援してくれるようになってきている。その中でロボットがなぜ必要なのかについて考えていく。

### 勝手に動く機械

#### パッシブな機械からアクティブな機械へ

ロボットの必要性について考えるにあたりまずは道具について考える。道具は人類が生活を便利にするために長年発明してきたものである。人は道具を使って何かをするという目的をあらかじめ持っており、人の使用目的に従って道具は受動的に役割を果たす。たとえばコンピュータは、文章を書いたりメールをチェックしたりする用事があるときに使用する。遠くへ行きたいとき、車に乗り、料理をするとき包丁で食材を切る。道具を使用する目的が使用開始前に人側に存在する。

受動的な道具と対照的な存在として、勝手に(能動的に)動く機械がある。能動的に働く機械は現在の人工知能や情報通信技術によって本格的に私たちの生活の中に入り込もうとしている。上記でコンピュータを受動的な道具として紹介したが、実際にはコンピュータや携帯端末が何らかの新しい情報を取得するとユーザに能動的に知らせることができる



ので、純粋に受動的な道具とは呼べない側面がある。コンピュータは、人工知能や情報通信技術を駆動する本体そのものでもあり、受動的に人から使われる道具から、能動的に人に働きかける存在へと変わりつつある。またロボットはショッピングモールやホテル内を自律的に動き回り、時にロボットの側から人に話しかけ、能動的に人に情報を伝えるようになる。さらに自動運転が搭載された車はルートを手勝手に考えて人を能動的に運び、IoT 機器によって知能化された部屋は人が快適に過ごせるように能動的にサービスをするようになる。

勝手に動く機械や機器は、それらを使用する意思やもしくはかかわりたいという意味が人側になくとも能動的に人に働きかけてくる可能性がある。勝手に動く機械や機器に囲まれた生活において人が快適に過ごすためには、機械や機器にどのような仕組みが必要なのか。人工知能や情報通信技術によって誕生する機械の能動性に対する人の理解の仕方の中に「ロボットは AI にとって本当に必要である」の答えが潜んでいる。ロボットの必要性について考察する上で重要な、動くものに対する人の捉え方について本題に入る前に解説する。

## 意図スタンス

Daniel C. Dennett は、人が動くものを見るときに認知バイアス（人が物事を解釈したり判断したりするときに起きる偏った傾向）として物理スタンス・設計スタンス・意図スタンスの3種類があることを示した<sup>1)</sup>。幾何学図形を使った動きの例でこの3種類の違いを説明する。直線や円、三角形といった幾何学図形を用いたアニメーションを想像すると、この3種類の動きの解釈が分かる。たとえば、画面に斜めに直線が引いてあり、その上を黒丸が線に接触しながら上から下の方向へ移動したり、黒丸が上下に放物線を描きながら移動する場合、人は、この動きを物理法則に従って動いていると理解する。これを物理スタンスと呼ぶ。アニメーションなので本当

に物理法則が働いているのではなく、単なる動きに人が物理法則を見出している点に認知バイアスとしての本質がある。

黒丸が一定の距離を左右方向に行ったり来たり、一定の周期で繰り返し動くのを見ると、人はそこに設計された機械仕掛けの存在を感じる。これは設計スタンスと呼ばれる。歯車やテコ、クランクなど機械的な機構が組み合わされることによって生じる動きであり、設計者がなんらかの動きを実現しようと機械機構を組み合わせた結果であると解釈するものである。

小さな三角形が左から右へ移動する後ろから同じ大きさの黒丸が追従して移動している場面で三角形が停止して、黒丸の方へ振り返ったとする。今度は三角形が黒丸の方向へ移動し始めて、ちょっと遅れて黒丸も逆方向へ移動し始める。人はこのようなアニメーションを見るとそこにストーリーを付けたい。この場合は三角が黒丸に跡をつけられていることに気がついて、黒丸を追い返したと理解したくなる。三角は黒丸を追い返すという行動を「意図的に」行ったように見える。この行動解釈を意図スタンスと呼ぶ。人は何か目的を持って動いているかのように見えるものを見るとその背後の意図を想定して解釈する傾向がある。

## ロボットの必要性

### 意図スタンスを促すロボット

ロボットは元来能動的に動くものであり、その形状も人型ではなくとも人を意識させるものであったり、ロボットを生き物に近い存在として捉えた解釈でロボットの動きや発言を人が理解することを可能にする。ロボットの行動はその背後に何らかの目的があると人に感じさせやすく、人はロボットに対して意図スタンスをとる。人工知能や情報通信の技術によって自律的に判断してさまざまな形で人にサービスを行う機器が人と直接やりとりする際のユーザ

インタフェースの在り方を考えると、ロボットが持っている、人に意図スタンスをとらせる能力が重要になる。

ロボットの行動に対して人が意図を想定するということは自然なことであり、あえてなぜ取り上げるとかの疑問に思う人がいるかもしれない。しかしこの至極当然なロボットの能力が人工知能にとってのロボットの必要性そのものなのである。自律的に判断し動く機器がとる行動が複雑化したり、機能が増えたとしても、人の意図スタンスを促すことで、意図や目的といった高い抽象度で機器の行動を人は理解でき、機器に対して指示できたり、機器とコミュニケーションできたりする。人に意図スタンスをとることを促せることが「ロボットはAIにとって本当に必要である」ということの意味なのである。

意図スタンスを促すことができれば、近い将来発生するであろう次に説明する状況にも対処できる。自律的に判断する機器は独自のロジックで状況を判断しさまざまな行動（やサービス）を行う。そのすべての行動や機能を、スイッチで動く受動的な道具と同じ形でユーザーに見せてしまうと自律性の向上や機器間の連携の組合せの複雑さに応じてユーザーと機器とのやりとりする内容も複雑化してしまう。人と機器のやりとりの複雑さがユーザーの扱える範囲を超えてしまう恐れが高い。機器の機能をそのままユーザーに見せて設計スタンスをユーザーにとらせるよりも、意図スタンスをとらせることで人は機器の動作をより理解することができ、お互いのやりとりをより単純化できる。

ロボットの導入は、人工知能や情報通信技術によって自律して複雑化した機器をどのように人に見せるべきかという問題に対する回答である。そして導入の本質は機器に対して人の意図スタンスを促せる点である。この提言は、自律性の高い機器は意図スタンスを促すデザインを採用するべきであると言いつつ換えることもできる。

## メンタライジングのパワー

意図スタンスを促すデザインを考える上で人が行うメンタライジング (Mentalizing) の役割を考えることは大変重要である。メンタライジングとは相手（人間や、動物、ロボットなど意図スタンスをとりやすい相手）の行動に対して人が想定する心のモデルである。心のモデルというと漠然としてしまうのでもう少し厳密に定義すると、相手の振舞いを人が見るときに相手がなぜその行動をとったのかを理解するための相手の内部状態である。具体的には、相手の行動の背後に人が見出す行動の目的・信念（相手が現在見えているもの、存在を信じているもの）・知識・感情である。人が意図スタンスで相手の動きを解釈する過程をより詳細に説明するものである。意図スタンスを促すデザインを考えることは、機器の状態を人に理解させるためのメンタライジングの要素それぞれをどのように人に抱かせるか考えることと同義であるともいえる。

意図スタンスを促すデザインを考える参考材料として、人が人工物に対してメンタライジングする現象について調べた研究について紹介しておく。Baker<sup>2)</sup>らは、駐車場に来る屋台のトラックにお昼ご飯を買いに行くエージェントの動きを題材に人のメンタライジングの課題を扱った(図-1)。図-1の黒三角はエージェントで、駐車場には2カ所の屋台トラックが来る場所があり、駐車場の真ん中には壁があるのでエージェントが最初に立っている場所からは1台のトラックは見えるがもう1台は見えない状況となっている。また、屋台のトラックには、韓国料理(K)、メキシコ料理(M)、レバノン料理(L)の3種類が日替わりでやってくる。図-1は一番上から順番に時間経過を表している。

図-1のエージェントの動きに対して意図スタンスによる解釈を与えると次のようなストーリーで解釈できる。エージェントはお昼ご飯を買う目的で歩き出した。この時点で観測している人にとってエージェントが何を食べたいかは分からない。しかし



エージェントから見えるのは韓国料理なので韓国料理を買いに行っている可能性が少し高いと予想できる。観測者の予想は、エージェントが韓国料理のトラックの前を素通りした瞬間、レバノンかメキシカンが食べたいとの推測に変化する。壁の向こう側のトラックが見えた瞬間エージェントが向きを変えて韓国料理の方へ行くのを見ると、エージェントが食べたかったのは、メキシコ料理で次に韓国料理が食べたかったということが推測できる。前提となる文脈が用意されると、人はアニメーションを意図スタンスで解釈して、エージェントの心のモデルから行動を理解しようとする。

Bakerらは、心のモデルをエージェントの目的・信念（相手が現在見えているもの、存在を信じているもの）に絞って心のモデルを用意し、エージェントの行動を理解するコンピュータを作った。さらに用

いたモデルを使って人と同じようにコンピュータが図-1の場面を解釈できるかについて調べた。つまり、人がエージェントの行動を観測して理解するのと同じように、コンピュータがエージェントの行動を観測して理解できるかどうかを調べた。実験の結果は、コンピュータが、エージェントの行動から直接何を食べたいのかを推測するよりも、エージェントの心のモデルを想定して推定した方がより高い精度で推定できることを示している。

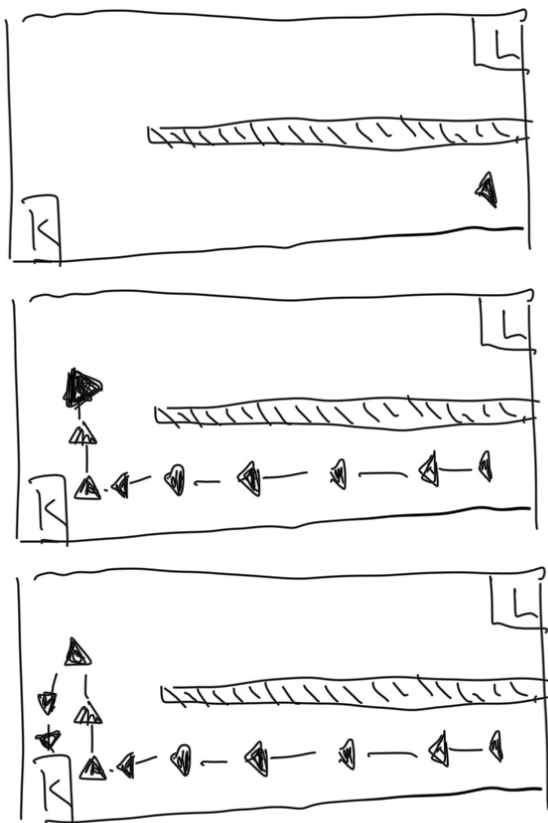
Bakerらの研究は、自律的な機器のユーザインタフェースとしてロボットを用いることで、人が機器に対してメンタライジングを行い、意図の抽象レベルで機器を理解し、機器を使うようになる可能性を示している。また機器が人からどう思われているかも計算できる可能性も示唆しており、自律的な機器が今後どのようなユーザインタフェースをまとって行くべきなのかを考える上での重要な知見であると考えられる。

## AIのインタフェースとしてのロボット

### どこに使えるの？ 使われていくの？

本稿の最初の章でロボットのやっていることは、道案内や情報検索、商品の購入、整理券の発券などコンピュータや携帯端末で十分できることばかりであると述べた。意図スタンスやメンタライジングのメリットがあったとしても他の情報機器で十分なタスクにあえてロボットを用いることは、エンタテインメント性以外に大きな理由があるのか疑問がある。意図スタンスやメンタライジングの効果が真に役立つ領域について考えてみる。

現状のコミュニケーションロボットの最大の問題は既存の情報通信機器で扱えるタスクの焼き直しが多数を占めているところである。道案内や情報検索、商品の購入、整理券の発券といった例を始めとして情報通信機器が得意とするタスクはユーザのリクエストに対してほぼ一瞬でサービスを提供可能な（提



■ 図-1 お昼ご飯を買いに行くエージェントのメンタライジング

供すべき) タスクである。たとえば道案内であれば行き先を入力したら即座に道順を地図で示してほしい、情報検索もすぐに検索結果の一覧を返してほしい。ほぼ一瞬で解決可能なタスクでは、タスク遂行に関する時間的なプロセスが皆無であり、意図スタンスやメンタライジングで可能となる意図の理解が役立つチャンスはほぼないのである。

一方で現実世界のタスクには時間的な過程を経て初めてタスクが完了するものがたくさんある。掃除、購入物の配達、ホテルや病院での運搬、商品購入のコンサルティング、個人の生活アシスト、自動運転などさまざまな例があり、その中に自律的に問題解決する機器が導入されていくものも今後多いと考えられる。さらに、人から機器へ一方的に命令するだけでなく、機器の方から人へ依頼したり、時には人と機器が協調してタスクを遂行したりする場面が考えられる。意図スタンスを促すことで機器に対するメンタライジングを人に起こさせ、時間的な経過の中で機器が何を行っているのかを人に意図レベルで理解させることができる。また人が機器の意図を理解することを通して、人と機器が協調してタスクを遂行できるようになる。

時間ボリュームのあるタスクを扱うロボットや、人と協調しながらタスクを行うロボットの研究はすでに存在する。今後、機器の自律性が高まることでますます自動化されていく領域であり、ロボットが真に役立つターゲットに絞って研究開発することが重要である。

## 意図して行う行動

機器が人から意図スタンスで理解されることの有効性は機器が時間ボリュームのあるタスクを扱っているときに現れる。意図して行う行動と時間ボリュームのあるタスクの間関係も本質的な構造なのでここで解説しておく。

Choen と Levesque<sup>3)</sup> は、合理的に判断して行動するエージェントを題材に意図と目的、信念の間の関

係について理論的体系化を行った。論文のタイトル Intention is Choice with Commitment (意図とは、約束のもとでの選択) が何を示すのか一見理解するのは困難である。簡単に要約すると次のとおりである。

人が意図しながらある行動を行う際には、自分の心(信念)の中で現時点で想像できる複数の未来の中から実現したい目標が成立している特定の(時間的)未来を選ぶ(choiceする)。ほかの未来ではなく、ある1つの未来の成立を意識しながら選ぶのである。目標を実現する未来を無意識のうちにただ選ぶのと、目標の成立を意識しながら未来を選ぶとの違いは、以下で説明する誓い(commitment)の存在に由来する。目標を実現したい意図を持つ場合、人はできる限りの努力を続けて自分の手で目標を実現できるように自分自身(場合によっては他人)に誓うのである。この誓いが簡単に破棄されるようでは、意図に基づく行動(選択した未来を実現する行動)とはならないので、(自分の能力で達成可能だと信じられる間)できる限りの努力を行い、(ほかの誰かに先を越されておらず、まだやられてないと信じている間)自分の手で実現しようとする。ただの目的の達成とはわけが違う。自分の意識に基づき目的を達成するさまがここには描かれている。誰かに先を越されて目標が実現されたら、目標は実現できていても目標を実現する意図が満たされたとはいえない。また選択した未来を自分の能力で現実のものにする確信が持てなくなったら(自分で未来を意図した形で実現できないので)、合理的なエージェントは目標の実現を諦めるのである。

意図とは、元々時間展開がある出来事において初めて定義可能なものであり、時間ボリュームのあるタスクを扱うユーザインタフェースを対象としてロボットを考えるべきであるという本質的な答えがここにある。目的と行動の間に生じる課題は人工知能分野でプランニングアルゴリズムとして長年扱われており、コンピュータが目的を達成する行動を自動的に選択することが実現されている。一方、意図の



抽象レベルでは、未来（タスク）を実現し続けようとする努力を維持する理由が表現されており、自律的に動く機器が時間ボリュームのあるタスクを遂行する過程の理解に繋がる情報となっている。

## 将来展望

人工知能や情報通信技術によって発展する自律的な機器のユーザインタフェース実現に向けて、人の意図スタンスならびにメンタライジングを促すことの重要性を述べた。またロボットという用語を用いて、意図スタンスを促すことのできる人工知能のためのユーザインタフェースの可能性について述べた。しかし、本稿の最初で述べた通り、ここでいうロボットとは機械の体を持ついわゆる「ロボット」に限定するのではなく、コンピュータグラフィックで描かれるエージェントや、スピーカを利用した音声エージェントといった、自律的に判断して動き、言葉や時にはジェスチャなどを用いて人とコミュニケーションできる人工物全般を指している。意図的な行動と時間ボリュームのあるタスクの関係が対象になること、意図スタンスをとることを促し機器に対するメンタライジングを人に行わせることは、機械仕掛けの体の有無が問題の本質とはならず、広い意味での擬人化を機器に対していかに人がするようになるかという問題の方が重要である。

擬人化に関して Epley らが興味深い研究を行っている<sup>4)</sup>。Epley らは、人が何かを擬人化する際の3要件について研究している。1つ目は、人間らしさを感じさせるデザイン要素 (elicited agent knowledge)、2つ目は、ほかのエージェントの振舞いを理解・説明するモチベーション (effectance motivation)、3つ目は、社会的な接触の欲求 (sociality motivation) である。1つ目は、人間の外見とまではいかなくとも、人として解釈しやすくなるデザイン要素 (たとえば、音声言語の使用もそれにあたる) の必要性を言っている。2つ目は、意図スタ

ンスとして解釈しやすくなる振舞いをシステムが示す必要があることを言っている。3つ目は、人は寂しくなると擬人化しやすくなるという知見であり、今回の本稿の主題とは関係がない要素である。

なんとなく可愛いからという理由やエンタテインメント性の向上の理由からロボットを導入するのは、ロボットはただのおまけにすぎない存在ということになってしまう。自律的な機器と人を繋げるユーザインタフェースを実現していく上でロボットを導入することは本質的に意義のある取り組みである。ロボット導入の真の狙いを理解していただけると筆者としては本稿を執筆した「意図」が達成されたと大変嬉しく思う。意図スタンス・メンタライジング・擬人化、これらを促すデザイン要素は徐々に体系化されてきており、今後、人工知能が社会に実装されていく中でますます重要になると考えられる。

### 参考文献

- 1) Dennett, D. : Intentional Systems, Journal of Philosophy, Vol.68, pp.68-106 (1971).
- 2) Baker, C. L., Jara-Ettinger, J., Saxe, R. and Tenenbaum, J. B. : Rational Quantitative Attribution of Beliefs, Desires and Percepts in Human Mentalizing, NATURE HUMAN BEHAVIOUR, Vol.1, No.64, pp.1-10 (2017).
- 3) Cohen, P. R. and Levesque, H. J. : Intention is Choice with Commitment, Artificial Intelligence, Vol.42, No.2-3, pp.213-261 (1990).
- 4) Epley, N., Waytz, A. and Cacioppo, J. T. : On Seeing Human : A Three-factor Theory of Anthropomorphism, Psychological Review, Vol.114, No.4, pp.884-886 (2007).  
(2018年5月1日受付)

今井倫太 (正会員) michita@ailab.ics.keio.ac.jp

1992年慶應義塾大学理工学部電気工学科卒業。1994年同大学院計算機科学専攻修士課程修了。同年、NTTヒューマンインタフェース研究所入社。1997年ATR知能映像通信研究所へ出向。2002年慶應義塾大学大学院理工学研究科後期博士課程修了。博士(工学)。2009～2010年シカゴ大学客員研究員。現在、慶應義塾大学理工学部情報工学科教授およびATR知能ロボティクス研究所研究員。2017年ドコモモバイルサイエンス賞社会科学部門優秀賞受賞。人型ロボットや自律エージェントと人のインタラクションの研究に従事。電子情報通信学会、人工知能学会、日本認知科学会、日本ロボット学会各会員。