

協調的パートナーを目指すインタフェースエージェント モデルの構築

Development of a Cooperative Interface Agent based on Team Cognitive Model

菅野 太郎 准教授 東京大学大学院工学系研究科

Tom Hope, 研究員, 独立行政法人産業技術総合研究所

Taro KANNO, Associate Professor, School of Engineering, the University of Tokyo

Tom Hope, Researcher, National Institute of Advanced Industrial Science and Technology

要旨: コンピュータ技術の発達に伴い便利な機械やソフトウェアが巷にあふれ、人間がこのような機械やソフトウェアを使う機会は益々増えている。これらの機械やソフトウェアの使い易さや安全性を確保するために、様々な人間-機械インタフェース研究がこれまでなされてきた。しかし従来の人と機械のインタフェースは機械を道具として捉えており、共通の目標を達成するための協調的なパートナーと捉える視点が欠如している。本研究では機械を協調的パートナーと捉え、従来の知覚・認知や行動意図を推論、推定に基づいた恩着せヘルプ型アプローチからさらに一步踏み込み、パートナーに必要とされる資質である、“状況の相互認識”、“意図の相互共有”の機能を持たせ、“自分の役割を見出し、阿吽の呼吸で人間を支援”できる機械の開発を目指すパートナー型インタフェース・モデルに関する研究を行う。

Abstract: Along with the introduction of advanced computer and software technology, we can find many useful tools and computer software anywhere in our daily life. Much effort has been put on making a machine more intelligent to enhance its usability and safety, but it still lacks the viewpoint of partnership from which we see such a machine as a partner working on common tasks. In order to realize more sophisticated human-machine interaction, it is necessary for a machine to become a true partner who can participate in cooperative activity, not by looking the user's behavior over his/her shoulder and imposing a self-righteous help on the user, but by proactively finding its role in their tasks and giving a nice help for the human partner. I believe the key requirement for such cooperation is that a machine possess a model for cooperative activity and can understand the human-machine cooperation (machine-in-the-loop cooperation). In this research, I will intend to construct a “partner interface” model that enables machines to share mutual awareness, team intention, and mutual belief in their tasks and the task environment in order to establish a true partnership between human and machine.

1. 研究目的

発電所の中央制御室や航空機のコックピット、航空管制、危機対応の現場など、高度な情報処理や判断を要求されるタスク環境では、タスクの複雑さや情報量の増加に伴い人間のタスク処理における認知的側面を支援するシステム・ツールの必要性が益々高まってきている。これまでも Associate Systems や Joint Cognitive Systems, Augmented Cognition、Human System Integration など様々なシステムコンセプトが提案され研究が続けられてきている。また日本では特に、擬人化エージェントを利用したユーザインタフェースや対

人ロボット、ペットロボット等の研究が盛んに行われており、人間社会をより豊かにする実用的、日常的なヒューマンエージェントインタラクション (HAI) の実現に期待が寄せられている。

人と機械 (エージェント) の協調性、親和性を高めるための一有力候補は人間-人間に見られるインタラクションの方法を HAI 設計に応用することであろう。この方向性を目指すべく、ユーザー状態 (意図やプラン、心的状態等) の理解に基づいた支援や機能設計に関する研究が数多くなされている。しかしながら多くの研究において期待するほどの interactivity は得られていない。これらの研究に共通する問題の一つは、協調的インタラクシ

ョンの概念が不十分、欠如しており、ユーザーの理解とそれに基づく機能設計に終始している点にある（エージェントが協調の外に置かれている：out-of-the-loop）。実際の人間-人間協調には相手の理解だけでなく、相手の自分に対する理解や、我々の理解、それらを共有する過程、それに基づく行為の織り交ぜが存在し（in-the-loop）、HAIにおいてもこの”We-ness”を生成するプロセスを積極的に扱わねばならない。本研究では、人間における協調の深淵メカニズムを、個人の認知プロセスと相互信念の観点から捉え、チーム協調における認知プロセスモデル（三層モデル）を提案する。

2. 研究経過

人間-人間における協調作業の観察、観察データを相互信念の観点から分析し、チーム構成員間での認知プロセスの共有・修正方法（インタラクションジェノタイプ）を抽出した。また、これらのモデルを人間-機械系に適応するためにエージェントモデルとして実装し、協調インタラクションのシミュレーションを行った。さらに、UA デザインのための認知モデル（Simplex2）と相互信念モデルを統合したインタラクションシステムの評価手法を開発した。具体的内容を以下に示す。

2.1. パートナー認知モデル

発達心理学における心の理論や、哲学における協調行動における意図に関する議論や reflexivity モデル、脳科学におけるミラーニューロンの発見、これらの理論や生理学的知見が示す人間の認知行動の特質は、人には再帰的に他人の心的状態を理解する仕組みが備わっているということである。我々はこのような再帰性を相互信念の概念を用いて人の様々な心的状態や認知プロセスに適用することで、図1に示すような協調行動における認知（チーム認知）の概念モデルを提案した。図1の各層は上から、1）主体となる人（エージェント）の認知（モデル）、2）相手の認知に対する信念（モデル）、3）相手の自分の認知に対する信念（モデル）、をそれぞれ表しており、各層には状況認識（SA）や知識、メンタルモデル、感情といった様々な心的状態や認知的コンポーネントを含み得る。各層には、例えば1層目では知覚-状態認識-意図形成-行動といった認知プロセス、2、3層の信念層では信念を獲得するための推論プロセスがあり、また、各層を参照・比較し、各プロセス・状態を修正、補完するといったメタ認知的操作も存在する。これらの1）相互信念構造と2）心的状態・認知コンポーネント、3）認知・推論プロセス、4）メタ認知操作、から概念モデルは構成され、

これらの各状態、プロセスの結果を基にインタラクションが駆動される様をモデル化した。

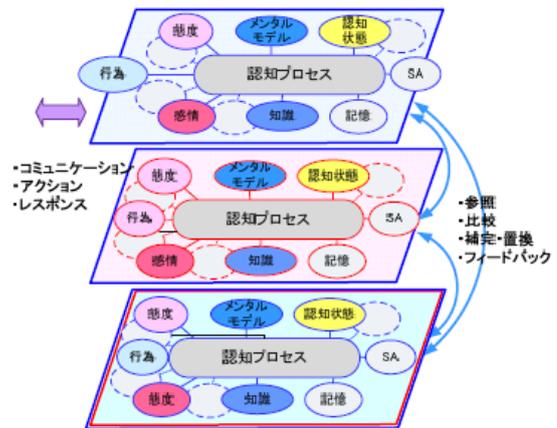


図1. パートナー認知モデル

2.2. インタラクション Genotype

HAI におけるインタラクション生成を考える上で発話内容や発話の機能（遂行動詞）といった観察可能な要素とそれを生成する理由・原因、機構を区別することは重要である。本研究ではそれぞれインタラクション Phenotype、Genotype と呼ぶ。概念図を図2に示す。上二段が Phenotype、下二段が Genotype にあたる。本研究では、二人組で協働タスクを行う実験を行い、実験中のビデオ・音声記録を行い、さらに実験後に実験中のビデオを被験者に見せ、インタラクション（コミュニケーション）の背後にある理由を相互信念の観点から説明させたり、観察者に背後理由を推論させたりすることでインタラクション Genotype の抽出・分類を行った。結果を表1に示す。

表1. インタラクションGenotype

Genotype		Phenotype (遂行動詞)
分類	理由	
自身の認知プロセスや相手の心的状態・プロセスの推論のため	- 必要・十分な情報や知識がない - 相手の心的状態・プロセスの推定に自信がない、など	Query Confirm
相手の認知プロセスや推論プロセスを助けるため	- とにかく共有しておく - 相手が必要・十分な情報がない - 前もって情報を提供しておく、など	Inform
相手の認知プロセスや推論を修正するため	- 齟齬の回避・回復 - 誤解の訂正、など	Inform Query Confirm

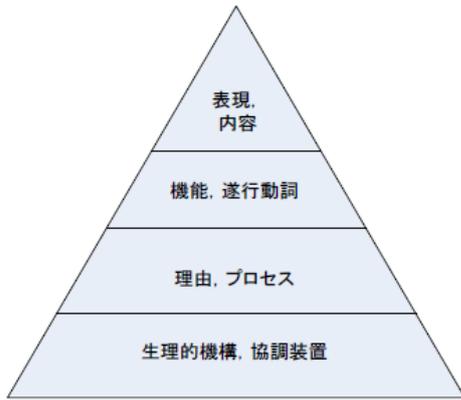


図2: インタラクション Genotype と Phenotype

2.3. モデルの実装

不確かで限定的な情報に基づく状況認識における人間の認知・推論を扱うために、認知・推論プロセス、つまり提案モデルの各層をベイジアンネットワーク(BBN)を用いて実装した。各3層BBNを編集することで知識分散等の個人差を表現できる。シミュレーションでは状況認識を想定した抽象タスクに関するBBNを作成した。各ノードは観測可能な徴候・事象およびその状態を、リンクはノード間の因果関係を表している。各ノードの確率は事象生起に対する確信度を表し、確率値で認知・推論プロセスを定量的に扱うことができる。また、思い込みや補完、整合性の評価といったメタ操作も各層BBNの操作によって実装した。

2.4. インタラクションシミュレーション

提案モデルの機能評価として二組のエージェントA, Bによる状況認識を例題とし、各エージェントが持つ個性と生成されたコミュニケーションによってチーム認知(状況認識共有)がどのように変化するかをシミュレートした。徴候の観察行動はシナリオに記述し、各エージェントが断片的に情報を収集しながら互いにコミュニケーションを行うことで二者間の共有状況認識が醸成されていく。評価指標として相互信念を考慮した正解度と共有度を考案した(式1, 2はその平均を表している)。Uは各エージェントの各層において生起認識されている状態ノードの集合を表している。表2に示すようなエージェントの個性をもつチームを定義し、それぞれの共有パフォーマンスを調べた結果を図3に示す。

チーム条件Bは、最も高い共有度を持つが正解度が高くない。これは常に相互に補完しあうことによって分散した知識を生かしていないためと考えられる。条件Cは総合的にパフォーマンスが高い。これらは個性の組合せによるチームパフォーマンスへの影響を表している。図4では別設定に

おける、2層目の信念の各エージェントの網羅性(B2: ピンク, A2: 水色)と正確さ(B2: 赤, A2: 青)の時間変化及び、生成されたコミュニケーションの意図と解釈を示している。注目する点は、ステップ14, 15で意図に沿った解釈を行った結果エージェントBの信念の網羅性が低下しているが、その後ステップ15, 16で、3層目を用いることで信念の共有度の低下を回復している点である。相互信念の有効性を示している。

$$accuracy = \left\{ \frac{U_{a1} \cap U_0}{U_0} + \frac{U_{b1} \cap U_0}{U_0} \right\} \times 1/2 \quad (1)$$

$$sharedness = \sum_{a,b} \left\{ \frac{U_{self2} \cap U_{partner1}}{U_{partner1}} + \frac{U_{partner1} \cap U_{self2}}{U_{self2}} \right\} \times 1/2 \quad (2)$$

表2. エージェントの性格

Team	説明
A	常に自分の認識(U ₁)を信じる
B	常に相手(U ₂)を信じる
C	詳細知識の保持者が正しいと判断

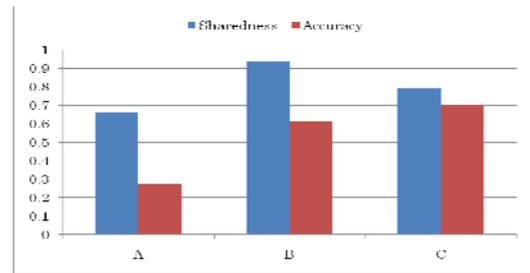


図3. 正解度と共有度

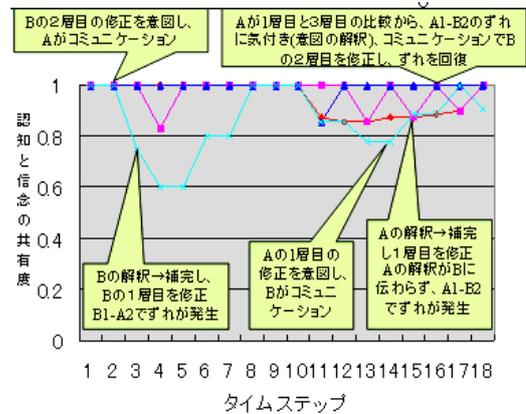


図4. インタラクションプロセス例

2.5. インタラクションシステム評価の開発

UA (Universal Access) デザインのためのユーザー

モデルである Simplex2 の 9 つの認知要素（知覚、フィードバック, Executive Function, 作業記憶, 長期記憶, メンタルモデル, 感情, 出力, 出力系列）と相互信念モデルを統合したインタラクションシステム評価マトリクスを開発した。各 9 つの認知要素に対し、相手の認知要素を推論・推定する際の自身の各認知要素を評価する 9×9 の質問票を開発した。“「対象システムを通して相手とインタラクションする際の〇〇を評価」”を評価票の基本の問いとし、これを通して包括的な評価項目を提供する。例として、相手の知覚を推定するために必要な自身の長期記憶に関する評価の質問項目をいくつか下に挙げる。

6.1 相手の知覚に関して

- 相手の知覚を推論するために何らかの事前知識や経験が必要か？
- 相手の知覚を推論するためにどのような事前知識や経験が必要か？

6.2 相手が得ている（操作に対するシステムからの）フィードバックに関して

- 相手が知覚しているフィードバックを推論するために何らかの事前知識や経験が必要か？
- 相手が知覚しているフィードバックを推論するためにどのような事前知識や経験が必要か？

6.3 相手のワーキングメモリに関して

- 相手のワーキングメモリ（注意や負荷など）を推論するために何らかの事前知識や経験が必要か？
- 相手のワーキングメモリを推論するためにどのような事前知識や経験が必要か？

3. 成果

本助成期間中に得た主な研究成果は以下の通りである。

3.1. パートナーエージェントモデル

認知実験から人間の協調の背後にある機構を抽出し、相互信念モデルを用いて表現することで協調における認知モデルを開発した。提案モデルを BBN を用いて実装し、概念モデルが計算モデルとして実現可能であることを示した。提案モデルを実層したエージェントを用いて、協調プロセスにおけるインタラクションシミュレーションを行い、提案モデルの機能評価を行った。人間の協調に特徴的な相手の心的状態を考慮したインタラクションの生成や、共有状態の再現ができた。

3.2. インタラクションシステム評価

UA デザインのための認知モデルである Simplex2 の 9 つの認知要素に対する相互信念獲得を評価することによって対象インタラクションシステムを評価する手法を開発した。

4. 今後の課題と発展

パートナーモデル開発とその計算機上での表現の実現可能性は示された。今後は具体的な課題を用いたインタラクションエージェントの開発・評価が必要となろう。既に開発過程で個別課題への拡張可能性を見据えたモデル実装を行ってきた。今後はインタラクションエージェントの汎用モデルとしてパッケージ化し、WEB上での配布等を検討している。

インタラクションシミュレーションは、インタフェースエージェントとしての応用のみならず、人間のチーム協調のプローブツールとしての応用可能性もある。人間の認知的特性等をシミュレーションで設定し再現することで、フィールドワークや実験では実現できない様々な仮定下でのパフォーマンス評価が可能となると期待できる。

インタラクションシステム評価のためのデザインマトリクスに関しては、英国のHCI研究者と連携し具体的なシステムへの適用を既に計画している。システム評価のみならず、文化依存性評価等への発展も検討している。継続的な研究と連携が今後の課題となる。

5. 発表論文リスト

1. Kanno T., Watanabe A., and Furuta K., 2009, Modeling and Simulation of Human Interaction based on Mutual Belief, Proc. Conf. Human Computer Interaction International, (to be appeared).
2. Kitahara Y., Hope T., Nonose K., Kanno T., and Furuta K., 2008, Developing an understanding of genotypes in studies of shared interaction, Proc. 2nd Int. Conf. Applied Human Factors and Ergonomics, CD-ROM.
3. Kanno T. and Furuta K., 2007, Interaction Strategies for Smart AugCog. Proc. 3rd. Int. Conf. Augmented Cognition/51th Annual Meeting of Human Factors and Ergonomics Society, 266-271.
4. Kanno T., 2007, the Notion of Sharedness based on Mutual Belief, Proc. 12th . Int. Conf. Human-Computer Interaction, 1347-1351.
5. 北原, 野々瀬, 菅野, 古田, 2009, ヒューマンインタフェースシンポジウム講演集, (発表予定) .
6. 菅野, Hope, 飯塚, 古田, 2008, 相互信念モデルに基づくインタラクションシミュレーション, HAIシンポジウム2008予稿集, 2A-2.
7. 飯塚, 菅野, 古田, 2008, チーム協調におけるコミュニケーションのシミュレーション, ヒューマンインタフェースシンポジウム講演集,855-860.