

目的意図と作業意図間の関係性ネットワークを用いた循環的推定 Alternating Estimation of Local Objective and Global Purpose by Relational Network

松本 麻見^{*1}
Matumoto Asami

大本 義正^{*2}
Ohmoto Yoshimasa

西田 豊明^{*2}
Nishida Toyoko

^{*1} 京都大学工学部情報学科

The School of Informatics and Mathematical Science, Faculty of Engineering Kyoto University #1

^{*2} 京都大学情報学研究科

Graduate School of Informatics, Kyoto University #2

This study is aimed at the realization of a robot that can perform a project-type task which is composed of various sub-tasks that have dependencies to achieve an overall global goal. In order to increase our familiarity with robots, it is required that they can perform such a task. In project-type tasks it is essential to set local objectives and a global goal while retaining the integrity and linkages between them. In the method proposed in this paper, human intention is described as a two-layer model which includes local objectives and a global goal and defines the relationship between them. Next, in order to confirm the usefulness of the proposed model, we evaluate and verify the overall satisfaction and stress level of participants, through a questionnaire and physiological indices. Participants collaborating with the agent possessing the two-layer model were found to have less stress during the task, and higher satisfaction.

1. はじめに

近年、拡張現実上の仮想エージェントやロボットなどが実生活の中に入って来ている。このような状況において、embodied conversational agents (ECAs) は、単なるユーザインタフェースではなく、将来的には人間の活動をサポートするパートナーのような存在になることが期待される。そのようなエージェントは、特定のタスクに特化したものではなく、大きな目標を踏まえた上で、関連する様々なタスクを横断的に、時には独立してタスクを遂行し、時には人間と共同でタスクを遂行しながら、サポートするものになると考えられる。

お互いに関係性のある複数のサブゴールをもつ作業を通じて、一つの大きな目的を達成するタスクは、日常生活の中にもたびたび登場する。例えば、部屋の模様替えや買い物などである。このようなタスクを、本稿ではプロジェクト型タスクと呼ぶ。プロジェクト型タスクにおいて、人間は最終的にどのような成果を得たいかという漠然としたイメージを持っている。また、その成果を得るためにはどのような行動をとれば良いかを考えながらサブの作業を行っている。一方で、必要であれば、最終的な目標を修正しつつ、個々の作業における目的を変更したり、位置づけを変えたりすることもある。こうした特徴を持つ、プロジェクト型タスクにおいては、全体の意図と個々のタスクの意図の関係性を踏まえて、お互いを柔軟に調整しなければならない。したがって、プロジェクト型タスクにおいては、タスク全体での意図(目的意図)と個々のタスクでの意図(作業意図)との整合性や連動性を確保しておく事が必要不可欠である。

本研究では、このようなプロジェクト型タスクを遂行できるエージェントを作成するために、タスク遂行のパートナーの作業意図と目的意図の整合性と連動性を確保して推定する手法を提案する。意図推定に関する研究として、横山ら[横山 2009]は他者の意図/行動予測モデルを用いてその意図の推論を行っている。また、周藤ら[周藤 2013]は、トップダウン手法とボトムアップ手法の組み合わせにより意図推定を行っている。しかし、これらの意図推定に関する研究では、推定するのは単一の意図であった。

しかしプロジェクト型タスクでは、単一の意図を推定するのではなく、互いに依存関係にありながらも別々に存在している二種類の意図を推定し、お互いに影響を与えながら矛盾を生じさせない形で柔軟に変更する必要がある。本研究では、作業者が重視するポイントを記述した重視要因として意図を二層で記述し、それらの間関係性を定義しておくことでそれぞれの意図を循環的に推定する手法の提案を行う。この手法を用いることにより、ロボットがプロジェクト型タスクにおいて適切に作業を行ったり、アドバイスをしたり出来るようになると期待される。

2. 循環的意図推定

プロジェクト型タスクを遂行するためには、個々の作業における意図(作業意図)と、共同作業全体の意図(目的意図)の間の、整合性や連動性を保つ必要がある。そのために、作業意図の決定には目的意図からの影響を、目的意図の決定には作業意図からの影響を、それぞれ受けることができる、行動の推定・提案手法を提案する。

2.1 目的意図と作業意図間の関係ネットワークの概要

関係性がある様々なタスクやインタラクションを通じて、全体として大きな共同作業を達成する、プロジェクト型タスクには、個々の作業を遂行する local フェーズと、それらを統合してできる全体像の検討を行う global フェーズが存在する。循環的意図推定では、それぞれのフェーズにおいて、目的意図と作業意図の間で影響を伝播させる。そのため、目的意図と作業意図の関係性を、関係ネットワークとして記述する。(図1)

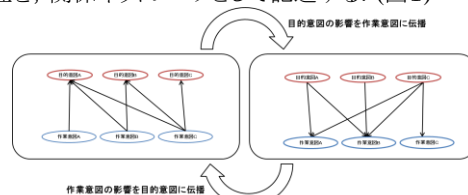


図 1: 関係ネットワークの概要

この図では、作業重視要因から作業意図を推定するレイヤー(local layer)と、目的重視要因から目的意図を推定するレイヤー(global layer)の 2 層になっていることが示されている。local

フェーズでは local layer の, global フェーズでは global layer の, それぞれの重視要因のノードが, 外部からの入力に基づいて重み付けを変更される. その後, システムの設計者によってデザインされたネットワークに基づいて, 抽象化された項目である各レイヤーの意図のノードの値を変更する. この意図のノードの値が, 協調する相手が持っている行動の意図と推測される. システムは, 推定された相手の意図のノードの値に基づいて, 協調的な行動の計画を行う.

本研究で提案するシステムでは, 作業と目的のレイヤーを, あらかじめ設計された構造で接続している. 各フェーズではそれぞれのレイヤーの意図を更新するが, 次の行動計画を行う前に, 更新された意図のノードから, もう一方のレイヤーの重視要因のノードに, あらかじめ設計された値を出力する. 入力を受けたレイヤーでは, 重視要因のノードの重みの変更に基づいて意図のノードの値が変更される. そして, もう一方のレイヤーの意図のノードから, 再び現在のフェーズで注目しているレイヤーの重視要因のノードに入力が入り, 意図のノードの値を更新する. こうして, もう一方のレイヤーからの入力によって, 意図のノードの値を更新することで, 互いの整合性・連動性を保つ.

この図では, 下から上へ影響が伝搬するように書かれているが, 図の一番下のノードと一番上のノードは同じものを示している. つまり, 全体として影響は循環するように設計されている. 図中の大きな矢印は, 目的意図から作業意図, 作業意図から目的意図への影響伝播を順次繰り返すことを示している. これにより, 目的意図と作業意図間の整合性・連動性が保たれる.

2.2 意図伝播のプロセス

本研究で提案するシステムを用いた協調作業における, それぞれの意図の伝搬プロセスを図2に示す. 図の左側が global フェーズ, 右側が local フェーズにあたる. 以下では, 共同撮影タスクを例に用いて概要を説明する.

global フェーズにおいて, エージェントがあるイメージに基づいたおすすめの動画を人間に見せると, 人間はそこからボールの動きのダイナミックさ等の目的重視要因を抽出する. 抽出した目的重視要因と自らの持っている映像のイメージとを比較し, 差分があればロボットに伝える. エージェントは人間の発言から目的重視要因を抽出し, 関係する目的意図を推定する. 推定された目的意図を達成するために行われるであろう作業重視要因に注目し, 実際に行われた作業と比較してどの要因を重視しているかを確認する. その後, 確認された作業重視要因に関係する作業意図を推定し, 作業意図が反映されているであろう目的重視要因を列挙する.

local フェーズにおいて, 人間から個々のシーンにおける細かい指示があった場合はそこから作業重視要因を抽出し, エージェントが自らの作業意図との比較を行う. そこで差分があればその内容を人間に伝える. 十分に差分が小さくなれば, 作業重視要因から作業意図, 目的重視要因, 目的意図, 作業重視要因の順に影響を上記と同様に伝播させ, その結果得られた作業意図が妥当なものであるか人間に確認を行う. そこから人間は作業重視要因を抽出し, 自らの作業意図との差分をエージェントに伝え, エージェントが作業重視要因を受け取った場合の処理に戻る, というループを繰り返す.十分に繰り返しが行われれば, 始めに戻って推定を続ける.

以上のようなプロセスを, 何回も繰り返し行われるタスク中に, 継続して行うことにより, 実際の作業中に変化する人間の意図を追跡しながら, その影響を適切に伝搬させることができると考えられる.

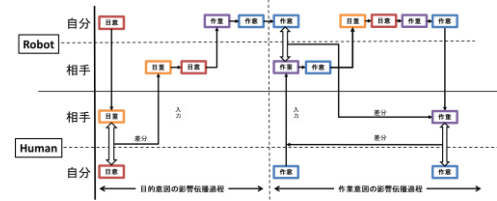


図 2: 提案手法を用いたインタラクションの流れ

2.3 循環的推定システム

本システムでは, 意図の循環的推定に加えて, 推定された重視要因の分布をふまえた編集映像列の出力, インタラクション中の話題決定, 次回撮影列候補の決定を行う.

詳細を図3に従って, 順次説明する. まず, 撮影フェーズで撮影された映像について, シーンごとにカテゴリを入力する. その入力を受け, 編集結果を出力する. 編集列が決定すれば, 推定した意図モデルと編集された結果から編集結果を評価する際のインタラクションにおける話題を決定・出力する.

次に, 実験参加者の発話中のキーワードや生理指標から目的[作業]重視要因が入力されると, 2.2 節の流れに従って影響の伝播を計算し, 意図モデルに反映する. 次回撮影列候補は推定された意図モデルに基づいて出力される. また, 意図モデルの変動に従って逐次グラフを出力する.

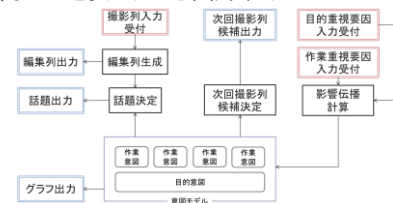


図 3: システム構成

3. 評価実験

本研究で提案する目的意図と作業意図の循環的推定がプロジェクト型タスク遂行中のストレス, 共同作業者の最終的な満足感に与える影響を検証するため, 実験を行った. 実験中は映像の撮影, および実験参加者の生理指標のリアルタイム計測を行った. 本実験では, 目的意図と作業意図が関係ネットワークで接続されたエージェント(伝搬エージェント)と, それぞれの意図が接続されていないエージェント(独立エージェント)との比較を行った. 伝搬エージェントは, 前の節で説明した関係ネットワーク利用した意図推定の伝搬を行った上で, 自らの撮影のプランニングと参加者への助言を行った. 独立エージェントは, 作業意図のみを参照して, 自らの撮影のプランニングと参加者への助言を行った. 実験後にエージェントへの印象と, 撮影された映像に関するアンケートを行った. アンケートは, 7段階のリッカート尺度と自由記述で回答してもらった. 影響の評価は, アンケートと実験中に計測した LF/HF の変化によって行った. タスクは人間とエージェントとのからくり装置の共同撮影タスクである. 実験参加者は情報学系の所属ではない大学生 20 名(実験群: 男性 8 名, 女性 2 名 対照群: 男性 8 名, 女性 2 名)である.

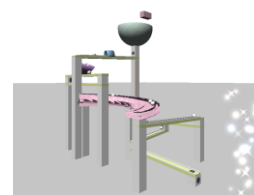


図 4: 仮想空間上からくり装置

3.1 実験環境

実験に用いるシステムは、図5のように、実験者が使用する撮影システム、実験参加者が見ている画面を管理するためのWOZ操作システム、音声通話システム、生理指標のリアルタイム計測システム、循環的推定システムにより成る。実験者は循環的推定システムと参加者のリアルタイム生理指標データをふまえ、音声通話システムを用いてエージェント音声のWOZ操作を行う。また、参加者は撮影システム上で、撮影・編集映像の視聴を行うが、参加者画面の適切なタイミングでの切替は、実験者が撮影管理用WOZ操作システムを用いて行う。

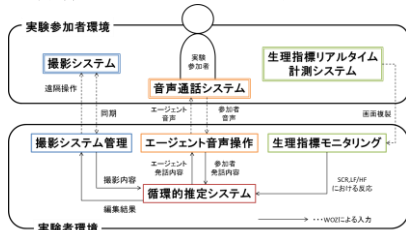


図 5 : 実験環境

3.2 実験手順

まず始めに、撮影環境や既存の映像作品、からくり装置やコントローラの操作方法を提示・説明し慣れてもらった。その後、生理指標の計測を開始し、共同撮影タスクを実行してもらった。実験では、撮影から映像出力までの流れを4回繰り返した。一回の流れは3つのフェーズに分割された。それは、実際の撮影を行う撮影フェーズ、撮影された映像の組み合わせによって最終的な映像を作成する編集フェーズ、次回の撮影のやり方を相談する計画フェーズである。最後にアンケートを記入してもらい実験を終了した。

3.3 実験で定義した関係ネットワーク

本実験では、予備実験における検討と『一人でも出来る映画の撮り方』[西村 2013]を参考にして図6のような関係ネットワークを構成した。

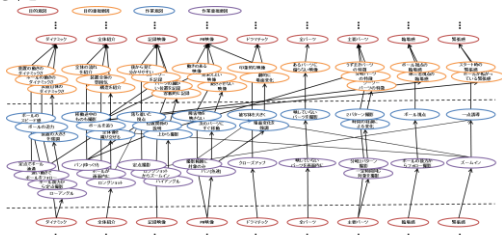


図 6 : 本研究で用いる関係ネットワーク

3.4 結果

(1) 生理指標に関する分析

本実験のタスクにおいて、参加者にかかっていたストレスを、LF/HF の値から分析した。これは、様々な先行研究において、LF/HF が人間の作業中のストレスを検出する手がかりになると報告されているためである([Nater 2006], [南谷 1999]など)。Nater らの研究において、ストレスがかかっている場合に 5.0 以上の値をとっていたため、LF/HF の値がそれを超えた際にストレスがかかっていると見なした。参加者ごとにタスクの遂行時間は異なるため、ストレスがかかっている時間をタスク遂行時間で割ったものをストレス反応時間割合と呼び、ストレスの指標とした。撮影フェーズ・編集フェーズ・計画フェーズを含めた全体と、各フェーズごとのストレス反応時間割合を、各サイクルごとに示したものをそれぞれ以下に示す。これらのデータは正規分布に従わなかったため、検定にはノンパラメトリックな手法を用いた。

まず撮影フェーズ・編集フェーズ・計画フェーズを含めた全体のストレス反応時間割合の推移を図7に示す。この図から、全体的に見ると実験群・対照群ともに反応割合が増加傾向にあることが分かる。これは、実験を長時間続けたことによるストレスに起因すると考えられる。1 サイクル目の値と、3 サイクル目の値をウィルコクソンの符号順位和検定にかけたところ、実験群は $p=0.33$ であり、対照群は $p=0.051$ であった。つまり、対照群のみ、1 サイクルから 3 サイクルにかけて LF/HF の値が変動していることが確認された。このことから、伝搬エージェントとインタラクションしている参加者は、時間の経過によるストレスの増加が抑えられていることが示された。

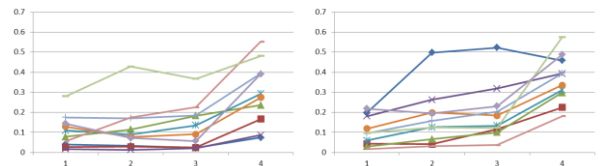


図 7 : LF/HF 反応割合(全体)

続いて、各作業フェーズにおけるストレスの受け方について、分析を行った。

撮影フェーズのストレス反応時間割合の推移を図8に示す。この図を見ると、実験群では、おおむね 0.1 以下の値で安定している一方、対照群では 0.2 以下の範囲で激しく変動していることが分かる。ストレス反応時間割合の変動幅が対照群の方が大きいかどうかをマンホイットニーの U 検定にかけたところ、対照群の方が大きな変動をしている傾向が示された ($P=0.094$)。これは、参加者が、独立エージェントとインタラクションを通じて仕上げたいイメージと実際の撮影との対応を把握出来ないために、撮影中にシーンごとの撮影手法について考えながら撮影を行う必要があったためではないかと考えられる。

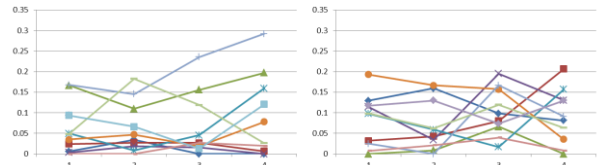


図 8 : LF/HF 反応割合(撮影フェーズ)

次に編集フェーズのストレス反応時間割合の推移を図9に示す。この図を見ると、実験群・対照群ともにほとんど 0.2 以下の値で横ばいになっていることが分かる。マンホイットニーの U 検定の結果、平均値に差は見られなかった。これは、編集フェーズにおいて映像を評価する際に共同作業者は客観的な視点で評価を行っているため、ストレスを受けにくかったと考えられる。

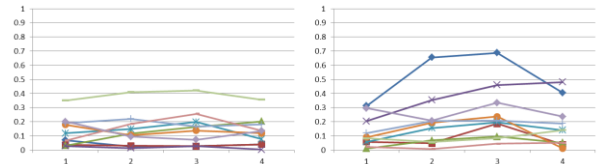


図 9 : LF/HF 反応割合(編集フェーズ)

最後に、計画フェーズのストレス反応時間割合の推移を図10に示す。ここでサイクル数が3になっているのは、最後の撮影の後には次回の計画フェーズを行わないためである。この図を見ると、実験群では2例を除き、ほとんど 0.1 以下の値で横ばいになっているが、対照群ではほとんどが 0.3 までの間を推移している。マンホイットニーの U 検定の結果、対照群の方が平均値が

大きい傾向があった ($p = 0.054$)。これは撮影フェーズと同様に対照群の共同作業者が、映像の全体イメージを考慮した各シーンでの撮影がどのようになるのかを自分でもよく分からないまま次回撮影について話し合っていることにより生じるストレスに起因していると考えられる。

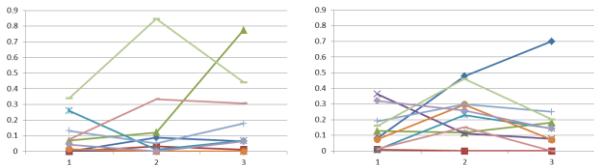


図 10 : LF/HF 反応割合(計画フェーズ)

(2) アンケート結果

参加者は各項目において、7 段階で評価した。設問ごとに、マンホイットニーのU検定を行った結果を図11に示す。

「エージェントが 1 つの映像に対して一貫したイメージを持っていると感じましたか」という問いにおいて、伝搬エージェントの方が有意に高い評価が得られた ($P=0.074$) ことから、エージェントの助言や作成された編集映像における一貫性が確保されたと考えられる。

「エージェントが意図を持って映像を提示していると思ったか」という問いにおいては、平均値に差があったものの (5.0 vs 4.2), 有意な差はなかった。これと上の結果から、エージェントは一貫した提案を行っているものの、自ら固執するものではなく、参加者の望むと考えたものを提示していると受け止められたと考えられる。

「最終的に仕上がった映像に満足しましたか」という問いに関しては、伝搬エージェントの方が平均で 0.8 ほど高くなったものの (5.1 vs 4.3), 有意な差はなかった。また、どちらのエージェントについても 4 以上であったことから、参加者はタスクの結果にある程度満足していたことがうかがえる。

「エージェントは人間らしく感じましたか」という問いに関して、伝搬エージェントの方が評価が高い傾向にあった ($P=0.074$)。この理由として、伝搬エージェントは、目的意図から影響を受けるため、各シーンにおいて参加者が最も好む撮影手法を採用する訳ではない。しかし、インタラクションを通して全体イメージと各撮影手法のすり合わせを行う過程で、エージェントが考慮している目的意図を感じ取り、より人間とインタラクションしている感覚になったと考えられる。実際、実験群の自由記述欄では「お互い意見を出して話し合いをしている実感があった」という回答が見られた。

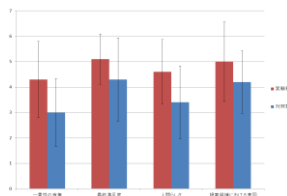


図 11 : アンケート結果

4. 議論

以上の結果をまとめる。LF/HF の分析から、参加者は伝搬エージェントとのインタラクションでは受けるストレスを低く抑えることができた。一方、独立エージェントとのインタラクションでは、作業の計画においては比較的高いストレスを受け、作業中においては状況に依存して受けるストレスが大きく変動している。このような差が生まれた原因としては、伝搬エージェントは、各フェーズにおいて参加者の意図を推定した際に、循環的にその影響

を伝搬させることで、作業意図と目的意図の間での整合性と連動性を確保できたためであると考えられる。アンケートの結果からも、参加者は伝搬エージェントの方がより一貫した態度を持っていると感じている。また、アンケートにおいても、参加者は伝搬エージェントの方がより人間らしいと感じており、このような感覚がストレスを低減させている一因と考えられる。

本研究において提案した循環的意図推定システムは、プロジェクト型タスクの遂行において有用に働くことが示唆された。このシステムは DEEP をベースに構築されているため、ユーザとエージェントとのインタラクションを繰り返すことで、お互いの目的意図と作業意図を共有していくことを目指している。しかしながら、今回の実験では、何度も言葉をやり取りするような活発な議論が生じる場面は少なく、将来的に、より積極的な議論を促すような仕組みを実装する必要があると考えられた。

また、実験後に行った関係ネットワークに関するアンケートでは、関係ネットワークのいくつかの接続部分において、それらの関係性を見直す必要があることが分かった。今後それらの部分を修正することでより精度の高い関係ネットワークが得られることが期待される。

5. まとめ

本研究ではプロジェクト型タスクを円滑に遂行するために重要である、タスク全体を通しての意図と各タスクにおける意図の整合性並びに連動性を確保できる、協調的な意図推定システムの構築を目指した。そのために、2 層で記述された作業意図と目的意図を、関係ネットワークを介して接続し、それぞれの意図を推定する際に、お互いの影響を伝搬させる手法を提案した。

この手法の有用性を検証するために、意図を伝搬させるエージェントと、独立に推定するエージェントのそれぞれとインタラクションを行いながらプロジェクト型タスクを遂行する評価実験を行った。生理指標やアンケートの結果から、提案手法は、作業意図と目的意図の整合性と連動性を確保し、実際の作業における人間のストレスを軽減することができる点で、有用であることが示された。

将来課題としては、タスク中のインタラクションを増加させることで、共同作業者にエージェントが議論の対象であると見なしてもらうことなどがあげられる。

参考文献

- [横山 2009] 横山 絢美: 協調課題における意図推定に基づく行動決定過程のモデル的解析, 電気情報通信学会論文誌, 2009.
- [周藤 2013] 周藤 沙月: テーマ表現提示による目標明確化機能を持つ対話型提案システムの実, 京都大学大学院情報科学研究科修士論文, 2013.
- [西村 2013] 西村 雄一郎: 一人でも出来る映画の撮り方, 洋泉社, 2013.
- [Nater 2006] Nater, Urs Markus and La Marca: Stress-induced changes in human salivary alpha-amylase activity? associations with adrenergic activity, Psychoneuroendocrinology, Elsevier, 2006.
- [南谷 1999] : ストレス・疲労にともなう心拍拍動—ニューラルネットによる自律神経活動の評価—, 電気情報通信学会技術報告, 1999.