

卒 業 論 文

題 目

行動情報と発話情報の組み合わせによる
信念と欲求の逐次的推測モデルの検討

指導教授

今井 倫太 教授

令和2年度

慶應義塾大学 理工学部 情報工学科

渡邊 悠太 (61720736)

論文要旨

学科	情報工学	学籍番号	61720736	氏名	渡邊 悠太
(論文題目) 行動情報と発話情報の組み合わせによる信念と欲求の逐次的推測モデルの検討					
<p>本研究の目的は、人間にストレスを与えること無い、人間と長期的に共存可能な対話システムの実現である。しかし、対話システムは相手の心的状態を考慮することなく対話をすることがあり、対話相手に不自然さやストレスを与えることがあるため、対話システムは相手の心的状態を推定することが重要である。そこで、本論文では心的状態推定アルゴリズム MIoM を提案する。MIoM は、行動情報と言語情報の両方を用いて心的状態を推定することにより、対話システムが人間に与える不自然さやストレスを軽減することの手助けとなる。本実験では、本研究で作成した、行動情報や言語情報を含むデータセットを用い、MIoM による心的状態の推定を行った。人間評価とアブレーションスタディにより、行動情報と言語情報の両方を心的状態の推定に活用することが有効であることが示された。</p>					

目次

1 序論	1
2 関連研究	3
2.1 行動情報から心的状態を推定する研究	3
2.2 発話情報から心的状態を推定する研究	3
2.3 単一情報による心的状態推定の問題点	3
3 提案	4
3.1 アルゴリズム	4
3.2 システム構成	4
4 評価	5
4.1 実験設定	5
4.2 実験手順	5
4.3 実験結果	7
5 考察	8
6 今後の課題	9
6.1 単語埋め込みモデルの検討	9
6.2 対話相手の発話に対する応答生成	9
6.3 実世界情報による心的状態の推定	9
7 結論	10
謝辞	11

図 目 次

3.1	MIoM のシステム構成	4
4.1	本実験における環境	6

表 目 次

4.1 人間による推定と推定モデルの相関	7
--------------------------------	---

第1章

序論

対話システムは、発話解釈と発話生成の両方において発展を遂げており、我々の生活に浸透しつつある。しかし、対話システムとの対話は不自然さやストレスを与えることも少なくない。本研究の目的は、人間に不自然さやストレスを与えることが無い、人間と長期的に共存可能な対話システムの実現である。

対話システムと人間との対話では、対話システムが対話相手の心的状態を推定し、それを考慮した対話を行うことが重要である。人間は、気分が落ち込んでいる対話相手の発話に対してネガティブな発話解釈をしたり、励ましの言葉をかけるように、対話相手の心的状態によって相手の発話の解釈を変えたり、自身の発話の内容を変えている。また、相手が知らないことについて詳しく説明したり、相手が好むことについて話を掘り下げる。対話システムにおいても、人間と同様に対話相手の心的状態に合わせて、発話解釈を変えたり、自身の発話の内容を変えることで、より自然な対話を実現することができる。つまり、対話相手に合わせた臨機応変な発話解釈や発話生成により対話における不自然さやストレスをなくすためには、対話相手の心的状態を推定することが重要である。

人間の心的状態を推定する研究は、人間の行動情報から心的状態を推定する研究と発話情報から心的状態を推定する研究が存在する。人間の行動情報から心的状態を推定する研究は、代表的には、Baker et al. の研究がある。Baker et al. は、環境の状態と環境中を移動する人間の行動や観測状況、信念をベイズ推論に適用し、環境中を移動する人間の信念と欲求を推定している。発話情報から心的状態を推定する研究では、代表的に高橋らの研究がある。発話から得られた事象を信念と捉え、考えられる欲求の候補を生成し、尤もらしい欲求を基に発話者の意図を推定する。また、心的状態を考慮することによる検索精度の向上ために、検索クエリへの入力から人間の心的状態を推定する研究も存在する。

従来研究では、行動情報のみから人間の心的状態を推定する研究や発話情報のみから人間の心的状態を推定する研究は存在した。しかし、行動情報と発話情報の両方から人間の

1. 序論

心的状態を推定する研究はない。従来研究における心的状態の推定は、行動のみが観測される場合や発話のみが観測される場合には有効であるが、行動と発話の両方が観測される場合において、行動による発話の解釈の変化や、発話による行動の解釈の変化を捉えることができない。つまり、行動情報と発話情報の相互作用を考慮した心的状態の推定ができていない。

本研究では、行動情報と発話情報の両方から人間の心的状態を推定するシステム MIoM を提案する。MIoM は、行動情報と発話情報の両方をベイズの定理に適用し、人間の心的状態を推定する。MIoM は、人間の心的状態の推定において、心的状態を一つに決め付けるのではなく、同時に複数保持し、やり取りの中でその可能性を動的に変えていく。MIoM は、行動情報と発話情報をもとに、複数保持した心的状態の可能性が大きく変動したタイミングまたは一定の時間、変動しなかったタイミングで人間に質問を提示し、質問の応答を発話情報として心的状態の推定に反映することで、発話情報による行動の解釈の変化や人間の行動による発話情報の解釈の変化を捉え、行動情報と発話情報の相互作用を考慮した心的状態の推定が可能となる。

本論文の構成は以下の通りである。第二章では、関連研究においてどのように人間の心的状態を推定していたかを述べる。第三章では、人間の行動情報と発話情報から心的状態を推定するシステム MIoM を提案する。第四章では、MIoM と単一の情報のみから心的状態を推定するシステムを用いて実験的に評価し、第五章では評価結果について考察する。第六章では、MIoM における今度の課題について述べる。最後に、第七章で本論文を締めくくる。

第2章

関連研究

- 2.1. 行動情報から心的状態を推定する研究
- 2.2. 発話情報から心的状態を推定する研究
- 2.3. 単一情報による心的状態推定の問題点

第3章

提案

3.1. アルゴリズム

3.2. システム構成

MIoM のシステム構成を図 3.1 に示す.

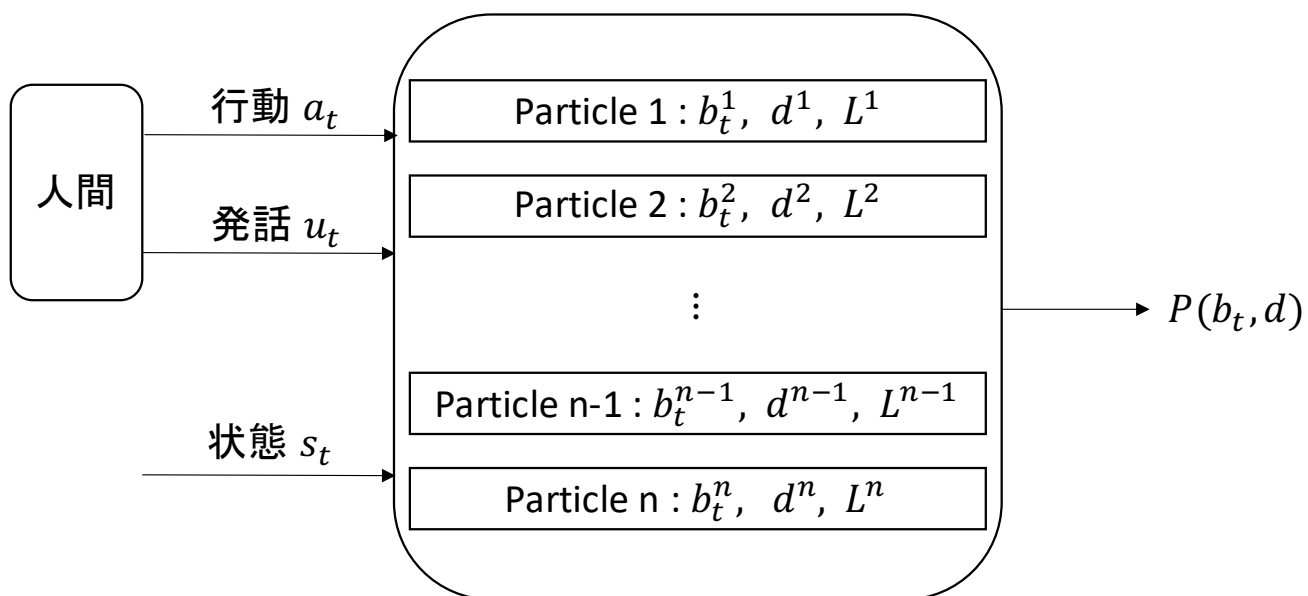


図 3.1: MIoM のシステム構成

第4章

評価

4.1. 実験設定

7×5マスで表現される環境中に日本食の屋台、イタリア料理の屋台、中華料理の屋台の3つのうち2つが存在する。環境中の学生は移動し、対話システムと対話をしながら食事を購入する屋台を決める状況を考える。学生は、いずれの屋台が出店しているかに関する事前知識は持たないものとする。学生の行動 a_t は上、下、左、右の4方向への移動とし、発話 u_t は対話システムから提示される食事に関する質問に対する学生の応答とする。信念 b_t は、どの屋台が出店していると考えているか、欲求 d は学生が3種類の屋台をどの順に好むかを表す。

4.2. 実験手順

図4.1に本実験における環境の一例を示す。本実験には、本研究で作成したデータセットを利用した。本データセットには、屋台の組み合わせとその位置を含む環境設定と、その環境設定で考えられる学生の行動および発話が含まれる。30人の実験参加者に、本データセットで指定された環境設定と行動および発話を提示し、環境中の学生の信念と欲求をそれぞれ7段階で推定させた。また MIOm と、行動情報と発話情報の一方のみを基に心的状態を推定する Unimodal Inference of Mind(UIoM) により信念と欲求の推定を行った。MIOm(action + utterance)、行動情報のみを基に心的状態を推定する UIoM(action)、発話情報のみを基に心的状態を推定する UIoM(utterance) の3つのシステムによって、環境中の学生の信念と欲求をそれぞれ7段階で推定した。実験参加者によって得られた推定結果と MIOm および UIoM によって得られた推定結果を比較し相関係数を算出した。

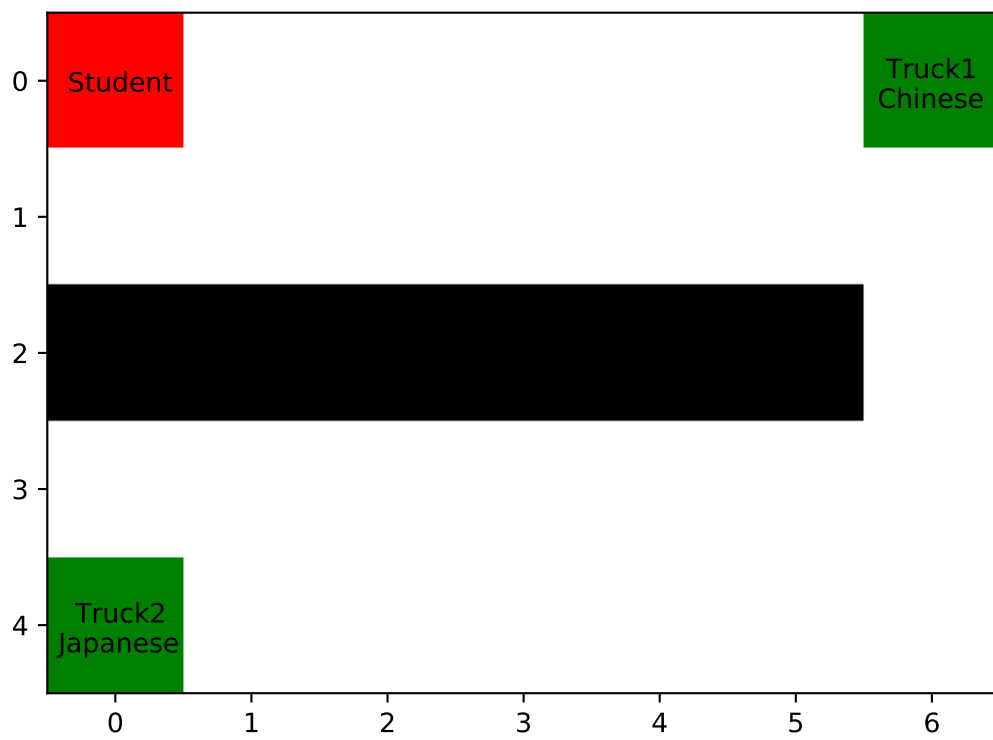


図 4.1: 本実験における環境

4.3. 実験結果

表 4.1: 人間による推定と推定モデルの相関

モデル	相関	
	信念	欲求
UIoM(action)	0.124	0.419
UIoM(utterance)	0.216	0.494
MIoM(action + utterance)	0.244	0.549

表 4.1 に，実験参加者による信念と欲求の推定結果と UIoM(action), UIoM(utterance) および MIoM による信念と欲求の推定結果との間の相関係数を示す．

表 4.1 より，MIoM は信念と欲求の推定の両方において UIoM(action) および UIoM(utterance) よりも強い相関を示した．また，いずれの推定システムにおいても欲求推定の相関が信念推定の相関よりも強いことがわかった．

第5章

考察

第6章

今後の課題

- 6.1. 単語埋め込みモデルの検討
- 6.2. 対話相手の発話に対する応答生成
- 6.3. 実世界情報による心的状態の推定

第7章

結論

本論文では，MIoMによりマルチモーダルな心的状態推定について検討した．実験の結果，行動情報と発話情報の両方を心的状態の推定に用いることが有効であることを示した．今後の展望としては，より実世界に近い環境設定や三次元の行動や幅広い発話を扱えるようにMIoMを拡張したいと考えている．

謝辞

本研究を進めるにあたり、研究の機会及び貴重なご意見を頂きました、
慶應義塾大学理工学部 今井 倫太 教授
に深く感謝致します。

論文の査読をして頂き、細部にわたって御意見を頂きました
理工学研究科博士課程2年 福地 庸介 氏
理工学部研究員 前川知行 氏
に厚く御礼申し上げます。

実験に御協力頂いた被験者の方々に心より御礼申し上げます。

最後に日頃から御指導、御協力下さいました今井研究室の皆様に心より感謝いたします。

令和3年1月

参考文献

- [Ameba] Ameba. アメーバブログ. <http://ameblo.jp/>.
- [AU 06] AU. EZ メール読み上げ. http://www.au.kddi.com/ez_mail_yomiage/index.html, 2006.
- [goo] goo. goo ブログ. <http://blog.goo.ne.jp/>.
- [Kazuhiko Shinozawa *et al.* 05] Junji Yamato Kazuhiko Shinozawa, Futoshi Naya, and Kiyoshi Kogure. Differences in effect of robot and screen agent recommendations on human decision-making. *International Journal of Human-Computer Studies*, pp. 267–279, 2005.
- [Livedoor] Livedoor. *LivedoorBlog*. <http://blog.livedoor.com/>.
- [W.Berger *et al.* 71] Kenneth W.Berger, and Gerald R.Popelka. Extra-facial gestures in relation to speechreading. *Journal of Communication Disorders*, pp. 302–308, 1971.
- [WordPress] WordPress. . <http://ja.wordpress.org/>.
- [xypoint 00] xypoint. *Nomad*. <http://www.xypoint.com/>, 2000.