

インタラクション再生モデルによるさりげない学習支援

青山 秀紀[†] 尾関 基行^{††} 中村 裕一^{††}

[†] 京都大学工学研究科, 〒 606-8501 京都市左京区吉田本町

^{††} 京都大学学術情報メディアセンター, 〒 606-8501 京都市左京区吉田本町

あらまし 本稿では、ユーザの状態に適応した支援を行うためのモデル “*Interaction Reproducing Model*” を用いたユーザ支援について報告する。このモデルは、人間のオペレータの操作による理想的なインタラクションのパターンを有限状態機械として学習しておくものであり、支援時には、現在行われているインタラクションが学習済みのどのインタラクションパターンに最も近いかを判断し、そのパターンを基にシステムを動作させることで適切な支援を実現する。これにより、ユーザの状態を直接認識することなく、ユーザの状態に合わせた支援を行うことが可能となる。我々は、この仮説を実証するために本モデルを使用した教示システムを試作し、ユーザの状態（講義に対する興味の度合い）に従って教示内容を変化させる実験を行った。以前の報告とはモデルの設定を変えた実験によりその精度を調べ、有効性を確認した。

キーワード HCI, さりげない支援, ユーザ状態の認識, パターン認識

An Ambient Way of Learning Support with the Interaction Reproducing Model

Hideki AOYAMA[†], Motoyuki OZEKI^{††}, and Yuichi NAKAMURA^{††}

[†] Graduate School of Engineering, Kyoto University,
Yoshida-Honmachi, Sakyou-ku, Kyoto-shi, 606-8501 Japan

^{††} Academic Center for Computing and Media Studies, Kyoto University,
Yoshida-Yoshida, Sakyou-ku, Kyoto-shi, 606-8501 Japan

Abstract This paper presents a learning support with the *Interaction Reproducing Model (IRM)* that adjusts computer supports to match the state of the user, and it consists of a set of Interaction Finite State Machines (I-FSMs). Each I-FSM is trained using actual interaction records, and it represents an ideal interaction pattern for a user state. It can choose an appropriate system action by reproducing the interaction pattern of the I-FSM most similar to the current interaction. We developed a prototype teaching system, and conducted preliminary experiments. The results by different configuration from the previous report show the efficacy of our framework.

Key words HCI, Appropriate Supports, User State Recognition, Pattern Recognition

1. はじめに

情報システム技術の発展に伴い、日常生活や仕事、学習などの場でいつでもどこでも人間を支援することのできる情報環境、ユビキタス環境の構築が進められている。このユビキタス環境においては、人々は従来のように能動的にシステムに働きかけるのではなく、その存在を意

識することなくその恩恵を享受できることが求められる。そこで、システムには人間の自然な振る舞いからその要求を認識し、自律的に支援することが期待される。その際、システムからの支援を一方的に押しつけるのではなく、ユーザの主体性や意欲を削ぐことがないよう、ユーザの意図や周囲の状況を認識し、それに応じて支援の内容や方法、タイミングの選択を行う必要がある。

このような背景から、我々は、**さりげない支援**を提供するためのシステム設計とその効果を探っている。ここでいうさりげない支援とは、ただ単に支援が控え目であるということではなく、ユーザの主体性を尊重して必要な支援を過不足なく行うというものである。例えば、学習や訓練のための教示システムについて考えてみよう。最も単純な支援は、ユーザがそれに満足しているか否かに関わらず、あらかじめ設定された動作を行うものとなる。ユーザがその内容を既に知っていたり、全く興味が無かったりしても、事前にプログラムされた手順に従って教示を行う。一方で、人間の優れた教師は、生徒の状態を認識して臨機応変に教示方法を選択する。このような支援を実現するためには、ユーザとその周辺の外面的状況だけではなく、ユーザの思考や興味といった内部状態を推定し、それに応じた適切な支援を選択する能力が必要となる。

しかし、ユーザの内部状態には直接的に観測できないものも多く、情報システムがそれを認識することは容易ではない。人間の教師の場合は、生徒の表出する様々な情報や既に持っている知識・経験から生徒の内部状態を推測していると考えられる。また、様子をみるだけでは推測が難しい場合には、生徒に積極的に働きかけることによって、生徒の状態をより正確に把握している。我々は、このような、マルチモーダル且つプロアクティブな認識を実現するための方法論を探り、適切な支援を提供する情報システムを実現することを目指している。

本研究では、そのためのモデルの一つとして“*Interaction Reproducing Model (IRM)*”を提案する(図1)。そのキーアイディアは、ユーザと熟練した支援者の間で行われる外部に表出された振る舞い(インタラクション)をモデル化することである。このモデルは、適切な支援のインタラクションパターンを記録しておき、現在のインタラクションに合わせてそのパターンの一つを選択して、それを再現するように振る舞うことで適切な支援を実現する。

本手法の特徴を述べる上で、ユーザの内部状態認識に関連する研究を挙げてみよう。Kapoorらは表情や姿勢変化パターンなどを特徴量とし、サポートベクトルマシンなどを用いて教示システム使用時の児童の感情を認識する手法を提案している[1][2]。また、駒谷らは、音声応答システムにおいて、ユーザの発話継続時間やその認識の成否などを特徴量とした決定木によって、ユーザの習熟度、知識レベル、性急度を判別するモデルを提案している[3][4]。その他、人間の内部状態の認識に関する様々な研究が行われている[5][6]。これらの先行研究に対して、

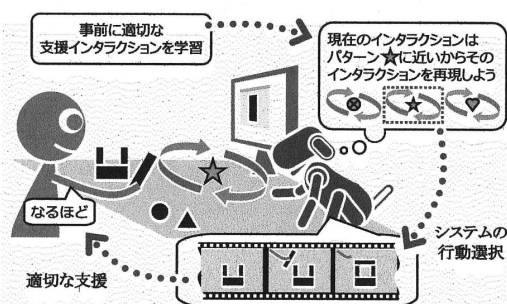


図1 Interaction Reproducing Model による支援

本研究は、ユーザの行動だけではなく支援側(支援者やシステム)の行動を含めてモデル化するところに特徴がある。また、ユーザの内部状態を直接認識するのではなく、手本となるインタラクションのパターンが再現されるようにシステムを動作させることが、ユーザの内部状態の認識を内包しており、同時にその状態に適したシステムの支援を決定することを可能にする。

本稿の構成は次のとおりである。まず2.で、ユーザの状態に適した支援モデルとはどのようなものかについて考察する。次に、3.で、その一つの実現案としてInteraction Reproducing Modelを提案する。4.では、このモデルを用いて試作した教示システムの構成について説明する。そして、5.では、まず、以前に行った実験で明らかになった問題点を解説し、それらを修正して行った評価実験とその結果を述べる。

2. ユーザの状態に合わせた支援モデル

人間同士のコミュニケーションでは、相手の様々な言動を観察し、時には自らの働きかけに対する反応をみることで、相手の内部状態を推測している。このような仕組みをモデル化することで、ユーザの内部状態を認識し、同時に適切な支援を選択することが本研究の目的となっている。

ここで扱う必要のある人間の内部状態には、興味、理解度、スキル、好き嫌い、やる気、気分など、様々なものがあるが、多くは抽象的な概念であり、直接的に物理量と対応づけたり、その振る舞いをモデル化したりすることは難しい。新しい試みとしては、脳波や心電などの生体信号を利用する研究が行われており、今後の発展が期待されるが、現状ではその関係は十分に明らかになってはいない。特に、支援システムに必要とされる内部状態(課題が難しいと感じている等)の推定には困難がともなう[7-10]。

このような問題に対し、我々の提案する“Interaction

Reproducing Model (IRM)”は、内部状態に応じた適切な「インタラクション」をモデル化するというアプローチをとる。つまり、内部状態を直接的にモデル化するのではなく、その内部状態の表出と考えられるインタラクションをモデル化する。ユーザの内部状態はユーザの様々な行動の違いとして表出され、支援者（または支援システム）の行動がそれに影響され、更にそれによってユーザの内部状態が影響を受ける。このような複雑な相互作用があるため、ユーザの行動だけでなく、システムの支援動作を含めたインタラクションのパターンのモデル化を考えることが鍵となる。このような方法をとることで、ユーザの内部状態のモデル化、その直接的な認識、その結果に応じた支援インタラクションの生成という個別の難問題を考える必要がなくなる。

インタラクションをモデル化するためには、時系列の状態を扱わなければならない。また、ユーザとシステムの二者の行動を同時に扱うことができるモデルでなければならない。時系列データを扱えるモデルには隠れマルコフモデル (HMM) や文脈自由文法など、様々なモデルが存在する。しかし、それらの多くは異なる二者の行動を同時に扱えるようには設計されていない。二者の扱いが可能なモデルとしては Coupled HMM [11] があるが、その構造や学習方法が複雑で、動作の解析や検証が困難である。このように、既存のモデルでは先に挙げた条件を満たす簡便なモデルが存在しない。

そこで我々は、二者のインタラクションを表現できるように有限状態機械を改良した新しいモデル “*Interaction Finite State Machine (I-FSM)*” を提案する。このモデルは、二者のインタラクションを自然な形で扱えるように設計されており、学習アルゴリズムも頻度による簡単なものである。その詳細については次章で述べる。

3. Interaction Reproducing Model

3.1 モデル構成

Interaction Reproducing Model (IRM) の概要を図 2 に示す。IRM は、複数の Interaction Finite State Machine (I-FSM) で構成されており、各 I-FSM は、ユーザの異なる状態（外部環境、内部状態）に対応したインタラクションのパターンをそれぞれ表現（記憶）している。

I-FSM は非決定性確率有限状態機械であり、パラメータとして状態遷移確率 A 、記号出力確率 B 、初期状態確率 π を持つ。I-FSM において、システムの行動は状態遷移 a_i に、ユーザの行動は記号出力 b_k に対応させる。ただし、この対応に必然性はなく、ユーザの行動を状態遷移に、システムの行動を記号出力と対応させることも

可能である。以下では前者の場合を説明する。このとき、状態遷移確率は、ユーザが特定の外部環境や内部状態にあるときのシステムの取るべき行動を確率の形で保持しており、記号出力確率は、同じくユーザが取ると期待される行動を保持する。

ユーザの状態が異なれば、インタラクションに表れるユーザの行動パターンは変化し、それはシステムの行動パターンにも影響する。これらの変化は I-FSM のパラメータの変化として表れるため、それを記憶（学習）しておくことで特定のインタラクションパターンを表現する I-FSM が構築できる。従って、区別したいユーザの状態の数だけ I-FSM を用意し、各々の状態のインタラクションのパターンを学習させることになる。

また、システムやユーザの行動 a_i 、 b_k には、動作や音声、「集中度」といった複合的な特徴など、多様な特徴を用いることができる。なお、このモデルでは、ユーザの行動とシステムの行動が交互に行われることを前提としているが、「何も行わない」という行動を加えることによって交互でないインタラクションに対応することができる。

インタラクションのような離散時系列信号の学習には隠れマルコフモデル (HMM) が一般的によく使われる。ただし、HMM では二つの信号源がある場合を直接扱うことができないため、前述のように二者間のインタラクションを扱う手法として Coupled HMM などが提案されている。HMM は状態遷移まで確率的に表現することができるため柔軟性の高いモデルとなっているが、本研究で用いる I-FSM はユーザの行動によって状態遷移が一意に決定される。そのため、I-FSM は Coupled HMM などより表現能力が劣る可能性がある。このような不利な点があるものの、I-FSM は Coupled HMM に比べて学習アルゴリズムが簡単で解析が容易であるという特徴を持つ。特に、本研究で対象とするような複雑な相互作用の起こるインタラクションの場合には見通しの良さが重要となるため、本研究ではまず I-FSM を用いることにした。なお、IRM の枠組みは I-FSM に依存していないため、インタラクションパターンを表現するためにより適した方法が見つければ、I-FSM と入れ換えることも可能である。

3.2 使用方法

IRM によってユーザの状態に適した支援を行う手順は次のようになる。まず、現在行われているユーザとシステムの間のインタラクションに対して、それに最も近いパターンを持つ I-FSM（学習済）を選ぶ。この処理を本稿では「インタラクションパターンの認識」と呼ぶこと

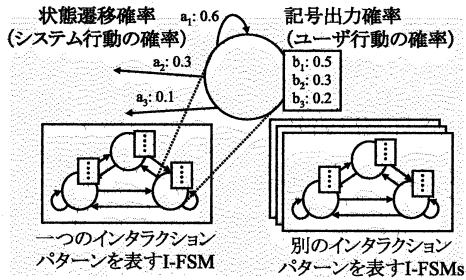


図 2 Interaction Reproducing Model

にする。次に、選ばれた I-FSM が記録しているインタラクションパターンに従ってシステムの行動を生成する。この操作をユーザの行動が起こる度に繰り返すことで支援を行う。以下、詳細について述べる。

インタラクションパターンの選択には、HMM による音声認識やジェスチャ認識と同様の方法を用いる。すなわち、過去の一定期間に行われたインタラクションの系列について、その系列がそれぞれの I-FSM によって生成される確率を求める。このとき、Coupled HMM では状態遷移が非決定的であるため確率計算に複雑なアルゴリズムが必要となるが、I-FSM では状態遷移を決定的なものとしているため^(注1)生成確率は記号出力確率の積によって容易に求められる。この再現確率が最大となる I-FSM (最尤 I-FSM と呼ぶ) が、現在行われているインタラクションを表すものとして選ばれる。選ばれた I-FSM の尤度が高いほど理想的なインタラクションが行われている状態に近いと考えられ、この I-FSM に沿った動作が続けるのが好ましいことになる。

なお、ここで述べた尤度の計算では、HMM の場合とは異なり、状態遷移確率は無視する。これは、後述するように、システムの行動が最尤 I-FSM の状態遷移確率によって決定されているため、ある I-FSM がしばらく選択されていた場合、状態遷移確率を尤度計算に含めるとその I-FSM の尤度が引き続き高くなってしまう。これでは、たとえユーザの行動が他の I-FSM の表すインタラクションパターンに従っていたとしても、適切な I-FSM が最尤となることが阻害されてしまう。これを防ぐため、尤度計算は記号出力確率のみを用いて行う。

このようにして選択された最尤 I-FSM を用いて、その時点のユーザの状態に適したシステムの行動を決定する。まず、ユーザとシステムのインタラクションから求められた最尤 I-FSM における現在の状態 S を求める。そして、その状態 S からの状態遷移確率 A に基づいて状態

遷移 a_i を決定し、それに対応するシステムの行動を実行する。

以上のように、IRM は、学習済みのインタラクションパターンの中から、現在行われているインタラクションに最も近いパターンのインタラクションを再現するようにシステムの行動を選択する。これによって間接的にユーザの状態を認識し、それに適応した支援を行っていることになる。

3.3 学 習

トレーニングデータの収集

まず初めに、I-FSM の学習に用いるためのデータ (ユーザとシステムの行動履歴データ) の収集を行う。将来的にはシステムを使用しているうちに自然に学習されるようにすることが考えられるが、ここでは人間のオペレータがシステムを制御しながら学習データを収集することにする。前述したように、IRM はこのトレーニングデータを再現するように動作するため、ここで収集するインタラクションは、システムがユーザを支援する際の手本となるものでなければならない。そのためには、ここで収集したデータを注目しているユーザの内部状態毎に適切に分類する必要がある。

インタラクションの分類

収集したトレーニングデータを分類 (内部状態と対応付け) する方法として、ユーザの自己申告による分類、観察者による分類、自動分類の三つの方法がある。以下、それぞれの利点と欠点をまとめる。

ユーザの状態 (特に内部状態) を自己申告してもらう方法は、手軽でわかりやすいが、申告内容の個人差が大きい。また、ユーザの状態は時々刻々と変化するため、一連のインタラクションが終わった後の報告では粒度が不十分である。しかし、インタラクション中に自分の状態を逐一報告させることはユーザに対する負荷が大きく、自然なインタラクションを阻害してしまう可能性がある。この問題を解決するために、生体信号によってリアルタイムに、且つ、より少ない個人差で答えを得ようという考えがある。しかし、生体信号によって人間の感情や思考などの内部状態を安定して得ることは未だ困難である [7-10]。

観察者 (特に、前述のオペレータが望ましい) がユーザの状態を判断する方法は、ユーザの自己申告に比べて個人差による問題が小さい。また、ユーザには負担をかけないため、一連のインタラクション中の変化にも追従できる。更に、被験者が年少であるなど、自身の内部状態の表現に支障がある場合にも用いることができる [1]。一方、他人の状態を外から見て判断することになるため、

(注1) : システムの過去の行動履歴がそのまま状態遷移の系列を表す。

その推定精度は観察者の能力に依存する。

インタラクションデータを何らかの基準に基づいて自動的に分類する方法は、前の二つとは異なり、ユーザの状態の推定を介さず、インタラクションパターン自体が持つ分布の特性によって分類を行う。この方法の最大の利点は、主観的な判断が入り込まず、個人差や観察能力の問題が起こらないことである。ただし、インタラクションデータは様々な観点でクラスタリング可能であり、それが設計者の狙ったユーザの状態に対応しているか否かについて十分に検証する必要がある。

後述する評価実験では、ユーザの自己申告による分類を用いた。また、観察者（オペレータ）による分類を同時に行い、その精度を確認した。自動分類に関しては、将来の課題とする。

学習アルゴリズム

学習によって、システムの行動の出現確率である状態遷移確率 A とその初期状態確率 π 、ユーザの行動の出現確率である記号出力確率 B を決定する。この学習は、対応する I-FSM からトレーニングデータが出力される確率を最尤推定量として、最尤推定法に基づいて行う。

具体的には以下のようなアルゴリズムとなる。

- (1) 学習対象のインタラクションデータにおいて、システムの行動 a_i が行われた頻度をそれに対応する状態 S_i の初期状態確率 π_i とする。
- (2) システムの行動 a_i に続いてシステムの行動 a_j が行われた頻度を状態 S_i から状態 S_j への状態遷移確率 A_{ij} とする。
- (3) システムの行動 a_i に続いてユーザの行動 b_k が行われた頻度を状態 S_i における記号出力確率 B_{ik} とする。

インタラクションパターンの表現と学習効率

インタラクションを正確に表現するためには、そこに表れた種々の行動（特徴）を網羅的に扱うことが望ましいが、扱う行動の数が増えると学習の効率が低下する（行動の数の二乗に反比例）ため、より多くのトレーニングデータが必要となる。そのため、支援の際に再現したいインタラクションパターンを吟味し、重要な意味を持つ行動に絞って学習する必要がある。なお、本実験では、筆者の一人が重要だと感じた行動を経験的に選択した。具体的には次章で述べる。

4. システム

本章では、5. で述べる評価実験のために試作した教示システムについて述べる。



図3 システム構成

4.1 システム構成

システム構成の概要を図3に示す。本システムは、モニタ、スピーカ、エージェントロボットで構成されており、システムが図と文章からなる講義スライドをモニタに表示し、解説文を合成音声で読み上げることで教示を行う。その際に、ユーザが講義内容に対して感じている難易度や興味の度合い（興味度）に応じて、講義の速度や内容を適切に変化させることを目的とする。

教示はモニタ・スピーカのみでも可能であるが、エージェントロボットが存在したほうがユーザがより多くの反応を示すことが予備実験で確かめられた。そこで、本実験ではエージェントロボット（AIBO ERS-7M2, Sony）を擬似的に教師として振る舞わせ、ユーザからの反応を促す。

4.2 ユーザ及びシステムの行動

磁気センサを用いて、「傾き」「頭の向き（エージェントを見ているか/モニタを見ているか/手元を見ているか）」「手の動きの大きさ」「姿勢の角度」を計測した。また、マイクを用いて発声の有無を認識した。これらの認識は「しているか」「していないか」の二値で行い、IRMへの入力とした。

システムの行動として、「次の説明文を読み上げる（モニタの表示はこれに同期する）」「傾く」「ユーザを見る」「モニタを見る/指さす」「ユーザに問かける（『良いですか』『分かりますか』からランダムで選択）」を用意した。学習時には人間のオペレータが、IRMによる制御時にはIRMがこれらの行動を選択し、システムを動作させる。

5. 評価実験

我々は、IRMによる支援の評価実験を行ってきた。それによって明らかになった問題点を改善し、再度評価を行った。

5.1 過去の評価実験の概要

過去の実験（以下、前実験と呼ぶ）では、ユーザが講義に対して感じている難易度を、「簡単」「ちょうど良い」「難しい」「理解できない」の4段階で認識し、それに応じて講義速度などを変更することで、ユーザの状態に適した支援を行うことを目的とした[12]。その結果、IRMによるインタラクションパターンの認識（ユーザの状態の推定）の精度は58%であり、隣接した状態に誤認識されたものを許すと79%の認識率となった。また、IRMによって制御された講義は、人間のオペレータが操作した場合には及ばなかったものの、講義の難易度によってあらかじめ定められた速度で行った場合よりも良い印象を得た。

5.2 改善点

本実験では、前実験の結果を踏まえて以下の点の改善を行った。

a) 認識対象とする内部状態の変更

前実験では、ユーザの感じている難易度に合わせて講義の速度を変化させたが、それに応じてユーザの感じている難易度を変化させることになる。この認識は人間にとってもやや困難であった。そのため、本実験では、認識の対象とする内部状態を、推定がより容易な興味度へと変更した。これは、トレーニングデータ取得時の手動制御オペレータの失敗がIRMの精度に影響するため、その影響を小さくした実験を行うためである。同じ理由から、認識の段階を「興味がない」「中程度又は不明」「興味がある」の3段階とした。また、認識対象とする内部状態の変更に伴い、支援の内容の調整も「講義速度の調節」から「講義内容の量の調節」に変更した。なお、講義内容は3段階の興味度に合わせてあらかじめ3種類用意している。

b) 内部状態の正解値の逐次取得

前実験では、一つの講義中にはユーザの状態はそれほど変化しないであろうと仮定し、事後申告によって一講義につき一つの内部状態を正解とした。しかし、手動制御オペレータの判断によると、一つの講義中にも数度の内部状態の変化があるように見うけられた。そこで、本実験では、より詳細なデータを得るため、被験者に講義中にボタンを押してもらうことで逐次正解値を得た。これによって、より精度の良いIRMの学習が可能になり、また、それによってより正確にIRMの評価を行える。

c) ユーザ毎のIRMの学習

前実験では、複数人の被験者から集めたデータによってIRMの学習を行ったが、インタラクションパターンに個人差が大きく、精度を落とす結果となった。そのため、

本実験では、各被験者毎にIRMを学習させた。

d) プローブ行動の導入

前実験で誤認識が起った状況では、正解 I-FSM の尤度と最尤 I-FSM の尤度の差が僅かであるケースが目立った。このような状況では、ユーザの行動を促す行動（以後プローブ行動と呼ぶ）をシステム側が行うことで、正しく認識ができる可能性がある。そこで、本実験では、プローブ行動を行うためのインタラクションパターンを用意した。具体的には、「中程度又は不明」のインタラクションパターンをこれに割り当て、「ユーザを見る」「ユーザに問いかける」の行動を多く行うように学習させた。また、最尤の I-FSM と次点のものの尤度差が小さい場合、ここでは最尤 I-FSM の尤度の8割を超える尤度を持つ I-FSM が他に存在する場合には、最尤の I-FSM ではなく、このインタラクションパターンによってシステムを行動させた。

e) I-FSM におけるユーザの行動とシステムの行動の入れ替え

前実験では、3. での説明とは逆に、ユーザの行動を状態遷移に、システムの行動を記号出力として扱ったが、3.1 で述べたように、I-FSM の状態遷移・記号出力には、ユーザの行動とシステムの行動のどちらを対応させても良い。しかし、前の状態遷移に対する記号出力は、前の記号出力に対する状態遷移よりもより強く依存しており、尤度計算における影響が大きい。そこで、「プローブ行動とそれに対するユーザの反応」がインタラクションパターンの認識により強い影響を与えられるように、前実験とは逆に、システムの行動を状態遷移に、ユーザの行動を記号出力に対応させた。

5.3 実験手順

このシステムを用い、まず、オペレータがシステムを手動で操作（手動制御）し、3人の被験者に4回ずつの講義を行った。次に、その結果をトレーニングデータとして学習させたIRMを用いた制御（IRM制御）によって、同じ3人に4回ずつの講義を行った。その様子を図4に示す。

なお、被験者には、システムがユーザの行動によって動作を変えることを知らせてあった。これを知らせていない場合には、被験者の動作が少なく、手動制御のオペレータにとっても被験者の内部状態の推定が困難であったためである。このことからわかるように、本システムのような支援システムの実用化には、CGやロボットによる擬人化インタフェースなど、人間の反応を引き出すしくみと併せて考えることも必要である。

正解とする内部状態は、講義中に被験者にボタンを押

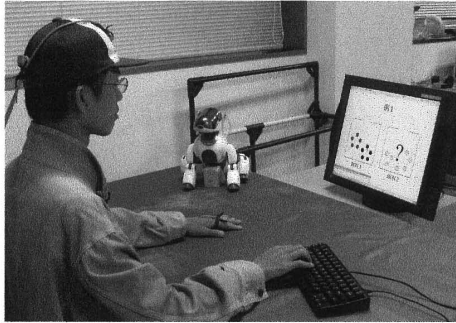


図4 実験風景

してもらふことで取得した。また、比較のため、オペレータによる被験者の内部状態推定結果も同時に記録した。さらに、印象評価のため、各講義の終了後には、「システムは適切に興味度を判断できていたか」「システムの行動は自然であったか」の2項目を5段階で評価してもらい、その回答を得た。

5.4 認識精度

インタラクションパターンの認識結果を表1に示す。ここで、認識率は、IRMによる認識結果がユーザによる内部状態の申告と一致していた時間の割合によって算出した。

IRMによるインタラクションパターンの認識率は66%であった。誤認識のうち、隣接した状態に誤認識されたものを許すと95%の認識率となった。本実験のオペレータによるユーザの状態推定の正解率が73%及び97%であったことを考慮すると、人間の推測能力に比べてそれほど劣らない判断をIRMで実現しているといえる。

条件が異なっているため単純な比較は出来ないが、本実験の66%という認識率は、前実験の58%よりも改善されている。I-FSMの尤度差が小さいときにプローブ行動を取るようにしたこと、より良いトレーニングデータが得られるように対象とする内部状態を変更し、さらに、ユーザに逐次正解を報告してもらったこと、その他5.2節で述べた改善がこの結果に繋がったと考えられる。

ただし、前実験と本実験で共通した誤認識の原因として、被験者の自己申告の信憑性の問題もある。前実験では19回の講義のうち2回、本実験では12回のうち1回の講義において、IRMによる認識結果は明らかに正しいだろうとオペレータに判断されたにも関わらず、被験者の申告が異なっていたことに依る誤認識があった。この点については、生体情報を援用するなど、今後の検討が必要である。

表1 インタラクションパターンの認識精度

被験者の回答 (正解)	IRMによる認識結果			認識率
	興味がない	中程度／不明	興味がある	
興味がない	319	159	80	57%
中程度／不明	325	1034	325	61%
興味がある	117	304	1144	73%
			計	66%

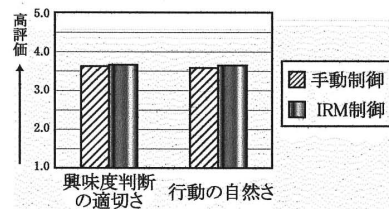


図5 ユーザによる支援システムの評価結果

5.5 支援の適切さの評価

支援の適切さに関する印象評価の結果を図5に示す。前実験では、IRMによる制御の印象はオペレータによる手動制御より劣っていたが、ここではその傾向は見られず、認識精度が上がったことによる改善が伺える。さらなる改善のためには、以下に挙げるような工夫が考えられる。

まず、今回の実験では十分な数のトレーニングデータを用意することが出来なかったため、収集した全てのインタラクションデータを用いてIRMの学習を行った。しかし、トレーニングデータ取得時（オペレータ操作時）の被験者の印象を調査し、評価が良かったときのデータのみを選んで学習すれば、平均して手動操作時よりも良い支援の実現も可能であると考えられる。

また、IRMは短期間のユーザの状態変化^(注2)に対応できないことも評価が悪い原因と考えられる。これに対し、IRMに入力する過去のインタラクションの長さを調整することで、認識の間隔を短くすることはできる。しかし、入力パターンを単純に短くしただけでは反射的な判断となり、認識結果が不安定になる可能性がある。そのため、各入力の重要度に応じて重み付けしたり、短い時間で判断するIRMと長い時間で判断するIRMを組み合わせるといった工夫が必要になるだろう。

6. まとめ

本稿では、ユーザの状態（外部環境・内部状態）に合

(注2)：スライド1枚分だけ難易度や興味度が異なるという場合が多かった。

わせて適切な支援を提供するためのモデル “Interaction Reproducing Model (IRM)” によるユーザ支援について報告した。IRM は、ユーザとシステムの間の適切なインタラクションをあらかじめ学習しておき、それをシステム使用時に再生することによって、ユーザの状態を直接推定することなく適切な支援を行うことができる。以前の実験により判明した問題点を修正し、提案モデルの有効性を確認するための評価実験を行った。その結果、インタラクションパターンの認識精度は 58% から 66% に改善された。同条件下での人間によるユーザの状態推定の正解率は 73% 程度であり、本モデルが十分に機能していることを示すことができた。

以上のように概ね良好な結果が得られたが、本研究はまださりげない支援に求められる機能のごく一部を扱った段階であり、5.4, 5.5 で挙げたような多くの課題を解決する必要がある。また、今後は、ユーザとシステムのインタラクションの中で自然に学習していく仕組みについても研究していく予定である。

文 献

- [1] A. Kapoor and R.W. Picard, “Multimodal affect recognition in learning environments,” *Proceedings of the 13th annual ACM international conference on Multimedia*, pp.677–682, Nov. 2005.
- [2] A. Kapoor, W. Bursleson, and R.W. Picard, “Automatic prediction of frustration,” *International Journal of Human-Computer Studies*, vol.65, no.8, pp.724–736, Aug. 2007.
- [3] 駒谷和範, 上野晋一, 河原達也, 奥乃 博, “音声対話システムにおける適応的な応答生成を行うためのユーザモデル,” *電子情報通信学会論文誌*, vol.J87-D-II, no.10, pp.1921–1928, Oct. 2004.
- [4] K. Komatani, S. Ueno, T. Kawahara, and H.G. Okuno, “User modeling in spoken dialogue systems to generate flexible guidance,” *User Modeling and User-Adapted Interaction*, vol.15, no.1-2, pp.169–183, March 2005.
- [5] S. K.D’ Mello, S.D. Craig, B. Gholson, S. Franklin, R.W. Picard, A.C. Graesser, “Integrating affect sensors in an intelligent tutoring system,” *Affective Interactions: The Computer in the Affective Loop Workshop at 2005 International conference on Intelligent User Interfaces*, pp.7–13, 2005.
- [6] C. Conati and X. Zhou, “Modeling students’ emotions from cognitive appraisal in educational games,” *ITS ’02: Proceedings of the 6th International Conference on Intelligent Tutoring Systems*, pp.944–954, London, UK, Springer-Verlag, 2002.
- [7] C.M. Jones and T. Troen, “Biometric valence and arousal recognition,” *OZCHI ’07: Proceedings of the 2007 conference of the computer-human interaction special interest group (CHISIG) of Australia on Computer-human interaction: design: activities, artifacts and environments*, pp.191–194, New York, NY, USA, ACM, 2007.
- [8] 眞溪 歩, “脳波認識とその応用,” *信学技報 NC2007-53*, 第 107 巻, p.109, Oct. 2007.
- [9] C. Puri, L. Olson, I. Pavlidis, J. Levine, and J. Starren, “Stresscam: non-contact measurement of users’ emotional states through thermal imaging,” *CHI ’05: CHI ’05 extended abstracts on Human factors in computing systems*, pp.1725–1728, New York, NY, USA, ACM, 2005.
- [10] T.M. Vaughan, W.J. Heetderks, L.J. Trejo, W.Z. Rymer, M. Weinrich, M.M. Moore, A. Kubler, B.H. Dobkin, N. Birbaumer, E. Donchin, E.W. Wolpaw, and J.R. Wolpaw, “Brain-computer interface technology: a review of the second international meeting,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol.11, no.2, pp.94–109, June 2003.
- [11] M. Brand, “Coupled hidden markov models for modeling interacting processes,” *Neural Computation*, Nov. 1996.
- [12] 青山秀紀, 尾関基行, 中村裕一, “ユーザの状態に適応した支援を行うためのインタラクションモデル,” *画像の認識・理解シンポジウム 2008, MIRU*, pp.121–128, 長野, July 2008.