

## 会話システムにおける表面的な処理を施す インタフェースの提案

平松 薫      山本 吉伸      安西 祐一郎

慶應義塾大学 計算機科学専攻

本稿では、ネットワークを介した対話でのユーザのブレイクダウンを避けるためのインタフェースを検討した。対話の相手が人間の場合と計算機の場合とを比較し、人間のキータイピングのシミュレーション(文字出力の速さとミスタイプ)と、あいづち(割り込み)のタイミングに着目した。そして、人間のキータイピングをシミュレートするフィルタとあいづちを行うプログラムを用意し、対話実験を行なった。その結果、対話システムにおいて、キータイピングのシミュレーションにより人間らしさを補えることを確認し、対話システムにおいてユーザのブレイクダウンを防ぐ可能性を示唆した。

## An Outside Processing on Dialogue System

Kaoru Hiramatsu      Yoshinobu Yamamoto      Yuichiro Anzai

Department of Computer Science, Keio University

This paper proposes an interface system to avoid users' "breakdown" on dialogue through computer network. When designing the interface system, we notice users' typewriting speed and response time on dialogue. We build an interface system which has a filter to simulate typewriting speed and response time. Then we make some experiments through computer network dialogue to show the effectiveness of the filter. According to the result of the experiments, we suggested that a filter could add humanity to the dialogue system and that it possibly reduce users' "breakdown".

## 1 はじめに

自然言語対話システムは、人間と計算機のインタラクションにおける有効な手段の一つであり、様々なものが提案されている [1][6]。現在のところ音声対話では認識などの問題から、主題の限定が行なわれているが、言語処理の面でも対話の主題の変化への対応や対話の主導権など課題は依然として多い。また、ユーザとシステムの間常に協調の関係が成り立っているとは限らないので、ユーザからの入力が対話システムの対応できる範囲からはずれる可能性もある。対話システムに対し例外的な入力があると、不自然な返答をする場合や無回答になる場合があり、これがユーザにブレイクダウン<sup>1</sup>を起す原因となる可能性がある。自然言語処理によってこのような例外的入力に対応し、対話の自然さを保つためには膨大な例外処理が必要になる。また、対話の相手を人間と認知しているか計算機と認知しているかによって、システムに対するユーザの態度に変化が出る可能性があり [7]、対話から得られる情報以外の影響も大きい。そこで本稿では、ユーザがブレイクダウンを起すことを防ぐインタフェース ABIS(Avoiding Breakdown Interface System) について検討する。

## 2 インタフェース ABIS の提案

本稿では、ネットワークを介したテキストのみによる対話システムを取り上げる。

### 2.1 設計及び実装

人間が人間と対話している場合と、対話プログラムと対話している場合を比較し、そのインタラクションでの相違点から、対話の自然さを補うインタフェースとして次の2つの要素に注目する。

1. 人間の入力速度とミスタイプシミュレーション  
→ 文字出力フィルタ
2. 対話中、キーワードを発見しだいその場であいづちを生成する  
→ 割り込み発話生成

以降、この二つについて説明する。

#### 2.1.1 文字出力フィルタ

ネットワークを介したテキストベースの対話システムでは、打ち込んだキーがそのまま相手の画面に反映される。

<sup>1</sup>Winograd の定義 [8] では、「習慣的・日常的で快適な「世界内存在性」が中断される瞬間」を指すが、本稿では特に対話の中断を指すこととする。

双方向からの入力が可能なことや対話の主導権を自由にとり発話できる点が、音声による会話に近いといえる。

人間のキー入力速度は、計算機への入力のうち一番遅いものに当たる。また、人間のキータイピングをそのまま反映している場合には、ミスキータイピングもそのまま相手の画面に反映される。これらの点が計算機上のプログラムと比べた場合の大きな差と考えられる。

そこで、出力状態の差が対話システムの自然さを損なう原因の一つであると考え、プログラムからの出力を人間による入力に近いものに変換するフィルタを作成する。

#### 予備調査

文字出力のフィルタを作成する際の予備調査として、実際に人間のキータイピングの速度とミスタイプ率を調べた。

表 1: 予備実験 (キータイプの速度)

被験者:	12名
調査方法:	タイピング練習ソフトの得点
結果 (入力速度):	217.3 文字/分 (3.62 文字/秒)
(ミスタイプ率):	4.73 %

日常的に計算機を使用する研究室の学生で調査したところ、表 1 の結果となった。キータイピング速度は毎秒約 4 文字程度で、そのうちミスタイプを約 4.5%含むというのが、平均的なユーザのキータイピングの状態として得られた。文字出力フィルタにおいて、次の要素を再現している。

- 人間のキー入力速度のシミュレーション
- ミスタイプのシミュレーション

このフィルタを通すことにより、文字出力速度が人間並の約 4 字/秒になるようにした。また、ミスタイプのシミュレーションでは、乱数を用いて約 4%の割合でミスタイプをするようにした。

#### 2.1.2 割り込み発話生成

人間同士の対話では、相手の発話中にあいづちをうつことが頻繁に見られる。一方、対話の主導権の移動が単純な対話プログラムでは、入力される文章が区切られた段階で解析を始めるものが多い。そこで本稿では、あいづち的要素をシミュレートし、キーワードからあいづちを生成する日本語による Eliza 型会話プログラムを設計した。このプログラムは perl で実装されていて、次の方針に従っている。

- 形態素解析や知識管理はしていない
- 会話の主題ごとにルールファイルを用意して、キーワード → リプライ という形で対話を進めていく
- 登録されているキーワードが相手の発話中に出現したら、その場で反応する

この対話プログラムは、ネットワークを介した対話システムを通して使用されることを前提としている。あいづちに関しては、意味や対話のながれに関する処理はしないで、ユーザからの入力からキーワードを探し、そのキーワードにのみ反応する。また、キーワードはルールファイル中に登録されている。

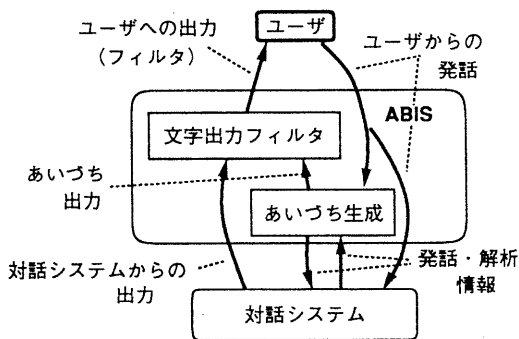


図 1: 対話システム-ユーザ-ABIS の関係

図 1は、本研究の ABIS の位置を示し、矢印が情報の流れを示す。あいづち生成部は、ユーザからの入力と対話システムの双方とインタラクションしながら発話を生成する。文字出力フィルタは、対話システムとあいづち生成部の両方からの出力をバッファリングし、人間のキー入力をフィルタの出力でシミュレートする。

### 3 実験

Loebner 賞大会 [3] でのプログラム評価方法に沿った方法で実験評価を行なった。対話相手として、プログラムと人間を用意し、被験者にネットワークを通して対話してもらい、対話を通して相手が人間かプログラムか当ててもらおうという形をとった。また、Loebner 賞大会 [3] とは違い、対話時間は 5 分を上限に相手が人間かプログラムか判断がつくまでとした。

### 3.1 手続き

被験者は、普段からキーボードを使用していて、キーボードに支障のない 12 人であった。

ネットワークで接続されている計算機を 2 台用意し、相手の操作状況がわからないように、互いに離れた場所に設置した。一方に被験者が座り、もう一方で実験のプログラムをスタッフが操作をした。被験者側に、ビデオカメラを設置し実験中の記録をとった。

被験者に対しては、次のように教示を行なった。

- 対話は、WS 上の phone<sup>2</sup> を通して行ないます。
- 相手は 6 条件で、人間かコンピュータです。
  - コンピュータは人間の真似をしますが、人間はコンピュータの真似はしません。
  - 例えば、コンピュータは打ち間違いをシミュレートしていますが、人間は故意に打ち間違いはしません。
  - また、人間がコンピュータの真似をしてわざと対話をつながらなくしたりはしません。
- 対話を通して、相手がどちらか当ててください。
- 決められた話題以外は、話すことはできません。
- 相手がわかった時点で対話をやめてください。  
(わからない場合も 5 分程度で対話は終了します。)

そして、実験スタッフの指示のもと、被験者は対話の相手としてプログラム 4 条件 (文字出力フィルタの有無 × (割り込み) あいづちの有無) と人間 2 条件の計 6 条件と対話した。その条件を、表 2 に示した。

また、対話実験は慶應義塾大学理工学部矢上校舎で行なうことを前提にし、各条件ごとに次のような異なる対話の主題を 6 つ用意した。(生協について、東急ハンズについて、秋葉原について、矢上近辺の食堂について、横浜について、図書館について) 対話の主題と対話条件の組によって判断結果に偏りがでないように、循環させて実験を行なった。

実験で検討する ABIS に関する要素 (出力フィルタの有無、あいづちの有無) によって被験者が相手を人間かプログラムか判断するように教示を行ない、さらに実験スタッフと対話プログラム両方に対して、対話の内容を次のように制限をした。

<sup>2</sup>本実験で用いたコンピュータネットワークを通した対話プログラム

表 2: 実験における対話相手の条件

実験条件	文字出力	あいづち
P1	無	無
P2	有	無
P3	無	有
P4	有	有
Ha	実験スタッフ A	
Hb	実験スタッフ B	
・ P1~P4	対話プログラム	

- 対話の主題が最初の 1 文でわかるようにし、対話の主題は対話を通じて変化しない  
(例: “あきはばらへのいきかたをおしえてください”  
→ 主題: 秋葉原)
- 対話の主題から離れた発話や質問には、実験者側 (スタッフ、プログラム) は答えない

この実験は、被験者が出力状態の差により人間かプログラムかを判断するため、対話を繰り返すことによる学習効果が大きい。そのため実施順による学習効果の差を少なくするために循環法により順序による差を相殺した。各条件について対話が終了後、被験者に対し次の項目を調査した。

- 各条件での対話時間
- ビデオによる被験者の対話中の発話と対話内容
- 対話のターン数
- 対話相手がコンピュータか人間かの判断とその判断理由 (内観報告)
- SD 法による評価 (5 段階評価, 14 項目)  
親しみやすい, (会話が) 自然な, 打ち間違いが多い, 反応が遅い, (相手が) 好意的, (会話が) 活発な, ふざけている, 会話が繋がらない, (相手が) 個性的, (出力が) 遅い, 冷たい, (文章が) 作為的, 感情的な, 子供っぽい

そして、6 条件すべてが終了した時点で、対話した相手を人間らしさで順位付けさせた。

#### 4 結果及び検討

実験から得た結果を各項目ごとにまとめ検討する。

#### 4.1 “人間らしさ”の順位

表 3 は各被験者がつけた相手 (人間、プログラム) に対する「人間らしさ」の順位である。教示において「対話相手を人間かプログラムか判断して下さい」とし、さらに話題に制限を加えたため、被験者の判断は相手が人間であってもプログラムと判断する場合が目立った。

表 3: 実験結果: 順位法

	P1	P2	P3	P4	Ha	Hb
平均	5.08	3.83	4.58	4.08	1.75	1.67
判定	0/12	0/12	0/12	1/12	6/12	5/12

・判定: 被験者の人間かプログラムかの判定

実験によって得た順位は、順位法によって検定した<sup>3</sup>。被験者のつけた順位は Friedman 検定<sup>4</sup>により一致していたことが示された。また、一致係数は 0.603 であった (図 2)。

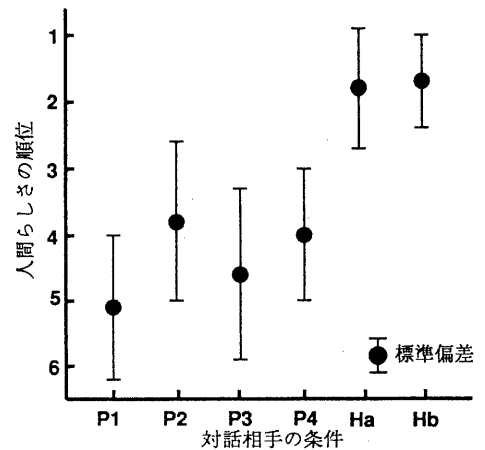


図 2: “人間らしさ”の順位

#### 4.2 5 段階評価

被験者に対して、内観報告と同時に行った SD 法による 5 段階評価で得られた結果をまとめた。また、それぞれの項目について正規化し、相関をとった。

<sup>3</sup> 順位和の平均値  $\bar{x} = 12 \times (6 + 1) / 2 = 42$

順位和の残差平方和  $S = \sum_{j=1}^k (x_j - \bar{x})^2 = 1520$

一致性の係数  $W = 12 \times 1520 / 12^2 (6^3 - 6) = 0.603$

<sup>4</sup> Friedman 検定:  $F_0 = 16.70, \phi_1 = 4.83, \phi_2 = 53.17$ , 危険率 1% で被験者の判断は一致した

特に相関の高かったのは、「(会話が)自然な」-「会話が  
つながらない」( $|r_{4,11}| = 0.975$ )、「反応が遅い」-「打ち間  
違が多い」( $|r_{10,13}| = 0.924$ ) などであった。項目間の  
相関値の絶対値が0.7以上のものをピックアップしたとこ  
ろ、質問項目を3つの集合に分けることができた(表4)。

表4: 相関の高い項目のグループ分け

集合 A	相手に対してプラスの印象の項目
質問項目	1 親しみやすい, 4 (会話が) 自然な 12 打ち間違が多い, 2 反応が遅い 10 (相手が) 好意的, 8 (会話が) 活発な
集合 B	相手に対するマイナスの印象の項目
質問項目	3 ふざけている, 11 会話がつながらない 6 (相手が) 個性的, 14 (出力が) 遅い 7 冷たい, 13 (文章が) 作為的
(集合 C	全体的に分散の小さかった項目)
(質問項目	9 感情的な, 5 子供っぽい)

集合Aと集合Bは対比的な集合であり、それぞれの項  
を比較すると負の相関を持っている。対話システムにおい

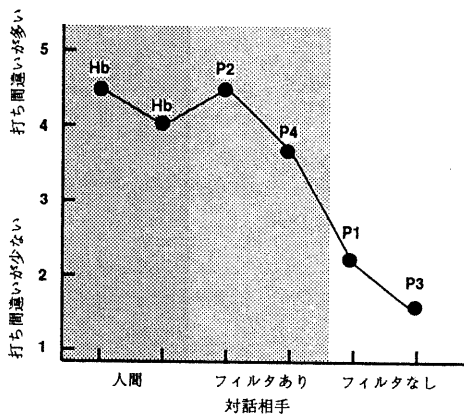
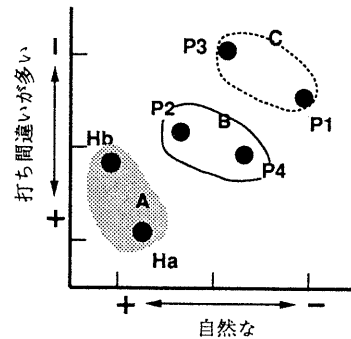


図3: 「打ち間違が多い」の評価

て、文字出力フィルタ(対話相手 P2/P4)に対する評価は、  
人間(Ha,Hb)に対する評価とほとんど差がなかった(図  
3)。これは、ABISの文字出力フィルタが、人間の打ち間  
違いに近い出力を再現していることを意味している。対  
話の相手に対してのプラスの印象を示す集合Aには、「打  
ち間違が多い」という項目も同類として含まれ、「親し

みやすい、自然、好意的」と「打ち間違が多い」が同等  
に評価されている。キータイピングのシミュレーションが  
対話に与える効果に、対話の自然さを補う要素が含まれ  
てブレークダウンの抑制に役立つと考えられる。

SD法による結果の各項目間のうち、「自然な」-「打  
ち間違が多い」の項目間には正の相関( $r_{4,12} = 0.614$ )が  
ある(図4)。また、「自然な」-「出力が速い」の項目間  
には、負の相関( $r_{4,14} = -0.608$ )がある(図5)。図4図5  
両方において、出力のシミュレーションをした(図4B, 図  
5ii)方がシミュレーションしていない場合(図4C, 図5iii)  
より、相手が人間の場合(図4A, 図5i)に近いことがわ  
かる。この図のグループ間の距離が、ABISのうち文字出  
力フィルタの効果になる。



(a) 「自然な」-「打ち間違が多い」

図4: SD法の項目間の相関(1)

### 4.3 内観報告

対話の主題を限定しても、被験者の発話に対して対話  
プログラムが反応できない場合が発生した。そのため、相  
手がどちらの場合も「話しがつながらない」など、会話の  
流れ自体に対する指摘が多かった。

キータイピングをシミュレートした場合は、記述とビデオ  
による対話中の記録から、2パターンに分けることがで  
きた。実験が始まってすぐに対話プログラムが反応できず  
に無反応になった場合、被験者はすぐにプログラムだと判  
断をしている。対話がある程度進んだ段階で無反応にな  
った場合には、被験者から「あ、考えてる」と相手を人間と  
してとらえていると思われる発言や「途中まで相手が人間  
ではと思った」との記述があった。人間かプログラムかの  
判断も無反応状態になってもすぐには判断が行なわれな

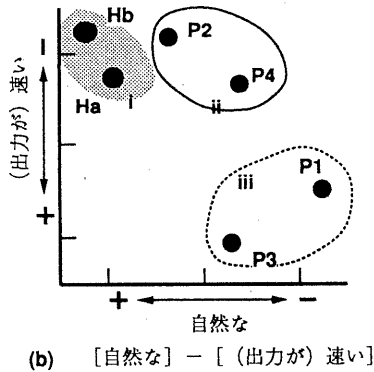


図 5: SD 法の項目間の相関 (2)

い。最終的な判断は違ったが、実際は相手がプログラムの場合も被験者が対話中に相手が人間ではないかと思った点は注目され、その時点まではブレークダウンが起きていなかったことを意味する。被験者の中には本実験で用いた文字出力フィルタの打ち間違いのシミュレートに対して、「文字出力速度が一定で不自然」といった指摘も一例あった。しかし、打ち間違いのシミュレートをした場合には、対話の内容から判断されている場合が多く(被験者 12 人中 11 人)、出力から人間かプログラムかを判断するのは難しかったようだ。打ち間違いをシミュレートしない場合には、「出力があまりにも速い」として、人間が相手の場合の文字出力速度と対話プログラムの出力速度が、明らかに違うとの指摘が多く(12 人中 9 人)あった。文字出力をシミュレートしない場合、対話の内容からではなく、文字出力のスピードにより判断が下されていたことがわかる。

判断基準として、あいづちなど割り込みに関するものを指摘するものは、文字出力でのフィルタに比べ少なかった。そのうち、あいづちのシミュレートをしている場合の指摘で対話相手について、「反応の速い時と遅い時が極端」という記述があった(12 人中 2 人)。対話プログラムは、対話中に登録されているキーワードが出てきた時点で反応し、その場で返答をするようになっている。キーワードはある一定のシナリオを想定し登録されているために、対話の流れが想定したシナリオに近い場合には、反応するキーワードが増え反応が頻繁になる。逆に、対話の内容がシナリオから離れると反応が極端に減ってしまうなど、本研究で製作した ABIS のうちあいづちのシミュレーションは用意した対話の流れに依存するところが大きかった。

#### 4.4 まとめ

本稿で提案したインタフェース ABIS を、実際の対話システムにおいて実験、評価し、結果を統計的にまとめた。その結果、提案したインタフェース ABIS のうちキータイピングのシミュレーションは有効であることがわかった。あいづちに関しては、内観報告にもあったように反応にばらつきが出てしまった。対話におけるタイミングの要素は重要であることから、さらに分析が必要と思われる。

#### 参考文献

- [1] 伊藤 克亘, “音声対話システム”, 電気情報通信学会技術報告, SP92-38, pp. 23-30, 1992
- [2] 大沢 一郎, “人間と対話する即応エージェントのモデル”, MACC 1993
- [3] 大沢 英一, “コンピュータは人間をだませるか -チューリングテストをめぐる-”, 情報処理, Vol. 34, No. 3, pp. 280-283, 1993
- [4] 佐藤 理史, “対話からの知識獲得に関する研究”, 文部省重点領域研究「音声対話」C 班会議資料, PASD-83-C-6, 1993
- [5] 中村 敏枝, “間(ま)の感性に関する心理学的研究”, 電気情報通信学会技術報告, HC92-71, 1993
- [6] 森屋 裕治, “対話予測を利用した音声による観光案内対話システム”, 電気情報通信学会技術報告, SP92-121, pp. 43-50, 1992
- [7] 山本 吉伸, “コミュニケーションの面白さ -しりとりはなぜおもしろい-”, 情報処理学会研究報告, HI 47-13, 1993
- [8] T. Winograd, 平賀 譲 訳, “コンピュータと認知を科学する”, 産業図書, 1989