

## 遺伝的アルゴリズムを用いた帰納的学習による 音声対話システムの有効性について

木村 泰知† 荒木 健治‡ 桃内 佳雄† 橋内香次†

†北海道学園大学大学院工学研究科 ‡北海道大学大学院工学研究科

†〒064-0926 札幌市中央区南26条西11丁目

‡〒060-8628 札幌市北区北13条西8丁目

E-mail: kimura@ai.eli.hokkai-s-u.ac.jp {araki,tochinai}@media.eng.hokudai.ac.jp  
momouchi@eli.hokkai-s-u.ac.jp

あらまし 本稿では、実際の対話例から応答を行うための学習手法について述べる。日常対話は雑談のような自然に行われる対話が多い。このような対話は、様々な話題があり、非文法的な発話も行われるため、予め与えられたテンプレートによる学習手法では対話例からのルール獲得は困難である。これらの問題を解決するために、我々は遺伝的アルゴリズムを用いた帰納的学習による音声対話システムを提案する。本手法は実際の音声対話例からルールの獲得を行い、交叉・突然変異・淘汰・帰納的学習を用いて適切な応答の生成を行う。本システムの評価実験により、正応答と準応答の合計が66.3%から76.1%に増加した。その結果、本手法が雑談に対して有効な手法であることが確認された。

キーワード 自然言語処理, 学習, 帰納的学習, 対話処理, 遺伝的アルゴリズム

## Effectiveness of Spoken Dialogue System using Inductive Learning with Genetic Algorithm

Yasutomo Kimura† Kenji Araki‡ Yoshio Momouchi† Koji Tochinai†

†Graduate School of Engineering, Hokkai-Gakuen University

‡Graduate School of Engineering, Hokkaido University

† S26-W11, Chuo-ku, Sapporo 064-0926, JAPAN

‡ N13-W8, Kita-ku, Sapporo 060-8628, JAPAN

E-mail: kimura@ai.eli.hokkai-s-u.ac.jp {araki,tochinai}@media.eng.hokudai.ac.jp  
momouchi@eli.hokkai-s-u.ac.jp

Abstract In this paper, we describe a learning method to perform the response from actual dialogue examples.

Our ordinary dialogue has spontaneous dialogue that relates to enjoyment of the topics and the dialogue process. The learning method using templates is very difficult to apply to a chat, which includes many varieties of the topics. To resolve this problem, we propose a method of Spoken Dialogue using Inductive Learning with Genetic Algorithm. Our proposed method generates a suitable response using crossover, mutation, selection, and inductive learning. The total of a correct response rate and a similar correct response rate increased from 66.3% to 76.1% using our proposed method. Therefore, we confirmed that our method is an effective method for a chat.

key words natural language processing, learning, inductive learning, dialogue processing, genetic algorithm

## 1 はじめに

最近、音声対話に関する研究や娯楽商品が注目を集めている。古くから対話の研究は行われ、古典的な対話システムに ELIZA がある。ELIZA [1] は精神病患者の情報収集のために作られ、使用単語の制限を行わず、自然な対話を行うことができる。しかし、曖昧な応答が多いという問題がある。ELIZA とは対照的な対話システムに SHRDLU [2] がある。SHRDLU は積み木の世界に限定した対話で、高い精度で応答が行うことができる。しかし、対象範囲の拡張は困難で、使用単語も限定されている。また、中川らの観光案内の中で協調的な応答を導入した音声対話システム [3] がある。しかし、応答を行うために予め観光案内の十分な準備が必要で、使用単語も限定されている。

日常対話はタスク志向 [4] ばかりでなく、雑談のような自然に行われる対話 [5] が多く存在する。雑談のような対話は話題が限定されず、対話自体を楽しむことが多い。さらに、何気ない対話のやりとりから、元気づけられたり、思いがけないことに気づかされたりすることがある。我々は、雑談のような対話が重要であると考え、加藤らは ELIZA を音声対話に拡張し、学習機能を追加したシステムを開発した [6]。しかし、加藤らの手法は枠組みが予め与えられ、制限が厳しい学習である。雑談は、様々な話題があり、非文法的な発話が行われるため、予め与えられたテンプレートによる学習手法では対話例からのルール獲得は困難である。我々は従来より機械翻訳に遺伝的アルゴリズムを用いた帰納的学習 (GA-IL) [7] を適用してきた。そこで、上の問題を解決するために我々は遺伝的アルゴリズムを用いた帰納的学習による音声対話システムを提案する。我々は本手法を "Spoken Dialogue system using Inductive Learning with Genetic Algorithm" (GA-ILSD) と呼ぶ。GA-ILSD は実際の音声対話の例からルールの獲得を行う。GA-ILSD で獲得されたルールが不足し、適切な応答ができない場合、ELIZA による応答を行う。これを ELIZA 応答と呼ぶ。ELIZA 応答

はユーザとシステムの対話を継続し、対話例を獲得するために行う。GA-ILSD は獲得したルールに対して交叉・突然変異・淘汰・帰納的学習を用いて応答を行う。これを GA-IL 応答と呼ぶ。本システムは、対話を継続することによって GA-IL 応答が次第に増加する。我々は、本稿で GA-IL の導入の評価について記述する。我々は音声対話システムに GA-IL を適用することによって、正応答と準応答の合計の割合が 66.3% から 76.1% へ向上することを確認した。それより、本手法は雑談に対して有効な手法であることを確認した。

## 2 処理過程

### 2.1 概要

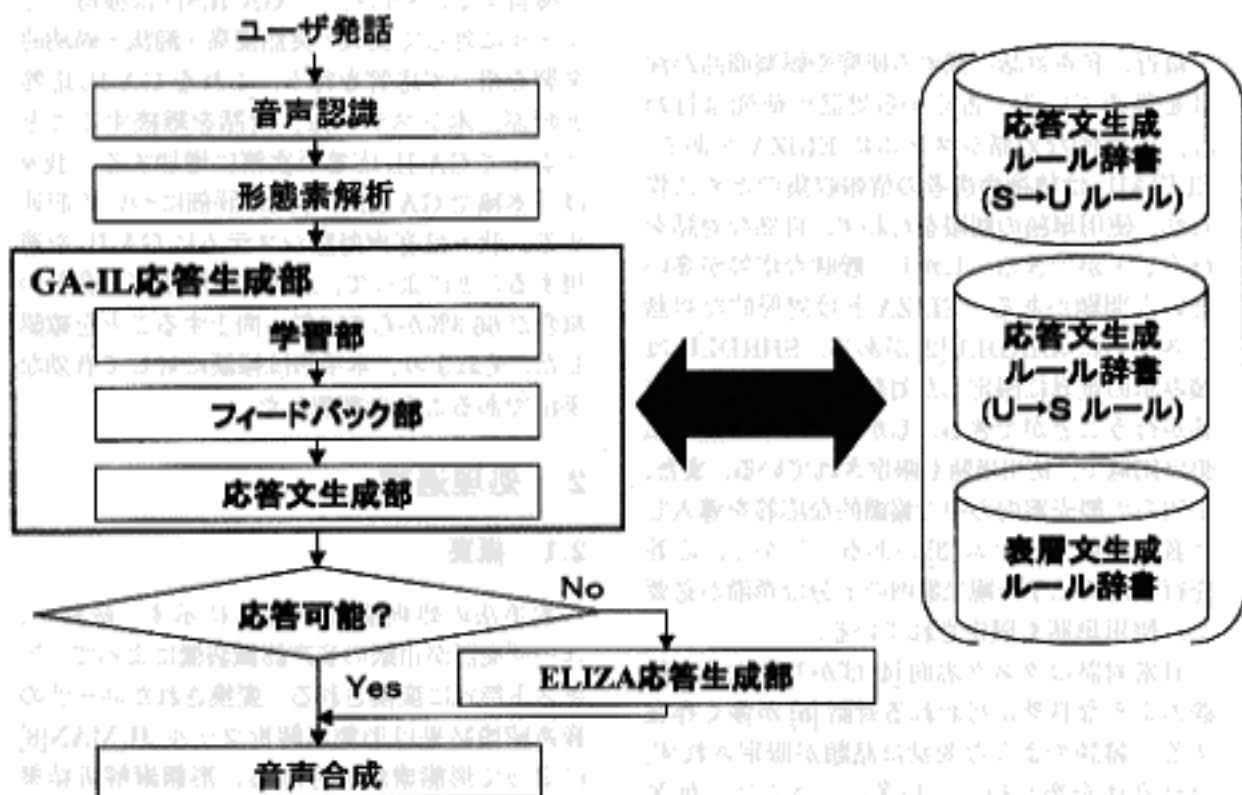
本手法の処理過程を図 1 に示す。最初に、ユーザ発話が市販の音声認識装置によって、テキスト形式に変換される。変換されたユーザの音声認識結果は形態素解析ツール JUMAN [8] によって形態素解析される。形態素解析結果は、品詞情報と区切り位置を決定するために利用される。GA-ILSD は GA-IL 応答と ELIZA 応答を用いて応答文を生成する。GA-IL 応答は対話例を利用し、学習部、フィードバック部、応答文生成部により応答を行う。学習部では対話例からルール獲得を行い、交叉・突然変異・帰納的学習を行う。フィードバック処理では、ユーザの発話から前回のシステム応答が正応答か誤応答かを判断し、その情報を利用して淘汰も行う。応答文生成部で獲得されたルールから最も適切な応答を行う。適切な応答ができない場合は、ELIZA 応答が行われる。最後に、市販の音声合成ツールを使用して音声による出力を行う。

### 2.2 GA-IL 応答生成部

GA-IL 応答生成部は学習部・フィードバック部・応答文生成部からなる。

#### 2.2.1 ルール獲得

図 2 は "GA-IL" によって対話例から応答に必要なルールを獲得する。ルールには応答文生成



### システム応答

図 1: 処理過程

ルールと表層文生成ルールの2つの種類がある。応答文生成ルールは対話例からシステムとユーザのやりとりを自立語列の対として獲得したものである。応答文生成ルールの使用法は、システムがユーザ発話の音声認識結果と応答文生成ルール辞書中の各ルールの左側を比較し、適切なルールであれば、そのルールの右側を使用して応答を行う。システムが対話例から応答文生成ルールを獲得する際には、さらに2通りの対の獲得方法に分けることができる。これらをU→SルールとS→Uルールと表記する。U→SルールとS→Uルールに分けることで対話例から効率よくルールを獲得できる。U→Sルールはルールの左側にユーザ発話の音声認識結果の自立語列を持ち、右側にシステムが応答した自立語列を持つ。S→Uルールは逆の順序の対である。表層文生成ルールは、システムが応答するために選択した応答文生成ルール

の右側の自立語に対して、付属語を加えて自然な応答を行う。全てのルールは単語列の他に品詞情報も保持する。システムの「休日にしたことは」に対して、ユーザが「テニスをしてました」と発話した場合、システムは左側に「休日した こと」、右側に「テニス してました」と品詞情報をもつS→Uルールを獲得する。

### 2.2.2 帰納的学習

システムは図3のように帰納的学習によりルールを生成する。処理の流れを以下に示す。

1. 学習対象ルールの選択  
GA-ILの獲得したルールから共通部分と差異部分を持ち、差異部分を両側に1つのみ持つルールのペアを選択
2. ルールの生成

- 差異部分のみを抽出したルール

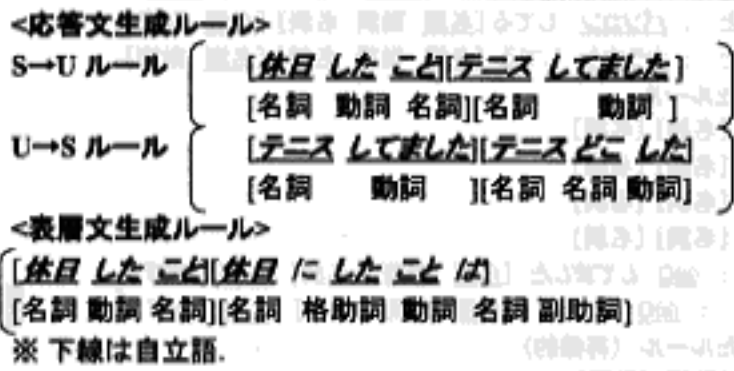
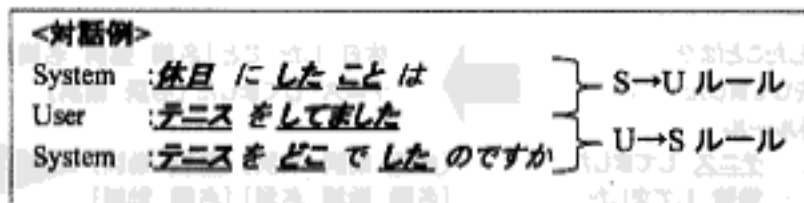


図 2: 対話例からのルール獲得

- 差異部分を変数に置き換えたルール

3. 繰り返し

このような場合、システムは図3の例のように対話例から(a)の4つの応答文生成ルールを獲得する。(a)の4つのルールから新しい6つの(b)のルールが存在する。さらに、変数を含んだルールからも(c)のようにルールを生成することができる。

2.2.3 遺伝的アルゴリズム

新しいルールが交叉・突然変異によって獲得される。獲得方法を以下に説明する。

1. 初期集団生成  
ユーザの発話に対して適用することができるルールを選択。選択されたルールを初期集団と呼ぶ。
2. 適応度の計算  
評価関数によってルールの適応度が計算される。  

$$\text{評価関数 (\%)} = \frac{\text{正応答度数}}{\text{正応答度数} + \text{誤応答度数}} \times 100 \quad (1)$$
3. 選択  
選択の詳細は2.2.4で記述する。
4. 交叉

2つルールを比較したときに品詞情報の並びが両側で2つ以上連続して同一であるルールのペアを選択し、図4のように同一品詞の中で一樣交叉を行う。

5. 突然変異  
突然変異率2%でランダムに行う。

2.2.4 フィードバック部

システムは予め与えられた誤りキーワードによって前回に使用されたルールの評価を行う。ユーザ発話に誤りキーワードが含まれている場合は前回の応答に使用されたルールに対して誤応答度数に1加え、含まれていない場合には正応答度数に1を加える。システムはこれらの正応答度数と誤応答度数を使用して、適応度を計算する。淘汰の条件は評価関数(1)において、75%以下のルールである。

2.2.5 応答文生成部

システムは応答文生成ルールと表層文生成ルールから応答を行う。システムは、ユーザ発話の音声認識結果と応答文生成ルール辞書にあるルールの左側を比較する。比較した結果、応答文生成ルール辞書中に適切なルールが存在した場合、そのルールの右側を使用して応答を行

<対話例>

[System] 休日にしたことは?

[User] テニスをしてました

休日 した こと [名詞 動詞 名詞]

テニス してました [名詞 動詞]

(a)対話例から直接獲得ルール

休日 した こと : テニス してました [名詞 動詞 名詞] [名詞 動詞]

学校 した こと : 勉強 してました [名詞 動詞 名詞] [名詞 動詞]

部屋 遊ぶ こと : パソコン してる [名詞 動詞 名詞] [名詞 動詞]

友達 遊ぶ こと : カラオケ してる [名詞 動詞 名詞] [名詞 動詞]

(b)帰納的学習により獲得したルール

休日 : テニス [名詞] [名詞]

学校 : 勉強 [名詞] [名詞]

部屋 : パソコン [名詞] [名詞]

友達 : カラオケ [名詞] [名詞]

@Q した こと : @Q してました [名詞 動詞 名詞] [名詞 動詞]

@Q 遊ぶ こと : @Q してる [名詞 動詞 名詞] [名詞 動詞]

(c)帰納的学習により獲得したルール (再帰的)

した : してました [動詞] [動詞]

遊ぶ : してる [動詞] [動詞]

@Q @? こと : @Q @? [名詞 動詞 名詞] [名詞 動詞]

図 3: 帰納的学習

う。選択された応答文生成ルールは自立語列からなるため、表層文生成ルールを使用して助詞等を含んだ自然な応答を生成する。応答に使用するルールの選択方法を優先順位の順番で以下に示す。

1. 一致率が 65% 以上で最も高いルールを選択 (65% 未満は ELIZA 応答生成部による処理)
2. 一致率が同じ場合、適応度が高いルールを選択
3. 適応度が同じ場合、最も新しく獲得されたルールを選択

2.3 ELIZA 応答生成部

システムが適切な GA-IL 応答を行うことができない場合に、ELIZA 応答が行われる。ELIZA 応答生成部は対話を継続し、対話例を獲得することが目的である。この処理で、システムはユーザ発話の結果を読み込み、予め与えたキーワードの検索を行う。もし、ユーザの音声認識結果にキーワードが含まれていた場合、キーワードに関連したルールにより応答文を

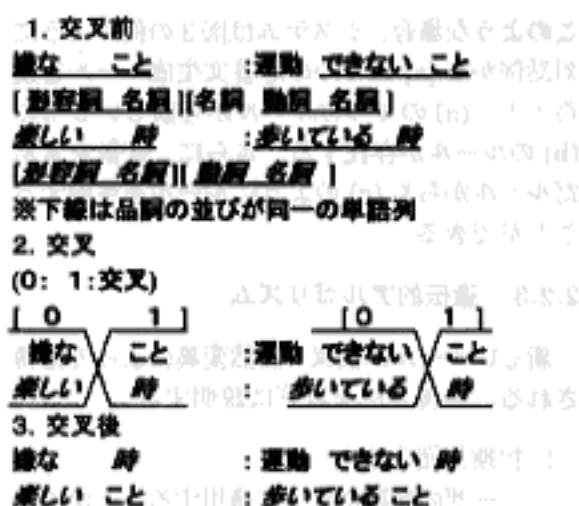


図 4: 遺伝的アルゴリズム

生成し、含まれていない場合にはいくつかの条件よりユーザ発話を引用した応答文や、対話の継続を促す応答文を生成する。今回の実験で ELIZA 応答に使用されたキーワードの個数は 196 個である。



### 3 評価実験

#### 3.1 評価方法

##### 3.1.1 初期条件

GA-ILSD が十分な辞書を予め持っていれば、システムは ELIZA 応答生成部を使用せず GA-IL 応答生成部のみで応答を行うことが可能である。しかし、我々は初期条件を一定にするために、GA-IL 応答生成部の辞書を空の状態から実験を始めた。

##### 3.1.2 システム応答評価の難しさ

音声対話による雑談の研究を行う際の障害は、対象問題を設定することが難しいことであり、次のようなことが考えられる。

1. どんな雑談を行うか？
2. 正解は？(どのように評価するか)

雑談は日常対話である。対話の中でどのような位置にあるかを考える必要がある。簡単に考えると、対話はタスク志向の対話と明確な目的をもたない対話に分けることができる。雑談は後者であり、タスクの達成及び、目的の達成によって評価を行うことはできない。応答の評価を行うには、ユーザが自然でかつ適切な応答と判断するか否かである。これは非常に曖昧な評価になる恐れがあり、ユーザによる揺れが大きいと考えられる。しかし、雑談には様々な話題があり、その話題も変化しやすく、正しい応答を定義することが困難である。たとえば、同じ質問に対する応答を考えた場合、一度正しかった応答が繰り返し使用しても正しいとは限らなく、人によっては全く異なる応答を求めていることがある。従って、何をもって正解の応答とすればよいかもはっきりしなくなる。このように、雑談は正解の応答が客観的に明確であるような問題とは異なった扱いが必要である。このため、当面の間は各応答の判断を被験者が主観的に行うことにする。被験者がシステムの応答に対して行う評価を正応答、準応答、誤応答、の3つに分類した。

1. 意味的に正しい応答

表 1: 応答の評価例

正応答	
U	あなたテニスをしなんでしょうね
S	テニスについて何か話してくれませんか
準応答	
U	その話はやめましょう
S	やめましょうのですね
誤応答	
U	もうやめようか
S	やめようって有名なの

- (a) 表現が自然・正応答
  - (b) 表現が不自然・準応答
2. 意味的に誤りである応答・誤応答

この評価を被験者が行うことにした。表 1 は被験者が行った応答評価の例を示す。

#### 3.2 ELIZA との比較

##### 3.2.1 実験方法

本手法に基づくシステムは、GA-IL 応答生成部と ELIZA 応答生成部をもつ。この実験では、GA-IL 応答生成部を除いた ELIZA と GA-ILSD を比較し評価を行う。一人の被験者が 2 つのシステムに対してそれぞれ対話実験を行なう。被験者はシステムが 1000 回応答を行うまで繰り返す。GA-ILSD の応答の評価は以下の 6 つに分類することができる。

- ① ELIZA 応答生成部から生成された正応答
- ② GA-IL 応答生成部から生成された正応答
- ③ ELIZA 応答生成部から生成された準応答
- ④ GA-IL 応答生成部から生成された準応答
- ⑤ ELIZA 応答生成部から生成された誤応答
- ⑥ GA-IL 応答生成部から生成された誤応答

ELIZA の応答の評価は GA-IL 応答生成部を行わないため、以下の 3 つに分類される

- ① ELIZA 応答生成部から生成された正応答
- ③ ELIZA 応答生成部から生成された準応答

表 2: 比較実験での音声認識結果

	cor	sim	err
ELIZA	61.6%	8.1%	30.3%
GA-ILSD	57.0%	9.5%	33.5%

cor: 正音声認識率

sim: 付属語以外は正しい音声認識率

err: 誤音声認識率

表 3: 比較実験における応答評価

	①	②	③	④	⑤	⑥
E (%)	43.3	-	23.0	-	33.7	-
G (%)	39.3	12.0	19.5	5.3	18.2	5.6

E: ELIZA

G: GA-ILSD

⑤ ELIZA 応答生成部から生成された誤応答

### 3.2.2 実験結果

ELIZA の有効な応答率 (①③) を図 5 に示す。そして、GA-ILSD の有効な応答率 (①②③④) を図 6 に示す。

### 3.2.3 考察

表 3 からわかるように、GA-IL を適用することで、正応答と準応答の合計の割合が 66.3% から 76.1% に向上した。表 2 からわかるように、GA-ILSD の実験における音声認識率は ELIZA の音声認識率より低いにもかかわらず、有効な応答が上回った。我々は本手法が雑談において有効であることを確認した。しかし、今回の実験では一人の被験者を対象としたため、複数の被験者を用いた実験を行う必要がある。それ故に、我々は次の実験で、複数被験者による GA-ILSD の個人差を評価した。

表 4: 複数被験者実験での音声認識結果

	A	B	C	D	E
cor(%)	53.5	51.7	48.9	37.5	61.2
sim(%)	3.4	4.6	10.1	13.1	7.1
err(%)	43.0	43.6	40.9	49.4	31.5

表 5: 複数被験者実験での有効応答

被験者	①	②	③	④	合計
A	19.3	13.6	20.5	3.3	56.7%
B	21.4	14.1	16.4	6.5	58.4%
C	36.0	1.4	25.9	2.8	66.1%
D	22.7	7.8	14.5	7.2	52.2%
E	14.1	1.4	38.1	9.4	63.0%

## 3.3 複数被験者による GA-ILSD の評価

### 3.3.1 実験方法

被験者は音声対話システムを扱ったことのない学部生 5 人が行う。それぞれの被験者は、GA-ILSD が 500 応答行うまで対話を続けた。対話は 1 週間以内で行う。我々はその 5 人の被験者をそれぞれ A, B, C, D, E と呼ぶ。

### 3.3.2 実験結果

この実験での、それぞれの被験者の音声認識結果を表 4 に示す。そして、各被験者による GA-ILSD の有効な応答率を表 5 に示す。

### 3.3.3 考察

GA-IL の有効応答 (②+④) が被験者 C を除いた被験者に対して 10% 以上であった。我々は、辞書の初期状態が空であることを考えると、500 応答の間で 10% は非常に高いと考える。被験者 C の場合、GA-IL の有効応答 (②+④) が少なかった。被験者 C はシステムが誤った音声認識によって、システムが誤応答をした場合も、その誤応答に対して合わせて対話を継続したため、似たような表現や内容が繰り返されることが少なかった。GA-ILSD は、被験者

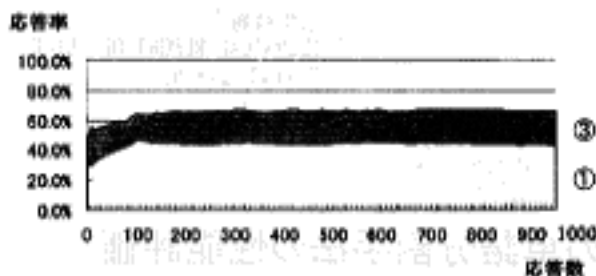


図 5: ELIZA

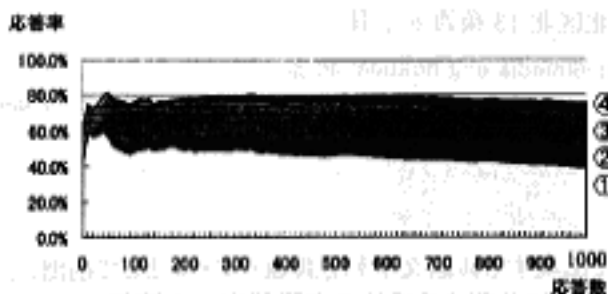


図 6: GA-ILSD

によって、有効な応答を行うまでに必要な対話例及び、時間にずれが生じる。被験者Cの場合、システムが500応答を行うまでの対話例では、学習を行うには少なかったと考えられる。被験者Cのような特別な場合を除けば、複数の被験者に対してGA-ILSDを用いて有効な応答を行うことができた。

#### 4 おわりに

本稿で、GA-ILSDを適用した音声対話システムの手法を提案した。評価実験の結果、本手法を用いた音声対話システムは、正応答と準応答の合計が76.1%であった。考察を行うためにELIZAと比較した結果、ELIZAは正応答と準応答の合計が66.3%であったため、GA-ILSDを適用することにより9.8ポイントの向上を確認した。さらに、複数の被験者に対して実験を行った結果、被験者に対してGA-ILSDの有効な応答を行えることを確認した。これより、我々は本手法が雑談に対して有効であることを確認した。

今後の予定は学習部とフィードバック部の改

良を考えている。学習部では、帰納的学習において共通部分の抽出を行う。フィードバック部では、システム応答に対してユーザが判断を行う処理の追加を行う。

謝辞 なお、本研究の一部は文部省科学研究費(No.09878070, No.10680367)及び北海学園大学ハイテク・リサーチ・センター研究費による補助のもとに行われた。

#### 参考文献

- [1] Weizenbaum, J. : ELIZA - A Computer Program for the Study of Natural Language Communication Between Man And Machine, Communications of the Association for Computing Machinery, Vol.9, No.1, pp36-45(1966).
- [2] Winograd, T. : Understanding Natural language, Academic Press, (1972).
- [3] 伊藤 敏彦, 小暮 悟, 中川 聖一: 協調的応答を備えた音声対話システムとその評価, 情報処理学会論文誌, Vol.39, No.5, pp1248-1257(1998).
- [4] R. W. Smith and S. A. Gordon: Effects of Variable Initiative on Linguistic Behavior in Human-Computer spoken Natural Language dialogue, The Association for Computational Linguistics, Vol.23, No.1, 141-167(1966).
- [5] 岡田 美智男: 何気ない行為を科学する, bit, Vol.30, No.12, 共立出版 (1998).
- [6] 加藤 誠巳, 谷 洋介, 桐越 孝之, 菅田 裕紀 "音声認識・音声合成を用いた Eliza 型 3DCG 会話ロボット"茶飲み友達"に関する検討", 情報第 57 回全大, 6C-3(1999-10).
- [7] 越前谷 博, 荒木 健治, 桃内 佳雄, 棚内 香次: 実例に基づく帰納的学習による機械翻訳手法における遺伝的アルゴリズムの適用とその有効性, 情報処理学会論文誌, Vol.37, No.8(1996).
- [8] 黒橋 慎夫, 長尾 真: 日本語形態素解析システム JUMAN version 3.61 (1999).