

ペットの反応を利用した言語表現手法に対する性能評価

山川 裕也[†] 荒木 健治[†]

[†] 北海道大学大学院情報科学研究科 〒060-0814 札幌市北区14条西9丁目
北海道大学大学院情報科学研究科
E-mail: †{yamakawa,araki}@media.eng.hokudai.ac.jp

あらまし 近年、ペットが人間にとって重要な役割を担うようになってきており、ペットに関連する様々な技術が研究されている。動物の言語についての研究も行われており、飼い主が考える動物の意思を人間の言語で表現することによって、動物とのより良い信頼関係が築けると考えられる。そこで、我々はこれまでペットとして飼育される動物を対象とした対話処理手法を提案している。本手法では、人間の発話と動物の反応を一組の入力として、出力となる飼い主が考える動物の意思を言語で表現する。入力と出力の組から応答規則を獲得する方法として帰納的学習を用いる。本稿では、評価実験を行った結果から提案した手法の有効性と、これまで1人であった被験者を増やし、複数の被験者による性能の評価について述べる。

キーワード 言語表現, 帰納的学習, 対話処理, ペット動物

Performance Evaluation for Spoken Expression Method Using Reaction of Pet Animals

Yuya YAMAKAWA[†] and Kenji ARAKI[†]

[†] Graduate School of Engineering, Hokkaido University
Kita 13 Nishi 8, Kita-ku, Sapporo-si, Hokkaido, 060-8628 Japan
E-mail: †{yamakawa,araki}@media.eng.hokudai.ac.jp

Abstract Pet animals are important for human being. It is desired to express animals' intention whose an owner thinks with human language. We propose a spoken dialogue processing method for pet animals. We use Inductive Learning for dialogue between an owner and his/her pet animal. The system based on this method acquires response rules from input that owner's utterance and reaction of animals, and express animals' intention whose an owner thinks by human language. In this paper, we confirmed the effectiveness of our method and the performance evaluation of several experimental subjects.

Key words Spoken Expression, Inductive Learning, Dialogue Processing, Pet Animals

1. はじめに

近年、動物は人間が精神的に良好な状態を保つために重要な役割を果たしている。人間同士のコミュニケーションが疎遠になり、孤独な状況に陥りやすい現代社会において、人間の心理的な支えとして動物の果たす役割は大きく、その重要性は今後益々増加していくと考えられる。

実際に動物を飼うことは容易ではないが、動物を扱った様々なツールが販売されている。言葉を話す玩具としてファービー [1] が1998年に販売され、1年程で全世界1200万台という驚異的な売上を達成した。また、家庭用自律移動ペットロボットとしてアイボ [2] が1999年に販売された。アイボは高価にも関わらず反響が大きく、ペットが人間にとっ

てどれほど重要か示した。(注1)

またペットに目を向けると、ドイツの家庭で飼育されているあるボーダーコリー種の犬は生後10ヶ月から物の名前を覚え始め、覚えている単語数は200語に上ったという事例も報告されている[3]。この犬の言語獲得能力は人間の3歳児並みであると言われている。

2002年には犬の言葉を翻訳するツール[4]が販売され、話題になった。多くの人がそのツールを買い、遂には猫の言葉を翻訳するツール[5]まで販売された。これらのツールは出力が曖昧なため研究という観点からは疑問が残るものの、実在の動物の鳴き声を人間の言語で表現する可能性を示した。

実在の動物は人間ではないため人間の言葉を話すことができない。実際に動物に人間と同じような思考が存在するのかは分からないが、飼い主が考える動物の発話というものは存在している。我々は飼い主が考える動物の発話を人間の言語で表現することによって、動物とのより良い信頼関係が築けると考えている。

飼い主が考える動物の発話を人間の言語で表現する方法として、我々は飼い主とペットとの対話に着目した[6]。人間の発話と動物の反応を一組の入力として、出力となる飼い主が考える動物の発話を言語で表現する。それにより、従来手法では困難であった飼い主が考える動物の発話を人間の言語で表現する手法の実現を目的とする。

2. 考えられる言語処理手法

動物の意思を推測し、人間の言語で表現する方法としては、動物の鳴き声を音声認識し、人間の言語に翻訳することがまず考えられる。動物の意思を推測し、翻訳によって言語表現するためには、動物が音声を発することによって多彩な意味を伝達している必要がある。人間の言語には文法が存在し、これによって有限の記号から無限の意味を表現できると考えられている[7]。動物にも人間の文法に近いものが存在する知見として、ある種の小鳥(ジュウシマツ)は複雑な産出規則に基づいて

歌を歌うことが報告されている。しかし、この小鳥の歌に見られる産出規則は様々なバリエーションの歌を歌うことを可能にするために存在し、複雑な意思を伝達することは不可能である[8]。このように現時点では、動物に関して人間の文法に匹敵するものを持つという事例は報告されていない[9]。

また、複雑な意味を音声によって伝達するためには、バリエーション豊かな音声を発することができる調音構造が備わっている必要がある。人間は二足歩行を行うことで調音に関わる骨格の構造に制限が無くなった結果、自在に調音を行うことが可能となっており、人間に近い知能を持つと言われるチンパンジーでも人間のようなバリエーションのある音声を発することは不可能である[10]。

さらに、鳴き声を発する頻度の問題もある。無口な人や雄弁な人がいるように、人間には個人差がある。同様に動物にも個体差があり、あまり鳴き声を発さない動物に対しては翻訳する機会が少なくなる。

以上のことから、動物の鳴き声を直接人間の言語に変換しても具体性を欠いた曖昧なものとなり、不満を感じるものになると考えられる。従って、動物の鳴き声のみを用いた翻訳という方法では言語表現手法として不十分である。

翻訳以外の可能性を探るため、我々は人間が何を手掛かりにして動物の意思を推測しているのかを調査した。実際のところ、人間は動物とのやりとりにおける動物の反応に何らかの内的状態が存在するものとして解釈を行っている[11]。身近な動物として犬や猫等のペットが挙げられる。飼い主はペット動物のしぐさや状態からペットの気持ちを推測して判断することができる。飼い主である人間はペットとの関係の中で人と共通する「心」の表情をペットの中に見出していると考えられる。人間の解釈したペットの意思は多種多様なものとなるため、飼い主の解釈したペットの意思をペットの応答とすることで言語表現が可能であると考えられる。実際には飼い主が解釈したペットの意思が本当にペットが思っていることであるとは限らないため、厳密にペットの意思を表現しているとは言えないかもしれない。しかし、言葉が話せる人間同士の

(注1): 後継機では人間の言葉を話すような仕様になっているものもある。



図 1 本システムのイメージ

表 1 未知記号列からの対応関係抽出の例

入力 1	$\alpha \theta \sigma \psi \delta \lambda \vartheta$	
入力 2	$\Xi \Sigma \psi \delta \Upsilon \Phi \Theta$	
セグメント 1	$\alpha \theta \sigma$	$\Xi \Sigma$
セグメント 2	$\psi \delta$	$\psi \delta$
セグメント 3	$\lambda \vartheta$	$\Upsilon \Phi \Theta$

対話においても相手の話した内容を聞いた人が解釈しなければならないため、お互いの話した意思が正確に伝わるわけではない。従って、人間が解釈した意思をペットの応答文とする対話処理が現時点で実現可能なペットに対する人間の満足度を向上させる方法であると考えられる。本システムのイメージを図 1 に示す。

3. 対話処理の概要

3.1 帰納的学習

帰納的学習とは、我々が提案した実例からそこに内在する規則を獲得する方法であり、共通部分、差異部分を再帰的に抽出するという処理を行う [12] [13]。

表 1 に示すように、入力 1、入力 2 のような対応関係を有する 2 つの未知記号を人間が見た場合、人間はまず 2 つの記号列に共通な部分を検出すると考えられる。次にその両側の差異部分を出現順に対応付け、表 1 のセグメント 1、2、3 のような対応関係を考える。このようにして抽出されたセグメントから表 2 に示すような 3 つのプリミティブを得る。以上のように実例より共通部分と差異部分を多段階に抽出することにより知識を獲得する。

表 2 セグメントからのプリミティブ抽出の例

セグメント 1	$\alpha \theta \sigma$	$\Upsilon \Phi \Theta$
セグメント 2	$\theta \sigma \gamma \mu$	$\Phi \Theta \Sigma$
プリミティブ 1	\aleph	Υ
プリミティブ 2	$\theta \sigma$	$\Phi \Theta$
プリミティブ 3	$\gamma \mu$	Σ

3.2 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処理手法

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処理手法 [14] では、過去の 2 つの対話例を比較して字面上の共通部分と差異部分を見つけ出す。その組み合わせを応答文生成ルールとしてルール辞書に登録し、すでに登録されている辞書中のルールと比較を行い、共通部分と差異部分を抽出することによって再帰的にルール獲得を行う。

このことにより、比較的少数の事例から応答を行うことができ、例外的な事例に対してもそれをルールとして保持することで対処することが可能となる。

3.3 対話処理手法

本手法の対話処理部分の流れを図 2、対話例からのルール抽出例を図 3 に示す。入力文から日本語形態素解析ツール JUMAN [15] を用いて自立語列が生成され、文中の自立語がルール抽出に利用される。日本語の形態素解析では自立語と付属語を分けて考えることが多い。付属語は自立語と組み合わせ初めて意味を持つため、単独では意味を持たない。応答文生成ルールは微妙な表現の変化にも頑健な処理を行うため、単独で意味を持つ自立語のみを用いる。また、自立語列と入力文の組から応答文復元ルールが生成される。この応答文復元ルールは、応答文生成部において出力された自立語列から付属語を付加して、完全な文として出力するために用いられる。

応答文生成部では、入力された人間の発話と動物の状態の組に対して、応答文生成ルール辞書を参照して応答文を生成する。本手法で入力となる動物の状態を以下に示す。

- (1) 動かず (飼い主を見て)
- (2) 回避行動
- (3) 身体接触

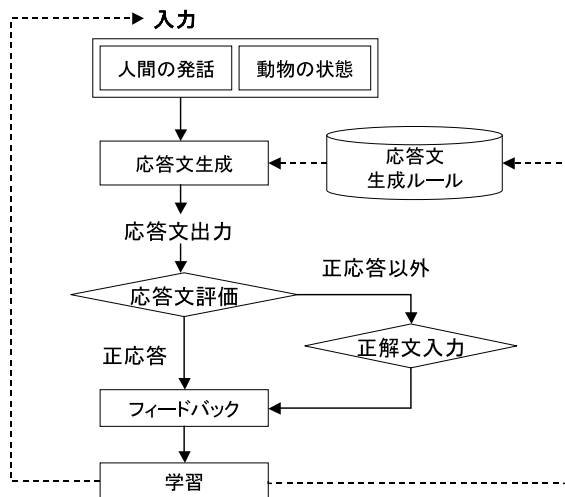


図 2 対話処理の流れ

- (4) 接近
- (5) 強い反応(モノを噛む, つかむ等)
- (6) 命令等に対する反応
- (7) その他(フレームアウト)

これらの状態はそれぞれの状態を複合して入力することが可能である。利用する動物の状態については[11]を参考にした。対話例から抽出されたルールには、対話スコアと状況スコアが記録される。対話スコア、状況スコアはそれぞれ、応答文が対話として成り立つ度合い、応答文が動物の状態と対応している度合いを判断する尺度である。状況スコアは動物の状態毎に値を保持している。これらのスコアは応答文を生成する際にルールを選択するために行う尤度計算に利用される。入力された人間の発話は自立語列に変換され、その自立語列と辞書中のルールの入力文テンプレートを参照して、使用するルールを決定する。自立語が完全に一致するルールの組み合わせが存在しない場合は、自立語列の一致率が50%以上で最も一致率が高いルールの組み合わせを選択する。

使用可能なルールの組み合わせが複数存在する場合、以下に示す評価式で尤度計算を行い、スコアが大きくなる組み合わせを採用する。

$$\text{尤度} = \sum_{\text{使用するルール}} [(\alpha * DS) + \beta * \sum ST]$$

ここでの DS は対話スコア、ST は状況スコアを示し、係数はそれぞれ $\alpha = 0.5$, $\beta = 1.0$ とする。

応答文生成時に使用可能なルールが複数存在する場合のルールの決定例を図4に示す。この例では、「こっちに来てちょうだい」という入力に対して、ルール1とルール3、ルール2とルール3という2つのルール選択候補が存在する。入力時の動物との距離が近い場合は尤度計算の際に近距離の状況スコアが利用される。その結果、ルール1とルール3の組み合わせが応答文生成に利用される。入力時に動物との距離が遠い場合はルール2とルール3の組み合わせが利用される。この例では説明の都合上状況スコアを近距離と遠距離のみとしたが、実際には本手法で利用している動物の状態全てに対して状況スコアが記録され、ルール選択の際に利用される。

評価基準、フィードバック方法を表3に示す。出力された応答文に対してはユーザが評価を行う。出力された応答文が正応答以外の場合はユーザが校正処理を行う。

フィードバック処理部では、応答文生成部で出力された応答文に対して行ったユーザの評価に基づいて、使用したルールに対してフィードバック処理を行い、対話スコアと状況スコアを増減させる。その結果、次回以降のルール選択においてより有効な応答文を出力する。

学習処理部では、出力された応答文に対してユーザが校正を行った際に、その校正された応答文と応答文生成ルール辞書とを比較して、共通部分と差異部分に注目して帰納的学習によるルール抽出を行う。その際に、構成された応答文を実例ルール、機能的学習により抽出されたルールを抽出ルールと呼び、それぞれのルールの初期状態を以下のようにする。

・実例ルールの初期状態

$$DS=1, \text{入力に対応する } ST=1$$

・抽出ルールの初期状態

$$DS=0, ST=0$$

このようにすると、初期状態においては信頼度が高いと考えられる実例ルールが優先的に選択されやすくなる。

	入力	出力
対話例	今日のおやつはおいしいですか	今日のジャーキーは大好きです
対話例2	今日ごはんおいしい	今日骨付き肉 大好きです



実例ルール	今日 おやつ おいしい: 今日 ジャーキー 大好きです
抽出ルール1	今日 @ おいしい: 今日 @ 大好きです
抽出ルール2	ごはん: 骨付き肉
抽出ルール3	おやつ: ジャーキー

図3 ルール抽出の例

ルール	DS	ST(近距離)	ST(遠距離)
1 @1/1/来て/ちょうだい:@1/1/行き/ます	2	3	-1
2 @1/1/来て/ちょうだい:@1/1/行き/たく/あり/ませ/A	1	-1	3
3 こっち:そっち	2	2	2

動物との距離が近い場合

動物との距離が遠い場合

使用ルール	評価式	使用ルール	評価式
1と3	$(2+2)*0.3+(3+2)*1.0=6.2$	1と3	$(2+2)*0.3+(-1+2)*1.0=2.2$
2と3	$(1+2)*0.3+(-1+2)*1.0=1.9$	2と3	$(1+2)*0.3+(3+2)*1.0=5.9$

出力: そっちに行きます

出力: そっちに行きたくありません

図4 ルール決定例

表3 評価基準とフィードバック方法

正応答	応答文として意味的に正しく、動物の状態にも対応する応答文 (DS と入力に対応する全ての ST の値を 1 増加)
誤応答	応答文として意味的に正しいが、動物の状態と対応しない応答文 (DS を 1 増加, 全ての ST の値を 1 減少)
文生成ミス	明らかに成り立たない文を生成した場合 (DS と対応する全ての ST の値を 1 減少)

4. 評価実験

前述の手法を搭載したシステムによる応答文評価のシミュレーション実験を行い、ペットの反応を利用した言語表現手法についての評価実験を行った。4.1 では実データを用いてこのシステムの性能を評価している。4.2 ではこのシステムを複数の被験者が評価を行い、それぞれのユーザに適應させた場合のシステムの性能を評価している。

ここでの評価基準、フィードバック方法は表3とする。初期状態を一定にするため、両実験ともルール辞書の初期状態は空の状態から実験を始めた。

4.1 実データを用いた応答文評価実験

実験の準備段階として、ペットの犬とその飼い主である理系大学院生とのやり取りをビデオカメラで撮影した。後日その映像を基に対話データとなる入力文、動物の状態、不正解時の正解文を第一

表4 実データによる応答文実験結果

正応答	71.7%(828)
誤応答	24.1%(278)
文生成ミス	2.94%(34)
未応答	1.30%(15)
発話数	1155
ルール数	670

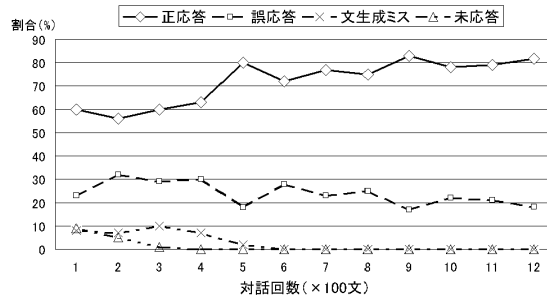


図5 実験の推移

表5 飼い主の発話例

ハル、おいで
ハル、お座り
ハル、お手
ハル、伏せ
ハル、待て

表6 4.1 で実際に使用した実験データの例

・ 飼い主の発話の例
飼い主の発話: ハル、こっちおいで
動物の状態: 接近
出力された応答文: はい、何でしょう
・ 誤応答の例
飼い主の発話: ハル、待て
動物の状態: 後退
分かりました、待ちます

著者が人手により作成した。飼い主からペット犬へは主に注意喚起、疑問、行為誘発、接近要求の順に発話されているというデータがある [11]。このうち動物の反応を得やすいということから、飼い主には主に接近要求、行為誘発発話を行うよう指示した。

総対話データ数は 1,155 発話になった。作成したデータを基にシステムが出力した応答文に対して第一著者が評価、不正解時の正解文入力を行った。主な飼い主の発話例と実際の実験データの例を表5、表6に示す。また、実験結果を表4、実験の推移を図5に示す。

表 7 被験者毎の実験結果

被験者	正応答	誤応答	文生成ミス	未応答	対話数	ルール数
A	77.7%(366)	17.6%(83)	0.85%(4)	2.97%(14)	471	164
B	55.8%(236)	37.6%(159)	0.71%(3)	5.91%(25)	423	141
C	66.0%(281)	29.3%(125)	0.47%(2)	4.23%(18)	426	111
D	63.1%(280)	33.3%(148)	1.80%(8)	1.80%(8)	444	270
E	53.0%(251)	31.9%(151)	5.06%(24)	9.92%(47)	474	422
F	44.7%(209)	50.9%(238)	2.14%(10)	2.35%(11)	468	560
G	44.1%(189)	53.6%(230)	0%(0)	2.33%(10)	429	173
H	62.1%(363)	34.5%(202)	0.17%(1)	3.25%(19)	585	120
I	65.0%(281)	29.9%(129)	0.69%(3)	4.40%(19)	432	198
J	68.2%(319)	28.6%(134)	0.21%(1)	2.99%(14)	468	170
平均(合計)	60.0%(2775)	34.9%(1599)	0.77%(56)	3.33%(185)	461(4620)	212(2329)

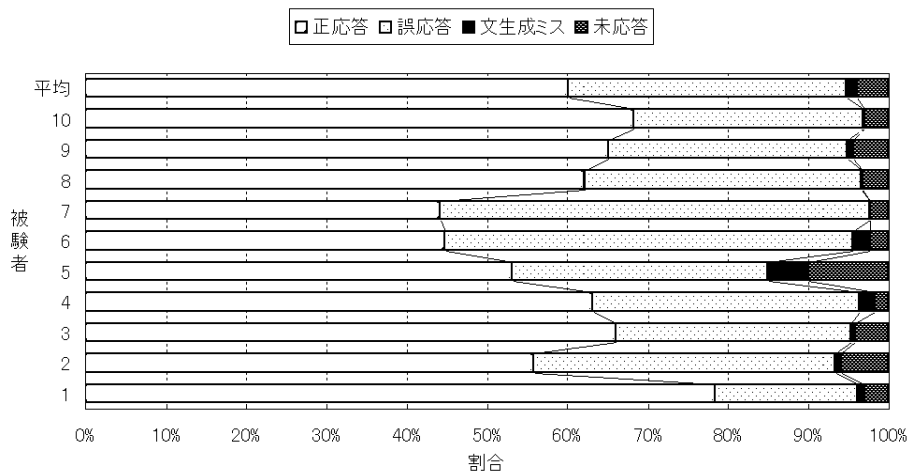


図 6 被験者毎の応答の割合

結果から対話が進むに従って、正応答の割合が増大し、誤応答、文生成ミス、未応答が減少している。最終的に正応答は80%前後まで上昇した。このことから学習の効果が現れていることが分かる。

ルールについては、対話数と比較してルール数が少ない。これは飼い主が犬に話し掛ける場合、短い発話が多くなり発話のパリエーションが少なくなるためである。動物を相手にする場合、対話からのルール獲得が人間を相手にする場合に比べ難しい。今後はシステムが少ないルールからでも正確な応答を行うことができるように改良する必要がある。

4.2 複数被験者による応答文評価実験

次に本システムが被験者が変わっても有効かどうかを確認するため、上記のシステムを用いて計10名の理系大学生、大学院生(男性:7名、女性:3名)の被験者による応答文評価実験を行った。入力文

や満足する応答というのは飼い主に応じて異なるため個人差が出てくる。その個人差に対応するために被験者自身が入力データを作成した。実験には犬及びその犬の飼い主が必要であるため実験環境を整えることが困難である。そのため被験者には4.1と同様の状況を想定して貰い、犬の飼い主になったつもりで入力文を作成するよう指示した。4.1と同様に入力文として主に接近要求、行為誘発発話の作成を指示し、被験者それぞれがデータを作成した。

評価基準は表3を基にするが、細かい基準についてはすべて被験者の主観により決定した。また、入力となる動物の状態は4.1と同様だが、(7)その他(フレームアウト)は除いた。

被験者毎の実験結果を表7、被験者毎の応答の割合を図6、実際に作成した飼い主の発話例を表8、

表 8 飼い主の発話例

お手してごらん	こっちこっち
もう少し待って	、お預け
、ボールをとって	、お回り
ごはんだからおいで	散歩に出かけるよ
、ジャンプ!	トイレに入りなさい
ハウスに入って寝なさい	、ダッシュ
あっちへ行ってなさい	お父さん呼んできて

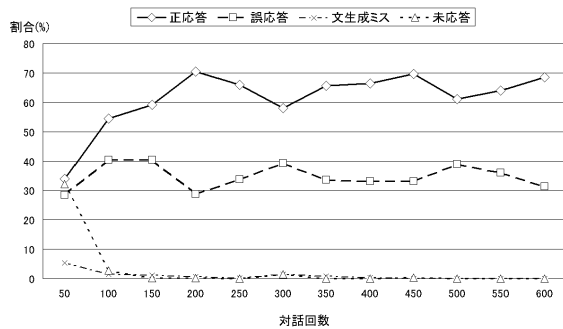


図 7 結果の平均データの推移

結果の平均データの推移を図 7 に示す。

図 6 が示すように正応答率は最低 44.1% から最高 77.7% と被験者によって最大 30% 以上の差があり、文生成ミス、未応答数については飼い主からの発話が多いほど増加する傾向にある。これらは飼い主に応じて満足する評価の基準、発話の種類が異なるためであると考えられる。また、表 8 から分かるように飼い主は様々な言葉でペットに話し掛ける。このように個人差が大きいことから、一般的に誰にでも対応できるとは言えない。現時点では飼い主毎に学習を行い、個人に適應させた形で使用する必要がある。

5. ま と め

本稿では、飼い主の発話とそれに対するペットの反応から帰納的学習を用いて飼い主の考えるペットの応答文を出力する手法について述べ、システムが作成した言語表現に対する性能評価を行った。4.1 の実験の結果から学習が進むに従って、正応答率が 20% 程上昇しており、ペットの状態に応じたより具体的な応答が可能になることを確認した。また、個人に適應させた出力を行うことができるかを確認するために、複数の被験者により評価を行い、正応答率に最大 30% 以上の差があるように個

人差はあるが、個人に適應させることである程度の応答が可能であることを確認した。今後の課題としては、実環境下で複数被験者に対して同様の実験を行うこと、データ作成のように今回は人手で処理した部分をシステムがリアルタイム処理をできるようにすることが挙げられる。但し、リアルタイム処理を行うためには困難な問題が数多く存在する。また、シミュレーションに限ればもっと様々な状況を想定して実験を行うことができる。さらに、システムに帰納的学習以外の手法を用いて比較実験を行うこと等が挙げられる。

文 献

- [1] http://www.hasbro.com/pl/page.viewproduct/product_id.13347/dn/tigertoys/default.cfm/
- [2] http://www.e-kasuga.co.jp/home/etc_aibo/aibo.htm
- [3] Juliane Kaminski, Josep Call, and Julia Fischer: "Word Learning in a Domestic Dog: Evidence for "Fast Mapping"", SCIENCE, vol 304, 11 JUNE 2004.
- [4] <http://www.takaratoys.co.jp/bowlingual/>
- [5] <http://www.takaratoys.co.jp/meowlingual/>
- [6] 前田宏幸, 荒木健治, 柁内香次: "コミュニケーション支援を目的とした帰納的学習によるペットを対象とした対話処理手法の性能評価", 信学技報 NLC2003-26 ,pp.7-12(2003-11)
- [7] 岡ノ谷一夫: "小鳥たちの歌の<文法>-アナロジーに基づく言語の期限の生物学-", 日本音響学会誌 54 巻 6 号 pp.457-465(2001-07).
- [8] 岡ノ谷一夫: 言語の発生, 渡辺 茂 (編) "心の比較認知科学", ミネルバ書房 (2000).
- [9] 渡辺茂: "動物に文法はあるのか", 「言語」, 大修館書店, 2002 年 4 月号.
- [10] 正高信男: "言葉の誕生 行動学からみた言語起源論", 紀伊国屋書店 (1991).
- [11] 藤崎亜由子: "人はペット動物の「心」をどう理解するか: イヌ・ネコへの言葉かけの分析から", 発達心理学研究, 2002, 第 13 巻, 第 2 号, 109-121.
- [12] K.Araki and K.Tochinai: "Effectiveness of Natural Language Processing Method Using Inductive Learning", Proceedings of the IASTED International Conference ARTIFICIAL INTELLIGENCE AND SOFT COMPUTING, pp.295-300, May, 2001.
- [13] 荒木健治: "自然言語処理とはじめ - 言葉を覚え会話のできるコンピュータ -", 森北出版, 2004.
- [14] 木村泰知, 荒木健治, 柁内佳雄, 柁内香次: "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処理手法", 電子情報通信学会論文誌, Vol.J84-D-II, No.9, pp.2079-2091(2001).
- [15] 黒橋禎夫, 長尾真: "日本語形態素解析システム JUMAN", version3.61.