

[サーベイ] 言語モデルと音声対話

井ノ上直己†

†KDD 研究所

〒 356-8502 埼玉県上福岡市大原 2-1-15

あらまし 音声認識技術の飛躍的な進歩の結果、人がコンピュータと自由に対話できる時代が現実見を帯びてきた。人がコンピュータと自由に対話するためには、音声認識・音声合成技術だけでなく、言語理解技術、音声対話処理技術が必須である。このような技術を利用した音声対話システムは航空機や列車の発車時刻・料金検索などのタスクを対象に、今までに多くの研究がなされてきた。しかし、実用化レベルに達しているものは、まだほとんど無いのが現状である。そこには、困難な多くの課題が残されており、今後ますます研究を進める必要がある。本稿では、現状の言語理解技術、音声対話処理技術について概観する。

キーワード 言語モデル, 音声対話システム, 対話モデル

Language Model and Spoken Dialogue

Naomi INOUE†

†KDD R&D Laboratories

2-1-15 Ohara Kamihukuoka-shi, Saitama 356-8502

Abstract As speech recognition accuracy is improving, it is expected that we can have a conversation with system spontaneously. In order to realize human-system conversation, spoken language understanding and spoken dialogue management is necessary as well as speech recognition and speech synthesis. Although many spoken dialogue systems are developed for air travel reservation and train timetable query, few systems are deployed and used by real users because of many issues. In this paper, I present studies relevant to spoken language understanding and spoken dialogue management.

key words Language Model, Spoken Dialogue System, Dialogue Model

1. はじめに

音声対話システムの研究は、1970年代の初めにARPAのプロジェクトとして米国で開始された音声理解の研究に溯ることができる。このプロジェクトでは、音声処理と言語処理とを統合して音声認識を行うことに焦点が置かれ、HARPYやHEARSAY等のシステムが開発された。しかし、当時は音声処理技術、言語処理技術だけでなく、コンピュータ技術そのものもまだ未熟であり、大きな成果が上げられなかった^[1]。1980年代に入り、音声認識において統計的手法の有効性が明らかになり、コンピュータ技術の進歩とあいまって、音声認識の研究は特定話者を対象とした孤立単語認識から不特定話者を対象とした連続音声認識の研究へと進歩してきた。さらに、近年では日常会話と同様の自由発話認識が活発に研究されるに至っている。このような背景から、音声により人間とシステムとが自由に対話できる音声対話システムへの期待が高まってきており、各国で多くの研究が行われている。しかし、実用化レベルに達した音声対話システムは、まだほとんど無いのが現状である。本稿では、音声対話システムの構築のための言語理解および音声対話処理に関し、近年の研究動向について概観する。

2. 言語モデル

音声対話システムにおける言語モデルの役割は、入力された音声を理解して対話システムを正しく動作させることである。しかし、従来から研究が行われてきた音声対話システムは、データベース検索が主なタスクであり、入力文の理解とは主に検索に必要なスロットを埋めることであった。そこでは、構文構造や意味構造などの深い構造を理解する必要性は重要視されていなかったと考えられる¹。そのため、言語モデルは、文法的に正しい文だけを認識し、認識処理における探索空間を小さくして認識性能を向上させることに利用されている。

本稿では、言語モデルに関しては概略のみを示すが、さらに詳しい解説は文献^[2]等を参照されたい。

2.1 音声処理と言語処理の統合

従来、言語モデルとしてCFGや有限状態オートマトンなどが利用されることもあった^{[4][5][6]}。このような文法を利用することにより、認識結果として構文構造の出力も可能であった。しかし、音声対話における発声バリエーションが大きいため、限定されたタスクであっても入力される文を受理できるCFGを作成することは困難であるという問題もあった^[2]。一方、N-gramモデルは、システムが受理できる文

¹ 音声翻訳などのように、構文構造や意味構造が音声認識後の処理で必要になるタスクもある。

のカバレッジも大きく、探索空間が小さくできるため、言語モデルとしてN-gramが良く利用されている。また、単語N-gramだけでなく単語の属するカテゴリを用いるクラスN-gram^[7]や可変長N-gramの研究も行われている^{[8][9]}。このようにN-gramが認識性能向上に寄与できるようになった背景には、対話コーパスの充実があげられる^{[10][11][12][13][14]}。

上記は、1文内において言語モデルを利用する認識手法であるが、次発話を予測することで探索空間を小さくして認識性能の向上を図ることも提案されている。例えば、プラン認識モデルを次発話の予測に利用する方法^[15]や次発話内の発話内行為を予測する方法^[16]が提案されている。

2.2 頑健な言語処理

対話音声では未知語の出現や非文法的な表現など書き言葉を対象とした言語処理では扱ってこなかった多くの現象が生じることが知られている^[17]。音声対話システムもこれらの現象は取り扱える必要がある。まず、未知語の処理として、システムが扱う文法とは別に、任意の音韻あるいは音節の組み合わせを許す文法を利用する方法^[18]がある。最近では、“filler fragments”と呼ばれる単語よりも小さい単位(subword)を辞書に登録し、通常の既知単語と同様に扱う方法^[19]やこの手法に単語クラスも考慮する方法^[20]が考案されている。

一方、小規模なタスクではキーワードスポテイング手法を用いてキーワードを検出し、その後に言語処理を行う方が効果的である。このような場合でも、キーワード以外の部分にGarbageモデル^[21]を利用する方が高い精度を得ることができる。

N-gramは局所的な言語モデルであり、隣接する単語間の関係を示している。これに対し、もっと長いスパンを考慮する手法としてトリガモデルが提案されている^[22]。

3. 対話モデル

音声対話システムを実現するためには、利用者の発声した内容を理解し、その内容に対して適切に応答しなければならない。具体的な処理手法は、音声対話システムが適用されるタスクに依存すると考えられるため、本章では対話システムを構築する際に考慮しなければならない項目を対話理解および発話の行為に関する理論にそって示す。

3.1 対話の公準

話し手が聞き手へ情報を伝達するために、下記の守るべき4つの基本原則^([23]から引用)が哲学者Griceによって提唱されている^[24]。

- (1) 質の公準：
根拠のある事実や真であると思うことを告げる。
- (2) 量の公準：
必要最低限の情報を伝える。
- (3) 関係の公準：
話し手と聞き手は互いに関連した事柄を話す。
- (4) 様式の公準：
明確に，簡潔に，順序立てて話す。

システムと人間との対話においても同様にこれらの基本原則は守るべきである^[25]。

3.2 語用論

対話は，話し手と聞き手が置かれた状況によってその表現が変わってくる。発話表現の意味をその発話が行われた状況との関係で捉えるのが語用論である。状況に従って，話す必要がないものは省略されたり，「それ」「あれ」と行った指示詞を利用することで明示しない，といったことが生じる。入力された発話内容を理解するためには，対話システムでは省略の補完や照応の解決が必要となる。省略の補完や照応などに関する種々の研究が今までなされてきている（例えば，^{[26][27]}）が，このような解析には，一般知識を利用した推論が必要になる場合もあり，現状ではごく限られた分野に限定せざるを得ない。

3.3 発話行為

話し手は何らかの意図を持って発話という行為を行うと考える。その意図と発話表現とを結び付ける理論として Speech Act と呼ばれる理論がある^[28]。文献^[28]では，発話の行為には以下の4つの段階が存在し，この段階に従って発話されると捉える。

- (1) 発話行為 (utterance act)：
語（形態素，文）を発話すること
- (2) 命題行為 (propositional act)：
命題を表現すること
- (3) 発話内行為 (illocutionary act)：
陳述，主張，命令，約束，依頼などを行うこと
- (4) 発話媒介行為 (perlocutionary act)：
発話内行為が聞き手の行動・思考・信念などに対して及ぼす結果。例えば，話し手が情報伝達を行うことにより，聞き手を納得させたり，自覚させたり，励ましたりすることに相当する。

上記のように，相手への命令や依頼などの意図は発話内行為として相手に伝達される。また，話し手の意図と発話表現とが直接結びつかず，文脈や言

語外知識に基づいて間接的に言外の意味を伝える行為を間接発話行為と呼ぶ。

対話システムが入力を受け付ける場合，入力された発話表現から話し手の意図（発話内行為）を推測する必要がある。一般に，1つの表現に対応する意図（例えば，依頼）はその場の状況や対話者の社会的地位等の外的要因により変化するため，適切に発話意図を決定する処理が必要になる。この発話意図を決定する手法として，入力文の文末表現と韻律情報とを組み合わせて利用する手法も提案されている^[29]。

3.4 信念

話し手と聞き手が対話をしている状況では，お互いが世界の状況に関して信じている内容があり，これを相互信念と呼ぶ。発話はこの相互信念に影響を及ぼすと考える^[30]。そのため，相互信念を管理する枠組みが必要になる。この相互信念は，発話の理解や発話内容の決定の際に参照され，矛盾があるか否かがチェックされなければならない。

例えば，文献^[41]の例を示すと，対話履歴は，データベース検索に必要なスロットの値が既に確認 (verify) されたか否かの情報とともにシステム内部で管理されている。このような管理状況は“system belief”と呼ばれ，文脈的な観点からの一貫性 (consistency) チェックや次発話の語彙の絞り込み等に利用されている。

3.5 プラン

一般に，人間がどのようにして発話を行うかを考えると次のようになる。まず欲求を持ち，この欲求を満たすために何をすべきかを決定し，この決定内容（意図）を実現するためにどのような行為を行うべきかのプランを立て，このプランに従って発話を行う（プランニング）。対話システムにおいてプランニング処理を実現するためには，システム内部で用意しなければならない知識の多さからごく限られた特定の分野に限定せざるを得ない。

逆に，対話システムが利用者からの入力を理解するためには，利用者のプランを認識する必要があると考えられる。このプラン認識モデルとして種々提案されており^{[31][32][33]}，これらの特徴としてタスクに依存するプランと依存しないプランとに分類している点があげられる。利用者からの入力に対する応答を生成するため，システムでは条件や出力表現を記述し，この記述に従って動作するが，第4.1節に示すように，この記述は通常スクリプト言語を利用して行われる。スクリプト言語を利用する場合でも，タスクに依存する処理と依存しない処理とを明確に分類しておくことは有効であると考えられる。

表 1 実際利用されている対話システムの比較

Domain	Language	Vocabulary Size	Average	
			Words/Utt	Utts/Dialogue
CSELT Train Timetable Info	Italian	760	1.6	6.6
SpeechWorks Air Travel Reservation	English	1000	1.9	10.6
Philips Train Timetable Info	German	1850	2.7	7.0
CMU Movie Information	English	757	3.5	9.2
CMU Air Travel Reservation	English	2851	3.6	12.0
LIMSI Train Timetable Info	French	1800	4.4	14.6
MIT Weather Information	English	1963	5.2	5.6
MIT Air Travel Reservation	English	1100	5.3	14.1
AT&T Operator Assistance	English	4000	7.0	3.0
Air Travel Reservations(human)	English	?	8.0	27.5

3.6 協調的応答

文献^[34]では、人間同士の対話を分析した結果、協調的な対話として下記の3種類の応答が重要だと述べられている²。

(1) Corrective Indirect Response:

入力された内容が事実と反する場合はそれを訂正する発話を行うこと。

(2) Supportive Indirect Response:

質問されたことに直接答えるだけでなく、付加情報を加えることで相手の理解を助ける発話を行うこと。

(3) Suggestive Indirect Response:

質問に対する答えが見当たらない場合などに、相手の意図を汲み取り関連情報を提示すること。

対話システムにおいて、利用者からの入力に対する応答として上記の応答を考慮することも重要である。例えば、データベース検索タスクにおいて、入力条件を満たす検索結果が多い場合は、さらなる絞り込み条件をシステムが利用者に提示する方法がある。これは、上記の(2)に相当すると考えられる。また、上記(3)の例として検索結果が1件も見あたらなかった場合に代案を提案する方法が考えられる^[35]。また、今までにはあまり考慮されていない上記(1)の点も協調的応答を生成するという観点から重要である。

3.7 入力確認

利用者とシステムとの対話では、発話のタイミングが重要となる。対話システムでは相手の発話を

² ただし、直接の応答で誤りを訂正するよりも、話し手の推定を正す間接的応答の方が協調的であるため、“*Indirect Response*”となっている。

理解し、何を発声するかを決定し、実際に発話を行うという処理が必要であるが、この処理に時間がかかったのでは対話にならない。なぜなら、黙っているという行為も1つの発話内行為であるため、本来とは別の発話内行為をシステムが遂行したことに相当するからである^[23]。

従って、システムが即座に回答できない場合は、相づちやうなずきなどにより、相手にちゃんと聞いているのだということを示す必要がある。そのため、相づちやうなずき機能を実装した対話システムの試作例もある^[36]。逆に、システムからの質問に対して利用者が回答を返さなかった場合、システムはこれを検出しなければならない。つまり、利用者からの入力タイミングを計測し入力が無かった場合のタイムアウト処理もシステムでは具備することが重要である^[37]。

さらに、音声認識性能が100%の性能が得られないことを前提とすると、音声認識結果の確認が何らかの形で行われなければならない。しかし、内容の確認のための対話が頻繁におこるようでは、やはり利用者には煩わしいものとなり、受け入れられない可能性がある^[38]。認識結果の確認方法として、認識した結果を利用者に質問して確認するよりも、システムから認識結果を含めた応答表現を出力し、利用者が特に誤りを正さなければ確認された見なす手法(implicit verification)が有効であったと報告されている^[39]。

4. 対話システム

4.1 実装方法

前章に示したように、音声対話処理には種々の考慮しなければならない項目がある。対話モデルの定義として、プラン認識などの理解のモデルを指すことも可能であるが、現在最も多く音声対話システ

ムが利用されているタスクに対しては、利用者からの入力に対して応答を行うモデルを対話モデルと定義できる。多くのシステムは対話フローを記述する枠組みとして各システム毎に開発された script 言語を利用している[40][41]。

4.2 対話システムの比較

現在世界中で音声対話システムの開発が行われている[42][43][44][45][46][47][48]。表1(文献[3]から引用)にこれらのシステムの比較を示す。表中の上から3システムは実用システムであり、次の6システムは実験室レベルのシステムである。また、最後の行は人間同士の対話の場合を示している。また、1発話内の単語数が少ない順に並べられており、この表からは、1発話内の単語数が少ないシステムほど実用化されているのが分かる。このことは、システム主導型の対話システムであったり、音声認識結果の確認方法として陽に質問を行う方法が1発話当りの単語数を少なくしていると考えられる。

4.3 対話システムの評価

対話システムの評価方法として、システム内部の各処理モジュール毎に性能評価を行う方法が考えられる。しかし、対話システム全体としての性能評価も重要である。航空機チケットの予約タスクなどを対象とした対話システムでは、タスク達成率(チケットの予約ができた割合)や1対話当たりの要した時間、ターン数などの測定が可能であり、これらで評価することができる。

さらに、対話システムの性能評価として Glassらは下記の新たな指標で評価することを提案している[49]。

(1) Query Density

1query 当たり に 伝 え ら れ た 平 均 属 性 数 (伝 え ら れ る 情 報 の 割 合)

(2) Concept Efficiency

利用者が自分の意図を伝えるのに要した平均ターン数

5. おわりに

近年の音声認識技術の進歩に伴って人間とシステムとが自由に対話できる世界が現実見をおびてきた。今までデータベース検索タスクを対象に、多くの対話システムが開発されてきたが、実用的なシステムはまだわずかである。本稿では、音声対話システムを構築する際に考慮しなければならない項目について概観した。しかし、まだまだ多くの研究が行われておりサーベイとしては十分ではないかもしれ

ない。さらに、詳しくは他の文献(例えば、[50]など)を参照されたい。

参考文献

- [1] 田中穂積: “音声対話の特質と言語処理技術”, 人工知能学会誌, vol.12, no.1, pp.18-23, 1997.
- [2] 中川聖一: “音声認識研究の動向”, 信学論, vol.J83-DII, no.2, pp.433-457, 2000.
- [3] V.W.Zue, J.R.Glass: “*Conversational Interface: Advances and Challenges*”, Proc. of the IEEE, *Special Issue: Spoken Language Processing*, pp.1166-1180, August 2000.
- [4] 北研二, 川端豪, 斎藤博昭: “HMM音韻認識と拡張LR構文解析法を用いた連続音声認識”, 情処学論, vol.31, no.3, pp.472-480, 1990.
- [5] N.Inoue, I.Nogaito, M.Takahashi: “A Linguistic Procedure for an Extension Number Guidance System”, IEICE Trans. on Inf. & Syst., vol.E76-D, no.1, pp.106-111, 1993.
- [6] Fernando C.N. Pereira and Michael D. Riley: “*Speech Recognition by Composition of Weighted Finite Automata*”, AT&T Technical Report, TR 97.18.1, 1996.
- [7] T.R.Niesler and P.C.Woodland: “*A Variable-Length Category-Based N-Gram Language Model*”, Proc. of ICASSP-96, pp.164-167, 1996.
- [8] I. Guyon and F. Pereira: “*Design of a Linguistic Post-processor Using Variable Memory Length Markov Models*”, International Conference on Document Analysis and Recognition, pp.454-457, 1995.
- [9] S.Deligne and F.Bimbot: “*Langage Modeling by Variable Length Sequences: Theoretical Formulation and Evaluation of Multigrams*”, Proc. of ICASSP-95, pp.169-172, 1995.
- [10] 江原暉将, 小倉健太郎, 篠崎直子, 森元逞, 樽松明: “電話またはキーボードを介した対話に基づく対話データベースADDの構築”, 情処学論, vol.33, no.4, pp.448-456, 1992.
- [11] 財団法人機械システム振興協会: “音声の知的処理に関する調査研究”, 1991.
- [12] V.Zue, S.Seneff, J.Glass, J.Polifroni, C.Pao, T.Hazen, and L.Hetherington: *Jupiter: A Telephone-Based Conversational Interface for Weather Information*, IEEE Trans. on Speech Audio Processing, vol.8, no.1, pp.85-96, 2000.
- [13] M. Eskenazi, A. Rudnickey, K.Gregory, P.Constantinides, R.Brennan, C.Bennett, and J.Allen: “*Data Collection and Processing in the Carnegie Mellon Communicator*”, Proc. of Eurospeech99, pp.2695-2698, 1999.
- [14] E. den Os, L.Boves, L.Lamel and P.Baggia: “*Overview of the ARISE Project*”, Proc. of Eurospeech99, pp.1527-1530, 1999.
- [15] 山岡孝行, 飯田仁: “文脈を考慮した音声認識結果絞り込み手法”, 情報処理学会, 自然言語処理研究会, 90-NL-78-16, 1990.
- [16] M.Nagata and T.Morimoto: “*An Information-Theoretic Model of Discourse for Next Utterance Type Prediction*”, 情処学論, vol.35, no.6, pp.1050-1061, 1994.
- [17] 白井克彦, 竹沢寿幸: “音声対話処理”, 人工知能学会誌, vol.9, no.1, pp.48-56, 1994.

- [18] K.Kita, T.Ehara and T.Morimoto: "Processing Unknown Words in Continuous Speech Recognition", IE-ICE Trans., vol.E74, no.7, pp.1811-1816, 1991.
- [19] D.Klakow, G.Rose, X.Aubert: "OOV-Detection in Large Vocabulary System Using Automatically Defined Word-Fragments as Fillers", Proc. of Eurospeech99, pp.49-52, 1999.
- [20] K.Tanigaki, H.Yamamoto, and Y.Sagisaka: "A Hierarchical Language Model Incorporating Class-Dependent Word Models for OOV Words Recognition". Proc. of ICSLP2000, vol.3, pp.123-126, 2000.
- [21] G.Wilpon, R.Rabiner, C.Lee and R.Goldman: "Automatic Recognition of Keywords in Unconstrained Speech Using Hidden Markov Models", IEEE Trans. ASSP38, no.11, pp.1870-1878, 1990.
- [22] R.Rosenfeld: "A Maximum Entropy Approach to Adaptive Statistical Language Modeling", Computer Speech and Language, vol.10, pp.187-228, 1996.
- [23] 角田達彦: "対話のモデル", 自然言語処理 (長尾真編), 岩波書店, 1996.
- [24] H.P.Grice: "Logic and conversation", in Syntax and Semantics III, (Ed. P.Cole and J.L.Morgan), Academic Press, pp.41-58, 1975.
- [25] 住田一男, 浮田輝彦: "質問応答システムにおける会話の自然性に関する考察", 信学論, vol.J70-D, no.11, pp.2287-2293, 1987.
- [26] H.nakaiwa and S.Ikehara: "日英翻訳システムにおける用言意味属性を用いたゼロ代名詞照応解析", 情処学論, vol.34, no.8, pp.1705-1715, 1993.
- [27] M.Walker, M.Iida, and S.Cote: "Centering in Japanese Discourse", Proc. of COLING90, 1990.
- [28] J.R.Searle: "Speech Acts: An Essay in the Philosophy of Language", Cambridge University Press, 1969. (言語行為"坂本百大, 土屋俊訳, 1986).
- [29] A.Kurematsu, Y.Shinoya: "Identification of Utterance Intention in Japanese Spontaneous Spoken Dialogue by Use of Prosody and Keyword Information". Proc. of ICSLP2000, vol.1, pp.98-101, 2000.
- [30] 山梨正明: "対話理解の基本的側面", 「対話行動の認知科学的研究」研究会論文集, 1984.
- [31] J.F.Allen: "Recognizing intentions from natural language utterances", in Computational Models of Discourse, (Ed. M.Brady and R.C.Berwick), The MIT Press, pp.107-166, 1983.
- [32] 飯田仁, 有田英一: "4階層プラン認識モデルを使った対話の理解", 情処学論, vol.31, no.6, pp.810-821, 1990.
- [33] D.J.Litman and J.F.Allen: "A Plan Recognition Model for Subdialogues in Conversations", Cognitive Science, vol.11, pp.163-200, 1987.
- [34] S.J.Kaplan: "Cooperative Responses from a Portable Natural Language Database Query System", in Computational Models of Discourse, (Ed. M.Brady and R.C.Berwick), The MIT Press, pp.167-208, 1983.
- [35] 伊藤敏彦, 中川聖一: "音声対話システムにおける協調的応答", 情処研報, SLP10-19, 1996-3.
- [36] 平沢純一, 川端豪: わかってうなずくコンピュータの試作, 信学技報, NLC97-54, 1997-12.
- [37] 伊藤克亘, 速水悟, 田中和世: 時間の扱いを考慮した対話システム制御手法, 信学技報, SP95-117, 1996-1.
- [38] 中川聖一: "音声対話システム", 情報処理学会, 第52回全国大会, シンポジウム論文集, 1996.
- [39] L.Lamel, S.Rosset, J.L.Gauvain, S.Bennacef: "The LIMSI ARISE System for Train Travel Information", Proc. of ICASSP99, 2240, 1999.
- [40] S.Seneff, R.Lau, and J.Polifroni: "Organization Communication, and Control in the GALAXY-II Conversational System", Proc. of Eurospeech99, pp.1271-1274, 1999.
- [41] V.Souvignier, A.Kellner, B.Rueber, H.Schramm, and F.Seide: "The Thoughtful Elephant: Strategies for Spoken Dialogue Systems", IEEE Trans. Speech Audio Processing, vol.8, pp.51-62, 2000.
- [42] L.Lamel, S.Rosset, J.L.Gauvain, S.Bennacef, M.Garnier-Rizet, and B.Prouts: "The LIMSI ARISE system", Proc. of IVTTA, pp.209-214, 1998.
- [43] R.Billi, R.Canavesio, and C.Rullent: "Automation of Telecom Italia Directory Assistance Service: Field Trial Results", Proc. of IVTTA, pp.11-16, 1998.
- [44] D.Jurafsky, C.Wooters, G.Tajchman, J.Segal, A.Stolcke, E.Fosler, and N.Morgan: "The Berkeley Restaurant Project", Proc. of ICSLP-94, pp.2139-2142, 1994.
- [45] S.M.Marcus, D.W.Brown, R.G.Goldberg, M.S.Schoeffler, W.R.Wetzel, and R.R.Rosinski: "Prompt Constrained Natural language - Evolving the Next Generation of Telephony Services", Proc. of ICSLP96, pp.857-860, 1996.
- [46] E.Levin, S.Narayanan, R.Pieraccini, K.Biatov, E.Bocchieri, G.Di Fabbriozio, W.Eckert, S.Lee, A.Pokrovsky, M.Rahim, P.Ruscitti, and M.Walker: "The AT&T-DARPA Communicator Mixed-Initiative Spoken Dialog System", Proc. of ICSLP2000, vol.2, pp.122-125, 2000.
- [47] A.I.Rudnick, C.Bennett, A.W.Black, A.Chotomongcol, K.Lenzo, A.Oh, and R.Singh: "Task and Domain Specific Modelling in the Carnegie Mellon Communicator System", Proc. of ICSLP2000, vol.2, pp.130-133, 2000.
- [48] V.Zue, S.Seneff, J.Polifroni, M.Phillips, C.Pao, D.Goddeau, J.Glass, and E.Brill: "PEGASUS: A Spoken Language Interface for On-line Air Travel Planning", Speech Communication, vol.15, pp.331-340, 1994.
- [49] J.Glass, J.Polifroni, S.Seneff, V.Zue: "Data Collection and Performance Evaluation of Spoken Dialogue Systems: The MIT Experience", Proc. of ICSLP2000, vol.4, pp.1-4, 2000.
- [50] (Ed.)B.H.Juang and S.Furui: "Special Issue on Spoken Language Processing", Proc. of the IEEE, August 2000.