

オンライン言語処理モデルにおける自己修復文の解析手法

伊藤 雅浩 佐川 雄二 大西 昇

名古屋大学大学院 工学研究科

〒464-01 名古屋市千種区不老町

入力単語をその出現順序にしたがって逐次処理を行なうオンライン言語処理モデルにおける、自己修復（言い直し）文の解析手法を提案する。対話をオンラインに処理することは、処理効率や認知的見地から見ても有効であるが、自己修復のように話者自身が直前の発話を翻す発話を行なっている場合に問題が生ずる。本手法では名詞句、格、動詞句を扱う3つのバッファを用意し、それら3つの相互作用によって解釈を進める。そして、それぞれのバッファ独自の方法で自己修復をオンラインに検出し訂正する。試行的にインプリメントした結果、コーパス中の自己修復文の約80%について自己修復を正しく検出、訂正した。

キーワード：音声対話処理, オンライン言語処理, 漸進的解釈, 自己修復

An On-Line Processing Model Coping with Self-Repaired Utterance

Masahiro Ito, Yuji Sagawa, Noboru Ohnishi

School of Engineering, Nagoya University

Furo-cho, Chikusa-ku, Nagoya, 460-01, Japan

We propose a method to cope with self-repaired utterance on an on-line language interpretation model, which processes input words incrementally. From both computational and cognitive points of view, incremental dialogue processing seems to have a lot of benefits. However, spontaneous speech includes so many disfluencies, such as self-repair, that it needs to be interpreted again, when the speaker changes his/her intent or corrects his/her previous utterance. The method interprets input sentence by using three buffers; Noun Phrase buffer, Case buffer and Verb Phrase buffer. Then we can find self-repaired utterance at the moment it occurs. As a result of implementation, we could detect and correct about 80% of the self-repairs in human dialogue.

keywords: spoken language processing, on-line language processing, incremental interpretation, self-repair

1 はじめに

人間と計算機との自然な対話を実現するためには、音声レベル、言語レベル、知識・概念レベルの処理を密接に関連させて、一体的に行なう必要がある[1]。その場合、入力とは従来のようなキーボードによるものでなく、音声によるものが望まれる。しかし、話し言葉は書き言葉に比べて自由度が多く、時間的制約がある。したがって話し言葉には、省略、倒置、言い直しなど、書き言葉を対象にしたシステムでは扱いにくい現象が多い。そのため、音声入力による自然言語処理システムには、音声対話文の文法、音声対話のモデルに基づいた処理が求められる。

音声対話のモデルに基づいた処理の一つにオンライン処理が提案されている。この処理は文を最後まで読み込むことなく、漸進的に解釈を行なう手法である。このアプローチは計算論的にも効率的で、心理学的にも人間の解釈過程を模倣するモデルとして重要視されている[2]。しかし、上述したように、音声入力には言い直し（自己修復）が多く含まれている。これは話者自身が生じた誤りを自ら訂正する現象で、オンラインで処理するためには、自己修復であることを聞き手のモデルにおいて確実に検出し、そこまでに構築されていた解釈を正しく訂正することが要求される。

従来、漸進的処理による対話理解の研究[2,3]や自己修復文を解析する研究[4-7]が別々に行なわれてきているが、漸進的に対話を解釈する際においても自己修復文の問題は不可避である。そこで、本研究では自然な発話に含まれている不適格な文の分析に基づいてオンライン言語解析モデルで自己修復を含む文の解析手法を提案する。

2 オンライン言語処理モデル

オンライン言語処理モデルとは、入力文（もしくは音声）を文の終りを待つことなく、語単位や1文字単位でその出現順序にしたがって順次処理を行なうモデルのことである。

2.1 オンラインモデルの特長

計算機において言語を漸進的に解釈することにはさまざまな特長がある。例えば、人間は入力音声を逐次理解してゆくことによって長い文の解釈を可能にしている。このように人間の処理過程を模倣するシステムを作成する上では、オンラインによる処理が有効である。

さらに、大きなまとまりを一度に処理するよりも処理が簡略化され、速度向上にも貢献することになる。一般に話し言葉では文の境界があいまいで、文の長さも長くなりがちである。しかし、オンラインに解釈することにより、文の長さに依存しない処理

が可能になる。

また、われわれは普段意識することなく、会話を円滑に続けていくためにさまざまな言語的・非言語的方略を使用している[8]。その一つに「あいづち」が挙げられる。これは、聞き手が発話の意味のまとまりをとらえた時点でそのことを話者に伝え、相手の発話を促す効果があると考えられる。さらに必要に応じて話し手の発話を中断して質問したり、別の意見を述べたりして発言権をフレキシブルに交替している。オンラインに対話を処理すると、こうした現象にも応用することが可能である。

2.2 オンラインモデルの問題点

しかし、問題点としては全体を見渡した処理を行なわないためにシステムのロバスト性を損なう可能性がある。

(1) 日本の人口が100万人を越えた年

(2) 日本の人口が100万人を越えた都市

例えば、上の例文のように(1)では「日本の」は「人口」を修飾するのに対して、(2)では「都市」を修飾している。これを正しく理解するためには「年」「都市」という2つの名詞のもつ意味やさまざまな性質を比較することによって可能になる。

しかしこの文をオンラインに処理させるとどちらも「日本の」の次に「人口」が読み込まれた時点で「日本の」は「人口」を修飾するという解釈に陥りやすい。したがって、最終的に「都市」が入力された時点でそれまでに生成されている解釈を再構成する必要がある。一般に、これらの2つの発話を実際に聞いてみるとポーズ、イントネーションといったさまざまな韻律情報が存在しているために、こうした情報の統合によって構文的あいまい性の解消に役立つものと思われる。

この(1),(2)のようなあいまい性を持つ例文は非オンラインに解釈する場合でも膨大な知識や処理が要求される困難な処理であることが予想される。さらに、音声対話においてはこのような構文的、その他のあいまいさの問題に加えて、話者自身の言い直しによる解釈の再構成が必要になる。こうした自己修復という現象は発話の度合いが自由発話に近くなるほど増加しており、村上[9]によると、対話文の2%から4%に含まれているという。また、本研究で用いた対話コーパス文の中でも約30秒に一度の割合で出現しており、自己修復文の適切な解析は必要不可欠である。

以下に例を示す。

(3) あの、朝、次の日の朝、見送ったんです。

(4) とってもいい、けんけー、経験をね、しているわけですよ。

(5) そちらの事務局の方でみて、面倒をみていただける

例文(3)では、「朝」と言った後に「次の日の」を挿入して精緻化を行なっている。例文(4)では、「経験」と発話すべきところを誤って「けんけー」と言ってしまったために話者自身が言い直している。また、例文(5)では一度目の「みて」という動詞が発話された後に「面倒を」を付け加えたために「みて」が二度発話されて冗長になっている。

こうした自己修復は話者自身の不適格性を補おうとして発話されたものであるが、この自己修復文そのものが未知語や冗長な発話といった新たな不適格性を生み出してしまふ。そのため、音声対話を理解する際には自己修復を含む文を確実に理解するメカニズムが求められている。

3 解析手法

そこで、適切な自己修復の検出と訂正をオンライン処理によって実現するために次のようなモデルを提案する。

3.1 解析の概要

日本語に限らず、人間は音声言語の解析を行なう際には動詞を中心とした述語が鍵になる。しかし、日本語の場合は通常、動詞が文末に来るので動詞が入力されるまでに読み込まれた句は短期記憶に保存しておく必要がある。そして述語が読み込まれた時にそれまで貯蔵してあった句を取り出して述語に対して適切な役割を与えてやることになる。よって、動詞が入力されるまでの一時的な記憶バッファを用意する必要がある。

そこで、解析の状態を3つのバッファを用いて表す。それぞれ、名詞句バッファ、格バッファ、動詞句バッファと呼び、データ構造はスタックになっている。それぞれのバッファは形態素1つ1つが入力される度に相互のバッファの内容を参照する。そして解釈がより大きなまとまり(チャンク)を構成することに名詞句バッファから格バッファへ、さらには動詞句バッファへと更新され、そのバッファが解析の結果を表すことにする。

まず、入力形態素を形態素解析の結果に応じてそれぞれどのバッファに送り込むかを決定する。例えば、名詞、代名詞などは名詞句バッファに送り込んで名詞句を構成する。「が」「は」「で」「を」などの助詞は格バッファに送られて後に入力される動詞の格を用意する。動詞、助動詞などは動詞句バッファに送られ、その時点で用意されている格バッファや名詞句バッファの内容から文章全体の意味を決定する。

例文(6)の解釈過程を図1に表す。

(6) 私は結論を言います。

input	NPbuffer	CASEbuffer	VPbuffer
私	私		
は		((HA)私)	
結論	結論	((HA)私)	
を		((HA)私)	
言い		((WO)結論)	[SAY((HA)私) ((WO)結論)]
ます			[SAY,declare((HA)私) ((WO)結論)]

図1: 例文(6)の解釈過程

3.2 自己修復文の再構成

自己修復文には中断前後において繰り返しの表現が多く現れている[4]。ただし、ここで言う繰り返しには表層的な同一語句の繰り返しから同一意味カテゴリのものでの繰り返しもある。さらに、

(7) 昨日は昨日で普通に行ったんです。

のように、同一語句が近い位置で発話されていてもそれが必ずしも自己修復であるとは限らない。そこで、上記の3つのバッファの中で会話文中に出現するさまざまな繰り返し表現の中から自己修復のみをオンラインに検出する方法として以下の手法を提案する。

規則1: 名詞句バッファ 自己修復を含む文では同一語句の繰り返しが頻繁に生じている。この冗長部分を削除することによって名詞句バッファ内における自己修復の検知、訂正が可能である。

(3) あの、朝、次の日の朝、見送ったんです。

それにより、上の例文は2つ目の「朝」が入力された時点で、1つ目の「朝」が削除されて名詞句バッファの内容が「次の日の朝」に更新される。

規則2: 格バッファ 格バッファの中で検出できる自己修復は同一役割の格の出現である。1つの文中に同じ役割を持つ格が2つ以上存在することは冗長なので重複した場合は前者を削除する。

(8) 部活を、バレーをやりたいのか

それによって上の例文は「バレーを」の「を」が入力された時点で、格バッファに2つの同一役割の格が存在することになり、既出の「部活を」が削除される。

規則3: 動詞句バッファ 同様に、動詞句バッファにおいても同じ動詞句が用いられた場合には、前者を冗長部分として削除することが考えられる。しかし、動詞句全体を削除してしまうと必要な部分まで削除してしまうおそれがある。また、同じ動詞を一

つの発話の中で意図的に何度も用いる場合もある。よって、冗長であると判断できるのは連続して同じ動詞が用いられている場合に限ることにし、その間に接続詞等の関係を表す品詞が挿入されている場合には冗長と見なさない。

さらに冗長であると判断されて前者を削除する場合でも、動詞句バッファ内のそれぞれの動詞句が持っている格すべてを削除すると、必要な情報が削除されるおそれがある。なぜなら、話者が言い直しを行なう場合には、必要な部分のみを言い直すことが多いので、動詞そのものを言い直した場合に誤りが発生する以前に発話した修飾語句も繰り返される可能性が低いからである。

そこで、動詞句バッファ内で近接する動詞（用言）が同一の場合は、それぞれの動詞が所有する格の結びを採用し、一つの動詞句に統合する。

(5) そちらの事務局の方でみて、面倒をみていただける

例文(5)では「みて」という動詞が2度使われている。しかし2つ目の「みて」が入力された時点で、1つ目の動詞句「事務局の方でみて」を削除してしまうと情報が欠落してしまうことになる。そこで、この規則によってそれぞれの動詞が所有する格の結びを採用することによって「事務局の方で面倒をみて」という動詞句が形成される。

規則4：未知語について 本研究においては暫定的に未知語は存在しないものとして、システムの持つ辞書に登録されていない語は、単語の断片や話者の言い損じであるとしてバッファの内容は更新されない。

規則5：活用語尾、助動詞のみの修復 これまで述べてきた方法で対応することができない現象が次に示すような述語の活用語尾、もしくは助動詞の言い直しである。

(9) 全部返していただき、けるのか

(10) 自分の存在感を認められ、られていない

このように話者は意図をなるべく効率的に伝えようとすることから、発話にいきづまった時に述語そのものの言い直しは行わず、必要最低限の言い直しにとどめる場合がある。述語（語幹）の言い直しが行なわれない場合、前述の動詞句に関する手法を用いることはできない。そこで活用語尾、助動詞の言い直しに対応する手法を考案する必要がある。

一般に日本語の動詞などでは、いわゆる助動詞がいくつか付いて複雑な語形を生み出している。動詞、助動詞はそれぞれ後ろに結び付く助動詞の種類に関わらず、常に同じ形を持つ「語幹」と後ろに結

び付く助動詞の種類に応じて変化する「活用語尾」から構成されている。

例文(9)では、語幹「いただく」の後ろに連用形の語尾「い」が発話された後、再び語尾「け」が発話されている。こうした場合には後から発話された語尾でオーバーライトする必要がある。

また、日本語動詞では言語化しようとする概念と実際に現れる表現との間で必ずしも一致するとは限らない。例えば、概念として「書く」、「丁寧」、「否定」、「並列」であっても

「書きませんでしたり」

とはならない。ここでは「丁寧」を削除しないと適切な日本語表現にはならない。助動詞の動詞や形容詞に対する結び付き方は図2のように線上の一定の位置に、それが現れているかどうかである[10]。

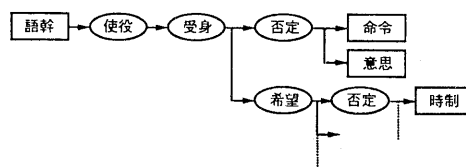


図 2: 助動詞の出現順序（抜粋）

この図は動詞の語幹から始めて右にたどっていく、楕円で囲まれた形態素はオプションで、拾っても拾わなくても良く、長方形で囲まれた形態素は必須なので必ず拾う。そしていずれかのパスをだどってゆけば可能な形態素の組合せが得られる。

この図を用いて、動詞の後に助動詞が入力される度にパスをたどっていく。そこで次の入力に到達不可能な助動詞であった場合には、そのパスを順次逆戻りして新しいパスを発見することによって再構成を行なうことができる。この手法は助動詞の言い換えだけでなく、単純な繰り返しにも対応可能である。

ここで今までに述べてきた再構成規則をまとめると、以下に示す6つである。

i) 名詞句の繰り返しに関する規則

名詞句の繰り返しがあった場合は前者を削除する。

ii) 同一役割の助詞に関する規則

同一役割の格が生じた場合は前者を削除する。

iii) 連続して同じ述語が使われている場合の規則

前者を削除するが、格は保存される。

iv) 未知語に関する規則

未知語は解釈に影響を与えない。

v) 活用語尾に関する規則

語尾の繰り返しは後者を前者で置き換える。

vi) 助動詞の接続に関する規則

不適切な助動詞の入力に対しては接続可能なパスを発見して再構成する。

3.3 具体例

ここでは前述の例 (5) を用いて解釈過程のシミュレーションを行なう。図 3 を参照。

i) 名詞「事務局」が入力される。

名詞句バッファに送られる。

ii) 助詞「で」が入力される。

名詞句バッファの内容を取り出して、「で」という格を格バッファに作る。

iii) 動詞「み」が入力される。

格バッファの内容が取り出されて、「見る」という動詞句を動詞句バッファに作る。

iv) 助動詞「て」が入力される。

動詞句バッファに送られて、意味を付加する。

v) 名詞「面倒」が入力される。

名詞句バッファに送られる。

vi) 助詞「を」が入力される。

名詞句バッファの内容を取り出して、「を」という格を格バッファに作る。

vii) 動詞「み」が入力される。

格バッファの内容が取り出されて、「見る」という動詞句を動詞句バッファに作る。しかし、ここで「見る」という動詞が連続的に用いられていることから先ほど述べた動詞句に関する規則が用いられ、1つの動詞句に統合される。

viii) 助動詞「て」が入力される。

動詞句バッファに送られて、意味を付加する。

input	NPbuffer	CASEbuffer	VPbuffer
事務局	事務局		
で		((DE)事務局)	
見て			[look((DE)事務局)]
面倒	面倒		[look((DE)事務局)]
を		((WO)面倒)	[look((DE)事務局)]
見て			[look((DE)事務局)] [look((WO)面倒)]
			[look((DE)事務局) ((WO)面倒)]

図 3: 例文 (5) の解釈過程

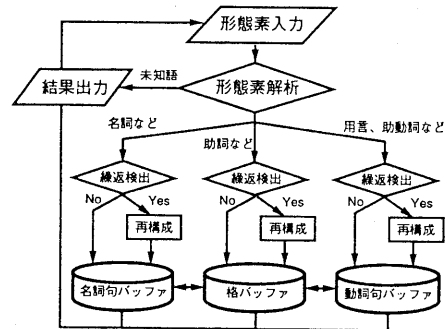


図 4: システム構成

対話コーパス 本研究で実際に使用した対話コーパスは NHK のラジオ番組から収録したものである。対話数は3つで、合計時間は44分間、内容は教育に関する電話相談である。

結果 コーパス中に含まれていた92例の自己修復文について解析行なった。結果は図5に示す通りである。

図に示す通り、合計74(80.4%)例の自己修復個所の訂正に成功した。

考察 検出および訂正に失敗したものには次のようなものが含まれる。

4 試作システム

この節では前節に述べた解析手法を試行的にインプリメントした結果について述べる。

概要 システムの概要は図4に示す通りである。形態素が一つ一つ入力される度に処理を行ない、その時の3つのバッファの状態を出力として表示する。

入力は形態素ごとに区切られたテキストをキーボードを通して行ない、その都度、3つのバッファの状態を出力している。

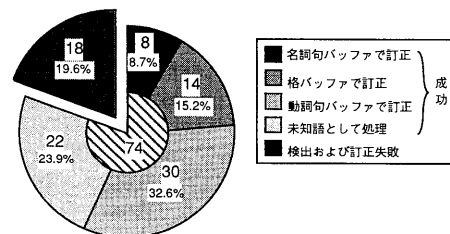


図 5: 解析結果

- 自己修復か並列なのかテキストからは判断できないもの（主に名詞が連続する場合）

(11) えー、木、あー、水曜日は出て

この場合は隣接する名詞句間のポーズといった韻律情報や、冗長語¹の有無が手がかりになると思われる。

また、会話文にはしばしば意図的に助詞を欠落させる現象がある[11]。この場合には、述語が読み込まれた場合に適切な述語に対する格を与えてやる必要があり、格バッファでの処理に再考を要する。さらに助詞の欠落によって名詞の連続が生ずることもあるので（例文(12)における「私」と「昨日」）、名詞が連続している場合にそれぞれの名詞の関係を考える必要がある。

(12) 私、昨日は行ったんです。

- 意図そのものを言い直しているもの

(13) そうです、あ、違います。

この場合には発話意図の矛盾を探す必要がある。このように、表層的な情報（例えば繰り返しや冗長語など）だけで発見できない現象については、深い意味処理などを行わなくてはならない。

その他の助詞の問題として複数の役割を持つ助詞の問題がある。例えば助詞「は」のようにさまざまな役割を持つ助詞については格フレームに適切な割り当ててやるなどの手法をとって、同じ助詞が用いられても格バッファで削除されるのを避ける必要がある。

(14) 私は太郎は死んだと思った。

また、例文(14)のように主語を表す格は連続して存在しても冗長であるとは限らない。会話文ではこのような複雑な構造を持つ文は稀であるが、主格については従来の文脈自由文法のようにスタックを用いて処理を用いて、最後に不要な格が残った場合に削除するなどの手法をとる必要がある。

5 おわりに

本研究では自己修復を含む文をオンラインに検出、訂正する手法を提案した。名詞句、格、動詞句という3つのレベルにおいて自己修復が発生した時点で検出を行なっている。結果は6つの規則を用いることによって、約80%の自己修復について検出、訂正を可能にしているが、表層的な文字情報だけでは判断できないものも存在している。

今後の課題としては、

¹ 発話権維持のために用いられる「えーと」「あの」といった特別な意味を持たない単語

- 韻律情報（ポーズ、ピッチなど）の有効利用。
Nakatani[6]らは、韻律情報を利用することによって90%程度の自己修復個所の検出に成功している。
- 他の解析手法（助詞落ちなど）との融合。
- 単文処理にとどまらず、文脈の利用。

などが考えられる。

謝辞 本研究を進めるにあたり、有益なコメントをいただいた名古屋大学大学院工学研究科の山村毅講師、および名古屋大学大型計算機センターの田中敏光助教授に感謝致します。

参考文献

- [1] 中川 聖一, 堂下 修司: 音声言語情報処理研究の動向と研究課題, 情報処理, Vol.36, No.11, pp. 1012-1019 (1995).
- [2] Jurafsky. D.: An On-Line Computational Model of Human Sentence Interpretation, *Proceedings of AAAI-92*, pp.302-308 (1992)
- [3] 松原 茂樹, 外山 勝彦, 稲垣 康善: 動的意味論を用いた漸進的談話解釈モデル, 言語処理学会第1回年次大会講演論文集, pp. 165-168, (1995)
- [4] 佐川 雄二, 大西 昇, 杉江 昇: 自己修復を含む日本語不適格文の分析とその計算機による理解手法に関する考察, 情報処理学会論文誌, Vol.35, No.1, pp.46-52 (1994)
- [5] Heeman. P.: Spoken Dialogue Understanding and Local Context, *Technical Report 523, University of Rochester* (1994)
- [6] Nakatani. C. and Hirschberg. J.: A Speech-First Model for Repair Detection and Correction, *Proceedings of ACL-93*, pp.46-53 (1993)
- [7] Shriberg. E., Bear. J. and Dowding. J.: Automatic Detection and Correction of Repairs in Human-Computer Dialog, *Proceeding of DARPA Workshop on Speech and Natural Language*, pp. 419-424 (1992)
- [8] 池田 裕, 池田 智子: 日本人の対話構造, 言語, Vol.25, No.1, pp.48-55 (1996)
- [9] 村上 仁一, 嵯峨山茂樹: 自由発話音声における音響的な特徴の検討, 電子情報通信学会論文誌, Vol.J78-D-II, No.12, pp.1741-1749 (1995)
- [10] 草薙 裕: LISPによる自然言語処理, 第5章 日本語動詞の解析, pp.113-134, 工学図書 (1988)
- [11] 丸山 直子: 助詞の脱落現象, 言語, Vol.25, No.1, pp.74-80 (1996)