

コストと報酬の関係に基づく自然な発話の解析

今一 修 松本 裕治

奈良先端科学技術大学院大学 情報科学研究科

1 はじめに

自然な発話には、省略、言い直し、倒置、間投詞など、さまざまな種類の不適格性がある。特に、間投詞、言い直しなどは話し言語に特有の現象であり、書き言葉にはあまり現れないものである。ATR 対話コーパス [10] を分析した結果によると、音声対話文において、言い直しは約 25%、間投詞は約 64% の文に含まれている。一方、キーボード対話文においては、それらはほとんど含まれていない。したがって、音声対話システムを構築する際には、これらに現象にも対処する必要性が生じてくる。

以前、我々は主に書き言葉における不適格性である制約違反や語順の違反などに対して、コストと報酬の関係に基づいた統合的な処理モデルを提案した [4]。ここでは、以下の 3 つのタイプの言語現象を扱っており、それぞれの現象に対して 1 つのモジュールが処理を担当している。

1. 制約違反
統語制約違反 (助詞欠落)、意味制約違反 (比喩・換喩・擬人化などで生じる選択制約違反)
2. 構造違反
倒置などの語順の誤り
3. 断片的な句
ゼロ代名詞や省略 (それ単独では解釈できず文脈情報を必要とするもの)

本稿では、このモデルを言い直しや間投詞を扱えるように拡張する手法について述べる。このモデルはモジュール性の高い枠組みであるので、上記の 3 つの現象を扱うためのモジュールを修正することなく、新たにモジュールを追加することにより拡張を行なうことができる。したがって、統合的な処理モデルを制御している中心的な原理である報酬最大・コスト最小の原理は変更する必要がない。新たに追加するモジュールも含めて、すべてのモジュールがこの原理に支配されて動作することになり、システム全体をみた場合、モジュール性が高く見通しのよいシステムとなっている。

2 関連研究

Hindle [3] は、自然な発話における言い直しを訂正する問題を取りあげ、Marcus の決定パーザにいくつかの規則を追加することにより、それらを扱う手法を提案した。Hindle の手法では、言い直しの位置を示す編集表現の存在を仮定しており、それが事前に音韻的な情報から検出可能であるとしているが、一般にこのような手掛りを検出する手法は確立されていない。

Nakatani ら [6] は、音韻情報を基に言い直しの位置を検出するモデルを提案している。ここでは、言い直しを検出する際の手掛りとして、単語間のポーズ、単語の断片の存在、同一単語の繰り返しなどが上げられており、Hindle の仮定をある程度支持している。

言い直し位置を示す編集表現の存在を仮定している Hindle に対して、このような編集表現の存在を仮定せずに、言い直しの検出と訂正を行なっている研究もある。

Bear ら [1] は、言い直しの検出と訂正を行なうためにパターンマッチング機構を用いている。この手法は 2 段階で構成されており、まず最初に語彙的パターンマッチング規則を用いて言い直しの候補を選択し、次に統語的、意味的、音韻的な情報を利用することにより、それらの候補の中から適切なものを選択する。この手法は、複数の言語情報を統合的に利用することによって、言い直しの処理をより頑健に行なうことができることを示している。

佐川ら [9] は、言い直しが原因で入力文の解析に失敗した場合、その入力文を言い直しを含まない文に変換することにより処理を行なっている。この手法では、言い直しの位置を適切に検出するために、語の反復、未知語、孤立語などの言語的な知識を利用している。

Heeman ら [2] は、言い直しのパターンを用いることによって、言い直しの検出と訂正を行なうアルゴリズムを提案している。言い直しパターンは訓練用のコーパスから抽出され、このパターンを用いて言い直しの検出と訂正が行なわれる。また、言い直しパターンによって検出された言い直しの候補の中から、誤った候補を除去するために、品詞のタグ付けに用いられるマルコフ

モデルを拡張した統計モデルを利用している。彼らは、現状の構文解析、意味解析が自然な発話の処理には不十分であるという理由から、統語的、意味的な情報を利用していない。

3 自然な発話

言い直し

言い直しとは、文の途中で発話を一旦中断して語句を訂正することである。例えば、例1では、「会議のこと」が「登録のこと」で言い直されている。

例1 会議のこと、あの一、登録のことで伺いたいのですが。

言い直しは、修復対象、冗長部分、訂正部分の3つの部分に分けることができる。修復対象は、言い直しを解釈する際に削除される部分である。修復対象の終了位置は、発話が一旦中断される地点と一致し、中断点と呼ばれる。冗長部分は、中断点から訂正部分の開始位置までの部分であり、ここには、ポーズや間投詞などが含まれる。訂正部分は、修復対象を置き換える部分である。例えば、例1では、修復対象が「会議のこと」であり、冗長部分が「あの一」であり、訂正部分が「登録のこと」となっている。

言い直しを処理する際には、修復対象と冗長部分を削除することが必要になるが、そのためには、中断点を適切に検出することが要求される。中断点の前後では、同じ語句の繰り返しや、類似した語句での繰り返しがあることが多く、前節で述べた従来の研究でも、これらの情報を利用しているものが多い。例えば、文献[1]では、1語の言い直しの中で、41%が同じ語句の繰り返しであり、23%が類似した語句の繰り返しであることが報告されている。

間投詞

間投詞は、「あの一」、「まあ」、「えーと」などの通常は意味を持たない語のことで、冗長語、つなぎ語などと呼ばれている。間投詞は、言い直しや言い淀みなどと同時に使われることがあり、言い直しの位置を検出する手掛りとなることが多い。文献[9]によれば、出現頻度の上位10種類で間投詞全体の約83%を占めていることが報告されているので、これらをあらかじめ辞書に登録しておくことにより間投詞を扱うことができる。

4 不適格文処理のための統合的な処理モデル

我々が以前提案した、不適格文処理のための統合的な処理モデル[4]は、制約緩和法[8]と部分解析法[5]を統合したモデルとして捉えることができる。

制約緩和法では、何らかの不適格性が原因で入力文の解析に失敗した場合、違反した制約条件を緩和することによって、入力文の受理を試みようとするが、緩和すべき制約の候補が複数存在する場合、どのようにして失敗の回復に最適な候補を選択するかが問題となってくる。効率良く不適格文を処理するためには、違反している制約の中から失敗の回復に有効なものを選択し、それを緩和することが必要となる。このモデルでは、緩和すべき制約を適切に選択するために、個々の制約違反に対して以下の2つの指標を考える。

コスト … その制約違反がどの程度のものか。
報酬 … その制約違反を緩和したときに、どのような構造が生成されるか。

コストは、その制約を緩和するときの不適格性の許容度を示しており、その値は小さい方がよい。報酬は、その制約違反を緩和することが、どれくらい失敗の回復に役立っているかを示す尺度であり、その値は大きい方がよい¹。

通常の緩和法では、制約違反によって解析が失敗すると、新しい句はつくられず、その制約を緩和したときに初めて、その句がつけられる。これでは、上記2つの指標、特に報酬、を適切に見積もることができないので、制約違反が起こったときでも不適格性を含んだ句をつくり、これに対してコストと報酬という2つの指標を割り当てている。コストは不適格性を含んだ句をつくる時に計算し、報酬は不適格性を含んだ句がどのような句かによって決定する。もちろん、この不適格性を含んだ句は、通常の解析には用いず、不適格文処理のときにのみ用いるので、不適格性を含んだ句をつくることによる通常の解析への影響はない。

不適格文処理は、通常の解析が入力文の処理に失敗したときに、部分解析結果の中から報酬が最大でコストが最小のものを選択することによって行なわれる。これが報酬最大・コスト最小の原理である。選択された部分解析結果の性質によって、モジュールA、モジュールBのいずれかが起動される。モジュールAは、主に

¹コストと報酬を基準とする考え方は、関連性理論[7]が関連性の程度条件として用いている文脈効果と処理努力から着想を得たものである。

制約違反を扱うためのものであり、選択された部分解析結果を通常の解析に使用させるものである。つまり、制約の緩和に相当している。

モジュール *B* は部分解析法の拡張と捉えることができ、語順の違反などの構造違反を扱う。部分解析法では、第 1 段階の解析が失敗した場合に、解析途中で得られた句や節などの断片を集めて、文全体の構造や意味を得ているが、モジュール *B* では、断片を集める際の情報として、語が持っている結合価に関する情報を利用している。選択された句の結合価情報を利用することにより、選択された句がどの句と結びつくかを知ることができ、結びつく句の候補を探索し、それらを結びつけることにより処理を行なっている。その際、新しく矛盾が検出された場合は、その結果に対してコストが付与されることになる。

モジュール *A* とモジュール *B* はともに報酬最大・コスト最小の原理に基づいて制御されているため、モデル全体として、モジュール性の高い、柔軟な処理機構を提供することができる²。

5 自然な発話への拡張

言い直しや間投詞を扱うためのモジュールとしてモジュール *B'* を新しく導入する。先にも述べたように、言い直しを適切に処理するためには、言い直しにおける修復対象と訂正部分の同定が必要となる。

言い直しにおいては、修復対象と訂正部分に同じ語句が現れることが多い。例えば、例 1 では、「こと」が修復対象と訂正部分に現れている。このような語句の繰り返しは、言い直しの検出・訂正において重要な手掛りとなる。さらに、音声対話の書き起こしテキストの場合、ポーズの位置に句点が置かれることが多いので、言い直しを検出するための手掛りとして、句点の情報も利用する。

言い直しの処理においては、修復対象と冗長部分の削除が必要である。そのためにモジュール *B'* では、スキップ操作を用いる。さらにスキップ操作に対してコストを考えることで、他のモジュールと同様にコストと報酬の関係に基づいて制御できるようにする。

そのコストはスキップの容易さによって定義される。先に議論したように、言い直しにおいては同じ語句や

類似した語句が反復されることが多い。したがって、修復対象と訂正部分に含まれる句との類似度をスキップの容易さの尺度として用いる。類似度は、統語的、意味的、音韻的な観点から見積もられるものであり、類似度が大きい程、スキップのコストが小さいと考える。さらに、スキップの容易さの尺度として、スキップする語句の長さも用いる。これは、間投詞や語の断片を適切にスキップするためのものである。このように定義されたコストは、スキップ操作によってつくられる新しい句に割り当てられる。

モジュール *B'* は、報酬最大・コスト最小の原理により選択された句がすでに解析で使用されているものであり、かつ、句点もしくは間投詞がその直前にあった場合に起動される。モジュール *B'* で用いる手続きは以下のようなようになる。

モジュール *B'* の手続き

1. 間投詞、挿入句

直前の語をその長さに応じたコストでスキップする。

2. 繰り返し

形態的な類似度を計算し、その類似度とその語句の長さに基づいてその語をスキップする。

3. 言い直し

修復対象と、訂正部分に含まれる句との間の形態的、統語的、意味的な類似度を計算し、その類似度とその語句の長さに基づいてその語句をスキップする。

4. 同格

意味的な類似度を計算し、その類似度とその語句の長さに基づいてその語をスキップする。

モジュール *B'* で、新しくつくられた句は、モジュール *A* が起動されたときに使用されることになる。モジュール *B'* もコストと報酬の基準に基づいて制御されているために、このモジュールを他のモジュールと協調的に動作することができる。

モジュール *B'* の例として、例 1 の文を考えてみる。通常の解析により、部分解析結果として、「会議のこと」、「あの一」、「登録のことで伺いたいのですが」、という 3 つの句が得られる。報酬最大・コスト最小の原理により、動詞句「登録のことで伺いたいのですが」が選択される。この句はすでに解析で用いられている句であり、かつ、直前に句点があるため、モジュール *B'* が起動される。モジュール *B'* は上記の手続きにより動作し、間投詞「あの一」がスキップされ、あるコストがスキップ操作の結果としてつくられる句に付与される (図 1)。こ

²このモデルでは、モジュール *A*、モジュール *B* の出力を受けとり入力文の適切な解釈を行なうモジュール *C* があるが、本稿に直接関係はないので説明は割愛する。

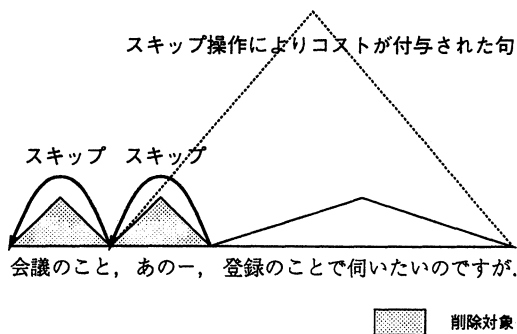


図 1: 例 1 の解析の途中状態

の句は通常の解析では使用されないため、再び不適格文処理が起動される。報酬最大・コスト最小の原理により、動詞句「あの一、登録のことで伺いたいのですが」が選択される。この句は通常の解析で用いられていない句であるため、モジュール A が起動され、この句が解析で使用される。しかし、解析はそこで再び停止し、不適格文処理が起動される。さらに、報酬最大・コスト最小の原理により、動詞句「あの一、登録のことで伺いたいのですが」が選択される。この句はすでに解析で用いられている句であり、かつ、直前に句点があるため、モジュール B' が起動される。モジュール B' は上記の手続きにより動作し、「会議のこと」を言い直しとしてスキップし、新しい句を生成する。以下、同様の処理を繰り返し、最終的にモジュール A が全体を覆う句「会議のこと、あの一、登録のことで伺いたいのですが」を選択することによって不適格文処理が終了する。

6 おわりに

本稿では、以前我々が提案した、不適格文のための統合的な処理モデルに新しいモジュールを追加することによって、言い直しや間投詞などの自然な発話に特徴的な現象を扱う手法について述べた。このモジュールも、統合的な処理モデルにおける他のモジュールと同様に報酬最大・コスト最小の原理に制御されており、モデル全体として、柔軟性の高い枠組みとなっている。

単純な例では、この新しいモジュールが他のモジュールと協調的に動作することは、確認しているが、解析の失敗の要因が一文内に複数個含まれるような複雑な例に対しては、スキップ操作のコストや類似度の尺度

を適切に見積もっていく必要がある。これに対しては、システムの大規模化を行なっていく過程において、コストの見積もりとして、どのような値が適切であるかを、実際のコーパスから学習していくことを考えている。

参考文献

- [1] J. Bear, J. Dowding, and E. Shriberg. Integrating Multiple Knowledge Sources for Detection and Correction of Repairs in Human-Computer Dialogue. In *Proc. of 30th Annual Meeting of ACL*, pp. 56–63, 1992.
- [2] P. Heeman and J. Allen. Detecting and Correcting Speech Repairs. In *Proc. of 32nd Annual Meeting of ACL*, pp. 295–302, 1994.
- [3] D. Hindle. Deterministic Parsing of Syntactic Non-fluencies. In *Proc. of 21st Annual Meeting of ACL*, pp. 123–128, 1983.
- [4] O. Imaichi and Y. Matsumoto. Integration of Syntactic, Semantic and Contextual Information in Processing Grammatically Ill-Formed Inputs. In *Proc. of IJCAI-95*, pp. 1435–1440, 1995.
- [5] K. Jensen, G. E. Heidorn, L. A. Miller, and Y. Ravin. Parse Fitting and Prose Fixing: Getting a Hold on Ill-formedness. *Computational Linguistics*, Vol. 9, No. 3-4, pp. 147–160, 1983.
- [6] C. Nakatani and J. Hirschberg. A Speech-First Model for Repair Detection and Correction. In *Proc. of 31st Annual Meeting of ACL*, pp. 46–53, 1993.
- [7] D. Sperber and D. Wilson. *Relevance*. Blackwell, 1986.
- [8] R. M. Weischedel and N. K. Sondheimer. Meta-rules as a Basis for Processing Ill-Formed Input. *Computational Linguistics*, Vol. 9, No. 3-4, pp. 161–177, 1983.
- [9] 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.
- [10] 江原暉将, 井ノ上直己, 幸山秀雄, 長谷川敏郎, 庄山富美, 森元逞. ATR 対話データベースの内容. Technical Report TR-I-0186, ATR 自動翻訳電話研究所, 1990.