

E_031

自動通訳システムのユーザインタフェースレベルでの統合 A User Interface-Level Integration Method for Multiple Automatic Speech Translation Systems

長田 誠也†
Seiya Osada

山端 潔†
Kiyoshi Yamabana

花沢 健†
Ken Hanazawa

奥村 明俊†
Akitoshi Okumura

1. まえがき

自動通訳システムの研究は数多く行なわれており[1][2][3]、これらのシステムの多くは、音声認識エンジン、機械翻訳エンジン、音声合成エンジンの3つのエンジンを持っている。音声認識エンジンの出力の正しさは、利用者が判断できるが、自由文機械翻訳エンジンの出力の正しさは、利用者が判断することは難しい。一方、機械翻訳エンジンには例文選択型翻訳エンジンがあり、翻訳精度が保障されているが、カバレッジが不十分であるという課題がある。

そこで、この2つの翻訳エンジンを統合することで翻訳精度とカバレッジを両立させて、更に使いやすくした統合方式が望まれている。本稿では、従来の統合方式と比べて利用者の操作を更に容易にするために、自由文機械翻訳エンジンを持つ自動通訳システムと、例文選択型翻訳エンジンを持つ自動通訳システムを、ユーザインタフェースレベルで統合する方法を提案する。

2. 従来手法

自由文機械翻訳エンジンと例文選択型翻訳エンジンの課題を解決するために、この2つの翻訳エンジンを統合した自動通訳システムが存在する[4]。このシステムは図1のように音声認識エンジンと音声合成エンジンと2つの翻訳エンジン（自由文機械翻訳エンジンと例文選択型翻訳エンジン）を持つ。また例文翻訳エンジンは、話し手が発話した内容に近い文を検索する例文検索エンジンと、検索した文の対訳を取得する例文翻訳エンジンを持つ。

このシステムは図1で示すように、話し手がこのシステムに発話してから、聞き手が翻訳結果を聞くまでに、最大3回の操作が必要になる。すなわち、1回目の操作はシステムに向かって発話、2回目の操作は音声認識結果を自由文機械翻訳エンジンで翻訳させるか例文選択型翻訳エンジンで翻訳させるかの選択、3回目の操作は2回目の操作で例文選択型翻訳を選択したときに、例文検索エンジンが出力した結果の中から1つを選択するか、もしくは例文の中に話し手が発話した内容が含まれていないときなどのときに、音声認識結果を自由文機械翻訳させるかの選択、の3回である。

さらに、この2回目の操作では、自由文機械翻訳か例文選択型機械翻訳のどちらの機械翻訳を選ぶかの手がかりがないため、利用者はどちらを選ぶか迷ってしまう課題も存在している。

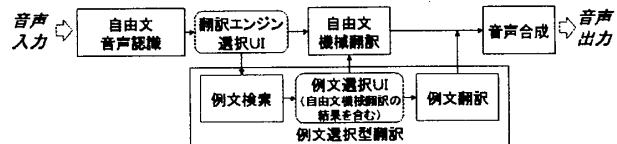


図1：従来システム構成図

このように、このシステムはシステムを構築するために必要なエンジンをモジュールレベルで統合し、その統合の結合部分をユーザインタフェースでつなぐという構成になっているため、利用者は発話するだけで翻訳結果が得られるようなシンプルな通訳システムのように操作することができない。このため、利用者はこの2つの翻訳エンジンの関係を含めた通訳システムの全体構成図を理解しておく必要があり、システムの内部状態が現在どこにあるかを意識しながら操作する必要がある。

3. 提案手法

本稿では、従来手法のように自動通訳システム内の各エンジンをモジュールレベルで統合するのではなく、図2のように自由文機械翻訳エンジンを持つ自由文通訳システムと例文翻訳エンジンを持つ例文選択型通訳システムをユーザインタフェースレベルで統合する方法を提案する。

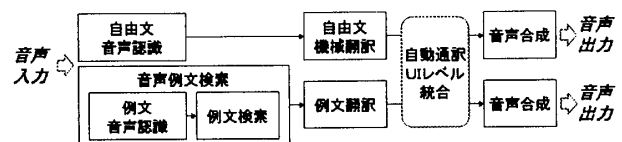


図2：提案手法の構成図

この提案手法では、図2における「自動通訳UIレベル統合」において、自由文機械翻訳の原文と訳文のペアと、例文翻訳の原文と訳文のペアを図3のように並べて表示する。図3における網掛けの選択肢が自由文通訳システムにおける入力文とその翻訳結果で、それ以外の選択肢が例文選択型通訳システムの例文検索結果とその対訳となっている。また太枠で囲まれた選択肢が現在選択されている選択肢で、利用者は上下キーを押すことでこの太枠を移動して自分が伝えたい内容を選択できる。

このようにすることで、このユーザインタフェースは、網掛けの選択肢は翻訳結果が正しくない可能性がある選択肢で、網掛けのない選択肢は原言語の文はユーザが伝えたい内容とは違うかもしれないけれども翻訳結果は保証されている選択肢であるということを利用者に示している。

† 日本電気株式会社 メディア情報研究所

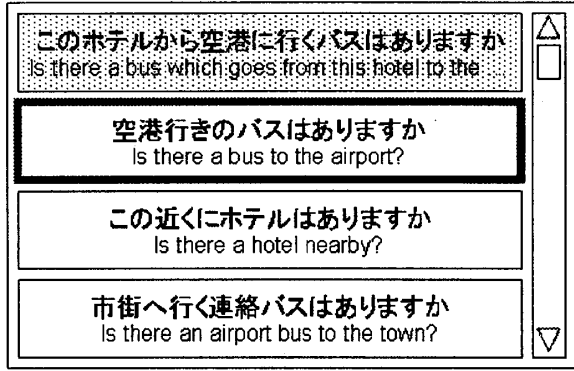


図3：自由文機械通訳と例文選択型機械通訳の統合UI

利用者は原言語の情報のみで選択肢を決定し、ただし網掛けの選択肢は翻訳精度が必ずしも100%でないことを意識すればよい。さらに、利用者はそれぞれの選択肢がどの翻訳エンジンの結果であるかを意識する必要もなくなり、操作が容易になる。また、システムの全体構成図を理解しておくことや、システムの内部状態を意識しておく必要もなくなる。

4. 実装

本提案手法を QVGA(320×240)の画面と十字ボタン、OK ボタン、キャンセルボタンを持つ携帯端末に実装した。ただし、ハードウェアの制約から、音声認識エンジンと音声合成エンジンを2つの通訳システムで共有している。

また QVGA の画面に図3の統合UIを表示し、このUIでの操作や音声認識の操作を含んだすべての通訳システムの操作を、上記ボタンのみでできるようにしている。

5. 評価

従来手法と提案手法で、ユーザの操作数について理論上の評価を行う。ただし、従来手法では最初の操作で自由文機械翻訳を選んだときに翻訳結果が正しくないときは処理が終了してしまうため、2つの手法で通訳の目的達成率が異なってしまう。このため、従来手法の自由文機械翻訳と翻訳エンジンの切り替えの順序を交換し、音声認識した後に自動的に自由文機械翻訳を行い、自由文機械翻訳の結果が正しいときは終了、正しくないときは例文検索を行うようにする(図4)。

ここで、音声認識率を x 、自由文機械翻訳率を y 、音声認識結果が正しい元での例文検索率を z 、音声認識結果が正しくない元での例文検索率を w とすると(音声認識結果が正しくないときの自由文機械翻訳正解率は0%としている)、従来手法と提案手法の平均操作数は次のようになる。

従来手法：

$$1 \ xy + 3\{x(1-y)z + (1-x)w\} + 2\{x(1-y)(1-z) + (1-x)(1-w)\} \quad (1)$$

提案手法：

$$2 \ xy + 2\{x(1-y)z + (1-x)w\} + 1\{x(1-y)(1-z) + (1-x)(1-w)\} \quad (2)$$

これにより、提案手法が従来手法よりも平均操作数が少なくなるときは、 $(2)-(1) < 0$ から $xy < 0.5$ となる。

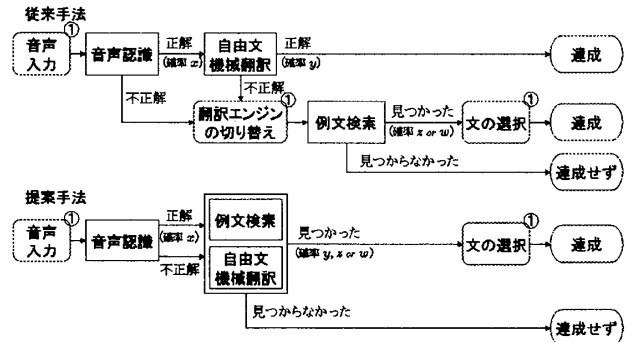


図4：従来手法と提案手法の操作手順 (利用者の操作を①がついたブロックで示す)

つまり、音声認識が正しくかつ自由文機械翻訳結果が正しくなる確率が50%以上のときは、従来手法の方が平均操作数が少なくなること示している。

ただし、この式ではすべての操作に対するコストが同じであるとしている。しかし、従来手法には自由文機械翻訳結果が正しいかどうかを利用者が判断するという操作が入っており、自動通訳システム利用者にとって、この操作のコストは高いと考えられる。このため、自由文機械翻訳の性能がある程度高いとしても、提案手法の方が利用者にとって使いやすい手法であると考えられる。

6. おわりに

本稿では、複数の自動通訳システムをユーザインタフェースレベルで統合する方法を提案した。従来の統合システムでは利用者がシステム全体を理解してシステムの内部状態を意識しながら使う必要があったが、提案手法では利用者がシステム状態を意識せずに使うことが可能である。また、それぞれの操作数を比較し、提案手法が従来手法よりも優位であることを示した。

参考文献：

- [1] Takao Watanabe, Akitoshi Okumura, Shinsuke Sakai, Kiyoshi Yamabana, Shinichi Doi, and Ken Hanazawa. 2000. An automatic interpretation system for travel conversation. In *Proceedings of ICSLP 2000*.
- [2] Jan Alexandersson, Norbert Reithinger and Elisabeth Maier. 1997. Insights into the Dialogue Processing of VERBMOBIL. In *Proceedings of ANLP*, pages 33-40.
- [3] Toshiyuki Takezawa, Tsuyoshi Morimoto, Yoshinori Sagisaka, Nick Campbell, Hitoshi Iida, Fumiaki Sugaya, Akio Yokoo and Seiichi Yamamoto. 1998. A Japanese-to-English speech translation system: ATR-MATRIX. *ICSLP-98*, pages 2779-2782.
- [4] 池田崇博, 安藤真一, 佐藤研治, 奥村明俊, 渡辺隆夫. 自由文通訳と例文選択型通訳を統合した自動通訳システム, FIT2002, LE-1