

英語短文穴埋め問題自動生成システム MAGIC の試験的公開

－創造的学習支援への期待を乗せて－

【ポイント】

名古屋大学大学院情報科学研究科の渡邊豊英教授、小尻智子助教を中心とする研究グループは、NTT コミュニケーション科学基礎研究所との共同研究により、英語短文穴埋め問題を自動的に生成するシステム MAGIC (Multiple-choice Automatic Generation system for Cloze questions)を開発しました。この度、名古屋大学では MAGIC の実用性などを評価するため、これを試験的に公開します。

入力英文から自動的に問題文章を選別できること、単語の品詞・位置に関わらず空欄を決定できること、誤選択肢を自動的に生成できることなどの特徴があります。本試験システムは従来の提案されている方法と比べて、TOEIC 問題を生成する実験ではその問題一致率がおおよそ 2 倍でした。また、従来の学習支援システムが予め用意された問題を順次提示し、解答させるという受動的な学習形態であることに対して、学習者が自らから好みの英文を用いて問題を作成し、解答できるという能動的な学習形態を実現できます。

(公開 URL) <http://www.watanabe.ss.is.nagoya-u.ac.jp/ja/>

【背景】

学習者が個々に学習を進める形態として個別学習[1]があります。この個別学習に利用する問題は教師などの専門家によって作成されています。近年、e-Learning の普及により、インターネットなどの情報ネットワークを介して、コンピュータが学習を支援する形態が加速化、拡大しつつあります。このような環境では、個々の学習者の学習目的に適した大量の問題が必要となっています。特に、選択肢などの問題は短時間で解答が可能である、クイズ的に解答が可能であることなど、学習問題としては最も標準的な出題形式となっているため、利用度が大きく、需要は高いと思われます。

【研究の内容】

英語短文から単語の穴埋め問題を作成する技術を開発しました。そのアプローチは専門家が穴埋め問題を作成する際に、経験に基づいて学習目的に応じた問題文を選択し、空欄や誤選択肢を決定していることに基づいています。専門家の問題作成に対する知識を獲得するために、既に存在する穴埋め問題から、穴埋めの単語、文法に関する統計情報を機械学習[2]して、システムに反映させています。

システムの処理は図 1 に示すように、大きくは 3 つの処理手続きから構成されています。最初に、入力された英文から問題文として適切な文を選択します(問題文の選択)。次に、選択された文に対して空欄部分を決定します(空欄部分の決定)。最後に、空欄部分に対応する誤選択肢を選出し、選択肢を生成します(誤選択肢の作成)。問題文の選択では優先度学習[3]により予め学習した優先度情報に基づいて穴埋め問題文の候補を優先度とともに出力します。空欄部分の決定では穴埋め問題文の候補に含まれる単語を系列ラベリング[4]により、予め学習した空欄の統計情報に基づいて空欄部分を決定します。誤選択肢の作成では誤選択肢候補を生成する規則に基づいて、辞書などを用いて誤選択肢となる単語を選択します。選択した単

語をインターネットで検索し、その検索に失敗した単語を誤選択肢として設定しています。

図 2 はその MAGIC の対話画面を示しています。

【成果の意義】

学習者が入力した英文から自動的に穴埋め問題を作成する方法、及びその試験システムを実現しました。これは学習者が自ら興味がある英文を選んで学習できるという側面を持ち、従来の予め用意された問題を受動的に解いていくという学習過程とは全く異なった能動的な学習形態を提供することができます。さらに、知的教授システム[5]などの機能として連携させれば、問題作成において個人の学習ポイントに適合した指導を可能とする学習支援ができます。すなわち、より創造的な学習支援が可能となると思われます。

本システムの性能を評価するために、TOEIC[6]の英語短文穴埋め問題の穴埋め部分の単語(正解語)が補充された英文を入力とし、従来提案されている方法と比較しました。その結果、従来の同様なシステムに比べて、おおよそ 2 倍の一致率(TOEIC と同じ空欄箇所を選出した割合)で、空欄を確定しました。本方法がより精度の高い処理を実現していることがわかります。

【用語の説明】

[1] 個別学習

一人一人を対象に行う学習方法です。家庭教師、チュータなどが学習者とともに実施する学習ですが、独学など、自分一人で行う学習も広い意味では含んでいます。

[2] 機械学習

人間が様々な知的活動で自然に行っている学習過程において、それと同等な機能をコンピュータで実現させるための技術・手法のことです。標本集合を対象にデータ解析を行い、有用な規則、知識、判定基準などを抽出するために行われます。

[3] 優先度学習

複数の候補に対してそれぞれ優先度を算出して候補集合を相対的に分類する方法です。

[4] 系列ラベリング

単語列から固有な表現を抽出するために、それぞれの単語にタグ(印)を付けることをいいます。

[5] 知的教授システム

コンピュータが学習者モデル(学習者の学習理解の状態を表したもの)を用いて学習者の学習状況に応じて学習を指導できる枠組みを有した教師システムです。

[6] TOEIC (登録商標)

英語の統一試験

(参考資料: 図)

図1 システムの処理



図2 MAGICの対話画面

(a) 英文入力画面

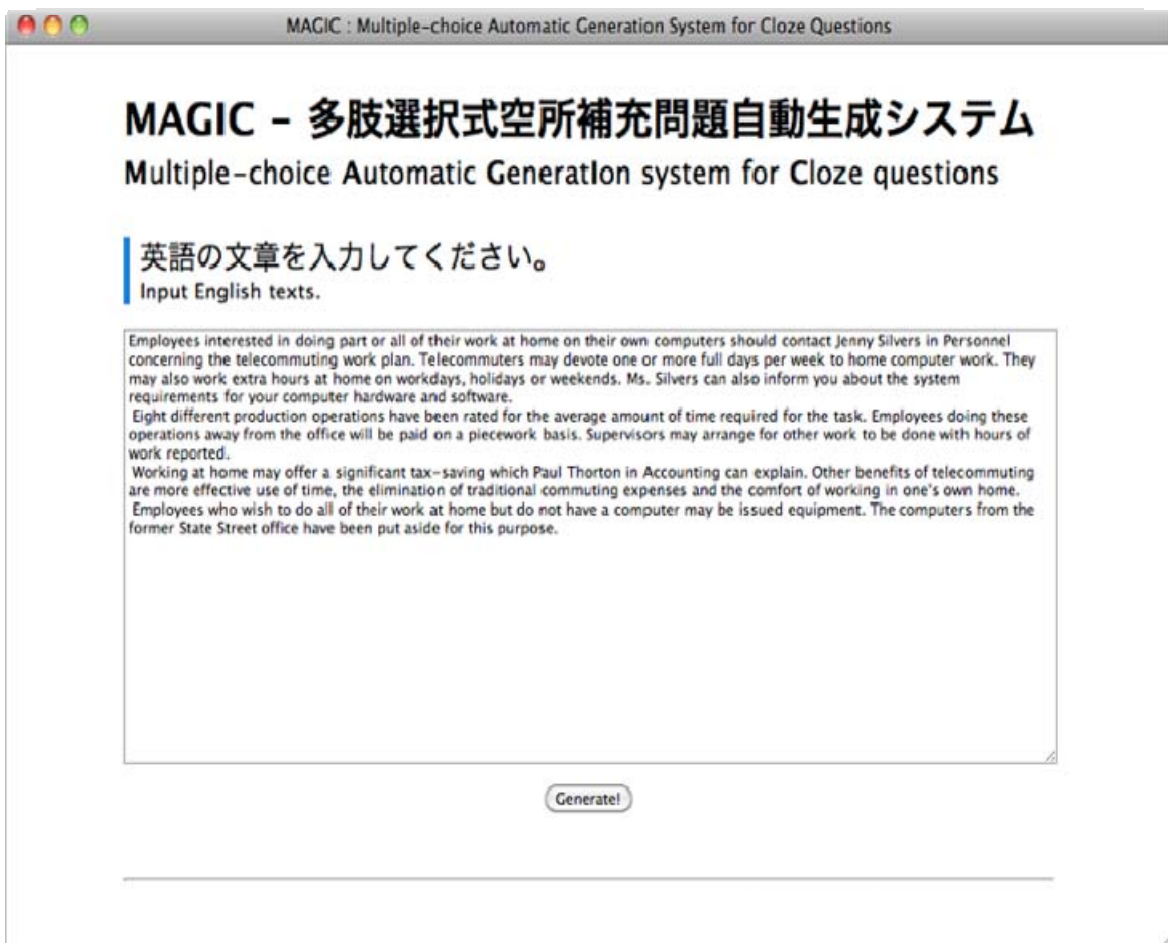


図 2 MAGIC の対話画面

(b) 問題生成の画面

