

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 22 日現在

機関番号：14301

研究種目：若手研究(B)

研究期間：2014～2015

課題番号：26730115

研究課題名(和文) 作業者育成を支援するクラウドソーシング実現のための統計的スキル推定技術の開発

研究課題名(英文) Statistical skill estimation methods for crowdsourcing with career development

研究代表者

馬場 雪乃 (Baba, Yukino)

京都大学・情報学研究科・助教

研究者番号：40711453

交付決定額(研究期間全体)：(直接経費) 2,100,000円

研究成果の概要(和文)：統計的な作業者スキル推定手法を構築した。特に英文和訳を対象にして、翻訳者の語彙知識を利用して翻訳能力を測る新しい確率モデルを提案した。英文翻訳作業時に、英文中の各単語を知っているか否かを作業者に答えさせ、これを語彙知識として用いる。提案モデルは、作業者が高い翻訳能力を持つ場合と低い翻訳能力を持つ場合で異なる語彙知識回答モデルから構成される。クラウドソーシングにて収集したデータに提案手法を適用し、我々が考案したスキル推定手法を用いることで、英文和訳に対する適切な翻訳者割当が可能となることを確認した。

研究成果の概要(英文)：We built a statistical skill estimation method. We particularly focused on English-Japanese translation jobs and proposed a method which incorporates translators' vocabulary knowledge for translation ability assessment. Our method involves just asking translators to tell if they know given words. Using this vocabulary information, we build a probabilistic model to estimate the translators' vocabulary and translation abilities simultaneously. We evaluated our method in a realistic crowdsourcing translation setting in which there is a great need to measure translators' translation ability to select good translators. The results of our experiments show that the proposed method accurately estimates translation ability and selects translators who have sufficient skills in translating a given sentence.

研究分野：機械学習

キーワード：クラウドソーシング スキル推定

1. 研究開始当初の背景

(1) クラウドソーシングとはインターネットを通じて不特定多数の人々に業務を委託する仕組みの総称である。必要なときに必要なだけ外部の労働力を獲得できるクラウドソーシングは、産業界でも学術界でも急速に利用が拡大している。クラウドソーシングの利便性を向上するための研究がデータマイニングの分野で行われている。

(2) 特に、成果物の品質管理手法が広く研究されている。クラウドソーシングではインターネット上で全ての取引がおこなわれるため発注者と作業者が一度も顔を合わせることがなく、発注者は能力や信頼性がわからない相手に仕事を依頼することになる。作業者から提出された成果物を検品し、一定の品質に達していない場合は別の人に依頼し直すことで品質を保証することはできるが、大量の仕事をクラウドソーシングで発注すると検品作業は困難となる。そこで、発注者による検品作業を省いても高い品質の成果物を安定して獲得できるようにする、統計的品質保証手法が研究されている。

(3) 統計的品質保証手法等のクラウドソーシングにデータマイニング技術を適用する研究の多くは、発注者の利便性向上を目的としていた。一方で、作業者の利便性向上はほとんど注力されていない。たとえば、作業者に対して「作業者が好みそうな仕事」を推薦する手法がようやく研究されはじめたに過ぎない。

2. 研究の目的

(1) 本研究では、作業者の利便性向上の一環としてクラウドソーシング上で作業者のスキル育成の実現を目指し、仕事を通じて作業者のスキルを向上させるための統計的技術を開発する。開発した統計的技術により、作業者の現在のスキルと作業者が目標とするスキルに応じて、適切な仕事を適切な順番で提示する支援システムの実現を目指す。この支援システムを利用することで作業者は、仕事に取り組み報酬を獲得しながら、スキルを向上させることができる。

(2) 目標スキルは作業者から与えられるものとして、本支援システムを実現するためには他に、作業者の現在のスキル・仕事の要求スキル・作業者が仕事に成功する確率という三つの情報が必要となる。これらの情報を使って、支援システムは仕事集合の中から下記二つの基準に従い適切な仕事を選ぶ。(a) 作業者の現在のスキルで成功可能な仕事(十分な品質の成果物を提出し報酬を得ることができる、難しすぎない仕事)、(b) その仕事に成功することで、目標スキルの到達に近づく仕事(目標スキルの獲得に適した、簡単すぎない仕事)。

3. 研究の方法

(1) クラウドソーシングにおける主要な仕事の一つである英文和訳翻訳に着目し、スキル推定のための確率モデルを構築した。作業者には翻訳作業実施前にテストを受けさせ、その結果を利用して翻訳能力を推定する。翻訳能力と語彙能力には関係があると仮定し、事前テストは翻訳テストと語彙テストの両方から成るものとした。翻訳テストでは、いくつかの英文が与えられ、それぞれを和訳するよう指示される。翻訳結果は、専門家によって採点される。語彙テストは翻訳テストの英文を利用して実施され、各英文中の単語の意味を知っているかどうかを回答させる。語彙テストは翻訳テストと異なり自己申告制であるため、専門家による採点を必要としない。専門家による翻訳テストの採点にはコストが掛かるため、翻訳結果は一部だけが採点されるものとした。我々が構築した確率モデルは、事前テストを受けた作業者が、新しい翻訳文を正しく翻訳する確率を推定する。その際、作業者は新しい翻訳文に対する語彙テストも受けるものとする。

(2) 確率モデルは、ある翻訳文に対するある作業者の、語彙テスト結果の生成過程を表現する。この生成過程では、翻訳文が与えられたときに作業者は、自身の翻訳能力が高く、文章の難易度が低いほど、高確率で正しい翻訳文を出力する。次に、自身の語彙能力が高く、各単語の難易度が低いほど、その単語を「知っている」と回答する。ただし、語彙能力は二つの異なるパラメータで表現されており、作業者がその翻訳文を正しく翻訳するときとそうではないときで、発揮する語彙能力が異なるものとする。文章の難易度は、その文章の特徴とそれに対する重みパラメータの線形和で表現する。文章特徴として、文章中の単語数、カンマの数、3-gram の perplexity (言語モデルに従う度合い) 等を用いる。単語の難易度も同様に、その単語の特徴とそれに対する重みパラメータの線形和で表現する。単語特徴として、単語の難易度インデックス、コーパス中での出現回数を用いる。図 1 に確率モデルを示す。ここで、 ψ_k, θ_k, ϕ_k がそれぞれ、作業者 k の翻訳能力、語彙能力(翻訳成功時)、語彙能力(翻訳失敗時)を表す。また、 s が文章特徴、 τ が文章特徴の重みパラメータ、 v_s が文章 s の難易度である。同様に、 w が単語特徴、 π が単語特徴の重みパラメータ、 μ_w が単語 w の難易度である。 z_{ks} は、作業者 k が文章 s を正しく翻訳するかどうかを表す二値の潜在ラベルであり、 y_{kw} は、作業者 k が単語 w を「知っている」と回答するかどうかを表す二値のラベルである。

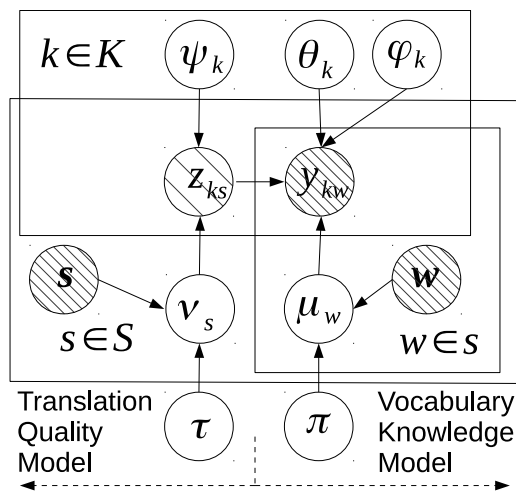


図 1: 翻訳能力推定の確率モデル

(3) 確率モデルのパラメータである、作業員 k の能力 (ψ_k, θ_k, ϕ_k)、文書・単語特徴の重み (τ, π) を事前の翻訳テストと語彙テストの結果から推定する。推定には、翻訳の正しさを潜在変数とした EM アルゴリズムを用いる。EM アルゴリズムは、E ステップと M ステップの二段階から成り、E ステップでは、確率モデルのパラメータを固定した上で潜在変数の期待値を計算する。M ステップでは、潜在変数の期待値を用いて確率モデルのパラメータの最適化を行なう。収束するまで、E ステップと M ステップを交互に繰り返すことでモデルパラメータの推定を行なう。M ステップにおいては、翻訳能力と語彙能力 (翻訳成功時) の値が近づくような正則化項を用いる。

(4) 実際に翻訳作業をクラウドソーシングで依頼し、実験用データセットを収集した。「Wikipedia 日英京都関連文書対訳コーパス」から 10 語以上の英文を抽出し、そこからランダムに 104 文章を選んだ。これらの文章に対して、55 人の作業員から合計 1,498 件の翻訳結果を獲得した。文章あたりの平均翻訳者数は 14.4 人であった。翻訳結果を専門家が採点し、その結果を使って翻訳の正しさの正解ラベルを作成した。採点は 5 段階評価で行われ、5 点の場合は「正しい翻訳」、4 点以下の場合は「誤った翻訳」だとした。さらに作業員には、翻訳元文章の各単語について知っているか否かを回答させた。

4. 研究成果

(1) 翻訳・語彙能力から成る確率モデルを用いることで、翻訳能力推定精度が向上し、クラウドソーシングでの翻訳作業員の雇用費用の削減につながる事がわかった。次のようなクラウドソーシングの利用方法を想定し、実験を行った：(a) ある翻訳元文章に対して、作業員を選び、翻訳作業を実施させる。(b) その作業員が正しい翻訳を返した場合はそこで

終了する。誤った翻訳を返した場合には、別の作業員を選び、同じ文章の翻訳を実施させる。(c) 正しい翻訳が得られるまで、a, b を繰り返す。この利用方法において、正しい翻訳が得られるまでに必要な作業員の人数を作業員雇用費用として、評価した。結果を表 1 に示す。

表 1: 作業員選択手法ごとの翻訳者雇用費用の比較。費用は、ランダム選択の場合を 1 とした相対値。

作業員選択手法	翻訳テストの採点率		
	1%	6%	13%
ランダム	1.000	1.000	1.000
既存手法	0.820	0.706	0.669
既存手法+語彙テスト	0.731	0.678	0.660
翻訳・語彙能力モデル (正則化項なし)	0.736	0.699	0.661
翻訳・語彙能力モデル (正則化項あり)	0.660	0.675	0.664

ここで「既存手法」は、翻訳元文章の特徴を利用して、「正しい翻訳」か「誤った翻訳」のどちらを返すか予測する分類器を作業員ごとに学習する。「既存手法+語彙回答」は、翻訳元文章の特徴に加えて、作業員が「知っている」単語の特徴の平均値も用いる。表 1 から、「翻訳・語彙能力モデル」に基づく選択手法が、最も少ない費用で正しい翻訳者を選択できていることがわかる。特に、翻訳テストの採点率が小さいとき、つまり翻訳テスト採点に費用を掛けないときに、「翻訳・語彙能力モデル」による翻訳者雇用費用は他の手法よりも小さい。このことから、翻訳・語彙能力を導入した確率モデルは翻訳能力を精度良く推定し、適切な翻訳者を選択できていると言える。

(2) 翻訳能力と語彙能力の関連付けは、翻訳能力の推定精度向上に有用である。表 1 の「翻訳・語彙能力モデル (正則化項なし)」と「翻訳・語彙能力モデル (正則化項あり)」を比較すると、後者の方が、翻訳者雇用費用が小さいことがわかる。後者は正則化項の導入により、翻訳能力と語彙能力の推定結果が近づくように制御することで推定の効率化を図っている。これにより、特に翻訳テストの採点率が小さいときに、翻訳能力の推定精度が向上した。

(3) 語彙テストの導入、既存手法の翻訳能力推定精度の向上にも有用である。表 1 の「既存手法」と「既存手法+語彙テスト」を比較すると、後者の方が、翻訳者雇用費用が小さいことがわかる。語彙テストの導入が、既存手法においても翻訳能力推定精度の向上につながった。

(4) 翻訳能力というある一つのスキルに絞っ

て研究を実施したが、翻訳テストと語彙テストを組み合わせることで翻訳能力の推定精度をさせるという点で本成果はユニークなものであり、人工知能分野の一流国際会議に採択されるに至った。国内においても、DICOM02015 優秀論文賞を受賞するなど高く評価された。このスキル推定手法は専門家による採点をほとんど必要とせず低コストで運用可能という、実用面でも大きなメリットがある。今回は翻訳スキルに着目し、翻訳テストだけではなく語彙テストの結果も組み合わせて翻訳スキルを推定する研究を実施したが、このように本来の目的スキルに関するテストと他のテストを組み合わせてスキル推定精度を向上させるというやり方は、翻訳以外のスキルにおいても有用と考えられ、汎用性が見込まれる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 1 件)

① Yo Ehara, Yukino Baba, Masao Utiyama, Eiichiro Sumita, Assessing Translation Ability through Vocabulary Ability Assessment, Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI), 2016 [査読有].

[学会発表] (計 4 件)

① 江原 遥, 馬場 雪乃, 内山 将夫, 隅田 英一郎, 語彙能力と翻訳能力の関係, NLP 若手の会 (YANS) 第 10 回シンポジウム, 2015 年 9 月, 和倉温泉ホテル海望 (石川県七尾市).

② 江原 遥, 馬場 雪乃, 内山 将夫, 隅田 英一郎, クラウドソーシング翻訳のコスト削減のための翻訳者割り当て手法の提案, 第 14 回情報科学技術フォーラム (FIT), 2015 年 9 月, 愛媛大学 (愛媛県松山市).

③ 江原 遥, 馬場 雪乃, 内山 将夫, 隅田 英一郎, クラウドソーシング翻訳のコストを削減する翻訳品質の事前予測モデル, マルチメディア、分散、協調とモバイル (DICOM02015) シンポジウム, 2015 年 7 月, ホテル安比グラウンド (岩手県八幡平市).

④ 江原 遥, 馬場 雪乃, 内山 将夫, 隅田 英一郎, 単語テストを利用した翻訳品質事前予測モデル, 2015 年度人工知能学会全国大会 (第 29 回), 2015 年 6 月, 公立はこだて未来大学 (北海道函館市).

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

[その他]

特になし

6. 研究組織

(1) 研究代表者

馬場 雪乃 (BABA, Yukino)

京都大学・大学院情報学研究所・助教

研究者番号: 40711453