

May 27, 2010

RT0907
Computer Science 5 pages

Research Report

Multi-lingual translation of short utterances of international spectators

Hitoshi Akimoto, Tetsuya Nasukawa, Hironori Takeuchi, Hiroshi Kanayama

IBM Research - Tokyo
IBM Japan, Ltd.
1623-14 Shimotsuruma, Yamato
Kanagawa 242-8502, Japan



Research Division
Almaden - Austin - Beijing - Haifa - India - T. J. Watson - Tokyo - Zurich

観客の声の翻訳知識自動獲得方法

秋本 仁志, 那須川 哲哉, 竹内 広宜, 金山 博

Multi-lingual translation of short utterances of international spectators

Hitoshi AKIMOTO, Tetsuya Nasukawa, Hironori Takeuchi, Hiroshi Kanayama

近年, オンラインゲームの発展に伴い, 様々な国のユーザーがゲームに観客として参加するケースが増加しており, 観客がゲームのイベントに反応して行う応援や非難などの発言が多言語化している傾向が見られる. このような場合に, 現状では観客であるユーザーにとって未知の言語を用いた他の観客の発言は, たとえユーザーの感情と同じことを意味していても, 多くの場合にノイズとして扱われてしまう. 一般的に, ゲーム中のユーザーの発言は省略語や会話表現であることが多く, 対訳辞書や機械翻訳を用いてユーザーの母国語に翻訳するのは困難である.

そこで本稿では, 翻訳に利用するための多言語コーパスを作成するために, 同じチームやプレイヤーを応援している観客が, 応援しているチームが得点したときなどの類似した状況でした発言は同じような意味であるだろうという仮説を利用して, 同じような状況で発言された他言語にまたがる同様の意味の語句を収集する手法を提案する.

1. はじめに

オンラインゲームの普及と発展に伴い, ゲームに参加するユーザーは実際にゲームをプレイするだけでなく観客の立場でチャット等を用いて応援や非難をしながら観戦を楽しむということも行われるようになった. また, ゲームコンテンツの多言語化に伴いゲーム中ではユーザーによって使用する言語が異なる可能性があり, その場合ユーザーが理解できない言語での発言は, そのユーザーが思っていることと同じ意味であってもほとんどの場合ノイズとして扱われてしまっている.

他言語を理解するための技術として, 対訳辞書を用いたり, 機械翻訳を用いたりする方法があるが, 一般的にオンラインゲームのチャットでの発言は会話表現や省略語, また Web やゲーム中で新たに流行している新語などが非常に多く用いられ, これらの発言に対して辞書引きや機械翻訳で得られた結果をユーザーに提示するのは精度の面で大きな問題がある.

そこで, 本稿では同じ立場で見ている観客は同じイベントを見たとき, 同じようなタイミングで同じような意味の発言をするという仮定に基づき, 類似した状況で発言された複数言語にまたがる語句を収集する手法を提案する. ここで, 同じ立場とは例えば同じチームや同じプレイヤーを応援している観客同士の関係などを示し, 類似した状況とは応援しているチームが得点した場合や失点した場合のように発言を誘発するイベントがゲーム中で起きたことを観客として見ている状況を示す. この手法で集められた, ある言語対の同じ立場の観客が同じシーンを見ながらした発言を翻訳に用いることで, 厳密な語の意味は異なるかもしれないが, 観客に翻訳による違和感を抑えながら観戦しているゲームの臨場感や没入感を高めることの助けになると考えられる.

なお, 本稿ではゲームに参加しているユーザーのうち, 試合などゲームを実際に操作しているユーザーをプレイヤーと呼び, 行われているゲームを観戦しながら他のユーザーとコミュニケーションをとっているユーザーを観客と呼ぶ.

2. 問題設定

2.1 提案手法の対象と目的

本手法ではゲームなどの観客がゲーム中のイベントに対して発する短い発言を対象として, 同様の外国語の発言と意味的に結びつけることで, 翻訳を可能にするための知識を獲得することを目的とする.

対象となる発言とそのタイプの例を以下に示す.

発言の例

- 声援
 - ・頑張れー，やれー，いけー
- 喜び
 - ・やったあ，よっしゃ，
- 感嘆
 - ・すげえ，さすがっ，素晴らしい
- 悲観
 - ・ダメだー，あーーーーー
- 非難
 - ・ひどすぎる，サイアクー
- 悲しみ・落胆
 - ・うわあ，やられたー，残念，orz...

例に示したような発言に対して，例えば日本語の「よっしゃ」と英語の「Good job!」や，「ダメだー」と「Noooo!」という発言を結びつけることで翻訳知識を獲得できる．

2.2 前提条件

全ての発言は，発言者の固有 ID 及び，発言者の使用する言語と紐付いているものとする．

また，全ての発言はシーン／イベントと紐付いているものとする．サッカーゲームの例を挙げるとシーンには，シュートシーン，ゴールシーン，ドリブルシーンなどが存在する．

2.3 問題の特徴

観客の発言はそれぞれのシーンを見たことによって引き起こされる．同じシーンを見たときでも，シュートを決めた選手のサポーターと，点を取られたキーパーのサポーターは意味的に異なる発言をすることが多い．例えば「シュートが成功したシーン」に起こる発言の例としてそれぞれ以下のようなものがある．

シーンに紐付いた発言の例

- ・素晴らしい，すげー（シュートを決めた選手／チームのサポーター）
- ・やった，いいぞ（シュートを決めた選手／チームのサポーター）
- ・やられた，ゲー（失点したキーパー／チームのサポーター）
- ・惜しいっ（失点したキーパー／チームのサポーター）

シーンごとの発言回数の分布は一定ではなく，発言が起こりやすいシーンとそうではないシーンが存在する．また同じシーンを見ている場合でも，観客の立場によって発言の起こりやすさは異なり，応援している選手が交代で出場した場面などのように，ある立場の観客にとっては特に発言が起こりやすいシーンが存在する．

2.4 既存手法

他言語にまたがるコーパスから翻訳対を獲得する代表的な手法として，対訳辞書を用いて各言語の文の意味的近さを調べることで対訳関係を推定する手法[1,2]がある．しかし，本手法が対象とするゲームの観客の発言のように「うわあ」と「うわあっ」が同じ意味として出現するような多様性に対して，網羅的に辞書登録して対応するのが困難であるという問題がある．

また，2.2 で述べたように，同じシーンを見たときに，ある立場の観客が喜んでいるシーンで，別の立場の観客が悲しんでいるという場合が頻繁に起こる．よって，2.1 で例に挙げたような「声援」，「喜び」などの発言のタイプを考えると，同じシーン／イベントに紐付いている発言でもタイプが異なることがあるため，翻訳対の獲得は簡単ではない．

3. 提案手法

2.3 で述べた問題の特徴を利用することで、本手法では発言者となる観客の立場を特定できれば、同じ立場の観客は同じシーンを見て同じ様な発言をすると仮定する。つまり、同じ立場の観客が同じイベントに対してする発言は、言語が異なっても同じ意味を表現している可能性が高いと考えられる。

また、同じ立場の観客は同じオブジェクトに関するシーン/イベントに反応すると考えられる。ここで、オブジェクトの例として、カーレースなら車、サッカーなら選手などが挙げられる。例えば、サッカーゲームを見ている場合には、応援している選手がドリブルで駆け上がる時に声援を送り、その選手がシュートを打ったときに歓声を上げ、シュートが外れたときに残念がるなどといった、シーン/イベントに対する発言がある。つまり、観客の発言がその観客の立場と見ているシーン/イベントによって起こるために、シーン/イベントによって発言のタイプが変わるだけでなく、立場ごとに発言を誘発するシーン/イベントと、発言が起こりにくいシーン/イベントが存在する。

本手法では、これらの仮定や特徴を利用することで、シーン/イベントに紐付いた発言の量と発言者の情報から観客の立場を推定し、観客の立場を利用することで、同シーン/イベントで起きた同じ立場の観客の発言を獲得することで、言語をまたがった意味的に近い発言の集合を獲得する。

以下に、本手法の処理の流れを示す。また、その後各 Step の詳細について述べる。

- Step 発言量の分布と画面の切り替わりからシーンを分割する
- Step 各シーンで起きた発言を正規化して、発言・発言頻度からシーン特徴ベクトルの作成
 - ・ 発言の語尾 "ー" の有無などわずかな差を吸収して正規形にする
 - ・ 辞書があれば正規化に利用
 - ・ 特徴ベクトルの要素は[0:1]正規化
- Step 各シーンを特徴ベクトルでクラスタリング
 - ・ クラスタ数を決定しない凝集型クラスタリングなどを利用
 - ・ 過去の結果から given なクラスタの特徴ベクトルがある場合は利用可能
- Step 各シーンクラスタ（類似シーン集合）におけるユーザーの立場推定
 - ・ 同シーンクラスタ内で各シーンの発言者の発言回数でユーザー特徴ベクトル作成
 - ・ 特徴ベクトルでクラスタリングできたクラスタにおける各ユーザーの立場推定
 - ・ クラスタリングについては Step 同様に行う
- Step 各言語について同クラスタ/同立場のユーザーの発言集合ペアの抽出

3.1 シーン分割 (Step I)

まず、画面の切り替わり情報や発言頻度などの分布を用いて、シーンの分割を行う。ここでは発言量が閾値以上の箇所のみシーンとして抽出を行う。

シーン分割の手法としては各シーンの画像の類似度を利用して自己組織化マップを用いた手法[3]などがあり、本稿ではシーンの分割は既存手法を用いて行うものとする。

3.2 シーン特徴ベクトルの作成 (Step II)

Step I で分割した各シーンにおいて、発言されたそれぞれの語の頻度を要素とする特徴ベクトルをシーン毎に作成する。それぞれの発言は、以下に挙げるような表層上のわずかな差は吸収してまとめ上げる。

- ・ 句読点・感嘆符・長音などの有無
- ・ 漢字/仮名書き・大文字/小文字などの差
- ・ 辞書を用いることができるならば同義語などの意味的に近い語句

また、特徴ベクトルの要素は、その中の最大の頻度のものが 1.0 となるように正規化する。

3.3 シーンのクラスタリング (Step III)

Step II で計算したシーンの特徴量を利用して、シーンのクラスタリングを行う。各シーン間の類似度はそれぞれの特徴ベクトルのコサインなどを用いて計算することができる。

また、例えばサッカーゲームなどの同じゲーム内で、それぞれのチャット履歴から発言集合を抽出したい場合などは、それまでの処理で得た既知のシーンクラスタを利用して、シーンをカテゴリ化することができる。

3.4 ユーザーの立場推定 (Step IV)

Step III にてクラスタリングされた、シーン数 2 以上で構成される各シーンクラスタについて、発言の有無のデータからユーザーをクラスタリングする。またそのとき、クラスタ (の重心) 間の距離 (乖離度) が十分あるかどうかで立場が明確に分かれるかを評価する。

ユーザーのクラスタリングは以下の手順で行うものとする。

1. 各シーンでの発言の有無という形式で各ユーザーのデータを表す (例 user1: {あり, なし})
2. 類似度がある閾値以上のユーザー同士 (またはクラスタ同士) をクラスタリングする (N 個のクラスタができる)
3. 作成されたクラスタ間の距離が十分離れているかどうかについて乖離度を計算して評価する。
4. 全てのクラスタ間の乖離度が閾値以上かつ、クラスタ数が立場数より多い場合は 2 に戻る。
5. クラスタ間の乖離度が閾値以下のものがあるか、またはクラスタ数が立場数以下になった場合は終了する。(終了条件)

ユーザーのクラスタリングが終了した状態で、クラスタ数が立場数以下の場合、各クラスタがユーザーの立場に相当する。また、クラスタ数が立場数以下にならない場合は立場推定が行われなかったものとする。

3.5 同クラスタ / 同立場のユーザーの発言集合抽出 (Step V)

シーンのクラスタリング、ユーザーの立場推定を行った結果を用いることで、各シーンクラスタリング中での同じ立場のユーザーの発言を、同様の意味である集合として抽出する。

このときに、違う立場と推定されたユーザー同士が同じ発言をしているような場合は、その発言を抽出対象から外す。また、シーンクラスタでの立場ごとの発言集合の類似度が閾値以上であったら、シーンクラスタ全体を抽出対象から外す。このようにすることにより、ゲーム開始時など立場に関係なく多くのユーザーが発言するシーンにおいて、ノイズが混ざることある程度防ぐことができる。

3.6 Step ~ で抽出済みのデータからのさらなる情報抽出 (Optional)

例えば、サッカーゲームの観客であるユーザーが応援するチームはゲーム中一貫しているように、立場はシーンごとに大きく異ならないという仮定を置くことで、あるシーンで推定されたユーザーの立場を、それ以外のシーンにも適応することができる。そのとき、あるシーンで同じ立場だったユーザー同士は他のシーンでも同じ立場であるとする。こうすると、各ユーザーについて、各シーンクラスタでの立場を要素とした立場ベクトルを作成することが可能になり、この立場ベクトルを用いてユーザーをクラスタリングすることで、これまで立場を推定できなかったシーンでの立場を仮定し、発言集合を抽出することができる。

ただし、このアルゴリズムを用いることで、出力結果の精度が落ちる危険性があるため、閾値の設定などは、Step II~V に用いる物より慎重に設定する必要がある。

4 . まとめ

本稿では、本手法ではゲームなどの観客がゲーム中のイベントに対して発する短い発言を対象として、同様の外国語の発言と意味的に結びつけることで、翻訳を可能にするための知識を獲得する手法を提案した。

この手法を用いることで、現行のゲーム中のチャットなどで起こるユーザーが理解できない言語での発言は、そのユーザーが思っていることと同じ意味であってもほとんどの場合ノイズとして扱われてしまうという問題の解決を助け、観客に翻訳による違和感を抑えながら観戦しているゲームの臨場感や没入感を高めることができると考えられる。

今後の課題として、色々な国のユーザーが参加するオンラインゲームや、国際的なスポーツの大会を観戦しながら書き込むことで意見を交わす掲示板などから実際のデータを取得し、提案手法の有効性を示す必要がある。

参考文献

- [1] P Resnik, NA Smith. "The web as a parallel corpus," Computational Linguistics, " Vol. 29, No. 3, pp.349-380, 2003.
- [2] L Shi, C Niu, M Zhou, J Gao. "A DOM tree alignment model for mining parallel data from the web," Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL, pages 489{496, Sydney, July 2006.
- [3] 波多野 賢治, 亀井 俊之, 田中 克己, "多段階自己組織化マップによるビデオ映像記述支援と類似シーン検索," 情報処理学会論文誌, Vol. 39, No. 4, pp 933-942, 1998.