

災害時情報への質問応答システムの適用の試み

風間 淳一 Stijn De Saeger 鳥澤 健太郎 後藤 淳 István Varga
情報通信研究機構

概要

我々が開発している質問応答システム「一休」を、東日本大震災関連のツイートデータに適用した。我々の目標は「宮城県で孤立しているのはどこ?」といった質問に回答することである。今回、想定外の災害に一休を適用する場合を想定して、最小限のコストでどの程度のシステムを構築できるかを試み、問題点の洗い出しを行った。開発の過程で、場所の曖昧性解消、イベントの場所や時間の認識の無視できない重要性が明らかになった。例えば「福島」は全国に50以上もあり、適切なひとつを選択する必要がある。また、「仙台駅のアーケードの柱にあるコンセントを開放しています」から、直接的な係り受け関係がなくとも「仙台駅で開放している」ことを認識する必要がある。今回の試みでは、場所・地名の処理に注目し、単純な手法を実装して予備的な評価を行ったが、今後本格的な技術・資源の開発が必要である。災害対応に必要なこれらの技術を整理し、今後の展望を述べる。

1 はじめに

今回の東日本大震災では、Twitterをはじめとした情報システムの情報交換手段としての可能性が強く認識された一方で、大量に飛び交う情報から迅速、正確に状況を把握することは大変難しいという教訓を得た。救援活動や復興支援においても、組織や個人間での情報共有が十分に行われず、多くの無駄や問題が生じている。我々は、こうした問題を解決するため、これまで培ってきた情報分析技術を用いて、より適切な状況把握・判断を行うための情報を提供する情報配信基盤技術の開発を始めている。

今回の震災で、正確な状況把握を阻んだ本質的な原因は、大量の情報を整理し、必要な人に必要な情報を届ける手段が十分に準備できていなかったことである。例えば、Twitterで発信される安否確認、救援要請のツイートは、明確な宛先がなく、救援を行う行政等に有効に届かなかった。一方で、大量に飛び交うデマも含めた未整理の情報が振り回された人も多く出た。そうした中、Twitterでは、ハッシュタグの使用に見られるように、ユーザが自ら適応して情報整理の手段を編み出そうとする行動も見られた。また、「sinsai.info」「Google パーソンファインダー」等のように、安否情報等の整理を人海戦術で行うという試みも行われた。これらの活動をサポートするため、ボランティアの研究者らにより、自然言語処理を用いてTwitter上の安否情報を整理することを旨とした「ANPLNLP」の取り組みが行われたが、開発の速度や多数のボランティアの組織化には課題があったことが報告されている [4]。

本稿では、我々の取り組みの第一歩として、質問応答システムを東日本大震災関連のツイートデータに適用した試みについて述べる。「質問」で表されたユーザの情報ニーズに対し、通常の検索のように単なるドキュメントの列挙ではなく、「答え」を端的に列挙する質問応答システムは、上で述べた情報整理の問題を解決する有力な手段の一つであると考えられる。今回の震災で明らかになったのは、刻々と変わる災害の状況の中で、その都度、安否情報、支援物資情報等々と、システム

を構築するのでは、状況変化に十分に対応できないということである。この震災からの復興の過程では、想定しきれない問題領域が新たに出てくる可能性もある。また、我々が備えるべき次の災害が、今回の震災と同じである保証はない。トピックを限ることなく質問に回答をする質問応答システムは、臨機応変な対応が必要とされる災害時にも有効であると考えられる。

以上をふまえると、「今回の震災と同様の状況に対応できる」ことを条件としつつ、質問応答システムの実用性を改善していくことが必要である。今回、質問応答システムとしては、我々が開発している「一休」を使用し、これを、東日本大震災に関連するツイートに対して適用した。災害時には、情報の取りこぼしが無いように再現率を保ちつつ適合率を向上させる必要がある。しかし、一休の適用の過程で、再現率を低下させる次のような問題が明らかとなった。

- **場所の非明示性** 一休をそのまま適用したのでは、「宮城県で孤立するのはどこ」といった、震災時でも通常時にも重要であると考えられる場所を含む質問にほとんど回答できないことが分かった。「一休」は後述するように係り受けパターンを元に回答を抽出するが、ツイートに限らず、一般的に、イベントが起きた場所がイベントを表す動詞等に明示的には係らないことが多いためである。
- **場所の包含性** 場所には包含性がある。例えば、仙台市が宮城県の中にあることを正しく認識し、それを処理する手だてがなければ、文中に「仙台市」と記述されていても、「宮城県」の質問には回答できない。
- **場所の曖昧性** 一部の地名は非常に大きな曖昧性を持ち、上記の包含性を扱おうとする場合に、特に問題となる。例えば、「福島」という地名は日本全国に57もあり、そこから正しい一つを選ぶ必要がある。

本研究では、これらの問題に対して、簡単な対処方法を実装して、一休を適用した。予備的な評価実験で、それらの対処方法による回答の量の増加と精度の向上を確認し、これらの問題の本格的な解決が重要であることを示す。

2 質問応答システム「一休」

一休は、スマートフォンからの音声入力を備えた、ユーザからの多様なトピックの質問に意外な回答まで含めて回答することができる質問応答システムである。^{*1}

一休では、入力の問題文^{*2}は、2つの名詞スロットを持つ肯定文の係り受けパターン（入力パターン）に変換される。入力パターンと言い換えの関係にあるパターン（言い換えパターン）が、De Saeger et al. [1]を元にした方法であらかじめ列挙されてデータベースとして格納されており（**言い換えDB**）、また、元となる文書群^{*3}からは、2つの名詞が共起する係り

^{*1} <http://www2.nict.go.jp/x/x161/>に「一休」の紹介ビデオがある。

^{*2} 具体的には、名詞1つと疑問詞1つを係り受けで繋ぐ比較的単純な構造の問題文

^{*3} 通常の「一休」では、Web 6億ページの文書。今回の研究では東日本大震災関連のツイート。

都道府県 市区町村 町域 郵便番号
郵便番号データ: 宮城県/亶理郡山元町/坂元 989-2111

↓ 郡などを分割
住所: 宮城県/亶理郡/山元町/坂元
↓ 辞書エントリの生成

地名文字列(検索キー)	住所
(宮城(県))(亶理(郡))(山元(町))坂元	宮城県/亶理郡/山元町/坂元
(宮城(県))(亶理(郡))山元(町)	宮城県/亶理郡/山元町
(宮城(県))亶理(郡)	宮城県/亶理郡
宮城(県)	宮城県

地名文字列内の () は、省略可能であることを示す。したがって、省略する場合しない場合の全ての組み合わせが登録される。

図1 地名辞書の生成

受けパターンがあらかじめ列挙されている(係り受けDB)。入力パターン、および言い換えDBから得られる言い換えパターンを、質問中の名詞でスロットを埋めた状態で、上記の係り受けDBにマッチさせ、空いているスロットを埋める名詞を抽出して、それを回答とする。回答のランキングは、上記の言い換え時のスコアに基づく。例えば、「宮城県で孤立するのはどこ」という質問文は、「Xで孤立するのはY」という入力パターンとなる^{*4}。その言い換えパターンには、例えば「XでYが孤立する」「XでYから動けない」などがある。係り受けDBに、「宮城県でA小学校が孤立する」「宮城県でB小学校から動けない」といったデータがあるとすると、「A小学校」「B小学校」が回答として抽出される。この例から分かる通り、一休は、言い換え認識を用いて様々な表現の差を吸収して回答を抽出することができる。

一休は、元々、億ページ単位のWebデータ中の複数の冗長な言及を集約して回答を抽出することを前提としたシステムである。今回利用したツイートは次節で述べるようにWebに比べて量が非常に少ないため、質問応答エンジンが上手く働かず再現率が低く抑えられてしまう可能性もある。我々は、既に1文中に回答がない場合でも機械学習や推論を用いて回答を抽出する技術[2][5]や、コアとなる言い換え・含意獲得の精度を向上させる技術[3]の開発も行っており、これらは上記の再現率の問題の軽減に寄与すると期待される。また、質問の形式も、上述の通り、現在はかなり限定されたものになっているが、現在複雑な質問文に対応するための技術の開発も行っている。今回は、これらの成果を取り込んだ新エンジンの利用は間に合わなかったが、今後、新エンジンを用いた設定での評価を行う予定である。

3 災害時情報への適用

3.1 ツイートデータ

ハッシュタグや含まれる単語から判断して東日本大震災に関連すると思われる約320万ツイート(3月11日から6月17日)を用いた。ツイートは、正規化、文への分割を行い、Jumanを用いて形態素解析を行った。

3.2 地名・場所辞書の作成

まず、日本郵便が公開している郵便番号データを利用して、地名辞書を作成した。郵便番号データからは、「都道府県/市区町村/町域」で表される住所の情報を取り出すことができる。そこから、可能な地名文字列から住所へのマッピングを

^{*4} 正確には、質問文からはいくつかのルールにより複数の入力パターンが生成される。例えば、これに加えて、「YはXで孤立する」なども生成される。

福島 → 福島県/福島県/福島市,三重県/桑名市/福島,兵庫県/三田市/福島,岡山県/倉敷市/福島,長野県/伊那市/福島,大分県/中津市/福島, ...
岐阜県/揖斐郡/揖斐川町/福島,京都府/京都市/上京区/福島町,青森県/南津軽郡/藤崎町/福島,福島県/会津若松市/河東町/福島

図2 「福島」に対する地名辞書エントリ

取り出した。ここで、可能な地名文字列とは、その住所を表し得る文字列であり、住所から簡単な方法により生成した(図1)。多少過生成となるが、場所の認識のカバレッジを上げるため、この方法を用いた。「都道府県/市区町村/町域」という住所の階層性は、先に挙げた場所の包含性に対処するための情報源となる。このようにして、2,486,545のエントリを持つ辞書(地名辞書)を得た(地名文字列-住所の対の数は5,129,162)。そのうち、84,633エントリが曖昧性のある地名だった。曖昧性がある場合の平均の曖昧性は32.2、曖昧性の最大は366であった。例えば、「福島」の曖昧性は57であった。図2に、「福島」に対する辞書エントリの一部を示す。

次に、上位下位関係抽出ツール^{*5}を用いてWikipediaから抽出した上位下位関係から、下位語が場所らしいものを取り出して利用した。これは、「学校」などの、郵便番号データには載っていないような場所にも住所へのマッピングを生成するためである。上位語が「(自治体名)の(*X)」というパターンにマッチする(Xは「施設」「学校」など)場合に、下位語が場所らしいと判断した。上位語中の(自治体名)を、地名辞書で検索して下位語に住所を付与する^{*6}。最終的に、255,273エントリを持つ辞書(場所辞書)を得た。

地名辞書と場所辞書をマージし、地名・場所辞書とした。辞書はトライ構築のライブラリであるmarisa-trieを利用して接頭辞検索ができるようにした。

3.3 地名・場所の曖昧性解消

地名・場所辞書は曖昧性が大きく、正しく曖昧性解消できなければ、場所に関する正しい回答の抽出は不可能である。将来的には、本格的な開発が必要であるが、本研究では、以下の単純な方法を用いた。情報が無ければ最も広範囲な地域を表す住所、直前に曖昧性解消された住所がある場合には、それと最も整合性のある住所が選ばれることを目指した手法である。

- **住所候補生成:** ツイートの各文に対し地名・場所辞書で最左・最長一致検索を行い、名詞句に住所候補を付与する。その際、名詞句全体がマッチしない場合でも、その範囲内で最左のマッチを選び、できるだけ住所を付与する。1文字の地名・場所は誤ったマッチである可能性が大きいいため、無視する。また、住所は、青森、岩手、宮城、福島、栃木、千葉の各県に限定する^{*7}。また、固有表現認識器をさらに用いて不要なマッチをさらに取り除くことも試みた。マッチした地名・場所が、固有表現認識で「場所」でも「組織」でもない場合にのみ取り除く。
- **曖昧性解消** ツイートの先頭から住所候補に対して以下の処理を行う。
 - 直前に曖昧性解消された住所がない場合: 候補のうち、県・郡・市(郡部の場合は町)部分が検索キーと一致するものを、この優先度で選ぶ。つまり、より

^{*5} 高度言語情報融合フォーラム ALAGIN (<http://alaginrc.nict.go.jp>) から入手できる。

^{*6} 他のヒューリスティクスも用いたが、紙面の都合上、省略する。

^{*7} 曖昧性解消技術が十分ではない場合、このような割り切りも必要であると考えられる。

```

final_generated = []
ツイート中の各文 s に対して:
generated = [""], bare = ""
s 中の各単位 u (名詞句又は語) に対して:
generated_next = [], bare += u
if uが地名・場所である名詞句:
generated の各 g に対して:
u の可能な文字列 u' (uも含む) に対して:
generated_next << g + u'
else if uがイベントを表す動詞等:
generated の各 g に対して:
generated_next << g + u
直前の場所 p の可能な文字列 p' (pも含む) に対して:
generated_next << g + p' + u
else:
generated の各 g に対して:
generated_next << g + u
generated = generated_next
if generated.size > 100: (生成の爆発を防ぐため)
generated の各 g に対して:
final_generated << g + (sの残りの部分)
generated = [bare]
if generated.size > 0: final_generated += generated

```

u の可能な文字列は、生成の爆発を防ぐため県・郡・市・町部を単に切り出したものとする
例)「宮城県/気仙沼市」の場合、「宮城県」「気仙沼市」のみ

図3 補完文の生成方法

広い地域レベルで検索キーと一致しているものを優先する。例えば、「福島」の場合には、最も広範囲な「福島県」が選択される。該当するものが無ければ、列挙された最初の候補を選ぶ。候補は文字列の長さで昇順にソートしてあるので、近似的ではあるが、できるだけ広い地域レベルでの候補が選択される。

- **直前に曖昧性解消された住所がある場合:** 直前で曖昧性解消された住所と県・郡・市(町)まで一致するものがあればそれを選択する。無ければ、県・郡まで一致するものを選択する。それも無ければ、県まで一致するものを選択する。該当が無ければ、曖昧性解消された住所がない場合と同じように候補を選ぶ。

3.4 イベントの場所の認識と補完文の生成

場所に関する質問に適切に答えるためには、地名・場所の曖昧性解消を行った後、文中の動詞などで表される各イベントが起きた/起きる場所を認識し、一体の係り受け DB の作成に反映させる必要がある。また、この時、場所の包含性も考慮する必要がある。ここでは、イベントの場所は直前に出現した地名・場所であるという仮定を置き、図3で示す方法を用いて、直前の地名・場所の可能な文字列に助詞「で」を加えた文字列を、イベントを表す動詞等の直前に補完して新たな文(補完文)を生成することとした。例えば、「気仙沼中学校へ避難しています」という文があった場合、辞書により「気仙沼中学校→宮城県/気仙沼市」が分かるとすると、「気仙沼中学校へ気仙沼中学校で避難している」「気仙沼中学校へ気仙沼市で避難している」「気仙沼中学校へ宮城県で避難している」という補完文が生成される。元の文と補完文は、係り受け解析器^{*8}で解析して、一体の係り受け DB 生成モジュールへ入力する^{*9}。補完を行うことにより、文の総数は、7,899,564 (約790 万文) から 168,076,416 (約 1.7 億文) に増加した。固有表現認識を考慮した場合は、111,231,829 (約 1.1 億文) に増加した。

曖昧性の解消、補完文の生成ともに、簡単なルールベースの

^{*8} <http://www.tkl.iis.u-tokyo.ac.jp/~ynaga/jdepp/> で公開されている高速な日本語係り受け解析器 J.DepP を使用する。
^{*9} 言い換えパターン DB は、通常の一休が使用している Web 6 億ページの文から作成したものを用いる。

質問: 宮城県で孤立するのはどこですか

回答例	抽出部分(抜粋)
グランドホテル	グランドホテルが孤立している可能性もあります
介護保険施設	仙台市若林区の津波被害で孤立した介護保険施設などを、
気仙沼市本吉町大谷地区	気仙沼市本吉町大谷地区が孤立しています！
宮城県南内陸部	情報です！宮城県南内陸部が孤立している。
宮城県花山村 東六郷小学校	宮城県花山村が孤立しているため 東六郷小学校が孤立したと聞いたのですが...
牡鹿半島	牡鹿半島が孤立している模様
石巻好文館	孤立した石巻好文館に避難者1600名が居るそうです。
鹿折唐桑地区	トンネルがつぶれていて鹿折唐桑地区が孤立している模様です。

表2 補完あり (NER) の場合の回答例

手法であるが、ツイート中の前方の文まで考慮しているため、ある程度グローバルな処理とすることができる。

4 評価

提案手法が、東日本大震災時の場所に関する質問に対してどの程度効果があるかを、表1に挙げる宮城県に関する32個の質問を用いて調べた。回答の精度は、各手法の回答上位50個(50個無い場合には回答数)の回答を(公平性のため手法が分からないようにマージし)、著者の一人が正解かどうかを判定して計算した。その際、回答が抽出された元のツイートを見て、実際に宮城県で起きている・起きた・起きる(またその可能性がある)という場合には正解とし、質問の回答として適当ではない場合、宮城県とは関係ない場合、イベントが否定されている場合などは誤りとした。

結果を表3に示す。表中 Pre は適合率、AP は正答の数を50と仮定したときの平均適合率である。補完を行わない場合には質問にほとんど回答できないが、補完を行うことにより、回答数が大幅に増加していることが分かる。補完を行わない場合、32個の質問に対する回答は合計で215であったが、補完を行う場合は、5,916であり、補完が非常に重要であることが分かる。補完は完璧ではないため、適合率は低下することが懸念されたが、質問(5)以外は、補完によってむしろ適合率が向上している。固有表現認識を利用した場合(表中「補完あり(NER)」)、その性質上回答数は多少減少する。適合率・平均適合率が上昇した質問もあれば低下した質問もあったが、適合率の平均は0.17から0.14へと低下した。平均適合率の平均は0.045から0.048へと上昇している。固有表現認識の有効性の確認にはさらに評価が必要であると考えられる。

表2に、補完あり(NER)の場合の「宮城県で孤立するのはどこですか」という質問に対する回答(正解と判定されたもの)と、抽出元のツイートの該当部分をいくつか示す。「Xが孤立している」「孤立したX」など質問文とは異なるパターンでの柔軟なマッチングにより回答が得られていることが分かる。さらに、例えば、回答中の「グランドホテル」は「孤立」に関係するツイートデータ中で1回も宮城県と共起しておらず、本システムでなければ、発見できなかった回答である。

- (1) 宮城県で孤立するのはどこですか (2) 宮城県でどこで充電できますか (3) 宮城県で何を提供していますか
 (4) 宮城県で何が発生していますか (5) 宮城県で何が不足していますか (6) 宮城県でどこが食料を提供していますか
 (7) 宮城県でどこに避難していますか (8) 宮城県の避難所はどこですか (9) 宮城県の通行止めはどこですか
 (10) 宮城県でどこが渋滞していますか (11) 宮城県でどこが安全ですか (12) 宮城県でどこが音信不通ですか
 (13) 宮城県で誰が行方不明ですか (14) 宮城県で誰が死亡しましたか (15) 宮城県で何が崩壊していますか
 (16) 宮城県で被害を受けているのはどこですか (17) 宮城県で被害を受けたのは何ですか (18) 宮城県で津波が襲ったのはどこですか
 (19) 宮城県の震度は何 (20) 宮城県で火災が発生しているのはどこ? (21) 宮城県でどこが停電していますか
 (22) 宮城県でガスが使えないのはどこですか (23) 宮城県で断水しているのはどこですか (24) 宮城県でネットが使えるのはどこですか
 (25) 宮城県で機能している病院はどこですか (26) 宮城県で自衛隊が入ったのはどこですか (27) 宮城県に行ける道はどれですか
 (28) 宮城県で発生している病気は何ですか (29) 宮城県で自衛隊の連絡先はどこですか (30) 宮城県で支援物資はどこに届けますか
 (31) 宮城県で誰が支援物資を届けていますか (32) 宮城県で汚染されているのは何ですか

表1 評価で使用した宮城県に関する質問

質問	補完なし				補完あり				補完あり (NER)			
	回答	正答	Pre	AP	回答	正答	Pre	AP	回答	正答	Pre	AP
(1)*	0	0/0	0.0	0.0	96	17/50	0.34	0.12	76	19/50	0.38	0.19
(2)*	0	0/0	0.0	0.0	8	2/8	0.25	0.017	8	2/8	0.25	0.017
(3)*	31	10/31	0.32	0.11	999	22/50	0.44	0.32	999	24/50	0.48	0.35
(4)*	70	5/50	0.1	0.068	999	20/50	0.4	0.18	999	20/50	0.4	0.18
(5)*	1	1/1	1.0	0.02	74	25/50	0.50	0.23	74	26/50	0.52	0.25
(7)*	0	0/0	0.0	0.0	9	7/9	0.78	0.10	11	7/11	0.64	0.078
(8)	0	0/0	0.0	0.0	38	8/38	0.21	0.041	36	7/36	0.19	0.039
(10)*	0	0/0	0.0	0.0	7	5/7	0.71	0.056	7	5/7	0.71	0.056
(11)	0	0/0	0.0	0.0	1	0/1	0.0	0.0	1	0/1	0.0	0.0
(14)	25	0/25	0.0	0.0	980	1/50	0.02	0.00087	861	1/50	0.02	0.00074
(15)*	0	0/0	0.0	0.0	1	1/1	1.0	0.02	0	0/0	0.0	0.0
(16)*	0	0/0	0.0	0.0	24	6/24	0.25	0.060	23	6/23	0.261	0.060
(17)*	4	0/4	0.0	0.0	227	16/50	0.32	0.119	202	17/50	0.34	0.16
(19)	15	3/15	0.2	0.055	482	10/50	0.2	0.15	464	10/50	0.2	0.153
(24)	28	0/28	0.0	0.0	999	0/50	0.0	0.0	999	0/50	0.0	0.0
(27)	41	0/41	0.0	0.0	999	2/50	0.04	0.012	999	2/50	0.04	0.012
(32)	0	0/0	0.0	0.0	18	1/18	0.056	0.005	14	1/14	0.071	0.005
計/平均	215	19/195	0.051	0.0080	5916	143/557	0.17	0.045	5773	147/550	0.14	0.048
(10)	106	16/86	0.14	0.020	2444	121/299	0.50	0.12	2399	126/299	0.40	0.13

表3 宮城県に関するいくつかの質問に対する評価。全ての手法で回答数が0のものは省いて表示している。また、最後の行は一体が得意とする(補完ありで適合率が高い)上位10個の質問(*で示した)に関する計/平均である。

5 今後の課題とまとめ

今回、災害時の情報配信基盤技術の開発の第一歩として、質問応答システム「一休」を東日本大震災時のツイートデータに適用することを試みた。今回の適用手法の単純さを考えれば、期待の持てる結果だと考えるが、実用とするにはさらに改良が必要である。今後、次に挙げるような点で改良を行っていく予定である。

- 2節で述べた通り、使用した一休は最新のものではない。今後最新のエンジンを使用してより詳細な評価を行う。
- 場所に関して今回提案した手法は、簡単なルールベースの方法であり、機械学習などを利用したより高精度の手法を開発する。また、「時間」にも、曖昧性以外は場所とほぼ同様の問題があり、それらに対処を行う。
- 場所や時間は、従来の照応解析・項構造解析ではあまり積極的には扱われてはいないが、従来の意味での照応解析、項構造解析を用いて動詞に係る項を補完することでも再現実率の向上が期待できるため、これらの適用を行う。
- ツイートに対する係り受け解析の精度を向上させる。特に、現在の解析では、助詞などの省略により精度が低下しやすいため、これに対処する。また、補完文の生成により不自然な単語列となり解析が誤る場合が多く、また処理

する文数が大幅に増加してしまうため、より直接的に係り受け構造を操作して補完を行うことを検討している。

謝辞 ツイートデータについて助言をいただいた榎剛史氏、また、J.DepP に関して助言をいただいた吉永直樹氏に感謝いたします。

参考文献

- [1] Stijn De Saeger, Kentaro Torisawa, Jun'ichi Kazama, Kow Kuroda, and Masaki Murata. Large scale relation acquisition using class dependent patterns. In *Proceedings of ICDM'09*, 2009.
- [2] Stijn De Saeger, Kentaro Torisawa, Masaaki Tsuchida, Jun'ichi Kazama, Chikara Hashimoto, Ichiro Yamada, Jong Hoon Oh, Istvan Varga, and Yulan Yan. Relation acquisition using word classes and partial patterns. In *Proceedings of EMNLP 2011*, 2011.
- [3] Julien Kloetzer, Stijn De Saeger, Kentaro Torisawa, Motoki Sano, Jun Goto, Chikara Hashimoto, and Jong Hoon Oh. Supervised recognition of entailment between patterns. 言語処理学会第18回年次大会, 2012.
- [4] Graham Neubig, Yuichiroh Matsubayashi, Masato Hagiwara, and Koji Murakami. Safety information mining - what can NLP do in a disaster -. In *Proceedings of IJCNLP 2011*, 2011.
- [5] Masaaki Tsuchida, Kentaro Torisawa, Stijn De Saeger, Jong Hoon Oh, Jun'ichi Kazama, Chikara Hashimoto, and Hayato Ohwada. Toward finding semantic relations not written in a single sentence: An inference method using auto-discovered rules. In *Proceedings of IJCNLP 2011*, 2011.