

論文の内容の要旨

ソーシャルタギングからの ことばが指し示す実世界対象の表現獲得

氏名 馬場 雪乃

本論文では、ことばが指し示す現実世界の対象物（実世界対象）を、実世界データを用いて表現するという問題に取り組んだ。ここでは、カメラやセンサーなどのデバイスを通じて現実世界から直接取得できるデータのことを実世界データと呼ぶ。特に本論文では、実世界データのうち写真（視覚データ）と場所（地理データ）をその代表例として取り上げ、「ことばが指し示す写真」、すなわちあることばの指示対象（例、「犬」）が写っている写真を獲得するという課題と、「ことばが指し示す場所」、すなわちあることばが指し示す場所を地理的な領域として獲得するという課題に取り組んだ。

この取り組みは、以下の三つの点で意義があると考えられる。第一に、これまで主に他のことばとの関係によって表現されていたことばの意味に対して、実世界データという別の側面からの解釈を与えることである。これにより、ことばの意味をより多角的に捉えることが可能となり、たとえばこれまでと異なる観点からことばの類似性を測ることで“mug”と“milk”の混同のような問題の解決に貢献すると考えられる。第二に、現実世界の情報と自然言語で表された情報とを意味的に関連づけられることである。たとえば、現実世界で見つけた名前のわからないものについての情報を検索する際などに、ことばの実世界データによる表現は利用することができるだろう。第三に、膨大な実世界データをことばを軸として構造化することである。人間とコンピュータのインタラクションの多くは自然言語によって行われており、実世界に関する情報をことばを介して効率的に検索・取得するためにこの構造化は意義があると考えられる。

本論文では、「ことばが指し示す写真」「ことばが指し示す場所」の獲得を行う機械的な手法を提案した。これにより、訓練データ作成などの人手を掛けずにことばが指し示す実世界対象を獲得できる。機械的に獲得するためには、ことばと写真、ことばと位置情報のデータ対が必要となる。このデータ対を取得するために、本研究ではソーシャルタギングと呼ばれる、World Wide Web (Web)上でコンテンツに対してタグ（コンテンツを説明するキーワード）

を付与する仕組みに着目した。ソーシャルタギングによって生成されたデータを用いて機械的に、ことばの指示対象を獲得することを目指した。ソーシャルタギングデータを利用する上では、ことばの曖昧性とノイズタグの問題がある。ことばの曖昧性については、曖昧性が解消された「ラベル」を用いて、そのラベルが指し示す対象を獲得するという方法と、ことばの曖昧性を許容し、指示対象を曖昧さを含んだ確率表現として獲得する方法をそれぞれ採用し、その実現手法を提案した。ノイズタグの問題に対しては、一つのコンテンツに与えられた複数のタグづけを統合したり、同じタグが与えられた複数のコンテンツを利用したり、外部データを利用するといった解決方法を提案をした。

本研究の貢献は以下となる。写真については、(1)曖昧性が解消されたことばである WordNet 上の名詞概念に対して、機械的に写真を割り当てる手法を提案した。(2)この手法により、人手を用いずに WordNet 概念が指し示す写真を収集し、また大量の写真を WordNet 概念を軸として構造化できるようになった。(3)特に、収集した写真は画像認識のための訓練データとして用いても効果的であることを評価実験で示した。場所については、(1)「特定の場所と潜在的に関連があることば」に対して地理的な領域を機械的に割り当てる手法を提案した。(2)割り当てられた場所が人間の認識に対してある程度合致することを評価実験で確認した。獲得した「ことばが指し示す場所」は、テキストデータに対する地理情報の自動アノテーションや、ことば同士の地理的な類似度計算・上位下位関係の獲得などに役立つことが期待される。