

多義語をちゃんと区別できるとコンピュータは賢くなるのか？

# マルチセンス分散表現を用いた質問応答システムの研究 Applying Multi-Sense Embedding to Question Answering

韓南琦, 宮尾祐介

## どんな研究？

- 世界中には表記や発音が同じでも、意味は全く違う言葉がたくさんある
- しかし人間は日常会話の中でうまくそれらの意味を分けて理解している
- コンピュータにも同じように多義語が区別できるようにしたい

## 何がわかる？

- 言葉の意味ごとの分散表現を計算するために必要なアルゴリズム
- コンピュータにその分散表現をうまく学習させるためのパイプライン
- 多義語をうまく区別する質問応答システム

人間は話を聞くだけでも、多義語の意味をちゃんと分けて理解できる

Q : How much did **Mercury** spend on advertising in 1993? (Wangら, 2007)



- 人間は辞書を引かなくても、一つの言葉の色々な意味をうまく区別できる
- コンピュータもこんな風に多義語の理解ができると、今より賢くなるのでは？

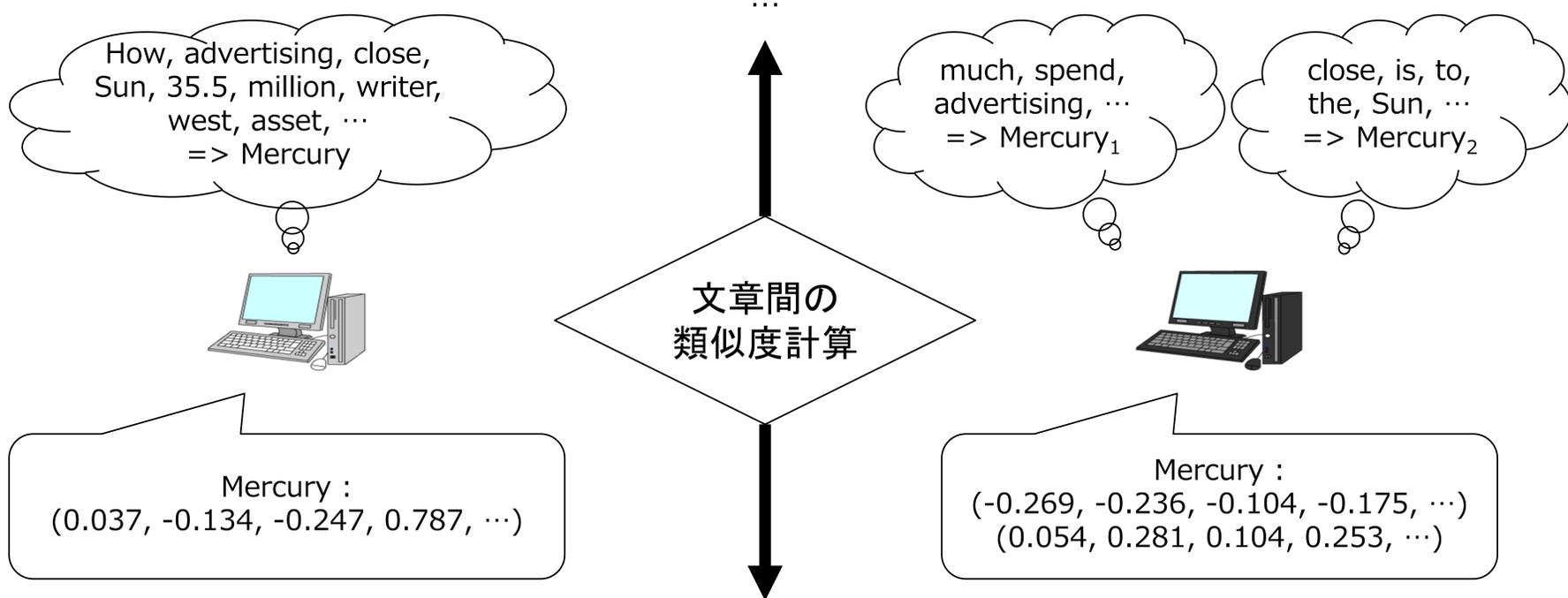
## 言葉の意味ごとの分散表現を計算し、学習する質問応答システム

同じ文脈で使われている言葉たちは、類似した意味を持つ傾向がある。(Harris, 1954)

既存の手法 (word2vec, GloVe, ...)

新しいマルチセンス分散表現の手法

問題 1 . How much did **Mercury** spend on advertising in 1993?  
問題 2 . How close is **Mercury** to the Sun?



- 答え候補 1 . **Mercury** is 35.5 million miles from the sun and completes one orbit in 88 days.  
答え候補 2 . Fuladi is a former staff writer for the **Mercury** news' west magazine.  
答え候補 3 . Last year the company spent pounds 12m on advertising, **Mercury** asset management has turned to radio to increase brand awareness.