

テキスト含意認識に有用な概念意味情報

横手 健一[†] 田中 翔平, ボレガラ ダヌシカ, 石塚 満[‡]

東京大学工学部電子情報工学科[†] 東京大学大学院情報理工学系研究科[‡]

1. はじめに

テキスト含意認識 (Recognizing Textual Entailment) とは、近年積極的に取り組まれている自然言語処理のタスクの一つである。これは、二つの文、T(Text)とH(Hypothesis)が与えられた時に、TがHを含意するか(TからHを推論可能かどうか)を判定する。このような言語の意味を推論する技術が確立すれば、より「知的な」テキスト処理が可能になるので、様々な分野への応用可能性が考えられている。

外部の言語資源、及びそれらに基づいたツールを併用して RTE システムを構築するのが一般的であるが、その中の一つの選択肢として、単語の意味的類似度を特徴量として学習しようというアプローチがある。これは、言語の意味的類似性に関する研究が進んでいて、その計測のためのツールが容易に手に入ることや、テキスト間の含意関係度と意味的類似度が完全には無関係ではないことが理由になっていると考えられる。しかし、両者は同義ではないため、含意関係の判定に利用する情報としては必ずしも適切ではなく、学習の過程でこの差異に基づくエラーが生じていると考えられる。

そこで本研究では、まず、意味的類似度から「含意関係度」を導く非線形変換を行い、その変換結果を特徴量として学習するという手法をとることで、判定精度を高めることを目指す。上記の変換関数の学習は、二つのテキスト対と、その含意関係の有無を二値で示しただけの一般的な訓練データでは適切でなく、独自の特徴量が必要である。一般的な訓練データから、その独自の特徴量を得るための抽出手法も、合わせて提案する。

2. システムに利用する言語情報

2.1 概念辞書に基づく単語間意味的類似度

WordNet に代表される概念辞書は、単語が

synset と呼ばれる同義語のグループに分類され、簡単な定義や、他の同義語のグループとの関係が記述されている。この情報に基づいて、単語間の意味的な類似度を定量的に計測する手法が多数提案されている。本研究でもその中のいくつかを採用し、システム内部では以下の様に 6 次元の得点ベクトルという形で表現する。

$$[\text{Term's Score vector}] : \mathbf{t} = (s_1, s_2, s_3, s_4, s_5, s_6)^T$$

ベクトルの要素と、計測手法の対応関係は、以下の通りである。

- 1: Jcn (Jiang-Conrath) Similarity
- 2: Path Similarity
- 3: Res (Resnik) Similarity
- 4: Lch (Leacock-Chodorow) Similarity
- 5: Lin Similarity
- 6: Wup (Wu-Palmer) Similarity

2.2 文脈に基づく単語の重要度

特定のテキストに対する含意関係度とは無関係に、テキスト中の単語が相対的に持つ重要度について考察する。

以下の例文で考える。

He teaches children about computers.

ここで、children と computers はどちらが情報として重要であろうか？

もし He が教師であることが既知であれば、teaches children は教師の仕事なので特に重要ではなく、computer を教えたという事実の方が情報として重要だと言える。他方、He が情報技術者であった場合、computer について教える事は重要ではなく、children を対象に教えたことが重要な情報ということになるだろう。

即ち、初めの問いには絶対的、普遍的な解があるわけではなく、テキストの前後の文脈に大きく依存する問題である。この様な、文脈に依存した単語の相対的な「重み」は、RTE の判定を行う際に有益な情報である。

3. 変換関数の決定とそのための特徴量抽出

意味的類似度から含意関係度を導く非線形変換は、具体的には特定の単語 x とテキスト T を入力として

Useful Conceptual Semantic Information for Textual Entailment Recognition

[†] K. Yokote, Univ. of Tokyo

[‡] S. Tanaka, D. Bollegala, M. Ishizuka, Univ. of Tokyo

$s_i = C_i(\text{semantic_score}_i(x, T))$
の様定義する。

また、 $\text{semantic_score}_1 \sim \text{semantic_score}_6$ 全てにおいて、異なる意味の類似度の計測方法を用いるものとする。

$C_1 \sim C_6$ を決定するためには、関数の各入力値に対して、出力となる目標値の設定が必要である。しかしながら、この値を RTE の標準的な訓練データから獲得するのは容易ではない。なぜなら、必要な情報は単語の T に対する含意関係度であるが、訓練データは二つのテキスト H、T の内容と、その含意関係の有無を示す二値のラベルしか持っていないからである。

テキスト含意というものは、意味的、構文的関係が複雑に作用して引き起こされる現象であるので、二つのテキスト間に含意関係があったとしても、その中のどの実体（単語）がどのような影響を及ぼしていたかの解析は、とても難しい問題である。本研究では、これをベースラインシステムを利用した探索と、ルールベースの特徴抽出アルゴリズムによって解いた。

1：一つの訓練データで注目する H の実体（単語）は一つとし、それは T に対する lexical matching を満たさなかった中で $\text{semantic_score}_i(x, T)$ が最大の単語 x のみとする。これは、複数の単語を考慮しようとする場合、その信頼度割り当てが難しいことと、後述する試行実験の総数が指数的に増大することを防ぐためである。

2：前述した x について、 $C_i(\text{semantic_score}_i(x, T))$ に 0 あるいは 1 を与えて、判定の変化を確認する。

この試行実験を行うことで、 $\text{semantic_score}_i(x, T)$ は、次の三つのクラスに分類できる。

【POSITIVE】： $C_i(\text{semantic_score}_i(x, T))$ を 1 とみなすと、システムは正しく含意判定を行うが、0 とみなすと、間違っただ判定を行う。

【NEGATIVE】： $C_i(\text{semantic_score}_i(x, T))$ を 1 とみなすと、システムは間違っただ含意判定を行うが、0 とみなすと、正しく判定を行う。

【ANY】： $C_i(\text{semantic_score}_i(x, T))$ にどちらの値を与えても、システムの判定能力は変化しない。（正しい判定は、正しいままで、間違っただ判定は、間違っただまま。）

3：1～2を全ての訓練データについて行う。このようにして分類されたクラスの中で、 C_i の決定に重要な情報を持つのは【POSITIVE】と

【NEGATIVE】の二つである。

そこで、【POSITIVE】に含まれる全ての $\text{semantic_score}_i(x, T)$ に対して、 $C_i(\text{semantic_score}_i(x, T))$ の目標値として 1 の特徴量を与え、一方で【NEGATIVE】には 0 を与える。

このようにして、入力(semantic score)と出力(1 or 0)の二値を持った訓練データを得る。これを用いれば、あとは一般的な線形回帰モデルを設定して、機械学習によって関数 C_i の決定が可能である。

この手法は、「システムの含意判定能力が改善されるかどうか」で目標値を定めているため、意味的類似度から含意関係度を導く変換関数の決定手段として適切だと言える。

4. 実験結果

得点ベクトル t の要素	意味的類似度を直に与えた場合の精度	変換結果を与えた場合の精度
s1	40.250	52.170
s2	34.682	49.738
s3	44.866	51.948
s4	21.294	51.752
s5	44.376	51.994
s6	23.362	44.638

(精度は、RTE-7 開発セットの Micro F 値[%]である)

上記表の様に、全ての要素について明確な上昇があった。また、このシステムを TAC RTE-7 (2011) に提出し、世界第二位の精度を記録した [yokote 11]。

TAC RTE-7 (2011) Result

```
(micro_average)
precision="46.84"
recall="43.58"
f-measure="45.15"
```

謝辞

本研究について多数のアドバイスを頂いた石川開様、土田正明様はじめ NEC 情報・メディアプロセッシング研究所の皆様に感謝致します。

参考文献

[yokote 11] K.Yokote, S.Tanaka and M.Ishizuka: **Effects of Using Simple Semantic Similarity on Textual Entailment Recognition**, RTE-7(Recognizing Textual Entailment) Workshop Note at Text Analysis Conf.(TAC2011), NIST, Gaithersburg, Maryland, USA (2011.11)