

文間関係認識のための構造的アライメント

後藤隼人[†] 水野淳太[†] 村上浩司[†] 乾健太郎^{†‡} 松本裕治[†]

奈良先端科学技術大学院大学[†] 東北大学[‡]

{hayato-g, junta-m, kmurakami, inui, matsu}@is.naist.jp

1 はじめに

一対の文が与えられたときに、一方の文を他方の文が内容的に含意または推論できるかどうかを同定することは、含意関係認識 (RTE) と呼ばれ、情報検索や質問応答などの情報アクセス技術にとって重要である。Pascal RTE Challenge[5]においては、含意、矛盾、不明 (Unknown) の3種類の関係に分類する課題として定義されている [6]。また、より多様な関係を扱う研究として、Cross-document Structure Theory(CST)[15]や、言論マップ [13] などが挙げられる。これらは文間関係認識と呼ぶことができ、RTE はその部分課題に位置づけることができる。

文間関係認識には様々なアプローチが提案されているが、アライメントを行うかどうかで大きく分けることができる。アライメントを行わない手法としては、既存の語彙知識を用いて言い換え可能かどうかを判定する手法 [4] や、文間で共通して用いられている単語を素性として機械学習を行う手法 [16] が挙げられる。しかし、これらは含意かどうかを判別する2値分類の問題となるため、より多様な関係を扱うことは難しい。

一方アライメントを行う手法の大まかな流れは、以下の通りである。

1. 解析 形態素解析、構文解析といった基本的な解析
2. アライメント 文間で対応する単語間の対応付け
3. 関係分類 これらの結果から文間の関係を判別

文間関係認識におけるアライメントは、文間のどの単語対を比較すれば文間の関係を分類できるのかを明らかにすることが目的であり、語彙知識や文構造に着目した手法がいくつか提案されている [10][11]。先行研究の多くは単語アライメントの情報を関係分類のための素性の一つに位置づけており、文構造や事実性といった情報と併せて関係分類を行う。しかし後述の通り、これは必ずしも得策ではないと考えられる。

そこで本研究では、単語同士のアライメントに加えて、単語間の依存関係の対応付けをとる処理をアライメントの中で行う方式を考える。これを局所構造アライメントと呼び、明示的に関係分類から切り離す。本稿では局所構造アライメントの詳細および予備実験の結果について報告する。

2 アライメントと関係分類

2.1 従来の単語アライメントの問題点

アライメントを用いた文間関係認識の先行研究の多くにおいて、アライメントとは文間で意味的に対応す

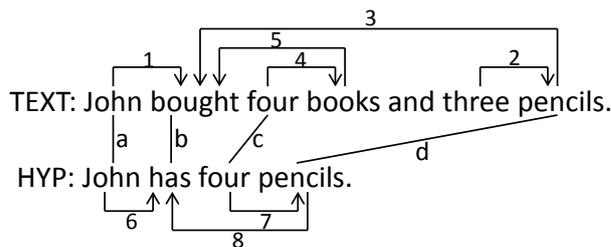


図 1: 単語アライメント

る単語間に対応付けをとることを指している。例えば図 1 における単語アライメントは a~d の単語間の対応付けを行うことである。しかし、この例における HYP 側の four と pencils の意味的・統語的關係は、それぞれの単語が対応する TEXT 側では成り立っていない。この問題に対し、Sammons ら [17] は、関係分類の段階で意味役割や共参照といった情報を加えることで対応することを試みている。しかし、これは以下の2点により得策ではないと考えられる。

1. 関係分類は、事実性や共参照といった多くの情報を取り扱う複雑な工程であり [8]、それに対しさらに単語間の依存関係の間の対応付けという操作を加えることは、問題をより複雑化している。
2. 単語間の意味的・統語的關係をとらえるのは、意味役割付与や共参照解析などに相当し、それぞれが一つの研究として成立している。それらを関係分類から切り離すことで、モジュール性の面で有利な可能性がある。

本研究では、単語アライメント間の構造的な対応付けを関係分類から分離し、局所構造アライメントとして明示的に行うことで、関係認識の複雑さを軽減する。

2.2 提案：局所構造アライメント

図 1 において、HYP 側の単語間の関係が TEXT 側で対応する単語間にもそれに対応しているかどうかを判別することが重要である。HYP の 6 の依存関係が表す意味的關係が、TEXT 側の対応する単語間でも成り立っている。一方で HYP の 7 の依存関係が表す意味的關係は、TEXT 側の対応する単語間で成り立っていない。これを区別する問題を考える。

より一般的には図 2 において、まず HYP における w_i, w_j はそれぞれ TEXT における $A(w_i), A(w_j)$ に単語アライメントされるとする。ここで、 w_i, w_j 間の意味的關係が $A(w_i), A(w_j)$ 間でも成り立っているかどうかを判別する問題を考え、これを局所構造アライメン

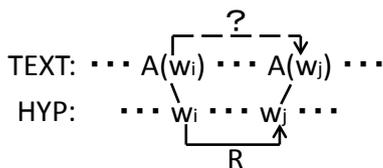


図 2: 局所構造アライメント

トと呼ぶことにする。

以下に図 1 における局所構造アライメントの適用例を示す。文間の単語アライメント (a~d) に対して、HYP における各依存構造 (6~8) が TEXT 中の構造に対応付けられるかどうかを判別する。

- 6 a, b の単語アライメントがあり、それぞれの文構造は 1 と 6 で対応しているので、1 と 6 の間に局所構造アライメントをとる。
- 7 c, d の単語アライメントがあるが、それぞれの文構造は HYP が 7 の係り受け関係であるのに対し、TEXT では構造的関係を持っていない。従って、7 に対して局所構造アライメントはとられない。
- 8 b, d の単語アライメントがあり、それぞれの文構造は 8 と 3 で対応しているので、8 と 3 の間に局所構造アライメントをとる。

このように単語間の対応付けだけでなく、単語間の依存関係の対応付けもとれたときに意味的対応付けもとれていると考えることができる。

ここで局所構造アライメントと呼んでいるのは、個々の依存関係ごとにのみアライメントをとるためである。単語間の構造は依存構造だけでなく、述語項構造なども考慮する。

2.3 関連研究

文間関係認識のためのアライメントとして最も単純な方法に、語彙知識を用いて最も似ている単語対に対してのみアライメントをとる手法があげられる [7]。それに対して、Sammons ら [17] や Harabagiu ら [8] は関係認識の際にアライメントの情報を素性の一つとし、さらに文構造や共参照といった情報も考慮しているが、明示的に文構造に対するアライメントは行っていない。

MacCartney ら [11] は単語アライメントの際に構造を考慮し、連続した単語列にアライメントされた場合はフレーズとしてアライメントしている。しかし、非連続の単語列については構造的に対応していてもアライメントをとることはできない。

一方で、多言語間ではあるが、統計的機械翻訳 (SMT) の分野では文間アライメントは重要な技術であり、構造も考慮した手法が提案されている。例えば、中澤ら [14] は文構造の類似度により、単語アライメントで扱うのが困難な、距離の大きな語順変化にも対応させている。しかし、これらの手法を直接利用することは以下の 2 つの理由により難しい。

1. SMT では大量の平行コーパスがあるため、生成モデルによって意味的に対応する単語対を獲得できるが、文間関係認識は平行コーパスに相当するものは存在してもごく少量である。
2. SMT ではアライメント対象となる 2 文は意味的に等価であるため、文法的に対応しない場合を除いてアライメントされない単語対は存在しないと仮定で

きる。しかし、文間関係認識が対象とするのは他方に付加情報がある場合や、そもそも全く関係のない (RTE での「不明」) 文対も処理対象に含むため、アライメントしないことが正解である場合がある。

3 プロトタイプシステム

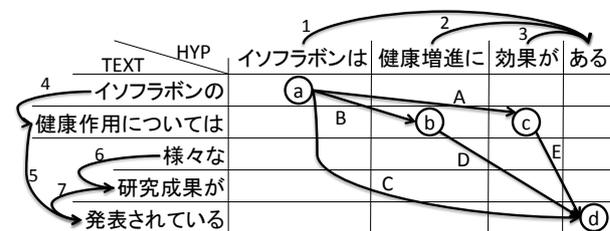


図 3: 局所構造アライメント

(セル内の “ ” は文節アライメントを、“ ” と “ ” の間のエッジは局所構造アライメントを示す)

局所構造アライメントでは単語間と、依存構造間に対応付けをとる。そのため、理想的にはこれらを結合して学習することが望ましい。しかし、そのためには大量のアライメント済みデータが必要であり、そういったデータをすぐに用意することは難しい。そこで本研究では、まず全体を単語アライメントと局所構造アライメントの 2 つの段階に分け、それぞれをルールベースでプロトタイプングすることを考えた。得られたシステムを利用して事例の収集、課題の性質の分析などを行い、どの部分を統計的学習で最適化するのが効果的かを検討するのが狙いである。本研究では、単語アライメントをとる単位を文節とする。ここでは便宜的に、文 A に対する文 B の関係を判断することとする。

3.1 意味的類似度に基づく文節アライメント

文節間のアライメントは以下の 3 種類の類似尺度に基づいて行う。便宜的に、文 A 内の文節 a と、文 B 内の文節 b との間に文節アライメントをとるかどうかを判断することとする。

1. 表層的な類似度
 - a と b で共通する名詞、動詞、形容詞、形容動詞のいずれかがある場合、または文節中の文字コサインが閾値以上の場合に文節アライメントをとる。
2. 語彙知識に基づく意味的な類似度
 - 以下の語彙知識を用い、意味的に対応する文節間にアライメントをとる。
 - 日本語 WordNet [1], 実体間関係知識 [18] a 中の単語の synset および hypernym に b 中の単語が含まれるかどうかを判断する
 - 例) 効果 - 作用
 - 事象間関係知識 [12], ALAGIN [9] 2 つの述語が意味的に対応するかどうか記述された知識で、a と b が定項を含めて対応するかどうかを判断する
 - 例) 防ぐ - 予防する
3. 構造的な類似度
 - 単一言語間のアライメントをとるためには、知識が不足していることが指摘されている [3]。そこで、係

り受け解析および、述語項構造解析を行い、2つの述語において、係り元の文節や項が一致する場合、それらの間にアライメントをとる。このとき、文節が一致しているかどうかは前述の2種類の類似度に基づいて判断し、2文節程度が一致する場合のみアライメントをとる。

しかし、係り受け解析における「係り先となる文節は1つのみ」という制約や、述語項構造解析の精度が十分ではないということから、この手法で十分な性能が得られない可能性がある。そのため、助詞で終わる文節は、それに続く数文節に対して関係があると見なし、係り受けなどと同様に一致するかどうかを判断する。

また、存在を示唆する動詞(ある/ない, 多い/少ない等)については、それらを人手でリスト化し、比較する2つの文節が存在/非存在の関係にある場合は、前述の条件(2文節程度が一致)を弱め、1文節が一致する場合でもアライメントをとる。

図3において、a~cは表層的・意味的な類似度からアライメントされる。dの[ある - 発表されている]は語彙知識には存在しないが、単語アライメントbおよびcがdに対して直接係っているという構造的な類似性に基づいてアライメントされる。

3.2 構造的類似度に基づく局所構造アライメント

次に、文節間の依存関係に対して対応付けを行う。これについても、対応する2文節のペアを入力として、対応付けされるかどうかを二値分類する分類器を構築することが望ましいが、プロトタイプでは以下のように判断する。

1. HYPとTEXTのそれぞれに対して、文節間の構造を調べる。
2. 双方の文節間に構造が存在した場合、それらの類似度を計算し、対応付けをとるかどうかを判定する。文節間の構造は様々な種類が考えられるため、人手でいくつかの構造を定義し、それぞれに異なる重みを付与し、類似度を計算する。以下に図3における適用例を示す。

a - c HYP側は直接関係は存在しないが、2文節とも同じ文節に係るという構造を持っており、TEXT側は係り受け構造になっている。従って、Aで指される局所構造アライメントをとる。

c - d HYP, TEXTともに係り受け構造となっているため、Eで指される局所構造アライメントをとる。

a - d HYP側は係り受けの関係にあり、TEXT側は文節1つ(健康作用については)を経由する構造となっている。従ってCで指される局所構造アライメントをとる。

このように文節を一つ以上経由することを許したりすることで、柔軟なアライメントをとる。同様の処理で、A~Eまでの5つの局所構造アライメントをとることができる。

最終的にHYP側の依存関係1~3はそれぞれ、以下のようにTEXT側の構造に対応付けられる。

- 1 アライメントCにより、依存関係4+5に対応付け
- 2 アライメントDにより、依存関係5に対応付け
- 3 アライメントEにより、依存関係5に対応付け

4 実験と評価

4.1 評価コーパス

関係認識を行う文対は、村上ら[19]に倣い、クエリに対する検索結果のテキストとする。クエリは表1に示す7種類を用意し、検索エンジンTSUBAKIで検索された文集合に対して提案手法を適用し、人手で正解判定を行った。正解データは、クエリ中の文節間の依

検索質問 (下線部がクエリ)	検索数
キシリトールは <u>虫歯</u> 予防に効果がある	95
イソフラボンは <u>健康</u> に効果がある	523
ステロイドは <u>副作用</u> がある	81
還元水は <u>健康</u> を守る	118
バイオエタノールは <u>環境</u> に良い	172
クローン技術は <u>規制</u> が必要だ	175
CO2は <u>温暖化</u> の原因である	905

存関係を局所構造アライメントがとらえられているかどうかを人手で判断し、各クエリごとに20文ずつ構築した。これらのデータの一部には村上ら[19]の定めた関係ラベルが付与されており、20文中に同意とその他が13:7程度になるようにサンプリングした。正解データを構築する際、クエリ中の依存関係にある文節対と、それに対応するテキスト中の文節対が意味的に対応するかどうかを併せて判断する。従って、文節アライメント単体の評価は局所構造アライメントによって行われていると考えてよい。

文節アライメントそのものの評価は、Brockett[2]が作成したRTE2006のためのアライメント正解データの構築基準が参考になるが、彼らは文脈を考慮して文間で出来るだけ1対1対応となるように単語アライメントを行うため、本研究の趣旨とは異なる。

4.2 実験結果と考察

4.1節で構築したコーパスに対して提案手法を適用し、その評価を行った。その精度と再現率を表2に示す。不明の関係にある文対における精度と再現率が、他の関係より低いのは、同意や矛盾の関係にある文対ではクエリ側の依存関係がテキスト側に含まれている可能性が高いが、不明の関係にある文対ではそうとは限らない場合が多いので、より難しい問題となっているためであると考えられる。

関係	同意	矛盾	不明
精度	54.38%	52.08%	47.08%
再現率	51.77%	51.04%	44.04%

4.3 エラー分析

エラー分析を行った結果、語彙知識の不足が大きな原因であることが分かった。例えば、次の文対においてそれが顕著である。

TEXT バイオエタノールは 環境に 良い
HYP バイオエタノールは 地球に 優しい

まず, [環境に - 地球に] は語彙知識にはないため, 文節アライメントをとることができない. 次に [良い - 優しい] は本プロトタイプでは構造的類似度に基づいて文節アライメントすべき文節対だが, [環境に - 地球に] に文節アライメントをとることができないために, 構造的類似度が低いと判断されてしまう. こういった問題に対しては, 知識不足を解消することで改善されると考えられる.

5 おわりに

本稿では, アライメントされた単語対の依存関係に対して対応付けを行う局所構造アライメントを提案し, そのプロトタイプシステムを構築した. そして, 人手で整備した評価コーパスにおいてその評価を行った. その結果, 同意の関係にある文対に対しては, プロトタイプシステムであるにもかかわらず 54.38% という精度が得られた.

今後はより大規模な評価を行うことで, 解決すべき課題を明らかにするとともにプロトタイプシステムの性能向上を図る. さらに, 構築されたシステムによって得られたアライメントデータを元に, 機械学習手法を適用していく.

謝辞 本研究は, (独) 情報通信研究機構の委託研究「電気通信サービスにおける情報信憑性検証技術に関する研究開発」の一環として実施した.

参考文献

- [1] Francis Bond, Hitoshi Isahara, Sanae Fujita, Kiyotaka Uchimoto, Takayuki Kuribayashi, and Kyoko Kanzaki. Enhancing the Japanese wordnet. In *The Joint Conference of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2009)*, 2009.
- [2] Chris Brockett. Aligning the RTE 2006 corpus. In *Microsoft Research Technical Report MSR-TR-2007-77*, 2007.
- [3] Alexander Budanitsky and Graeme Hirst. Evaluating wordnet-based measures of lexical semantic relatedness. *Computational Linguistics*, Vol. 32, No. 1, pp. 13–47, March 2006.
- [4] Peter Clark and Phil Harrison. An inference-based approach to recognizing entailment. In *Text Analysis Conference (TAC)*, 2009.
- [5] Ido Dagan, Oren Glickman, and Bernardo Magnini. The pascal recognizing textual entailment challenge. In *Proc. of the PASCAL Challenges Workshop on Recognising Textual Entailment*, 2005.
- [6] Danilo Giampiccolo, Hoa Trang Dang, Bernardo Magnini, Ido Dagan, Elena Cabrio, and Bill Dolan. The fourth pascal recognizing textual entailment challenge. In *Text Analysis Conference (TAC 2008)*, 2008. on-line Proceedings: <http://www.nist.gov/tac/publications/2008/papers.html>.
- [7] Oren Glickman, Ido Dagan, and Moshe Koppel. Web based textual entailment. In *Proc. of the First PASCAL Recognizing Textual Entailment Workshop*, 2005.
- [8] Sanda Harabagiu, Andrew Hickl, and Finley Laca-tusu. Negation, contrast and contradiction in text processing. In *Proc. of AAAI 2006*, pp. 755–762, 2006.
- [9] Chikara Hashimoto, Kentaro Torisawa, Kow Kuroda, Masaki Murata, and Jun'ichi Kazama. Large-scale verb entailment acquisition from the web. In *Conference on Empirical Methods in Natural Language Processing (EMNLP2009)*, pp. 1172–1181, 2009.
- [10] Andrew Hickl, John Williams, Jeremy Bensley, Kirk Roberts Bryan Rink, and Ying Shi. Recognizing textual entailment with lcc's groundhog system. In *Proc. of the Second PASCAL Challenges Workshop*, 2005.
- [11] Bill MacCartney, Michel Galley, and Christopher D. Manning. A phrase-based alignment model for natural language inference. In *Proc. of 2008 Conference on Empirical Methods in Natural Language Processing (EMNLP-08)*, pp. 802–811, 2008.
- [12] Suguru Matsuyoshi, Koji Murakami, Yuji Matsumoto, and Kentaro Inui. A database of relations between predicate argument structures for recognizing textual entailment and contradiction. In *Proc. of the 2nd International Symposium on Universal Communication (ISUC2008)*, pp. 366–373, 2008.
- [13] Koji Murakami, Eric Nichols, Suguru Matsuyoshi, Asuka Sumida, Shouko Masuda, Kentaro Inui, and Yuji Matsumoto. Statement map: Assisting information credibility analysis by visualizing arguments. In *Proc. of the 3rd ACM Workshop on Information Credibility on the Web (WICOW 2009)*, pp. 43–50, 2009.
- [14] Toshiaki Nakazawa and Sadao Kurohashi. Statistical phrase alignment model using dependency relation probability. In *SSST '09: Proceedings of the Third Workshop on Syntax and Structure in Statistical Translation*, pp. 10–18, Morristown, NJ, USA, 2009. Association for Computational Linguistics.
- [15] Dragomir R. Radev. Common theory of information fusion from multiple text sources step one: Cross-document structure. In *Proc. of the 1st SIGdial workshop on Discourse and dialogue*, pp. 74–83, 2000.
- [16] Han Ren, Donghong Ji, and Jing Wan. WHU at TAC 2009: A tri-categorization approach to textual entailment recognition. In *Text Analysis Conference (TAC)*, 2009.
- [17] Mark Sammons, V. G. Vinod Vydiswaran, Tim Vieira, Nikhil Johri, Ming-Wei Chang, Dan Goldwasser, Vivek Srikumar, Gourab Kundu, Yuancheng Tu, Kevin Small, Joshua Rule, Quang Do, and Dan Roth. Relation alignment for textual entailment recognition. In *Proc. of RTE 2009*, 2009.
- [18] Asuka Sumida, Naoki Yoshinaga, and Kentaro Torisawa. Boosting precision and recall of hyponymy relation acquisition from hierarchical layouts in wikipedia. In *Proc. of the 6th International Language Resources and Evaluation (LREC'08)*, 2008.
- [19] 村上浩司, 水野淳太, 後藤隼人, 大木環美, 松吉俊, 乾健太郎, 松本裕治. 文間意味の関係認識による言論マップ生成. 言語処理学会第 16 回年次大会発表論文集 PA2-22, 2010.