

基本文関係に分解した日本語含意関係認識アノテーション

Annotating Japanese Data Sets Broken Down into Basic Sentence Relation for Recognizing Textual Entailment

金子 貴美 戸次 大介 宮尾 祐介
Kimi Kaneko *1 Daisuke Bekki *1 Yusuke Miyao *2

*1 お茶の水女子大学大学院 人間文化創成科学研究科 理学専攻 情報科学コース
Advanced Sciences, Graduate School of Humanities and Sciences, Ochanomizu University

*2 国立情報学研究所
National Institute of Informatics

In recent years, Recognizing Textual Entailment has been studied extensively, aiming to establish semantic processing. To let a computer recognize Textual Entailment exactly, data consisting of monothematic Text - Hypothesis pairs and a methodology for the creation of such data sets are necessary, although such data and the methodologies for Japanese have been unavailable. Therefore, this paper proposes a methodology for the creation of specialized Japanese data sets for Textual Entailment, which consists of monothematic Text - Hypothesis pairs (i.e. pairs in which only one inference rule relevant to the entailment relation is highlighted and isolated). In addition, we compare our methodology with existing studies and consider validity of this methodology.

1. はじめに

近年、「意味処理」の実現に向けた研究を行うことが喫緊の課題となりつつある。1990年代以降、自然言語処理分野を牽引してきた統計的言語処理の技術が、意味の観点では今なお表層的な処理に留まっており、発展の停滞が予見されるからである。したがって、格解析や語義曖昧性解消などの様々な「意味処理」の研究が行われており、その1つとして、どの言語表現とどの言語表現が同じ情報を伝えているかの判断がある。その判断を行う技術が含意関係認識である。

含意関係認識とは、与えられた2文の間に含意関係が成り立つかどうかを自動認識する技術であり、文t1とt2の間の含意関係とは、人間が「t1が正しいとき、t2も(ほぼ)正しい」と判断する関係のことである*1。以下に例を示す。

例1) 含意関係認識の例

- ラベル=Y
- t1:川端康成は、「雪国」などの作品でノーベル文学賞を受賞した。
- t2:川端康成は「雪国」の作者である。

1行目の「ラベル」は含意関係の成立・不成立を表す。t1とt2の間に含意関係があれば「成立」を表す「Y」のラベルが、なければ「不成立」を表す「N」のラベルが付与される。

この判断を計算機が行うのは難しい。2文の間にはより基本的な含意関係が複数潜んでいる上、各関係は意味上の繋がりがああるため、計算機に個々の関係を正しく捉えて判断させるのは困難だからである。したがって、まずは文をより基本的な含意関係の連鎖として書き下した評価データを構築しなければならない。しかし、そのような日本語評価データとその構築手法が存在しないという問題がある。本論文では、上記のような問題を解決すべく、より基本的な含意関係の連鎖として書き下した日本語評価データの構築手法を提案する。また、提案手法により構築したデータを評価し、本手法の妥当性を考察する。

連絡先: 金子 貴美, お茶の水女子大学大学院人間文化創成科学研究科理学専攻情報科学コース戸次研究室,
〒112-8610 東京都文京区大塚 2-1-1, 03-5978-5789,
kaneko.kimi@is.ocha.ac.jp

2. 研究背景

Sammonsら[1]が、含意関係を正しく捉えるためには、より基本的な含意関係の連鎖として文を書き下す方法を確立させる必要があると指摘している。また、方法論の確立により、自然言語処理の他の既存手法をどう組み合わせることが出来るか、また、現時点で解決不能な問題は何か、が把握可能となると述べている。彼らは実際にデータを構築し、手法の妥当性も評価しており、構築したデータを用いて学習したシステムのほうが、原文データを用いたシステムよりも解析精度が高くなるかと報告している。また、[1]と同様に含意関係認識のために文を書き下した評価データの構築手法を提案した関連研究として、Bentivogliら[2]の研究もある。これらの手法の妥当性は、英文のデータセットを用いた場合しか評価されておらず、日本語データに適用した際の妥当性の判断が出来ない。したがって、日本語における妥当性は、改めて日本語の評価データを用いて判断する必要がある。

含意関係認識のための日本語データセットとしては、小谷ら[3]のデータのペアに1つしか現れないような人為的に構築されており、基本的な含意関係の連鎖として書き下すことや、基本的な含意関係が実際の文においてどう潜んでいるかの分析が出来ないため、手法の策定には不向きである。したがって、今回はRITE-2*2のバイナリクラスサブタスク(以下、「BCサブタスク」)の一部のデータを用いて評価データを作成しながら、日本語評価データ構築手法を策定することとした。RITE-2とは、含意関係認識の研究促進を目的とした評価型ワークショップである。与えられた文のペアの含意関係が成り立つか否かを判別するBCサブタスクをはじめ、4つのサブタスクから構成されている。BCサブタスクのデータの一部を用いた理由は、作成したデータをRITE-2のタスクで用いることで、認識精度の観点からも本手法の妥当性の検証ができるためである。

*1 自然言語処理でいう「含意」は、論理学でいう意味論的含意とは異なり、現象に対する一種のカバートームである。形式言語論でいう「含意」「前提」「会話の含意」を含む。

*2 <http://www.cl.ecei.tohoku.ac.jp/rite2/doku.php>

3. 提案手法

含意関係を基本文関係にブレイクダウンし、含意関係認識のための日本語評価データを作成する方法を考案した。基本文関係の一覧は表4に示す。基本的に書き換えていく度に意味が広がるよう書き下し、t1を徐々にt2に近づけるよう変換するとした。また、第1節で述べたように、2文の関係を自動認識させるために必要な基本文関係が書き出せるよう、t1とt2の書き換えは関係の繋がりを考慮して行うこととし、個々の現象にフォーカスし分析できるよう、書き下した各ペアでは、複数の関係が成立しないように書き換えることとした。

たとえば、先の例1のような文のペアが与えられたとする。この2文のペアを以下のように書き下し、関係ラベルを付与した。

例1') 例1を書き下した例

- [フレーズ:含意・前提] 川端康成は「雪国」などの作品の作者である。
- [集合・リスト] 川端康成は「雪国」の作者である。
- [一致] 川端康成は「雪国」の作者である。

一方、含意関係が成り立たない2文が与えられた場合の例を以下に示す。

例2) 含意関係が成り立たない場合の例

- ラベル=N
- t1:パルテノン神殿は、ドーリア式神殿の最高傑作と言える作品である。
- t2:パルテノン神殿は、ヘレニズム文化の影響下で建設された。

例2') 例2を書き下した例 a

- [修飾句削除] パルテノン神殿は、ドーリア式神殿の作品である。
- [主辞削除] パルテノン神殿はドーリア式神殿である。
- [フレーズ:含意・前提] パルテノン神殿は、ドーリア式神殿として建設された。
- [フレーズの不一致] パルテノン神殿は、ヘレニズム文化の影響下で建設された。

あるいは、書き換え易さを重視し、以下のように「含意関係が成り立たない場合の関係」ラベルの途中使用も良いとした。また、t1とt2が全く無関係で書き換えるのが難しい場合など、必要に応じて複数回使っても良いとした。

例2'') 例2を書き下した例 b

- [修飾句削除] パルテノン神殿は、ドーリア式神殿の作品である。
- [主辞削除] パルテノン神殿はドーリア式神殿である。
- [単語の不一致] パルテノン神殿は、ヘレニズム式神殿である。
- [フレーズ:含意・前提] パルテノン神殿は、ヘレニズム文化の影響下で建設された。

3.1 t2の変換

前述までの例で示したように、原則的には、t2より狭い意味をもつt1を変換してt1より広い意味をもつt2に近づけるように書き換えることとした。t2の書き換えも許すと、含意関係が成立しない場合であっても、t1とt2が示す状況(集合)が一致するよう書き換えることができってしまうからである。しかし、t2を変換したほうが判断が容易な場合もある。したがって、t1とt2が矛盾関係にあり、t2を書き換えてもt1の状況と重ならず、含意判定が変わらない場合に限り、t2を変換可能とした。以下に例を示す。

例3) t2の変換の例

- ラベル=N
- t1:ポイントカードには、クレジット機能付きとクレジット機能なしの2つの種類がある。
- t2:ポイントカードは現金専用でクレジット機能は付いていない。

例3') 例3を書き下した例

- [集合・リスト] ポイントカードには、クレジット機能付きもある。
- [フレーズ:同義・類義] ポイントカードはクレジット機能が付いているものもある。
- [t2の編集] ポイントカードは現金専用でクレジット機能は付いていない。
- [修飾句削除] ポイントカードはクレジット機能が付いていない。
- [モダリティの不一致] ポイントカードはクレジット機能が付いているものもある。

4. 評価

4.1 既存研究との比較評価

BCサブタスクのデータセットの一部に本手法を適用した。173組のデータを1名で書き下し、2名のアノテータがラベル付けを行った。アノテーションには、表4に挙げたラベルの他、該当するラベルがない場合のための「不明」ラベルを用いた。173組中、112組のデータは本手法の策定に用い、残りの61組について一致率と平均書き換え数、平均書き換え時間を算出した。今回扱ったデータにおける一致率は0.83であった。一致率は、以下の式で算出した。

$$\text{一致率} = \frac{\text{ラベル一致数}}{\text{全体数}^*3}$$

[2]の一致率の算出方法はDice係数*4であるため異なっているが、参考までに比較すると、本研究の一致率は[2]の、書き下した文毎の一致率0.78より若干高い値となっており、本手法の分類の妥当性は[2]と大差ないと考えられる。

表1に作成したデータ全体のスペックを示す。

表1: 作成したデータの仕様

		書き換えた文のペア		
		Y	N	書換総数 / 原文総数
原 文	Y (32)	116	-	116 / 32
	N (29)	96	29	96 / 29
	計 (61)	212	29	241 / 61

書き換え数は全体で212回*5であり、平均3.48回であった。これは[2]の平均2.98回よりも若干高い値となっている。また、「Y」ラベルの付与された組においては平均3.62回、「N」ラベルの付与された組は平均3.31回で、これらも共に、[2]の平均3.03回、平均2.80回よりも若干高い値となっている。日本語データであることや、複雑なデータが多いことを考慮すると、本手法は[2]相応の妥当性があると考えられる。

表2*6に[2]と本研究における関係ラベルの対応を示す。

*3 「一致」ラベルが付与される文は明らかであるため、「全体数」は「ラベル付けを行った文の数から一致ラベルを付けた文の数を引いた数」、「ラベル一致数」は「2名が付けたラベルが一致した文の数から一致ラベルを付けた文の数を引いた数」とした。

*4 Dice係数は、2つの集合の共通要素数を各集合の要素数の平均で割ったもので、アノテーションの一致率計算にも用いられる。

*5 ラベル付けを行った文の数から、含意関係の判断を行っただけで書き換えてはいない「一致」「不一致」のペア数を引いて算出した。

表 2: 既存手法 [2] と本手法における関係ラベルの分布

関係ラベル	文のペア数					
	既存手法 [2]			本手法		
	total	Y	N	total	Y	N
同義・類義	25	22	3	45	45	0
上位・下位	5	3	2	5	5	0
含意・前提	-	-	-	44	44	0
全体・部分	7	4	3	1	1	0
フレーズ：品詞の変換	9	9	0	1	1	0
共参照・照応	49	48	1	3	3	0
語順入れ替え	-	-	-	15	15	0
態の変化	7	5	2	7	7	0
修飾句削除	25	15	10	42	42	0
主辞削除	6	6	0	1	1	0
並列・従属節	5	4	1	14	14	0
集合・リスト	1	1	0	3	3	0
同格	3	2	1	1	1	0
関係節	1	1	0	8	8	0
時間の推論	2	1	1	1	1	0
空間の推論	1	1	0	1	1	0
数の推論	6	0	6	0	0	0
暗黙の関係	7	7	0	18	18	0
その他の推論	40	26	14	2	2	0
単語・フレーズの不一致	3	0	3	27	0	27
モダリティの不一致	1	0	1	1	0	1
時間の不一致	-	-	-	1	0	1
空間の不一致	-	-	-	0	0	0
数の不一致	-	-	-	0	0	0
“Demonymy”	1	1	0	-	-	-
“Statements”	1	1	0	-	-	-
合計	205	157	48	241	212	29

[2] の分類のうち、“Demonymy”, “Statements” は日本語の場合はあまり問題とならなかったため、本研究では除外した。一方で「語順入れ替え」「含意・前提」を追加した。表 2 より、[2] は本研究より「共参照・照応」が多く見られることや、本研究では [2] と比べ「含意・前提」「語順入れ替え」が頻発していることが分かる。これらから、言語の違いが関係ラベルの分布や分類の違いに現れているといえる。

また、書き換え時間は、1 組あたり平均 12 分であった。[2] の書き換え時間が 1 組あたり平均 15 分であることから、本手法は [2] と同程度の実用性をもつと考えられる。

4.2 関係ラベル毎の精度評価

2013 年 1 月 9 日-16 日の RITE-2 フォーマルラン*7において、15 チームが、本データを各自のシステムに認識させ、精度を算出した。表 3 に関係ラベル毎の平均精度一覧を示す。

表 3: 関係ラベル毎の平均精度と人間による誤分類数

関係ラベル	平均精度 (%)	文のペア数	人間による誤分類数
語順入れ替え	89.6	15	4
修飾句削除	88.8	42	0
集合・リスト	88.6	3	0
時間の推論	85.7	1	1
関係節	85.4	8	2
並列・従属節	85.0	14	2
単語：上位・下位	85.0	5	1
フレーズ：不一致	80.1	25	0
態の変化	79.9	7	2
単語：同義・類義	79.7	9	6
主辞削除	78.6	1	2
暗黙の関係	75.7	18	2
フレーズ：同義・類義	73.6	36	9
共参照・照応	70.9	3	1
フレーズ：含意・前提	70.2	44	7
単語の不一致	69.0	2	0
単語：全体・部分	64.3	1	1
フレーズ：品詞の変換	64.3	1	0
同格	50.0	1	1
空間の推論	50.0	1	1
その他の推論	40.5	2	2
モダリティの不一致	35.7	1	0
時間の不一致	28.6	1	1
合計	-	241	41

*6 [2] では、「単語」と「フレーズ」を分けずに分類しているため、表 2 では分けずに分類している。

*7 RITE-2 では、本手法により構築したデータを「ユニットテストデータ」として公開している。

「語順入れ替え」、「修飾句削除」が平均精度が 90%程度と高い精度を出した。また、元々の関係ラベル数が少ないものの、「集合・リスト」の精度も 90%程度となった。一方、精度が低かったラベルとして「時間の不一致」、「モダリティの不一致」、「その他の推論」「空間の推論」、「同格」が挙げられ、それぞれ 50%以下となった。また、「単語の不一致」、「フレーズ：品詞の変換」、「単語：全体・部分」は 50-70%の精度であった。ここで挙げた、精度が 70%以下のラベルは全て出現回数が 3 回未満であり、ラベルが 3 回以上出てきたものはすべて 70%以上の精度を出しているため、ラベルの出現頻度が精度に関連している可能性がある。出現頻度が低いラベルに関しても正確に認識するシステムを構築するにはどうすべきか、を今後検討する必要があると考えられる。

また、下記の例 4 のように、アノテータ 2 名により付けられたラベルが異なった文の数も関係ラベル毎に分類し、その数を「誤分類数」として表 3 に示した。

例 4) アノテータ A、B により付与されたラベルが異なる例
[書き換え前] 1991 年 12 月にソ連が崩壊し、2000 年 5 月 18 日にプーチンが露大統領となった。

- A: **[時間の推論]** ソ連崩壊後、プーチンが露大統領となった。
- B: **[フレーズ:含意・前提]** ソ連崩壊後、プーチンが露大統領となった。

表 3 より、「単語：同義・類義」「語順入れ替え」など一部を除き、人により判断が揺れやすく、誤分類が多いラベルの精度が、全体的に低くなっていることが分かる。このことから、人によりラベルの判断が揺れるケースと、計算機の判断が揺れるケースは相関関係があると考えられる。

5. まとめ

含意関係を基本含意関係・非含意関係にブレイクダウンした日本語評価データの構築方法を考案し、本手法に従って日本語評価データを構築した。構築したデータの一貫率、平均書き換え数、書き換え時間を既存研究と比較し、本手法の妥当性を示すと共に、言語による違いが関係ラベルやアノテーション結果の違いに現れることを指摘した。また、関係毎の認識精度の分析結果から、出現頻度と精度の関連性、人間の誤分類率と精度の関連性を示した。

今後は [4] で挙げた、ラベル付けの問題の検討、書き換えパターンの比較などの課題を検討していく予定である。

参考文献

- [1] Sammons M., Vydiswaran, V.G., Roth, D.: Ask not what textual entailment can do for you... In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. pp. 1199-1208 (2010).
- [2] Bentivogli, L., Cabrio, E., Dagan, I, Giampiccolo, D., Leggio, M. L., Magnini, B.: Building Textual Entailment Specialized Data Sets: a Methodology for Isolating Linguistic Phenomena Relevant to Inference. In Proceedings of LREC 2010, Malta (2010).
- [3] 小谷 通隆, 柴田 知秀, 中田 貴之, 黒橋 禎夫.: 「日本語 Textual Entailment」のデータ構築と自動獲得した類義表現に基づく推論関係の認識, 言語処理学会 第 14 回年次大会, pp.1140-pp.1143 (2008).
- [4] 金子 貴美, 宮尾 祐介, 戸次 大介.: 「基本文関係に分解した含意関係日本語評価データの構築」, 言語処理学会 第 19 回年次大会, pp.928-pp.931 (2013).

表 4: 基本文関係一覧表

基本含意関係		
関係ラベル	例	用法説明
一致		文同士が完全に一致する場合。作成した最後の文が t2 と一致するときに使う。
単語の書き換え		
同義・類義	紀元前 5 世紀→前 5 世紀, 各ボリス→諸ボリス, アケメネス朝(ペルシア帝国)→アケメネス朝ペルシア, 法廷で→法廷において	「単語」を(ほとんど)同じ意味の別の「単語」に書き換える。逆向きの書き換えが不可能な場合は除く。文脈によらず成り立つ同義表現を補う場合を含む。
上位・下位	最高裁判所裁判官→裁判官, 電子機器→コンピュータ	「単語」をその上位または下位(一方の語が指すものの集合がもう一方の語が指すものの集合を含む、is-a 関係)の「単語」に書き換える。
含意・前提	妨害した→止めた, 離婚した→結婚した	動詞や形容詞(イベントを表す名詞も含む)で、前者が成り立っている状況では必ず後者も成り立っている場合。
全体・部分	千代田区→東京都, 顔→口	物とその部品、のように「単語」で指されているものに物理的に含まれるものを指す、別の「単語」に書き換える。「上位・下位」に当てはまるものは除く。
フレーズの置き換え		
同義・類義	広がっていった→形成されるようになった, 西ヨーロッパを代表する国々→西欧諸国	「単語: 同義・類義」同様、書き換え元または先が「フレーズ」の場合。
上位・下位	イエニチェリ→オスマン帝国の常備軍	「単語: 上位・下位」同様、書き換え元または先が「フレーズ」の場合。
含意・前提	ノーベル文学賞を受賞した→作者である, 施政権を行使する→統治権を得た	「単語: 含意・前提」同様、書き換え元または先が「フレーズ」の場合。
全体・部分	パリ→西ヨーロッパを代表する国	「単語: 全体・部分」同様、書き換え元または先が「フレーズ」の場合。
品詞の変換	雑誌の差し止め→雑誌を差し止める	ある品詞(例: 名詞)での表現を、同じ単語の別の品詞(例: 動詞)に書き換える。
共参照・照応	これ→アケメネス朝(ペルシア帝国), (e は)勝利した→ギリシアの諸ボリスは勝利した	代名詞や省略された項(主語や目的語など)を、文中の別の句で埋める。または、文中で同じものを指す単語を別の単語で置き換える。
文構造の書き換え		
語順入れ替え	X が Y を Z した→Y を X が Z した, A は B である→B は A である	動詞、名詞、形容詞等に係る複数の句や節の順番を入れ替える。複数の句(「X が」「Y を」)が同じ句(「Z した」)に係る場合のみ該当。表層的に語順が入れ替わっていても、句・節以外の部分が入れ替わっている場合は、別のラベルが該当する。
態の変化	X によって Y が建設された→X が Y を建設した, X が Y を壊した→Y が壊れた	受身、使役、能格化などの構文変形を行う(または能動態に戻す)。
修飾句削除	一気に征服し→征服し	A → B という係り受け関係があるときに、A を削除する。
主辞削除	ドーリア式神殿の作品→ドーリア式神殿, 日本代表が勝利した→日本が勝利した	A → B という係り受け関係があるときに、B を削除する。B を削除しても意味が変わらない場合のみに使う。意味が変わる(逆向きの含意関係不成立)時は除く。
並列・従属節	A し、B した→B した A であるが B → A である	連用修飾や接続詞で 2 つ(以上)の節がつながっているとき、いずれかの節を残して別の節を削除する。助詞化した表現が修飾する場合を除く。
集合・リスト	「雪国」などの作品→「雪国」, バルカン半島, エジプト→バルカン半島,	「と」「,」「,」「など」などによりいくつかの名詞が列挙されているとき、そのうち 1 つを抽出する。
同格	ギリシアの都市国家(ボリス)→ギリシアの都市国家はボリスと呼ばれる	A(B)のような同格表現(A と B が同じものを指す)時、「A は B と呼ばれる」「A は B である」と書き換える。A(B とか)や A(B とか呼ばれる)のように、同格を示す表現が括弧内に明示される場合や、「A-B とかいう-」の様な、括弧に準ずる表現の場合も含む。単に A や B を削除する場合は、「修飾句削除」とする。
関係節	小アジア半島に建国したオスマン帝国→オスマン帝国は小アジア半島に建国した, さんまを焼くにおい→さんまを焼く, 後ウマイヤ朝はウマイヤ朝が再興した王朝→後ウマイヤ朝はウマイヤ朝が再興した	関係節+名詞で表されている名詞句を文に変換する。名詞が関係節の中に入る場合(内の関係)と入らない場合(外の関係)の両方共、「関係節」のラベルを付ける。関係節+名詞以外の部分も書き換える場合は、「関係節」ではなく「フレーズの書き換え」とする。
推論		
時間の推論	1931 年→1930 年代, 1821 年→19 世紀, A は 18 世紀末に発明された→A は 19 世紀には存在していた	年代・時間についての包含関係や、時間配置に関する推論が必要な場合。
空間の推論	ギリシア西南部やエーゲ海→ギリシア各地, A の上に B がある→B の下に A がある	地理の知識や空間配置に関する推論が必要な場合。空間配置を考慮する必要がなく、国名や地名に関する「上位・下位」や「全体・部分」の書き換えに相当する場合は除く。
数の推論	20 万とも 50 万とも言われるペルシア軍→ペルシア軍の大群	数学的推論や数に関する推論が必要な場合。時間に関する推論に該当する場合は除く。
暗黙の関係(主辞・修飾句挿入)	日本の軍隊→日本が保有していた軍隊, 日本が勝った→日本の代表チームが勝った	「A の B」「A からの B」など暗黙的に関係付けられた句の間の関係を明示化する。もしくは、「A」の(文脈から推測できる)暗黙の主辞や修飾句を補って「A の B」「B の A」「A という B」「B という A」とするような場合も含む。変換元より変換先の方が情報が多い(より詳細)な場合に使う。「上位・下位」「含意・前提」「主辞削除」などの逆向きの関係。補う語が元の文中に含まれる場合は除く。
その他の推論	津波は 20m に達した→津波は最大波高が 20m に達した	上記の関係以外で、一般常識に基づく推論が必要な場合。
基本非含意関係		
単語の不一致	小アジア半島⇔バルカン半島, ヘレニズム文化⇔ドーリア式	文の中で同じ場所にある「単語」が、お互い「同義・類義」「上位・下位」のような関係に無く、書き換えられない場合。
フレーズの不一致	オスマン帝国のスルタン⇔清の皇帝	「単語の不一致」同様、「フレーズ」が書き換えられない場合。
モダリティの不一致	到達した⇔到達しなかった, しなければならない⇔してもよい	動詞に続くモダリティ(主に助動詞)が異なり、書き換えられない場合。
係り受けの不一致	ドイツが⇔ドイツを, ギリシャによって⇔ギリシャに対し, X が Y を侵略した⇔(e が)Y を侵略した	係り受け関係にある 2 つの文節において、それぞれの内容語は同じだが、その関係が異なる場合。必要な項が存在しない場合も含む。
修飾の不一致	わずかに拡大した⇔大幅に拡大した	修飾句が異なり、書き換えられない場合。必要な修飾句が存在しない場合も含む。
時間の不一致	1921 年⇔18 世紀	時間表減に矛盾がある場合。
空間の不一致	A と B は重なっている⇔A の下に B がある	空間に関する表現に矛盾がある場合。単語やフレーズの不一致に該当する場合は除外。
数の不一致	3 か国⇔4 か国, わずか 3 か国⇔多くの国々	数値表現に矛盾がある場合。時間表現の場合を除く。