

# 言論間論理的関係コーパスの構築

村上 浩司 (NAIST), 増田 祥子 (大阪府立大学/NAIST), 松吉 俊 (NAIST), 松本 裕治 (NAIST), 乾 健太郎 (NICT)

## 1. 背景・目的

### 言論マップ生成課題

- ・ 言論の相対的な関係の把握
- ・ 言論間の論理的関係を解析し、個々の言論をノード、その間の関係をエッジとするグラフ

### 生成した言論マップの評価

- (1) 言論間の論理的関係の推論
- (2) 類似言論のクラスタリング

→ 言論間の論理的関係認識結果を評価するために評価コーパスが必要

言論マップで取り扱う言論 : Web上に存在する美文 (複雑な構造、省略) 任意の2文間に論理的関係ラベルを付与 (負例を含む)

### レベル別に考える

1. 単純命題: 述語項構造レベル : 2文間の述語と必須格の照合に問題を単純化
2. 文命題: 実際のWeb文書中の単文レベル : [仮定]条件節を対象外とした、美文
3. 複合命題: 実際のWeb文書中の複文レベル : Web文書中の完全なる美文

## 3. 言論マップ生成で用いる論理的関係

関係	ラベル	説明	例
類似	同義	AとBは同じ意味、もしくは近い意味を表す関係。言い換えも含む。	A: 再販制度のおかげで市場では価格競争が行われない。 B: 再販制度が廃止されると価格競争が激化する。
	類義	AとBが異なる主体で、述語がAとBで同義、似た意味を持つ関係。	A: ノルウェーは大西洋で93年からミンク鱈の調査捕鯊を行っている。 B: 日本は国際社会での協調性を重視し、調査捕鯊のみをおこなっている。
	同評価	価値判断系の「すべき」「すべきでない」という評価から、出来事の「よくないこと」といった発話者のものの見方、評価の仕方が表れている。	A: 既に死刑制度が立派に存在するにも関わらず、死刑に値するような残虐な犯罪が減った兆しが無い。 B: 日本やアメリカのように、死刑制度が残っている国ほど、犯罪が増えているんでしょう?
含意	含意	Bが成り立ればAも成り立つ関係。	A: キシリールは虫歯予防効果があります。 B: なぜキシリールは虫歯予防に良いのでしょうか?
	認識	BにはAの存在が認識されていることを示している。	A: キシリールには優れたむし歯予防効果があります。 B: 最近はキシリールガムによる、むし歯予防効果が目立っています。
	言明	BはAが存在することを表明している。	A: キシリールがお口の健康維持や虫歯予防にも効果を発揮します。 B: キシリールは虫歯の原因菌の感染を抑える効果が報告されています。
例示	例示	BはAを具体的に述べている。AとBは必ずしも同じ出来事とは限らない。	A: この死刑制度については国際的にも、日本国内においても賛否両論があります。 B: 日本政府は死刑廃止決議に反対し死刑制度存続を主張しています。
	詳述	Aの出来事をBが詳細化している。AとBは同じ出来事である。	A: グリーンピースが調査捕鯊を妨害した。 B: グリーンピースは日本の調査捕鯊船に対しては、スクューにロープを絡ませたり、薬品の入った瓶や発煙筒を投げ込むなどの違法行為を繰り返した。
	前提	Bの出来事が起こる前提となる出来事Aという関係。	A: 文部科学省は来年度から小6と中3の全員で、国語と算数(数学)の二教科の全国学力テストを行うそうです。 B: 全国学力テストは採点は全て依頼されている業者が行うらしい。
対比	対比	AとBの主体が異なり、述語が対立的意味を持つ関係。	A: 鯨油の必要が無くなったアメリカは捕鯊を止めた。 B: 日本は調査捕鯊を続けている。
	矛盾	主体が同じであるAとBの命題が真偽について言及している場合で、AとBのどちらかが偽であるような関係。	A: 条約上、IWCは調査捕鯊を許可・禁止する権限はありません。 B: IWCが日本の調査捕鯊中止勧告を決議する。
	対立	AとBの命題が話者による価値判断による場合、AとBの間で命題が相容れない関係。	A: 学校間の序列化を招く小・中学校の 全国学力テストに参加するべきでない。 B: 全国学力テストは賛成です。
対評価	対評価	「すべき」「すべきでない」という評価から、出来事の「よくないこと」といった発話者の事態の捉え方、評価の仕方が対立している関係。	A: 全国学力テストが学校教育を破壊する。 B: 全国学力テストも学力向上に生かすことが必要である。
	時間	継起	AとBの間に時間関係が存在し、Aの後にBが起きたと考えられる関係。
因果	結果	時間関係がAとBの間に存在し、Aが起きたことでBが起きたという因果関係が存在する関係。	A: 汚損のために書店に出荷できない書籍を各版元が謝恩価格で提供しました。 B: 書籍バージョンは、値崩れ誘発を防止する再販制度権限議論や、出版の専断論の観点から批判を受けた。
	根拠	Bの根拠・動機・理由になるAという関係。	A: 世に、まったく改換の情を見せない凶悪犯罪者がいる。 B: 制度としての死刑制度は、犯罪抑止には必要だ。
問題・対策	問題・対策	なんらかの問題・事態が発生したことを示すAとそれに対する解決策となるBという関係。	A: 主人はかゆみ等の皮膚疾患に悩まされていた。 B: ステロイド外用剤は、かぶれや湿疹などの多くの皮膚病に用いられる。

## 4. コーパス構築での考慮すべき点

### 1. 異なる文書中の任意の2文間に対する論理的関係のラベル付け

- ・ 関係が明示的に判断可能 : 同義、類義など
- ・ 関係が明示的に判断不可能 : 対比、例示、根拠、問題・対策など

#### 例)同一目的で手段が異なる場合

- A. ステロイドでアトピーを治す
- B. 漢方でアトピー性皮膚炎を治療する

“治す”の観点なら、類似  
“治し方”の観点なら、対比?

- ・ 人間 : 対比関係を付与可能 → どこかでそれらが直接対比されていたことを知っている
- ・ 機械 : 認識するためには知識が必要 → どこかの1文書で並列記載を認識?

### 2. 作業量軽減の必要性

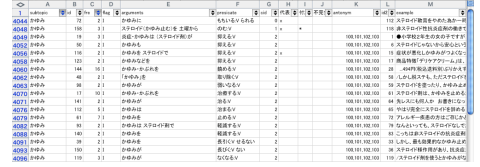
現在の作業 : クエリに関する文書集合中の言論を対象として任意のペアに関係を付与

- ・ 作業量は最大でO(n<sup>2</sup>) : 人間による確認にはきつい
- 作業量を軽減させる必要あり (ある程度対象にする言論ペアワイズを削減)

現実策 : 言論を述語項構造に変換しクラスタリングすることで、類似、対立を仮付与 (事象間関係知識により動詞間関係を認識し、項を共有する場合にクラスタ化)

- クラスタ内の言論は同義、クラスタ間についてその他の関係を付与
- 述語を固定し、項のバリエーションを調べる、またはその逆
- 作業量は幾分軽減
- クラスタリング結果に依存
- 複雑な関係には対応できていない
- 全く違う方法??

### 単純命題レベルの評価コーパス(例)



### 文命題レベルの評価コーパス(例)



## 5. まとめ

### 言論マップの評価コーパスの作成 課題

- ・ Web上の美文を対象
- ・ 多彩な言論間関係
- ・ レベルを換えたコーパスを作成中
- ・ 大規模での関係づけ
- ・ 関係の種類を精査
- ・ 効率的なコーパス作成手順のルーチン化

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意図性検証技術に関する研究開発」の一環として実施した。