

# パターン辞書を用いた意味解析 --- 感情や名詞述語文について

藤原竜樹, 徳久雅人  
鳥取大学

# 研究の目的

- 名詞述語文を解析するためのパターン辞書を作成
- 意味解析の実現

# 名詞述語文

ノドグロは境港産だ。

- 述語の品詞が名詞の文
- 断定や判定を行う文

# 名詞述語文の意味分類

(今田 2011)

生産物名詞：名詞述語文の主語



4通りの分類

属性叙述型

外延叙述型

範疇叙述型

その他

# 名詞述語文の意味分類

(今田 2011)

## 属性叙述型

主語:実体

述語:属性, 値

述語が主語の属性的概念を述べる

カレーは300円だ。

実体

値 (属性)

## 外延叙述型

主語:内包的概念

述語:主語の外延

述語が主語の外延を述べる

使った道具はドライバーだ。

上位

下位

## 範疇叙述型

主語:事物

述語:主語の範疇

述語が主語の帰属する範疇を述べる

パソコンはただの道具だ。

下位

上位

# パターンの作成

# 事例分析

- 以下の事例を分析
  - 今田の論文から得た例文(33文)
  - Wikipediaから収集した名詞述語文(36文)

- ・ 文の構造・構成の把握
- ・ パターン化でどの部分を変数化するか
- ・ どの単語をどのような情報として取り出すか

# パターン化の指針1

分類	指針	
構成	はが文が存在	
	主語, 助詞, 追加情報, 述語, および, 判定詞で構成	
共通	主語	パターンでの範囲は主名詞
		前に連体修飾節がある場合はフラグを立てる
	助詞	「は」, 「が」の2種類使用
		直後に「、」で区切る場合もあり
	追加情報	述語の前が連体修飾節が追加情報 事例によっては存在しない
	述語	述語の範囲は名詞のみ
	判定詞	収集した事例は文末に判定詞を含むので記述
		処理段階では判定詞では無い表現も存在
変数化	主語, 追加情報, 述語, 判定詞を変数化	
	現状処理しきれない文は該当部分を字面に残す	



# パターン化の指針1

分類	指針	
構成	はが文が存在	
	主語, 助詞, 追加情報, 述語, および, 判定詞で構成	
共通	主語	パターンでの範囲は主名詞
		前に連体修飾節がある場合はフラグを立てる
	助詞	「は」, 「が」の2種類使用
		直後に「, 」で区切る場合もあり
	追加情報	述語の前が連体修飾節が追加情報 事例によっては存在しない
述語	述語の範囲は名詞のみ 収集した事例は文末に判定詞を含むので記述	

冷蔵庫にある製品 は 国内向けのもの だった。

主語

追加情報

述語

判定詞

車体の強度部材に用いられる材料 は 鋼鉄 が 主流 で、

主語1

主語2

述語

判定詞

# パターン化の指針2

分類	指針	
属性叙述型	とり出される情報は、実体、属性、属性値	
	実体は現在のところ主語	
	属性、属性値は事例によって異なる	
	組み合わせ	属性が述語、属性値が追加情報
		属性が追加情報、属性値が述語
		属性が述語、属性値が追加情報と述語
		属性、属性値ともに述語
属性、属性値ともに追加情報と述語		
外延叙述型	とり出される情報は、上位語、下位語、追加情報	
	上位語は主語	
	下位語は述語	
範疇叙述型	とり出される情報は、下位語、上位語、追加情報	
	下位語は主語	
	上位語は述語	
	ただし、「～の一種」という述語自体が範疇叙述型を表す場合はその前の名詞が上位語	

# パターン化の指針2

分類	指針	
属性叙述型	とり出される情報は、実体、属性、属性値	
	実体は現在のところ主語	
	属性、属性値は事例によって異なる	
	組み合わせ	属性が述語、属性値が追加情報
		属性が追加情報、属性値が述語
		属性が述語、属性値が追加情報と述語
		属性、属性値ともに述語
		属性、属性値ともに追加情報と述語

車体の強度部材に用いられる 材料 は 鋼鉄 が 主流 で、

主語1
主語2
述語



車体の強度部材に用いられる 材料 は 鋼鉄 が 主流 で、

実体
属性値
属性

# パターン例

テーブルは大理石製である。



パターン：

/N1(は|が)[、 ]/yf[MT2]N3[MD4]

型名：属性叙述型

3つ組：実体 N1,属性 N3,属性値 N3

主語前：False

主語前のフラグは  
マッチの条件ではない

MT:体言の修飾語句  
MD:モダリティ表現  
その他の記述子(池原ら2004)

車体の強度部材に用いられる  
材料は鋼鉄が主流で、



パターン：

/N1(は|が)[、 ]/yf[MT2]N3が主流[MD4]

型名：属性叙述型

3つ組：実体 N1,属性 主流,属性値 MT2N3

主語前：True



46件のパターンを作成

# パターン例

/N1(は が)、 ]/yf[MT2]N3[MD4]	型名:属性叙述型	実体:N1	属性:N3	属性値:MT2N3	主語前:true
/N1(は が)、 ]/yf[MT2]N3[MD4]	型名:属性叙述型	実体:N1	属性:N3	属性値:MT2N3	主語前:false
/N1(は が)、 ]/yf[MT2]N3がN4[MD5]	型名:属性叙述型	実体:N1	属性:MT2N3	属性値:N4	主語前:true
/N1(は が)、 ]/yf[MT2]N3がN4[MD5]	型名:属性叙述型	実体:N1	属性:MT2N3	属性値:N4	主語前:false
/N1(は が)、 ]/yf[MT2]N3が主流[MD4]	型名:属性叙述型	実体:N1	属性:主流	属性値:MT2N3	主語前:true
/N1(は が)、 ]/yf[MT2]図[MD3]	型名:属性叙述型	実体:N1	属性:図	属性値:MT2	主語前:true
/N1(は が)、 ]/yf[MT2]N3[MD4]	型名:外延叙述型	上位語:N1	下位語:N3	追加情報:MT2	主語前:true
/N1(は が)、 ]/yf[MT2]N3[MD4]	型名:外延叙述型	上位語:N1	下位語:N3	追加情報:MT2	主語前:false
/N1(は が)、 ]/yf[MT2]N3[MD4]	型名:外延叙述型	上位語:N1	下位語:N3	追加情報:MT2	主語前:false
/N1(は が)、 ]/[MT2]N3[MD4]	型名:外延叙述型	上位語:N1	下位語:N3	追加情報:MT2	主語前:true
/N1(は が)、 ]/[MT2]N3[MD4]	型名:外延叙述型	上位語:N1	下位語:N3	追加情報:MT2	主語前:true
/N1(は が)、 ]/[MT2]N3[MD4]	型名:外延叙述型	上位語:N1	下位語:N3	追加情報:MT2	主語前:false
/N1(は が)、 ]/[MT2]N3[MD4]	型名:外延叙述型	上位語:N1	下位語:N3	追加情報:MT2	主語前:false
/N1(は が)、 ]/yf[MT2]N3[MD4]	型名:範疇叙述型	下位語:N1	上位語:N3	追加情報:MT2	主語前:true
/N1(は が)、 ]/yf[MT2]N3[MD4]	型名:範疇叙述型	下位語:N1	上位語:N3	追加情報:MT2	主語前:false
/N1(は が)、 ]/[MT2]もの[MD3]	型名:範疇叙述型	下位語:N1	上位語:もの	追加情報:MT2	主語前:true
/N1(は が)、 ]/[MT2]もの[MD3]	型名:範疇叙述型	下位語:N1	上位語:もの	追加情報:MT2	主語前:false
/N1(は が)、 ]/yf[MT2]製品[MD3]	型名:範疇叙述型	下位語:N1	上位語:製品	追加情報:MT2	主語前:false
/N1(は が)、 ]/yf[MT2]シンボル[MD3]	型名:範疇叙述型	下位語:N1	上位語:シンボル	追加情報:MT2	主語前:false
/N1(は が)、 ]/yf[MT2]メッセージ[MD3]	型名:範疇叙述型	下位語:N1	上位語:メッセージ	追加情報:MT2	主語前:false
/N1(は が)、 ]/yf[MT2]重金属[MD3]	型名:範疇叙述型	下位語:N1	上位語:重金属	追加情報:MT2	主語前:false
/N1(は が)、 ]/yf[MT2]場所[MD3]	型名:範疇叙述型	下位語:N1	上位語:場所	追加情報:MT2	主語前:false
/N1(は が)、 ]/[MT2]機構[MD3]	型名:範疇叙述型	下位語:N1	上位語:機構	追加情報:MT2	主語前:true
/N1(は が)、 ]/yf[MT2]構造[MD3]	型名:範疇叙述型	下位語:N1	上位語:構造	追加情報:MT2	主語前:false
/N1(は が)、 ]/[MT2]N3[MD4]	型名:範疇叙述型	下位語:N1	上位語:N3	追加情報:MT2	主語前:true
/N1(は が)、 ]/[MT2]N3[MD4]	型名:範疇叙述型	下位語:N1	上位語:N3	追加情報:MT2	主語前:false
/N1(は が)、 ]/yf[MT2]N3の(一種  1種)[MD4]	型名:範疇叙述型	下位語:N1	上位語:N3	追加情報:MT2	主語前:true
/N1(は が)、 ]/yf[MT2]N3の(一種  1種)[MD4]	型名:範疇叙述型	下位語:N1	上位語:N3	追加情報:MT2	主語前:false

⋮

# 運用手順

入力例：  
テーブルは大理石製である。

名詞述語文

名詞述語文  
パターン辞書

パターン照合

型名  
3つ組

判断条件による  
絞り込み

型名  
3つ組

候補：

型名：属性叙述型  
3つ組：実体…テーブル  
属性…大理石製  
属性値…大理石製

型名：範疇叙述型  
3つ組：下位語…テーブル  
上位語…大理石製  
追加情報…NULL

型名：外延叙述型  
3つ組：上位語…テーブル  
下位語…大理石製  
追加情報…NULL

# 判断条件

分類	条件
共通	字面を残したパターンがマッチしている
属性叙述型	述語の意味制約コードが全て抽象を表す
	述語の接尾辞が"数え方の辞典"に存在 <small>数え方の辞典 (飯田ら 2004)</small>
	述語の品詞が名詞の中でも用言性名詞や転成名詞
外延叙述型	主語が述語の意味制約コード上で親
	述語にカギ括弧が存在 (固有名詞とみなす)
	述語がアルファベットの単語 (固有名詞とみなす)
	主語が形式名詞 (主語名詞の具体性が低い)
	主語の前に連体修飾節があり, かつ追加情報が無い
範疇叙述型	述語が主語の意味制約コード上で親
	主語にカギ括弧が存在 (固有名詞とみなす)
	主語がアルファベットの単語 (固有名詞とみなす)
	主語の前に指示詞が存在
	主語と述語に共通する意味制約コードが存在



14の判断条件を作成(素性にも利用)

# 判断条件

分類	条件
共通	字面を残したパターンがマッチしている
属性叙述型	述語の意味制約コードが全て抽象を表す
	述語の接尾辞が"数え方の辞典"に存在 <small>数え方の辞典 (飯田ら 2004)</small>
	述語の品詞が名詞の中でも用言性名詞や転成名詞
外延叙述型	主語が述語の意味制約コード上で親
	<p>偽造投票用紙は本物と同じ薄い<b>緑色</b>だ。</p> <p><b>緑色</b>：意味制約コードでは「色彩」</p> <p>外航貨物船はおよそ<b>二千隻</b>だ。</p> <p><b>隻</b>：船の数の単位。数量を表す</p> <p>最初の自動車は<b>手作り</b>である</p> <p><b>手作り</b>：用言性名詞</p>
範疇叙述型	主語と述語に共通する意味制約コードが存在



14の判断条件を作成(素性にも利用)



# 実験

- ・ 使用データ
  - Wikipediaデータ  
(36文, 正解データ39件)
- ・ 手法およびテスト形式
  - ルールベース：判断条件のいずれかを満たした候補を最終出力とする  
クローズドテスト
  - 機械学習：判断条件を素性に候補の絞り込み  
2分割のクロスバリデーション  
(パターンと判断条件の作成はクローズド)
- ・ 正解の基準
  - 型名, および, 3つ組の一致で正解

型名の内訳

型名	正解データ数
属性叙述型	8
外延叙述型	2
範疇叙述型	29

# 結果

$$\text{再現率 } R = \frac{\text{一致数}}{\text{正解データ数}}$$

$$\text{適合率 } P = \frac{\text{一致数}}{\text{出力数}}$$

$$\text{F値} = \frac{2RP}{R+P}$$

手法	入力候補数	正解データ数	出力数	一致数	再現率 $R$	適合率 $P$	F値
ルールベース	100	39	60	39	1.00	0.71	0.83
機械学習	100	39	46	35	0.90	0.76	0.82

- ルールベースの再現率が1.00であることを確認
- 機械学習では出力が抑えられ適合率が増加



# 不一致例

## 理想解

これを 1 本のケーブルで接続できるように端子を 1 つにまとめたものが D 端子 である。  
上位語 下位語

マッチパターン

```
/N1(は|が)[、 ]/yf[MT2]N3[MD4]
```

型名：外延叙述型  
主語前：True

## 出力例

これを 1 本のケーブルで接続できるように端子を 1 つにまとめたものが D 端子 である。  
下位語 上位語

マッチパターン

```
/N1(は|が)[、 ]/yf[MT2]N3[MD4]
```

型名：範疇叙述型  
主語前：True

## 原因

判断条件の最適化不足

# まとめ

- 名詞述語文を解析するためのパターン辞書を作成
  - 46件のパターンを作成
- パターン選択は良好

# 今後の課題

- 規模の拡大
- ニコニコ大百科を用いたオープンテスト
- 主語なし文や体言止めなどの文の対応