

# 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

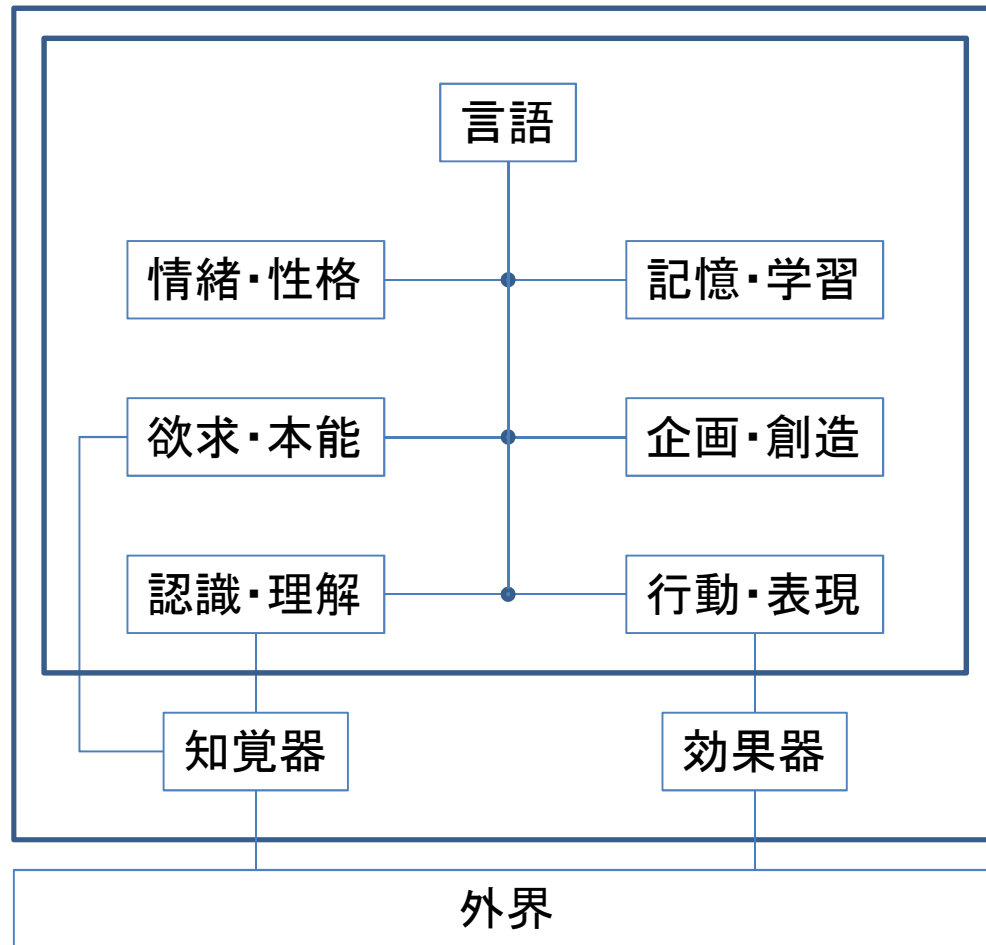
# はじめに

- **研究の目的**
  - 言語表現の意味処理
- **発表の概要**
  - 言語と意味
  - パターン辞書
    - 鳥バンク
    - 日本語語彙大系：感情推定用, 存在性理解用
  - 後半は藤原君

# 言語と意味の研究

- 岡田直之：語の概念の表現と蓄積，電子情報通信学会，コロナ社，1991.
  - 心のモデル：機能と領域，情報の抽象度と格納
- 意味をとらえる...
  - 心のモデルのどこかと連鎖関係にある.
  - 意味処理には心のモデルを備えたエージェントが必要
  - 真正面から取り組む.
    - 工学的モデルだが，実現まで少し距離がある.

# 心のモデル



AI 研究は, いずれかの領域

# 言語と意味の研究

- 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦: 日本語語彙大系, 岩波書店, 1997.
  - 意味処理 = 意味解析, 意味理解
    - 意味解析 = 表現から規定どおりに解釈
    - 意味理解 = それ以上に解釈 (池原JSAI誌 AIマップ2001)
- **意味をとらえる ...**
  - 意味解析の目的が決まっていれば, そのための情報を表現に対応付けた知識ベースを利用
  - 工学的アプローチ

# 機械翻訳を目的とする場合

- 日本語語彙大系

N1が N2に出る

[N1(529月, 530太陽, 531星)  
N2(388場所, 2610場)]



N1 come out in N2

15自然現象

N1が N2に N3から/より である

[N1(1190金銭 1194給与 1195報酬  
1196年金 1199料金)  
N2(3主体 1236人間活動)  
N3(3主体)]



N1 be provided for N2 by N3

16所有的移動

# 日本語語彙大系

- **日本語の用言述語の単文をパターン化**
  - 基本的用言 約6,000語
  - 約 14,800パターン
  - 日本語から英語への翻訳用のパターン対

# 鳥バンク

- 2002年1月～2007年3月
  - 池原悟：セマンティック・タイポロジーによる言語の等価変換と生成技術, CRESTプロジェクト.
- **日本語の重文複文のパターン化**
  - 日英対訳文対. 約12万文対
  - パターン対. 約23万パターン対



X-SPM

File Data

【文型パターン辞書】

レベル： 全 パターン数： 227,990

【検索キー】

日本語文： 5年かけて言語知識ベースを開発した。

消去 検索

kterm

```

**[time=0.20s]
@1
i
have
bet in five years      betted in five years      spent five years      run
in five years      lacked in five years
developing      exploiting      opening opening up
a language knowledge basis
.
@2
-----
-125.482505822728

```

Search Result

File Data

検索結果： 日本語パターン数 [延べ3334/異なり1612] 表示範囲：1 ~ 30 件目

順位：パターンID (適合率)

解析された部分  
適合パターン  
原文

1 : AD017255-00 (100.0%)

5年かけて言語知識ベースを開発した。

ADV1V2 N3 V4

日文型： /y</ADV1/ /V2(て/で) /ytk N3を/ cf V4. kako.

英文型： <N1 V2<sup>past</sup> to V4<sup>base</sup> N3<sup>obj</sup> .

日原文： みんな立ち上がって彼を迎えた。

英原文： All rose to welcome him.

1 : AM013863-00 (100.0%)

5年かけて言語知識ベースを開発した。

ADV3V4 N5 V6

日文型： /y</tkN1は> #1[ /cf ADV3] /f (V4/ND4をし) (て/で) /ytk

英文型： <I/N1> V4<sup>past</sup> #1[ADV3] to V6<sup>base</sup> N5 .

日原文： なんとか工面して授業料を払った。

英原文： I managed somehow to pay the tuition.

1 : AN071521-00 (100.0%)

5年かけて言語知識ベースを開発した。

N3 V4 N7 V8

意味属性縛り AND検索 重複カット

Search Result

File Data

検索結果： 日本語パターン数 [延べ194/異なり93] 表示範囲：1 ~ 30 件目

順位：パターンID (適合率)

解析された部分  
適合パターン  
原文

1 : AC032920-00 (100.0%)

5年かけて言語知識ベースを開発した。

VP1 NP3 V4

日文型： /y VP1(て/で) </yN2は> /NP3を/ cf V4. kako.

英文型： <I/N2> have VP1<sup>ed</sup> V4<sup>ing</sup> NP3 .

日原文： 1か月かけてこの計画を練り上げた。

英原文： I have spent a whole month working out this plan.

1 : AD014799-00 (100.0%)

5年かけて言語知識ベースを開発した。

VP2 VP3

1 : AF017135-00 (100.0%)

5年かけて言語知識ベースを開発した。

VP2 VP3

Selector for Semantic Attribute Con

縛りに使用する意味属性

- NI:日本語語彙大系 一般名詞意味属性
- NY:日本語語彙大系 用言意味属性
- IM:単文意味分類体系 名詞意味属性
- IV:単文意味分類体系 用言意味属性
- KR:講談社類語大辞典

縛りの強さ

- ◆ 同一または下位属性
- ◇ 1つ上位の属性と同一またはその下位属性
- ◇ 2つ上位の属性と同一またはその下位属性
- ◇ 3つ上位の属性と同一またはその下位属性
- ◇ 4つ上位の属性と同一またはその下位属性

絞り込み実行 キャンセル

Translation Result

英語パターン

<I|N2> have VP1<sup>ed</sup> V4<sup>ing</sup> NP3 .

翻訳結果

I have spent five years developing a language knowledge basis .

ITM Reboot

Bi-gram Tri-gram Translate

# 翻訳の方法

1. 入力文と日本語パターンの照合
2. パターン選択
  - マッチ数(字面優先...), 意味属性制約条件
3. 英語パターンの取得
4. 変数部への英訳挿入
  - WORD=単語辞書, PHRASE=日本語語彙大系(ALT)
5. 英単語n-gramによる選択
6. 英文の出力

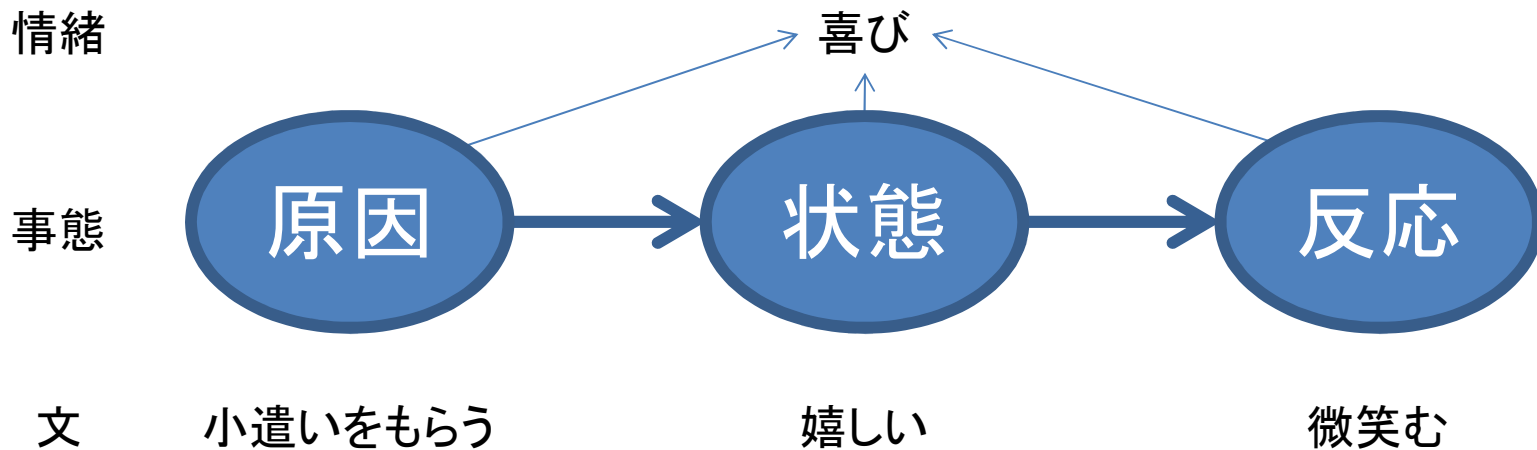
# 意味理解

- 感情(情緒)の推定
- 場所の理解 ... 開発中

# 情緒の推定

## • 事態と情緒

- 情緒の生じる原因となる事態が文に描かれる
- 情緒の状態が文に描かれる
- 情緒の表出や反応となる事態が文に描かれる



# 情緒推定用のパターン辞書

日本語語彙大系 用言パターンに, 情緒属性を追加

用言: 貰う  
パターン: N1が N2を N3に/から/より 貰う  
意味属性制約: N1(3主体) N2(\*) N3(3主体)



情緒名: 喜び  
原因: 獲得  
情緒主: N1  
情緒対象: N2  
条件: 目標実現・  
接近(N1, N2)

セット1



情緒名: 好ましい  
原因: 厚遇  
情緒主: N1  
情緒対象: N3  
条件: 対人関係・  
上下・接近(N1, N3)

セット2



情緒名: 悲しみ  
原因: 内的な不快  
情緒主: N1  
情緒対象: N2  
条件: 生理・  
乖離(N1, N2)

セット3

# 選択問題

- 目標実現・接近(人, 小遣い)  $\Leftrightarrow$  True
- 目標実現・接近(人, ゴミ)  $\Leftrightarrow$  False
- 生理・乖離(人, 病気)  $\Leftrightarrow$  True
- 目標実現・接近(まる子, 病気)  $\Leftrightarrow$  True  
マラソン大会の前日なので病気で休みたい



- 命題関数の充足を検査する問題
  - 目標実現・接近( $x, y; context$ )

多数の意見, 形容詞, 季節  $\rightarrow$  コーパスとSO-Scoreで解決.  
個人の事情, 特定の文脈  $\rightarrow$  開発中

# 場所の理解

- 例題

- 太郎が図書館で本を借りた。
- 太郎が家に帰った。

- 本はどこに有るか？

1. 本が図書館にある
2. 本が家にある
3. 本が太郎のところにある

# 動詞から直接いえること

タイミング	太郎が図書館で本を借りる	太郎が家に帰る
直前	太郎が図書館に存在する 本が図書館に存在する	太郎が図書館に存在する 本が太郎のところに存在する
最中	本が太郎のところに存在する	
完了	本が図書館に存在する	太郎が家に存在する

## パターン

$N1$ が  $L$ で  $N2$ を  $N3$ から 借りる



## 存在属性

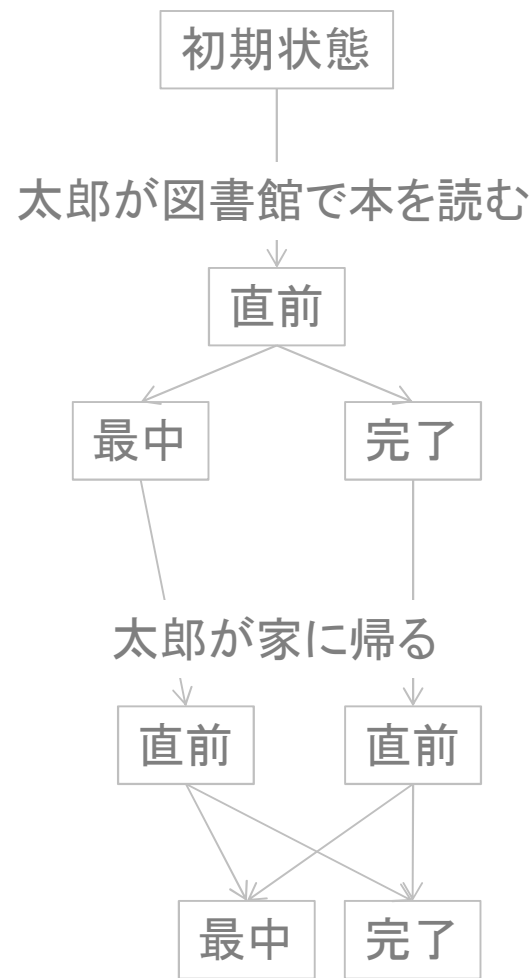
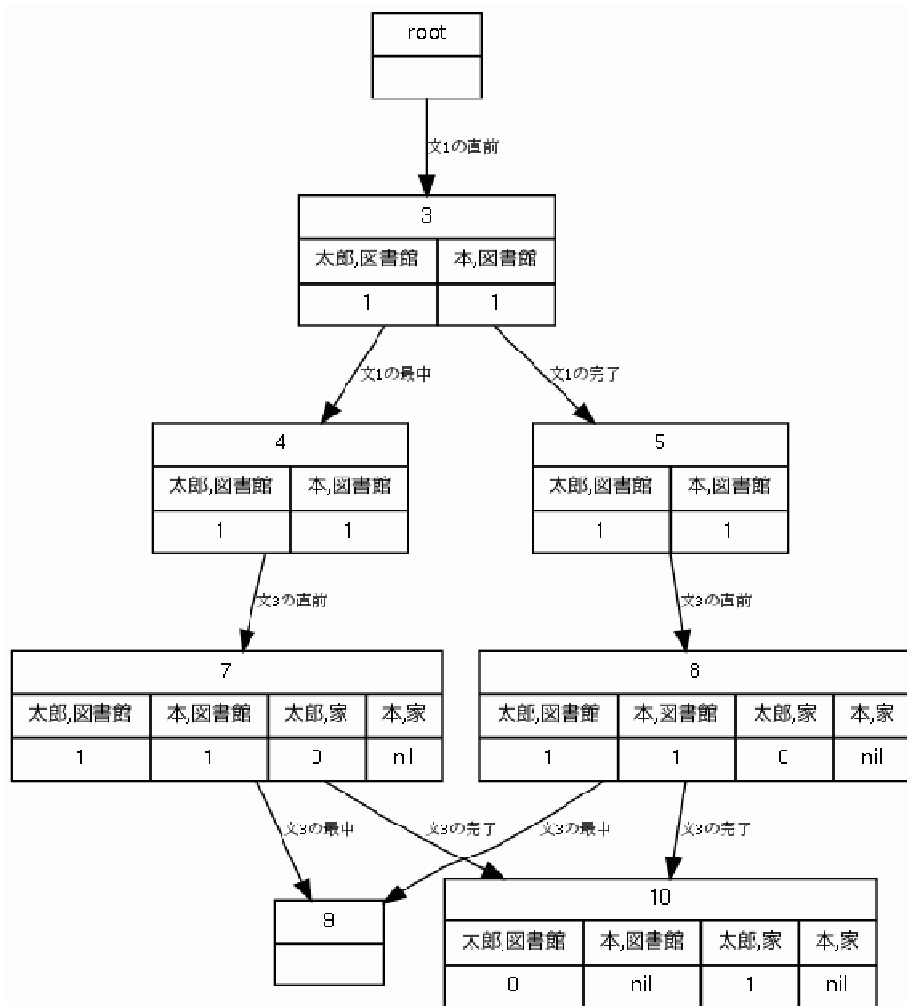
直前:  $N1$ が  $L$ に存在する  
 $N2$ が  $L$ に存在する  
 $N2$ が  $N3$ に存在する



# 常識的に補完すること

- 直前のタイミングでは，前文の状態を継承
  - 非移動の動詞では，位置を保持
  - 移動の動詞では，位置を非保持
  - 所有物は，移動の際，同時に位置が変化
- ：
- ：

# 実行の様子(開発中)



# 中まとめ

- パターンのマッチングで意味を処理する.

パターン辞書の基礎	カバー	応用
鳥バンク	日本語の重文複文 (述語2~3)	翻訳
日本語語彙大系	日本語の単文 (用言述語文)	翻訳 情緒推定 場所の理解

- 日本語の単文(名詞述語文)が未着手  
– 藤原竜樹くんへ