

事態オントロジー構築のための 知識獲得

阿部修也
奈良先端科学技術大学院大学

2006-09-14

1

目的

- ・コーパスから事態間の関係を獲得
- ・事態間の関係
 - 動機: 病気になる→病院へ行く
 - 目的: 病院へ行く→治療する
 - 部分: 寝る→躰をかく
 - 条件: 電車に乗る→切符を買う
 - 上位: 疾走する→走る
 - 結果、下位...

2006-09-14

2

関連研究

関係抽出の副産物 (Lin2001, Pantel2006)

Shakespeare wrote King Lear.
 ||
 Shakespeare is the author of King Lear. ⇨ "wrote" = "is the author of"

共起パターン (乾2004, Torisawa2006)

晴れていたため洗濯物が良く乾いた

2006-09-14

3

事態含意名詞

- ・事態含意名詞: 事態を含意する名詞
 - ・ 運動会→運動する、掃除機→掃除をする、勤務中→勤務する、犯人→罪を犯す
- ・事態含意名詞を含むパターン
 - 事態含意名詞と動詞(句)の共起
 - ・ 「掃除機でゴミを吸う」→「掃除する」と「ゴミを吸う」
 - 事態含意名詞と事態含意名詞の共起
 - ・ 「飛行機の離陸時間」→「飛行する」と「離陸する」

2006-09-14

4

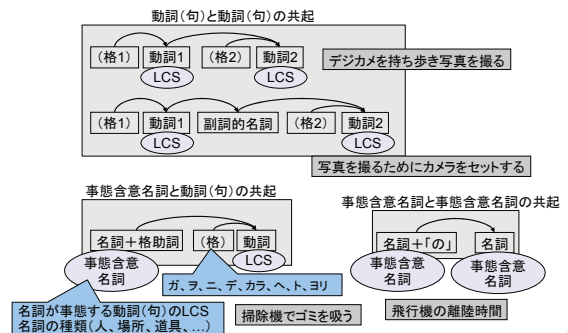
今日の話

- ・ 事態含意名詞を含むパターン
- ・ 事態含意名詞辞書の作成
- ・ 動作主性辞書の作成
- ・ 事態含意名詞を利用することの効果を予備調査

2006-09-14

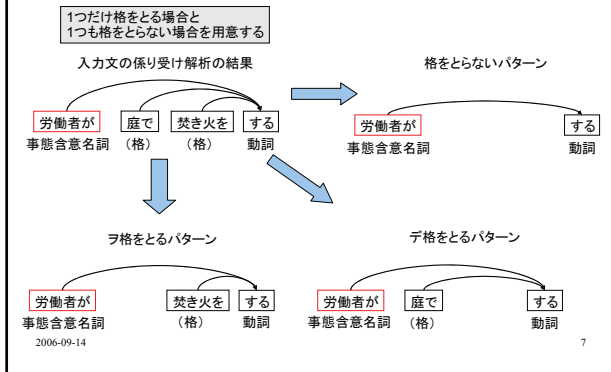
5

共起パターン



6

共起: 格の取り方



事態含意名詞を作成する方法

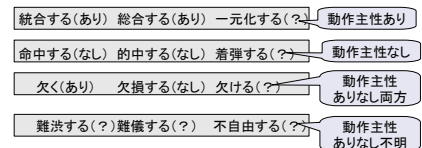
- サ変名詞+接尾辞→事態含意名詞
 - 「出張」+「先」→出張先
 - ・ 事態含意名詞が含意する事態: 出張する
 - ・ 事態含意名詞の種類: 場所
 - 事態含意名詞のクラス
 - ・ 事態含意名詞の動詞の部分→動作主性の有無
 - ・ 接尾辞→事態含意名詞の種類
 - 接尾辞と事態含意名詞の種類
 - 道具: 器、具、用、装置、機、書 (6種類)
 - 場所: 先、場、所、室、店、部屋、... (13種類)
 - 人: 人、員、業、者、作、業、者、者、... (19種類)
 - 事態: (接尾辞なし)、運動、活動、権、... (26種類)
 - 期間: 前、後、中、... (9種類)
- 2006-09-14 8

事態含意名詞の例

- 道具: マッサージ器
 - 場所: マッサージ部屋
 - 人: マッサージ師
 - 事態: マッサージ
 - 期間: マッサージ中
- 2006-09-14 9

事態クラスの辞書(動作主性)

- 最新のLCS辞書と分類語彙表
- 分類語彙表の類義語を用いてLCS辞書の内容を拡張
 - 動作主性が付与されていない語については次の規則で動作主性を付与



Web5億文コーパス

- Web5億文コーパス(河原,黒橋2006)
 - ・ 動詞と動詞の共起: 63,815,668
 - ため(に): 1,341,216
 - ・ 事態含意名詞と動詞の共起: 29,370,015
 - で: 1,732,431
 - が: 5,320,099
 - へ: 69,417
 - から: 433,635
 - に: 7,220,443
 - を: 14,544,104
 - より: 49,886
 - ・ 事態含意名詞と事態含意名詞の「の」の共起: 656,359
- 2006-09-14 11

スコア

- 基本的にはPMI
- ただし、事態含意名詞と動詞の共起が多い場合にバイアス

$$\log \frac{P(N \cap V)^2}{P(N)P(V)}$$

2006-09-14

12

事態タイプの制約

<事態含意名詞>で<動詞>

完納	で 公売	を 中止する
複合問題	で 実力	を 養成する
配列状態	で 塩輪群	が 出土する
転位	で 強さ	を 述べる
動語関係	で 物入り	になる
回復寸前	で 人工呼吸	が 止まる
破産者	で 復権	を 得る
開業後	で 税務申告	を 終える
保全計画	で 特別地区	に 係る
抜刀隊	で 切り込み	を 行う
研究中	で テーマ	が 飛び出す
自署	で 氏名	を 記載する
専攻間	で <数詞>志望	を 認める
印刷装置	で 有機材料	を する
マッサージ器	で	責める
賃借権	で 借地権者	が 設定する
積算価格	で 十分採算	が 取れる
横ばい	で	推移する
講習時間	で 技能	を 学ぶ

2006-09-14

13

事態タイプの制約

<事態含意名詞>で<動詞>

完納	で 公売	を 中止する
複合問題	で 実力	を 養成する
配列状態	で 塩輪群	が 出土する
転位	で 強さ	を 述べる
動語関係	で 物入り	になる
回復寸前	で 人工呼吸	が 止まる
破産者	で 復権	を 得る
開業後	で 税務申告	を 終える
保全計画	で 特別地区	に 係る
抜刀隊	で 切り込み	を 行う
研究中	で テーマ	が 飛び出す
自署	で 氏名	を 記載する
専攻間	で <数詞>志望	を 認める
印刷装置	で 有機材料	を する
マッサージ器	で	責める
賃借権	で 借地権者	が 設定する
積算価格	で 十分採算	が 取れる
横ばい	で	推移する
講習時間	で 技能	を 学ぶ

2006-09-14

14

<動作主性+場所>で<動詞(動作主性)>

勘定場	で 密	が 一括する
経由先	で Ethernet	を 選択する
納骨堂	で	を 献花する
転位	で 服	を 脱ぐ
脱衣所	で	を 脱ぐ
慰安所	で 性的強制	が 行う
治療室	で 治療法	を 行う
トレーニング室	で シェイプアップ	に チャレンジする
運転室	で	を 操縦する
加盟国	で 改正条約	を 批准する
従業先	で	を 仕事する
勘定場	で	を 一括する
精算所	で きっぷ	を 買う
理髪店	で	を 散髪する
焼却場	で	を 燃やす
研修室	で <数詞>回 定例会	が 開催する
加盟国	で	を 批准する
投票所	で 係員	に 申し出る
焼却場	で 焼却処分	を 行う

ノイズが少ないパターン(1/2)

- ・ 事態含意名詞と動詞の共起について比較的ノイズが少ないパターンを人手で確認した。

<事態含意名詞>
(<動詞,接尾辞>) + <格助詞> + <動詞(句)>

動作主性あり	が、ヲ、ニ、デ、カラ、ヘト、ヨリ	動作主性あり
名詞の種類(人、場所、道具、...)		

2006-09-14

15

ノイズが少ないパターン(2/2)

- <行為>+で+行為(動詞)
 - ・ 結婚で退社する、競売で落札する
- <行為者>+が+行為(動詞)
 - ・ 旅人が泊まる、歩行者が横断する、通行人が足を止める
- <機能的空間>+で+行為(動詞)
 - ・ 洗面所で歯を磨く、駐車場で車を止める、研究室で実験する
- <行為時>+に+行為(動詞)
 - ・ 運転中に事故を起す、撮影時にフィルムを取り出す、購入前に試す
- <道具>+で+行為(動詞)
 - ・ 洗濯機で洗う、洗濯機で脱水する、コピー機で複写する

2006-09-14

16

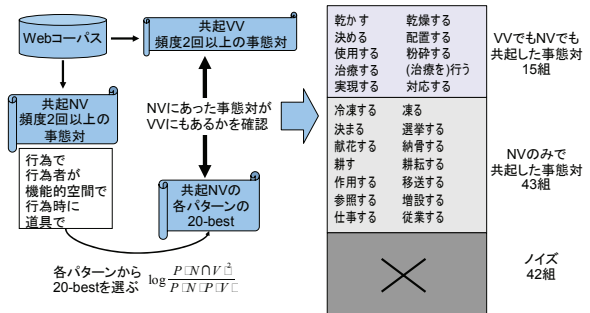
行為+で+行為(動詞)

マッサージ器	で	責める
転換装置	で 影響	を 及ぼす
掃除機	で	を 吸い取る
消臭装置	で 熱	を 加える
転換装置	で	を 及ぼす
乾燥機	で	を 乾かす
送風機	で 風	を 送る
人力装置	で 地上遠隔操作	を 行う
搬送装置	で 業界	を リードする
聴診器	で 音	を 聞く
注射器	で	を 注入する
加湿器	で 湿度	を 上げる
製氷機	で 水	を 作る
小便器	で 用	を 足す
除雪機	で 雪	を 飛ばす
コピー機	で	を 複写する
積算機	で 区実士休券	を 作る
洗面器	で お湯	を かける
洗面器	で	を すすぐ
測定器	で	を 測る

2006-09-14

17

事態含意名詞を用いて獲得できる事態対



2006-09-14

18

まとめ

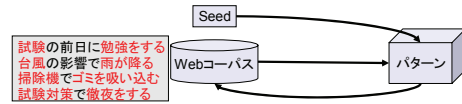
- ・ 動詞(句)と動詞(句)の共起だけではあらゆる事態の関係をカバーできない。事態含意名詞との共起を用いることで、一部の事態の関連をカバーすることができるようになる。
- ・ 事態含意名詞や動詞(句)の種類、共起パターンを適切に決めることで、ノイズを減らして事態間の関係を獲得できる可能性がある。

2006-09-14

19

今後

- ・ パターンをブートストラップ的に学習する方法を試みる。(Pantel2006)



2006-09-14

20