

言語処理を意識した 語彙概念構造の構築

岡山大学 工学部
情報工学科 竹内孔一

2005/3/23 COE 1

訂正

- p.34 L17 [x CONTROL [x MOVE FROM z]] 要求,求める
-> あがる

2005/3/23 COE 2

やってきたこと

- 岡山大学と国立情報学研究所
 - 漢語動詞を中心としたLCS辞書(約1300語)
 - 人手で構築
 - 言語処理での利用を設定して作成
 - 岡大 Webサイトで公開中
 - <http://cl.it.okayama-u.ac.jp> または「語彙概念構造」とGoogleで検索
 - 無料, 改変, 再配布自由, 営利/非営利は不問

なぜLCS辞書を作成したか どのように作ったか どう使いそうか

2005/3/23 COE 3

	A	B	C
1	接続	せつぞく	[]x CONTROL [BECOME []y BE A'
2	違反	いはん	[BECOME []y BE AT []z]
3	関連	かんれん	[]y BE AT []z]
4	発足	はつそく	[x=y CONTROL [BECOME []y BE A
5	移籍	いせき	[x=y CONTROL [BECOME []y BE A
6	競争	きょうそう	[]x ACT ON []y]
7	発売	はつばい	[]x CONTROL [BECOME []y BE A'
8	延期	えんき	[]x CONTROL [BECOME []y BE A'
9	急騰	きゅうとう	[BECOME []y BE AT [FILLED]z]]
10	解除	かいじょ	[]x CONTROL [BECOME []y NOT
11	公募	こうぼ	[]x ACT ON []y]
12	申告	しんこく	[]x CONTROL [BECOME []y BE A'

2005/3/23 COE 4

今日お話ししたいこと

- 言語処理におけるLCSの必要性
 - 言語処理の現状
- TLCSの構築
 - 性質を整理
 - 動詞に付与
 - 記述内容の整理
 - 問題点
- まとめ

2005/3/23 COE 5

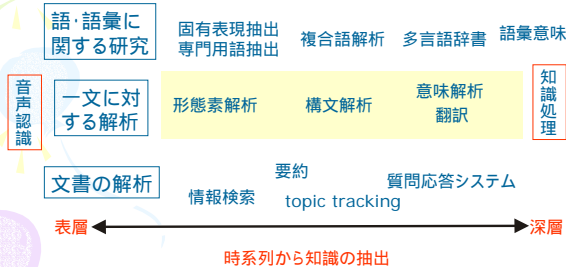
言語処理の必要性

- インターネットの普及
 - 電子化された文書の爆発的蓄積
 - アクセス可能な文書の内容
 - 技術情報, 店舗情報, 意見, 論文, 経験
 - メール, 辞書 (biology分野)
 - 期待
 - 知識は言葉による表現が中心
 - 言語処理による知識の利用

収集, 再利用

2005/3/23 COE 6

言語処理研究の全体図



2005/3/23 COE

7

言語処理のアプローチ

- 文の構造から こちらを重視
 - 構文解析->談話解析, RST
 - 文の構造を解析して文内の係り関係を解析
 - 文同士の関係を解析
 - 利点: 文の細かな意味を獲得(見込み)
 - 欠点: 応用により設計が変わる
- 単語の集合(文書群)
 - 単語, 複合語, 句を辞書登録しておき数え上げによる関係を利用
 - 利点: 規則外の関係を利用できる
 - 欠点: 構造を無視-> 大きく間違える

2005/3/23 COE

8

言語処理の現状

- 言語資源・ソフトウェア
 - 構文解析・形態素解析 (辞書も公開)
 - Cabocha, ChaSen (奈良先端 松本研究室)
 - KNP, JUMAN (東大 黒橋先生, 京大 佐藤研究室)
 - MSLR parser (東工大 田中・徳永研究室)
 - 電子化辞書
 - EDR(有料), IPAL 語の意味, 構文に基づく語間の関係, 角川類義語辞典
 - 英次郎 (日英翻訳辞書)
- 構文解析(係り関係の解析) → 意味 → 応用

2005/3/23 COE

9

言語処理の現状

- 構文解析
 - かなりの精度で1文の係り関係を提示(例: KNP)
- 格文法は
- 本質的に
- 統語規則と<P>
意味規則を<P> PARA
- 共存させた
- これより先の処理->意味
応用にあわせて
単語間の関係を再利用
可能な形で持っておく
- 文法であり、
日本語の
解析に
広く
用いられている。

2005/3/23 COE

10

LCSの必要性

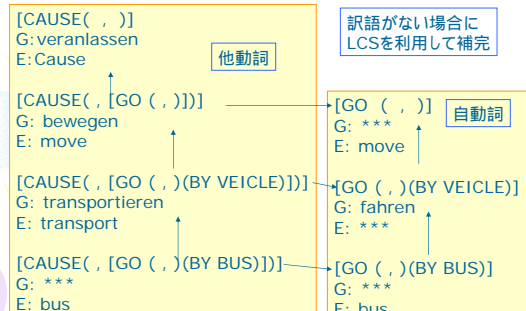
- 単語間の関係を整理
 - > 知識利用に貢献
- 必要な性質
- 網羅性がある体系
 - 構文解析と対応がある
- LCS (GL, FramNet)
- 表現の多様性の吸収例
「用いる」のLCS
[x CONTROL [y BE AT z]]
(y: 格文法, z: 日本語の解析)
同じグループの動詞・サ変名詞
「利用, 使う, 使用」
- 質問応答の例
Q. 格文法は[何に]使われていますか?
[x CONTROL [格文法 BE AT [日本語の解析]]] → A. 日本語の解析

2005/3/23 COE

11

網羅的な体系の利用例

- Dorr による英独LCS翻訳



2005/3/23 COE

12

言語処理とLCS

- 構文解析との対応
 - LCSの**構築**: 文の表現からLCSが推測できる
 - 完全に直感だけの意味辞書と異なる(**再現性**)
 - LCSの**利用**: 構文解析が成功すればLCSを通して抽象化した意味体系に変換できる
- 処理からLCSに求めること
 - 網羅性と全体系の提示, 構文解析との対応

2005/3/23 COE

13

LCSの構築

2005/3/23 COE

14

LCS構築の問題

- 発展中
 - より詳細な分析が続けられている
 - 記述体系が成長している
 - 様々なアイデアが盛り込まれている
 - 項構造, 使役, 自他交代, アスペクト分析
 - **粒度をどの程度で抑えるか** <-> 応用
 - 様々なLCS

2005/3/23 COE

15

LCS構築で目指すこと

- 盛り込むアイデアを整理
 - 粒度をそろえた体系作りの準備
- 網羅的なLCS体系を構築
 - 動詞・サ変名詞に対して付与 **深さより幅を優先**
 - 構築したときの判断基準を公開
 - 信頼度が利用者につかめるようにする
- LCS体系がどのように使えるか
 - 処理モデルを意識して構築

2005/3/23 COE

16

LCS構築のステップ

- LCS構築に向けての順序
 - 第一次 **今はここ**
 - 目標: 多くの動詞をみて体系を**網羅的に作る**
(理論では見つかっていないような動詞の振る舞いを探し整理する)
 - 第二次
 - 目標: **外の関係を明確にする** **アスペクト分析参照**
(多義のあるLCSに対してどのような条件でLCSが異なるかを記述する)

2005/3/23 COE

17

TLCSの構築

- 3つのアイデア **必要性からの説明**
 - 項構造
 - 必須項の決定とその役割
例) 編集する (x <y>) x: 動作主 y: 対象
 - アスペクト分析
 - 動詞の持つ時間変化に対する分析
→ **網羅性**: 変化する/しない
 - 概念の意味構造
 - 抽象的な事象構造を表現する
 - 例) 編集する
[x CONTROL [BECOME [y BE AT z]]]

2005/3/23 COE

18

項構造

- 問題
 - ↳ (1) 意味役割はいくつ? (2) どれが必須の名詞か?
- 提案
 - (1) Agent Theme Goal Source と限定
 - (2) 必須項の判定法
 - 必須項: 動詞に直接関係する名詞 → 態の入れ換え
 - (例) 文を英語に翻訳した/文が英語に翻訳された
道路を歩く/ *道路が歩かれ 必須項
 - (例) 風呂から上がる/風呂を」 必須項
東から風が吹く/*東に風が吹く

2005/3/23 COE

19

アスペクト分析

- 時間に基づく動詞の分類
 - ↳ Vendler の4つの分類をベースとした
 - 状態動詞 (維持, 位置) BE
 - 活動動詞 (操作, 行列) ACT
 - 到達動詞 (移動, 飽和) BECOME | MOVE
 - 達成動詞 (編集, 処理) CONTROL と BECOME
 - 動詞のとり名詞間の関係を記述
- テスト
 - 「10分間」「10分で」、「である」、「たくさん」

2005/3/23 COE

20

アスペクト分析

- 多義について
 - とり名詞によってアスペクトが異なる
 - (例) 杖をにぎる/おにぎりをにぎる
ピアノを弾く/ソナタを弾く
 - LCSの違い → 語義が違うのでは
 - にぎる(活動) [x ACT ON y]
 - にぎる(達成) [x CONTROL [BECOME [y BE AT z]]]
 - x, y, z の定義域を記述 → 外の関係

2005/3/23 COE

21

概念の意味構造

- 付与手順
 - 項構造, アスペクト分析
 - 項構造で得た名詞間の関係を意味述語で結びつける
 - 解釈で構築する → (テストはない)
- 意味述語
 - 11種類に限定
 - CONTROL, BECOME, MOVE, BE, AT, TO, FROM, WITH, NOT, ACT, ON
 - その他
 - 能格動詞 (x=y), 定義域の埋め込み (FILLED)

2005/3/23 COE

22

TLCS付与からの知見

- TLCSの付与
 - 項構造, アスペクト分析, 概念の意味構造
 - 約1300語に付与
 - 多くの動詞は設定した枠組み内
- 問題
 - 二格に関する問題
 - 概念意味の特異性

2005/3/23 COE

23

二格に関する問題

- 以下の3つを取り上げる
 1. HがGに 賛成/反対する
 - GがHに 賛成/反対される (受身で変わらない)
 2. HがGに 会う/ぶつかる
 - 受身はなく, HとGが等価に入れ替わる
 3. HがGに 感動する/悩む (心理動詞)
 - GがHを感動させる (受身なし)
- 問題点
 - 二格を項と認めるべきか 項である
 - 項ならば意味役割は何か? 構造はどうなるか

2005/3/23 COE

24

二格に対する分析

- HがGに反対する
 - 他動詞, 受身形がある
- 項構造: (Agent Theme)
- アスペクト分析: 活動動詞
- TLCS: [x ACT TO y]
- 問題点
 - [TO y] を積極的にサポートする根拠を見つけていない
 - [x ACT ON y]との違い **なぜ二格?**

2005/3/23 COE

25

二格に対する分析

- HがGに会う/ぶつかる
 - 受身形がない -> 非対格自動詞
- 項構造: (Theme (Goal))
- アスペクト分析: 状態動詞
- TLCS: [y BE WITH z]
 - HがGに会う/という -> GがHに会う/という

2005/3/23 COE

26

二格に対する分析(心理動詞)

- HがGに感動する/悩む
 - 受身形が無い → 非対格自動詞
 - 使役形で英語などと同じ格関係
 - 映画が彼を感動させた
 - 感動させる (英語などの分析)
- 項構造: (Theme (Goal))
- アスペクト分析: 到達動詞
- TLCS: [BECOME [x BE WITH [[FILLED]y BE TO z]]]

2005/3/23 COE

27

概念的意味の特異な場合

- XがYをZに 授かる/教わる/要求する
- 問題点
 - 「Zに(与える)」と同じ表現で方向が逆
- X (Agent/Goal) ←----- Z (Source)
Y (Theme)
- 英語では I received a letter from him.
- 分析対象
 - 授かる動詞 非対格動詞
 - 要求動詞 他動詞

2005/3/23 COE

28

授かる動詞

- 分析
 - 非対格自動詞 → Agent項が無い
 - 言い換え カラと言い換えられる
 - 太郎が先輩に/から数学を教わった
- 項構造: (Goal (Theme (Source)))
- アスペクト分析: 到達動詞
- TLCS: [BECOME [z BE [WITH [y MOVE FROM x TO z]]]]
- ポイント
 - リンキングルールで FROM x は二格
 - 授ける: [x CONTROL [y MOVE TO z]]

2005/3/23 COE

29

要求動詞

- 分析
 - 他動詞 → Agent 項がある
 - 使役構造「させ」と似ている
 - (要求)ファンが松井にサインを要求する
 - (使役)ファンが彼に編集させる
 - 使役の二格: 遠くから影響を与える Source
- 項構造: (Agent (Theme (Source)))
- アスペクト分析: 達成
- TLCS: [x CONTROL [y MOVE FROM z TO x]]]

2005/3/23 COE

30

要求動詞

- 問題点
 - サ変名詞を取り込み複雑な事象を記述
例) 監督が松井にトレーニングを要求した
 - さらに二格をとる
例) 監督が松井にトレーニングするように要求した

未解決

2005/3/23 COE

31

TLCSパターン

- | | |
|---------------------------------------|---------|
| 1 [x ACT ON y] | 操作, 運転 |
| 2 [x CONTROL[BECOME [y BE AT z]]] | 処理, 翻訳 |
| 3 [x CONTROL[BECOME [y NOT BE AT z]]] | 遮へい, 抑止 |
| 4 [x CONTROL [y MOVE z]] | 伝送, 伝搬 |
| 5 [x=y CONTROL[BECOME [y BE AT z]]] | 回復, 終了 |
| 6 [BECOME[y BE AT z]] | 飽和, 分布 |
| 7 [y MOVE z] | 移動, 遷移 |
| 8 [x CONTROL[y BE AT z]] | 管理, 維持 |
| 9 [x CONTROL[BECOME[x BE WITH y]]] | 認識, 予測 |
| 10[y BE AT z] | 関係, 位置 |
| 11[x ACT] | 会議, 行列 |

2005/3/23 COE

32

TLCSパターン

- 12[x CONTROL[BECOME [[NAME OF x]y BE AT z]]]
署名
- 13[x CONTROL[BECOME[x BE WITH [[FILLED]z BE AT y]]]
尊敬, 信用
- 14[BECOME[y BE WITH [[FILLED]y BE TO z]]]
感動
- 15[BECOME[z BE WITH [y MOVE FROM x TO z]]]
受ける, 授かる, 教わる
- 16 [x CONTROL [y MOVE FROM z TO x]]]
要求, 求める
- 17 [x CONTROL [x MOVE FROM z]]
上がる
- 18 [x ACT TO y]
賛成, 反対

2005/3/23 COE

33

TLCSの利用

- TLCSから取り出せる情報
 - 他動詞か自動詞か
 - 状態, 活動, 到達, 達成のどのタイプか
 - 必須項と表層格の結びつき
 - 外項か内項が取り出せる
 - 複合名詞解析

2005/3/23 COE

34

まとめ

- 言語処理からLCSの必要性
 - 構文解析の次の処理
- TLCS
 - 項構造, アスペクト分析, 概念の意味構造
- 付与
 - 二格に対する分析
 - 要求動詞に関する分析

2005/3/23 COE

35

Thank you!

2005/3/23 COE

36