

意味フレームを用いた 言語の知識構造への効果的な 結びつけ

黒田 航 井佐原 均

独立行政法人 情報通信研究機構

けいはんな情報通信融合研究センター

{kuroda,isahara}@nict.go.jp

電子情報通信学会 「言語理解とコミュニケーション」 研究会

11/04/2004

目標

- 📌 **Goal: 意味役割タグつきコーパスの開発**

- 📌 **Subgoal: 意味役割タグ体系の定義**

- 📌 Berkely FrameNet のアプローチを参考に

- 📌 **言語学と自然言語(情報)処理の融合**

- 📌 私は(いちおう)言語学者なので工学系の研究に関する
“勉強不足”は大目に見て下さい

- 📌 “以前, 似たような話があったけど, うまくいかなかった”
系の情報提供, 歓迎します

意味タグづけは可能か

- 率直に言って大変は大変だが, 理論的に不可能なわけではない (品詞タグづけが可能なのだから)
- 本質的な困難は意味タグ体系の定義
 - 意味型ベースの意味タグ体系 (e.g., WordNet) は面白くないし, 強力な推論機構がないと実用性に乏しい
 - 意味役割 = 意味フレームの要素を基本にしたタグ体系 (cf. BFN)

意味フレーム分析の基礎

- 📌 意味フレームとは幾つかのフレーム要素 (Frame Elements: FEs) からなる概念構造体
- 📌 狭い定義: 意味フレーム = “状況”の理想化
- 📌 広い定義: 部分/全体のような“構成関係”が定義しうるものなら何でもフレーム. ただし,
 - 📌 状況レベルのフレーム
 - 📌 個体レベルのフレーム (= モノの構成, 構造; ただし用途, 機能は含まず!)

分析例:“x が y を襲う”

● F10: 台風<<自然災害(大型)による被害の発生

● 大型の台風が九州(の人々)を襲った

● ??大型の台風がインパラの群れを襲った

● F01: 動物の捕食目的の攻撃

● 空腹のライオンがインパラの群れを襲った

● ???空腹のライオンが銀行を襲った

● F03: 銀行強盗<<資源強奪<<ヒトのヒトへの攻撃

● 覆面の男が銀行を襲った

● ??覆面の男が九州(の人々)を襲った

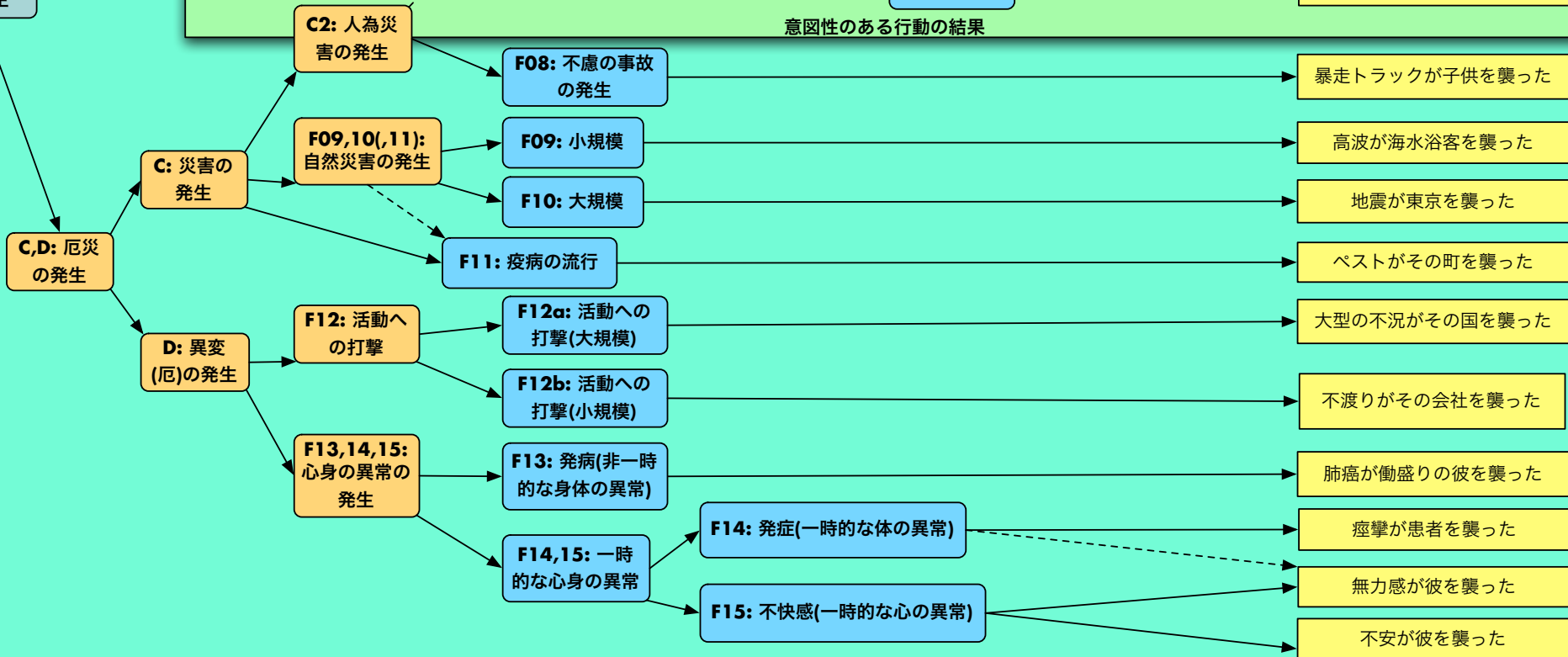
より抽象的 ←

→ より具体的

“XがYを襲う”の階層フレームネットワーク

最下位フレームは青で、事例は黄色で区別した

ROOT: Yからみた害の発生



意図性のある行動の結果

意味役割ベースのタグ体系

- 📌 意味役割 = フレーム要素名 (cf. 表層格フレーム辞書項目)
 - 📌 F06: 捕食のための攻撃 = {捕食者, 獲物, ...}
 - 📌 F03: 資源強奪 = {強奪者, 資源保有者, ...}
 - 📌 F10: 自然災害 = {台風, 都市, ...}
- 📌 意味フレームと(表層)格フレームは外延が一致する
 - 📌 両者は遺伝子型と表現型のような関係にある

意味フレーム分析の多次元化

📌 意味役割の多型性の問題

📌 名詞句の担う FE は一文に一つではない

📌 単純な意味フレーム分析 (cf. BFN) の限界

📌 多層意味フレーム分析 (Multilayered Semantic Frame Analysis: MSFA) の特徴

📌 一文内での意味フレームの統合を記述し、文の理解内容を記述する

📌 理解内容と知識構造との結びつきを重視

MSFA の思想

- 📌 統語記述と意味意味の思い切った切り離しが必要
 - 📌 順序に影響されない性質は、統語的性質ではない
 - 📌 統語から意味を考え、意味構造の記述可能性を“縛る”のは自滅的
 - 📌 “意味役割は少ないほどよい”と考えるのは統語中心の盲目
- 📌 記述内容の思い切った具体化が必要
 - 📌 理解内容を統語上の制約にあわせて矮小化せずに明示化するための記述フォーマットを提供

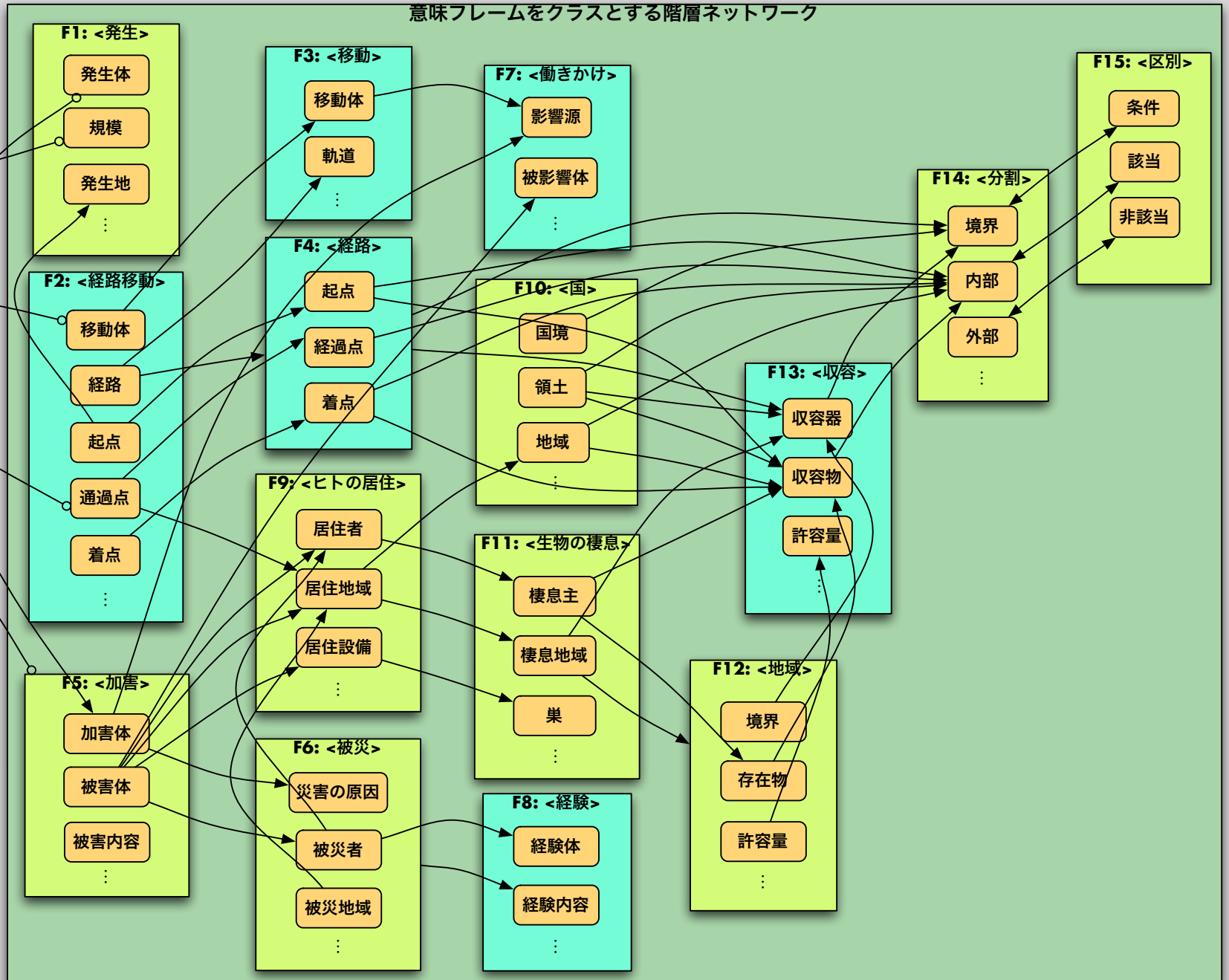
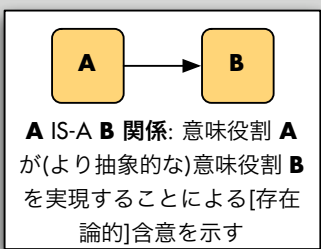
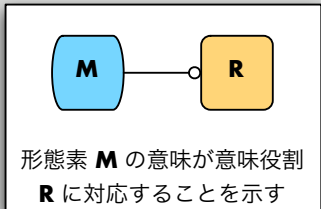
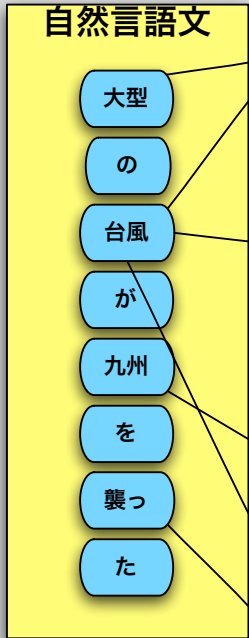
意味フレーム分析の多次元化

1. 意味フレームは言語的な単位ではないので、一文に幾つ結びついていても良い。ただし、
2. 文 s の形態素 m はどれも、少なくとも一つのフレーム f のフレーム要素 f_e を実現する
3. どの形態素も、異なるフレームごとに異なる意味役割をもっていてよい
4. どの形態素も、 f_e を実現しているフレームの数は多いほどよい

意味フレームの統合の効果

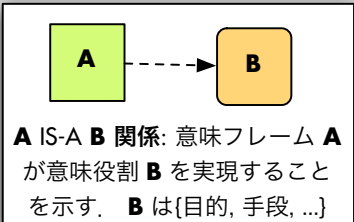
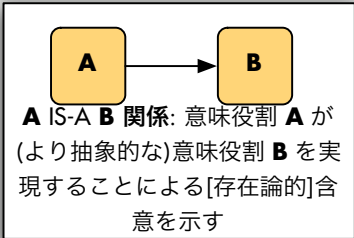
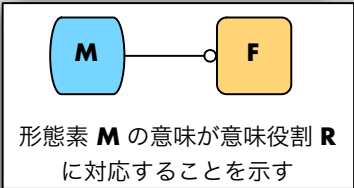
- 意味フレームの含意関係を利用すれば、全部を意味タグづけする必要はない
 - データベース内部での推論?
 - ただ、詳細に関しては大いに検討の余地アリ
- フレーム間の推論をどう表現するか?
 - フレームの体系をクラス/インスタンス階層で表現されたオントロジー体系と見なすと ...

意味フレームをクラスとする階層ネットワーク

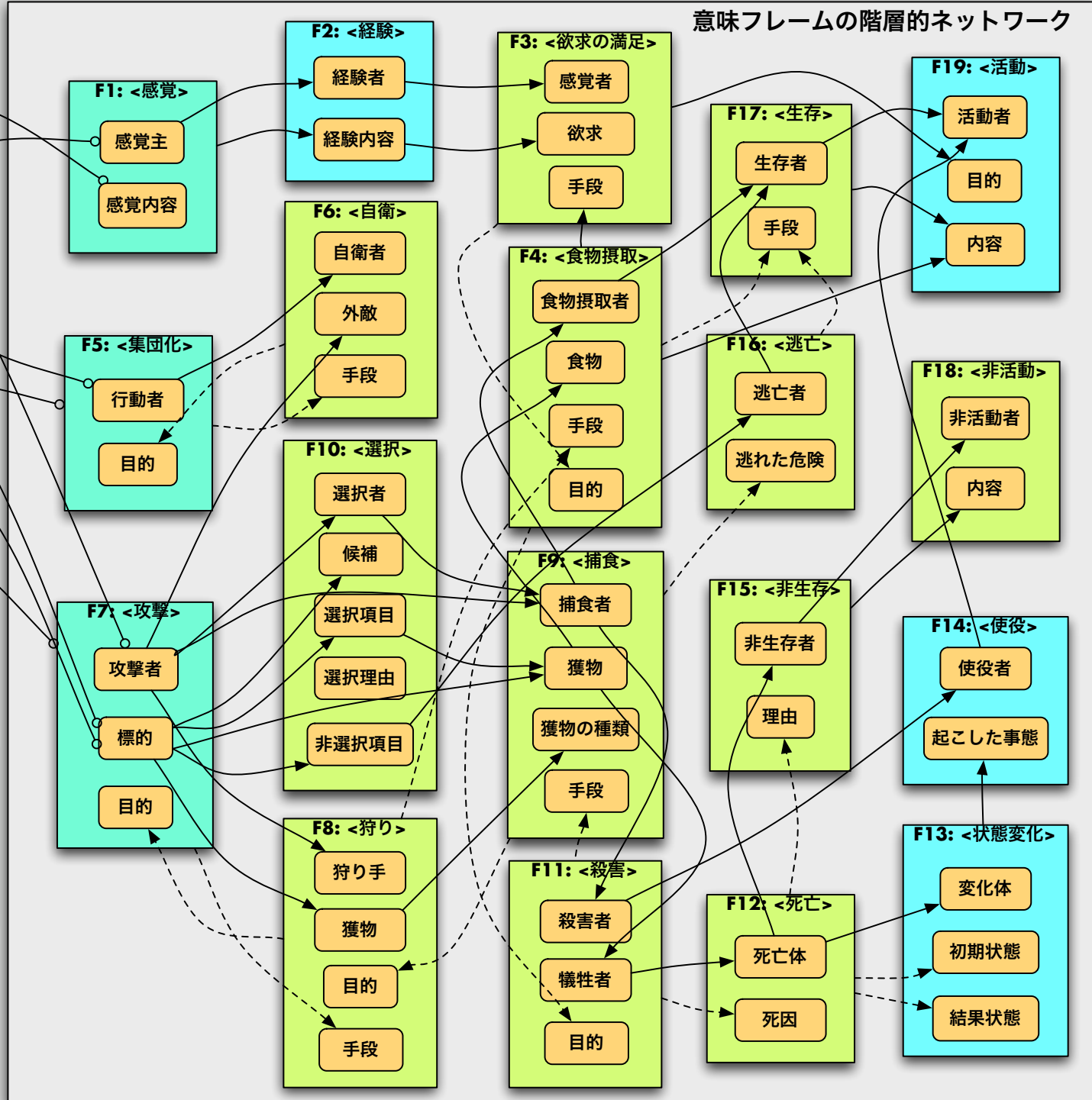


自然言語文

空腹
の
ライオン
が
インパラ
の
群れ
を
襲っ
た



意味フレームの階層的ネットワーク



課題

- 📌 作業手順の確立と、そのマニュアル化
- 📌 現実的に作業可能な(しかし手抜きでない)意味フレームの粒度、精度の見極め
- 📌 作業ツールの開発
- 📌 既存の言語資源 (EDR, IPAL, 格フレーム辞書) の流用による作業の効率化、利用価値の拡大
- 📌 機械学習を利用した(半)自動化は先の先

謝辞

黒宮 公彦

(大阪学院大学)

中本 敬子

(京都大学教育学研究科)

野澤 元

(京都大学人間・環境学研究科)

山梨 正明

(京都大学)