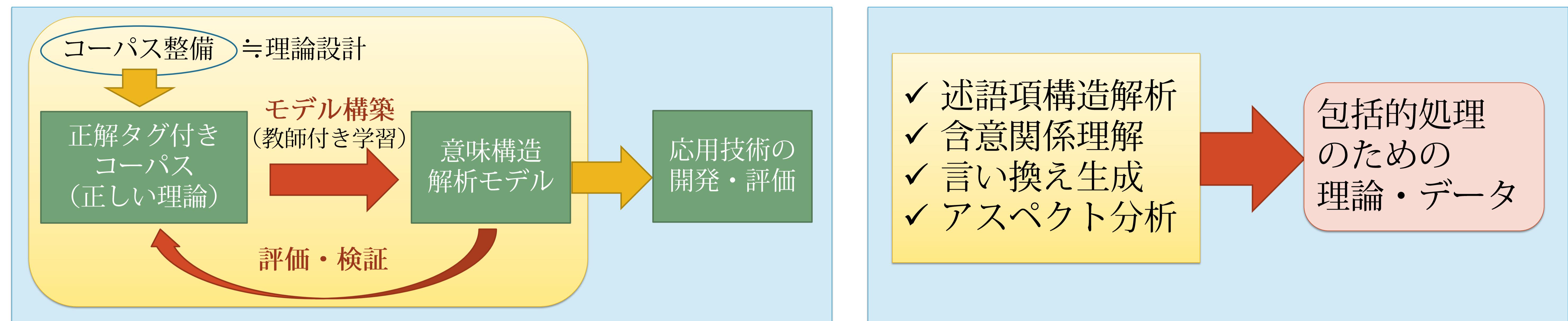


日本語動詞概念構造コーパスの設計

松林優一郎 宮尾祐介 相澤彰子（国立情報学研究所）

目的

効率的な意味解析技術開発サイクルの確立
要素技術を統合的に処理する理論枠組みを視野に



問題 & 挑戦

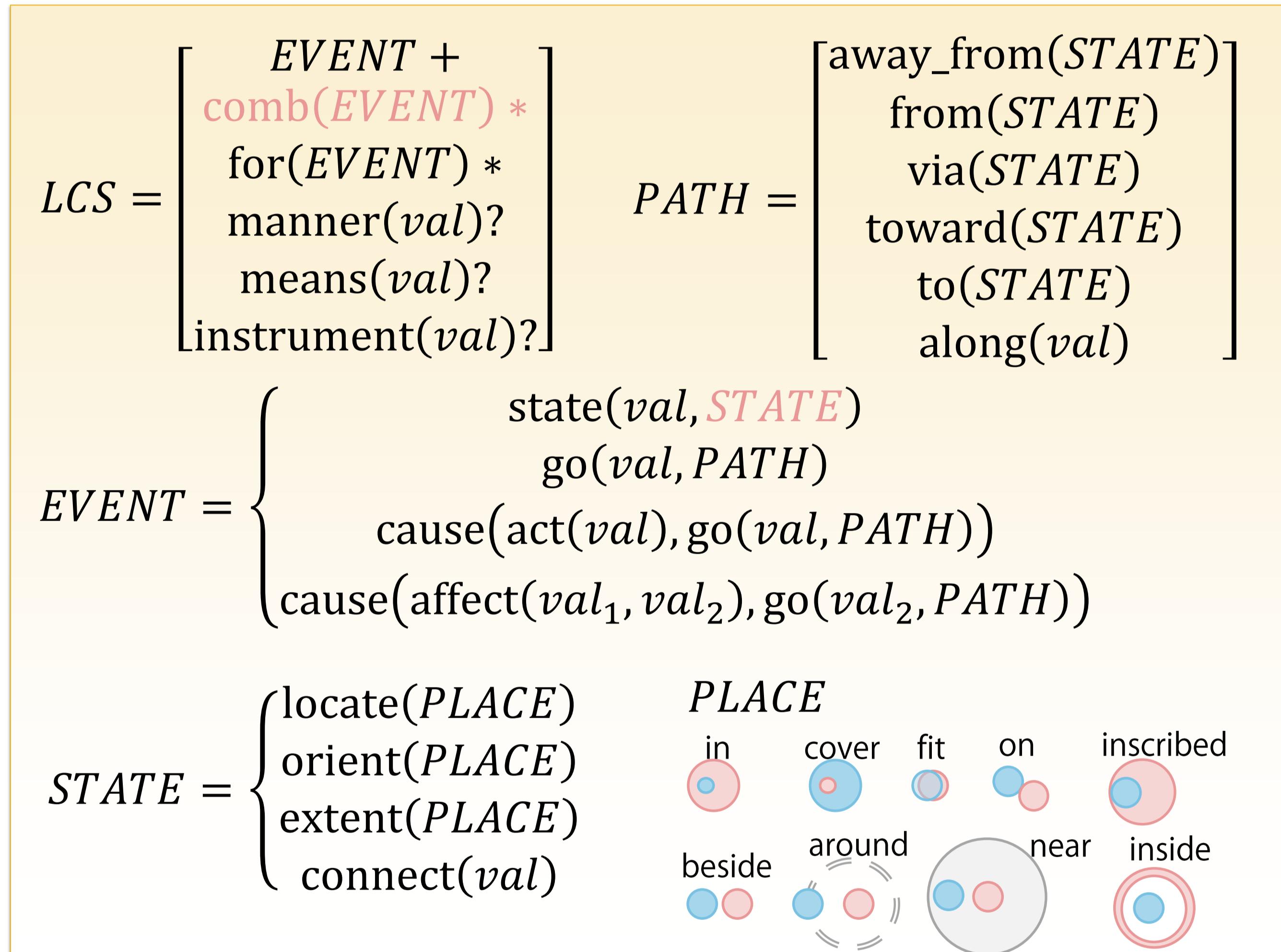
- 格助詞とその用法の対応関係には、曖昧性があり、動詞の意味と密接に関係している
→ 動詞ごとに、深層格や特定の用法の出現に関する分析が必要
- 一つの項は複数の意味役割を同時に持っている → 1フレーズ1タグ（従来記述法）では駄目
- 意味（構造）が与えられたら、含意関係を「計算できる」枠組み

方法

動詞概念構造辞書(LCS) + 項構造付与コーパス

LCS理論*を拡張して日本語動詞**の概念辞書を作成 (*Jackendoff, 1990, **事態性名詞含む)

述語の意味を実体や概念の相対位置として記述



買う.v

case: i=が, j=を, k=から, l=で,
duration:null

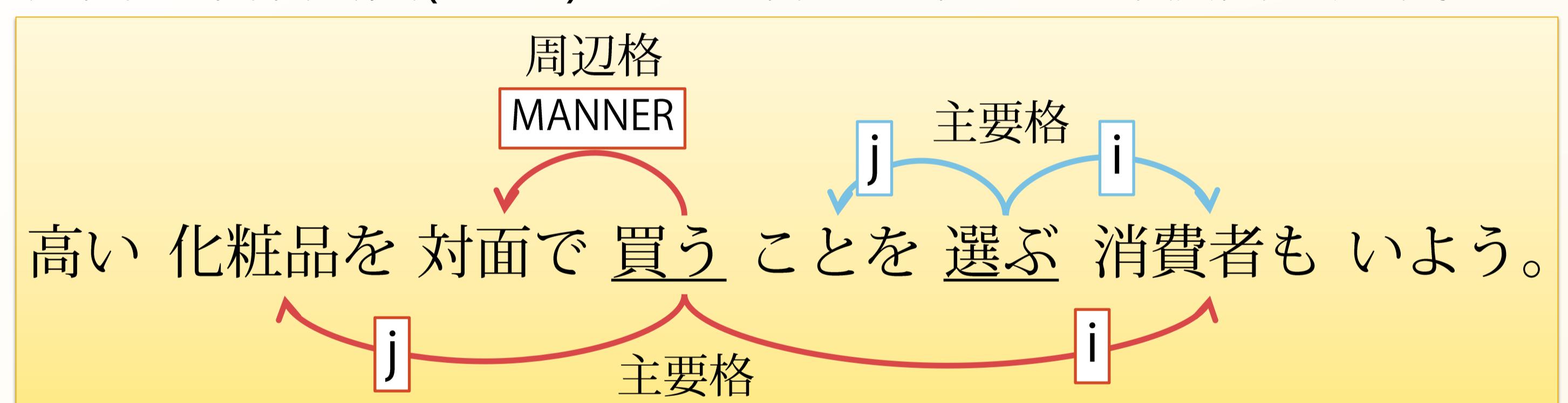
```
cause(affect(i, j), go(forward, [to(locate(in(i)))]))  
comb_parallel cause(affect(i, l), go(forward, [away_from(locate(in(i))), to(locate(at(k))))]))  
comb_parallel cause(affect(k, j), go(forward, [away_from(locate(in(k))), to(locate(at(i))))]))  
comb_parallel cause(affect(k, l), go(forward, [to(locate(in(k)))]))
```

京都大学コーパス（係り受け構造）の上部レイヤーとして項構造をタグ付け

主要格 (LCS内の変数) と周辺格 (述語独立な有限のタグセット) を区別し、それぞれを文にタグ付け

- 主要格 → LCSによって複数の役割が規定される
- 周辺格 → ラベルで表現される单一の役割

現代日本語文法(2009)を基に約30種の意味役割を定義



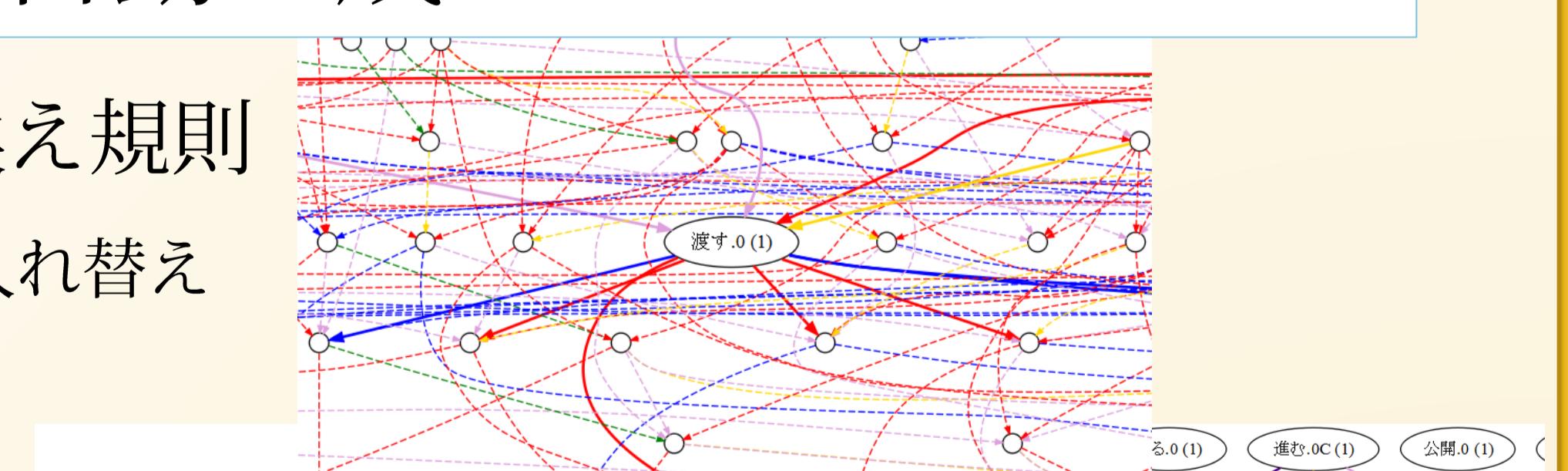
買う-i={ actor, agent(j), goal(j), agent(l), source(l) }
買う-j={ patient, theme }
買う-k={ agent(j), goal(l), agent(l), source(j) }
買う-l={ patient, theme }

構造式から導出される複数の役割

LCS上の書き換え規則を定義し、概念構造の関係グラフを自動生成

1 4種類の構造書き換え規則

- 主構造式とcomb式の入れ替え
- comb式削除
- 位置の上位概念化
- go, stateの取り立て
- pathの追加、削除
- 変数の単一化
- 変数の定項化 等々…



※画面は開発途中のものです。

現状と今後

- 約50語に対してLCSを記述済
- 200語程度書いたところでパイロット的に文タグ付けを開始
- 京都大学コーパスと同範囲の全テキストに対して語彙概念構造を割り当てタグ付けする予定