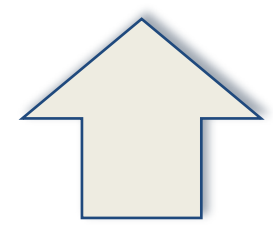


単語間関係に基づく 含意関係認識

水野淳太[†], 隅田飛鳥[‡], 乾健太郎[†]
[†] 東北大学 [‡] KDDI研究所

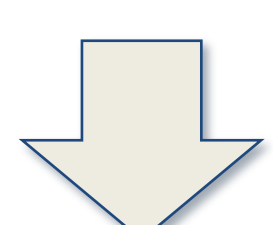
含意関係認識課題

t2: 川端康成は「雪国」の著者である



推論可能 → t1はt2を**含意**

t1: 川端康成は「雪国」などの作品でノーベル文学賞を受賞した

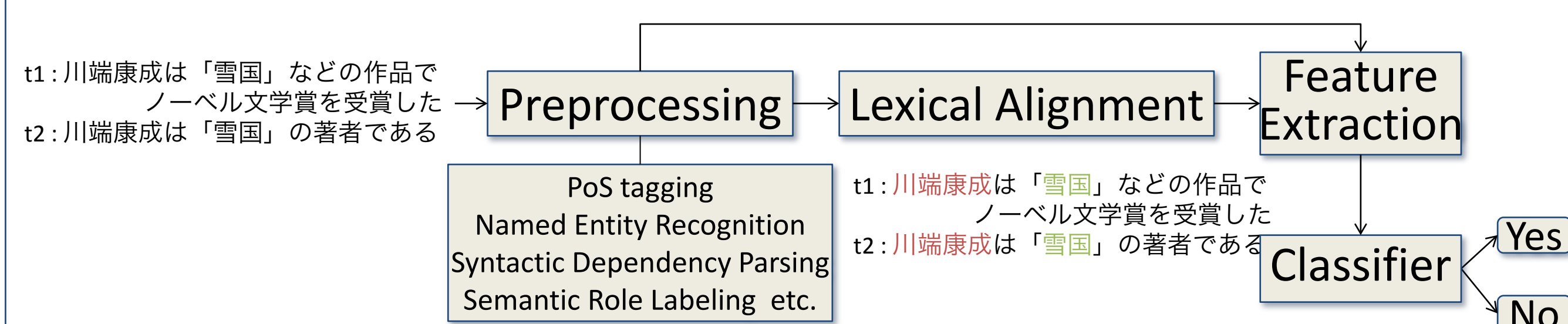


推論不可能 → t1はt2を**含意しない**

t2': 川端康成は「伊豆の踊子」の著者である

先行手法の問題点

アライメントに基づく認識モデル [Hickl 06, Sammons+09 etc.]



考慮されている単語間関係は依存関係, 述語項関係程度
→ 多様な単語間関係を取り扱う必要がある

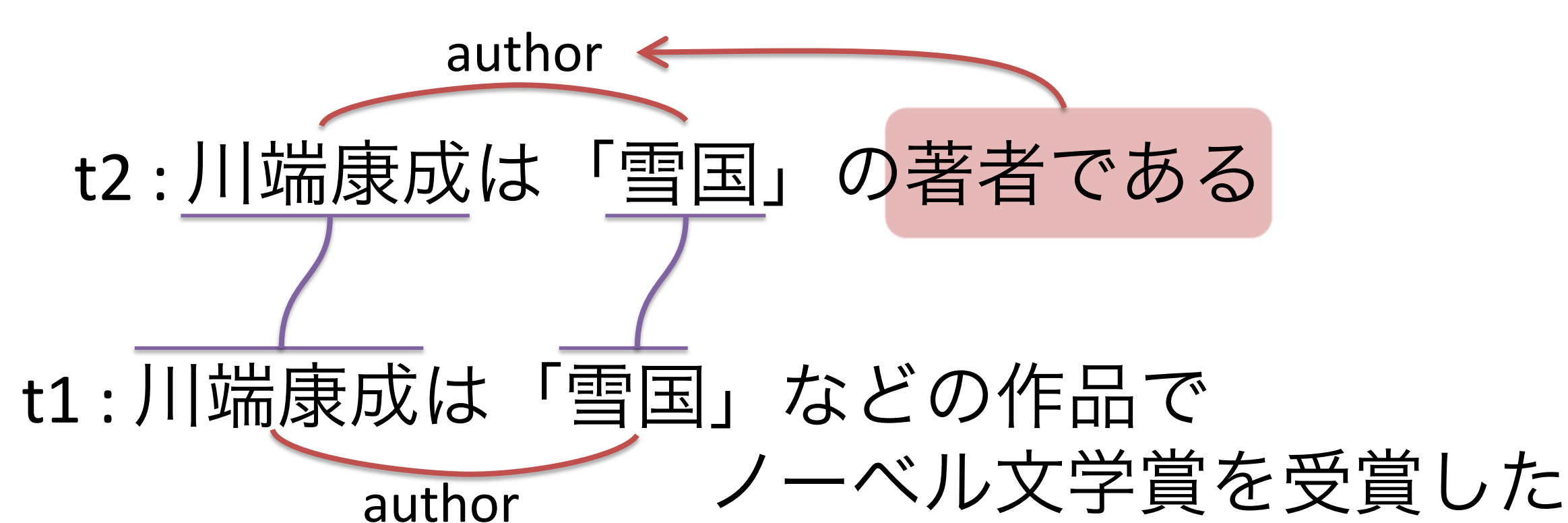
単語間関係に基づく含意関係認識

[水野+,11]参照

t2中の内容語がt1に多く含まれる→t1はt2を**含意**

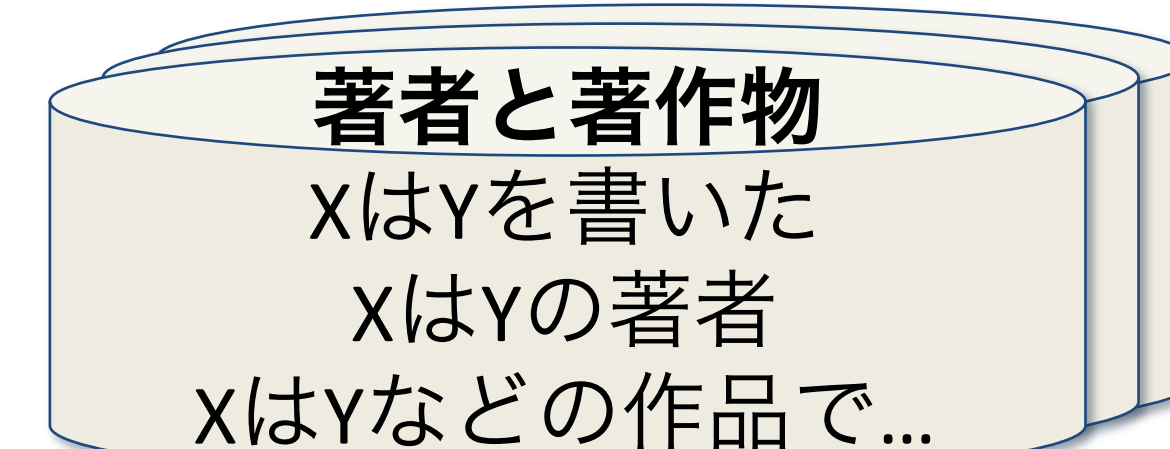
RITE2の開発データに対して, 表層, 意味および文構造の類似度に基づいて2文間の単語を対応付け, 被覆率によって分類
被覆率の閾値を0.6としたときの正解率 RITE2BC: 74.8% RITE2EXAMBC: 67.2%

非含意への誤分類事例 (被覆率が十分ではない含意事例)

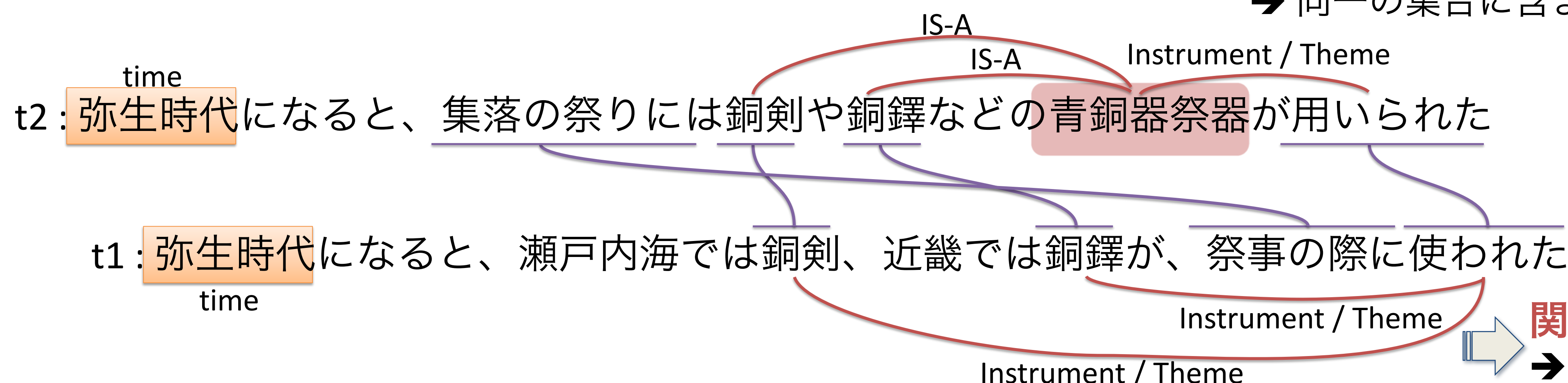


知識獲得技術を活用

1. 2単語間の関係を表すパターン集合を獲得



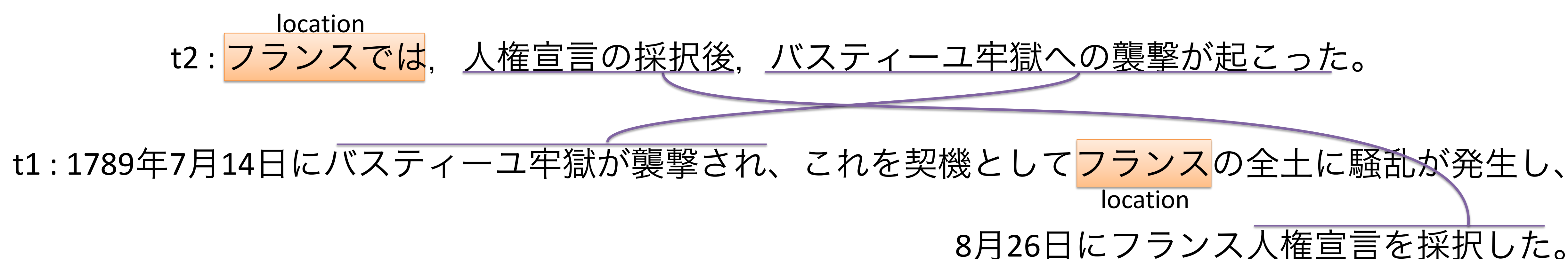
2. 文中の単語対が含まれるパターン集合を同定
→ 同一の集合に含まれていれば同じ関係を示す



関係の演算

→ IS-A + InstrumentとInstrumentの対応

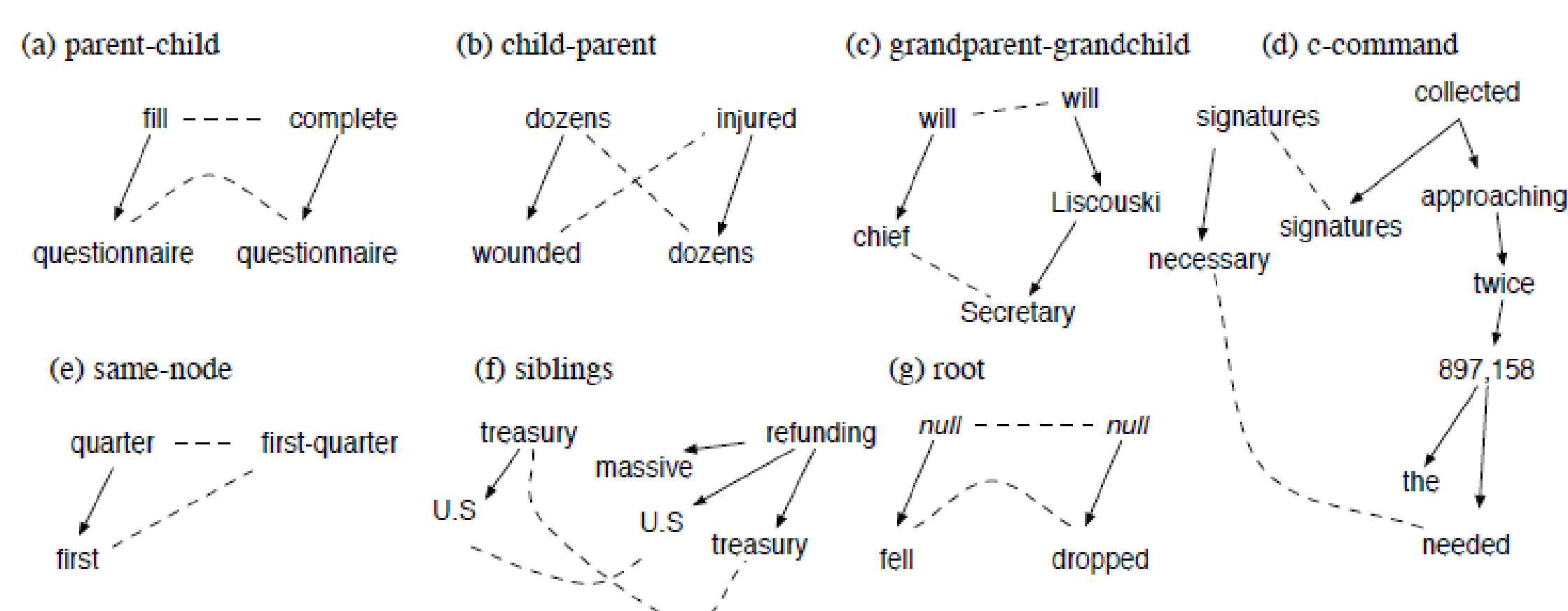
含意への誤分類事例 (被覆率が高い非含意事例)



⇒ イベント (採択, 襲撃) 間の時間的順序関係の認識 (SemEval2007 TempEval [Verhagen+, 07])

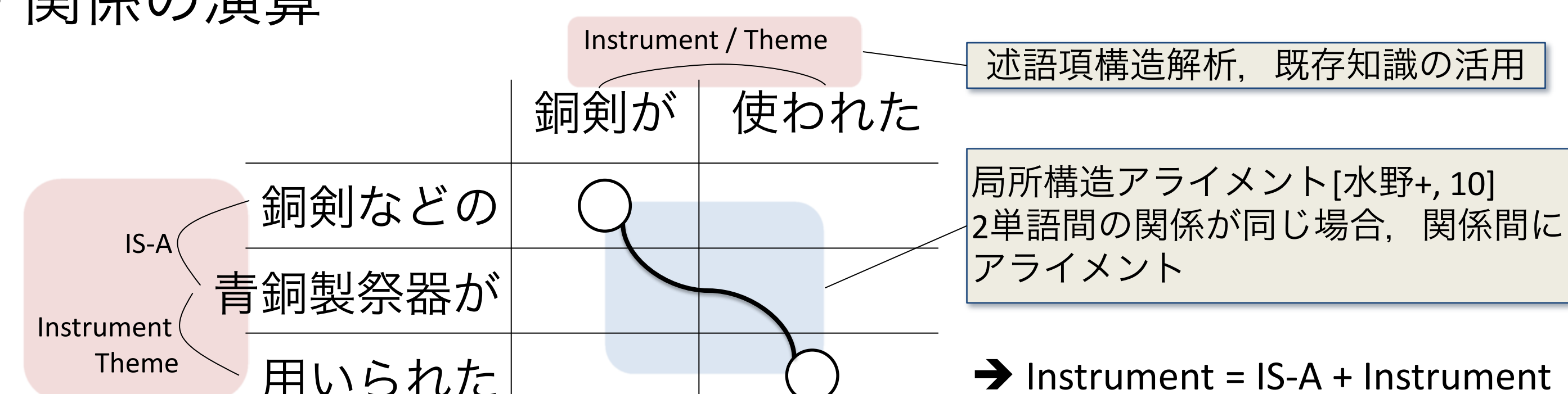
関連研究

- 文間で述語項関係の対応をとる [Shibata+, 12]
- 関係分類の素性として, 文内に含まれるエンティティ間関係を利用 [Rodrigo+, 08]
- 複数関係の演算 (COGEX@LCC)
- 言い換え認識 [Das+, 09]



今後の課題

- 事例分析
 - 開発データに対して述語項関係, 単語間関係を付与
 - 含意関係認識において重要な関係を列挙
- 既存の関係知識の活用
 - WordNet, ALAGIN 意味的関係抽出サービス, etc.
- 関係の演算



→ Instrument = IS-A + Instrument