

多様な言語表現に対する意味解析に向けて

東京工業大学 精密工学研究所

横野 光

(yokono@lr.pi.titech.ac.jp)

概要

多種多様な言語表現, 特に字義通りではない解釈を必要とする表現を我々はどのようにして解釈しているのか. 実際の解釈の過程から, そのような表現の処理に必要な資源について考察する.

字義通りではない表現の解釈

字義通りではない, とは言えその解釈には表層に現れている単語の何かしらの情報を必要とする.

ここで使われる情報とは表層上に出現しなくても解釈が可能な情報であると考えられる.

事例 1. “○○なう”

Twitter でよく使われる表現, 様々な形で使われるが, 特に “名詞 + なう” に着目する.

- ・ 場所 + なう (e.g. 京都なう) → その場所にいる
- ・ サ変名詞 + なう (e.g. 休憩なう) → 今, ~している (*)
- ・ 一般名詞 + なう (e.g. 昼食なう) → × 《名詞》している
昼食なう → 昼食を 食べている

↑これはどこから出てくるのか?

↑“昼食は食べるもの” という知識?

事例 2. 比喩表現 (楠見, 1995)

- ・ 暗喩
 - ・ 彼は羊だ → 彼は (羊のように) ○○だ.
“羊” の属性によって, “彼” の属性の一部を強調
- ・ 換喩
 - ・ 杯を傾ける → 酒を飲む
一連の動作からの推論 (スクリプト的な知識)
- ・ 提喩
 - ・ 花見 → 桜を見ること
上位 - 下位関係 (花 - 桜)

換喩以外については, 名詞が一般的にどう使われるか, などといった情報を記述すれば処理できる?

→ 名詞辞書の構築

参考文献

- ・ 寺井あすか, 中川正宣, “確率的概念構造に基づく比喩理解の計算モデル”, 認知科学, Vol.17, No.1, 2010.
- ・ 楠見孝, “比喩の処理過程と意味構造”, 風間書房, 1995.
- ・ James Pustejovsky, “The Generative Lexicon”, MIT Press, 1995.

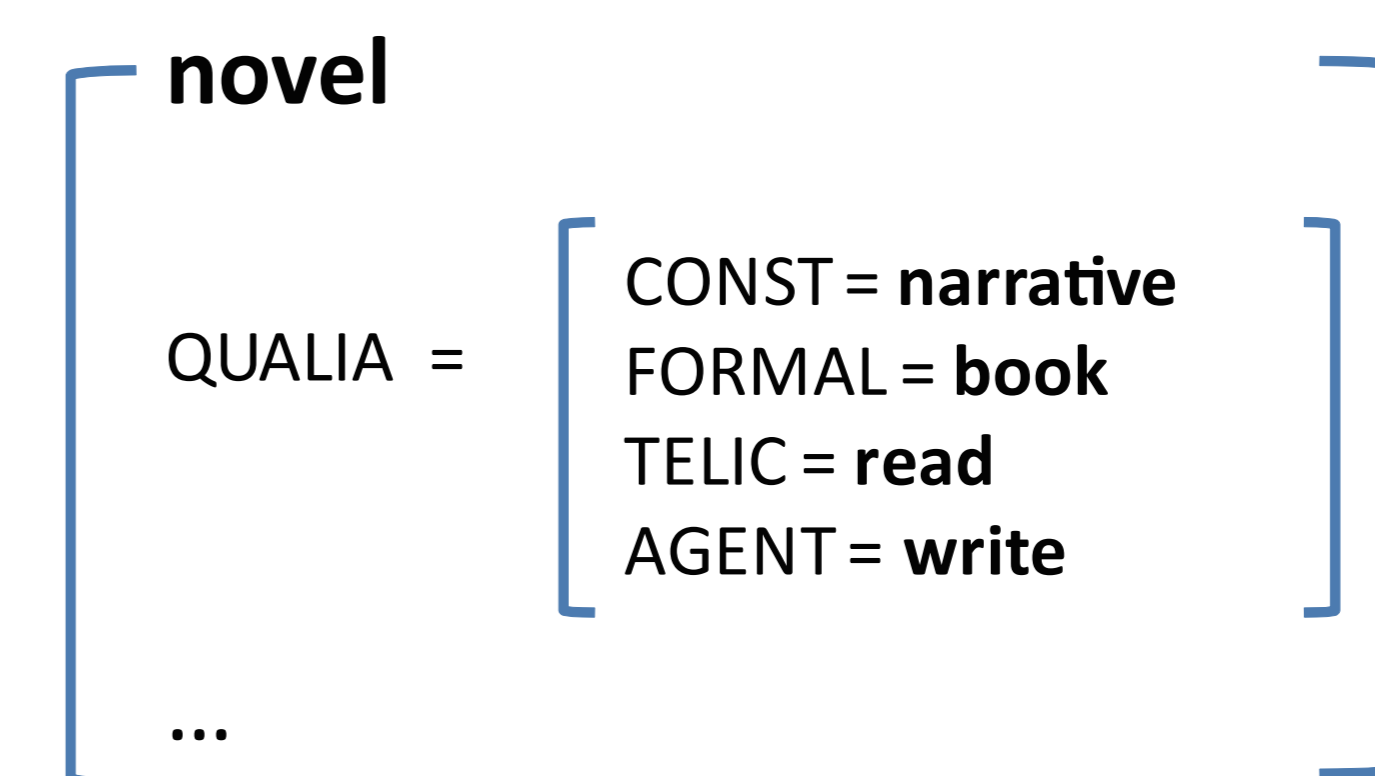
(*) “帰宅なう” だと “帰宅している” より “(今,) 帰宅した(ところ)” と解釈される.
この違いはその名詞が事象のどの点に焦点を当てているかということが関係すると考えられる.

名詞辞書の構築

- ・ 属性: (寺井ら, 2010) のモデル
- ・ 上位下位関係: 日本語語彙大系など
- ・ デフォルトの使われ方: クオリア構造

生成語彙論 (Pustejovsky, 1995) のクオリア構造

- ・ 構成役割 (Constitutive role)
対象とそれを構成する部分の関係
- ・ 形式役割 (Formal role)
より大きなドメインにおいてそれを区別する関係
- ・ 目的役割 (Telic role)
目的と機能
- ・ 主体役割 (Agentive role)
その起源や発生の要因



名詞 “novel” のクオリア構造 (簡略版)

どのようにして辞書を構築するのか

コーパスから自動的に獲得する

よく共起するものがデフォルトの意味を持つ?

→ 京大格フレームで調査

直後に出現する動詞を頻度でソート

“朝食”

動詞	格	頻度	格フレームのID
食べる	ヲ格	15062	食べる/たべる:動1
採る	ヲ格	9026	採る/とる?撮る/とる?執る/とる?取る/とる?摂る/とる?盗る/とる?捕る/とる:動1
済ませる	ヲ格	4663	済ませる/すませる:動1
取る	ヲ格	4391	取る/とる:動6
撮る	ヲ格	3462	撮る/とる:動1

……何となくそれっぽい

“机”

動詞	格	頻度	格フレームのID
向かう	ニ格	5568	向かう/むかう:動15
叩く	ヲ格	2193	叩く/たたく:動1
座る	ニ格	1944	座る/すわる:動1
拭く	ヲ格	1851	拭く/ふく:動3
並べる	ヲ格	1718	並べる/ならべる:動1

……そもそも, 机はどうするもの?

検討事項

- ・ どこまで人手を加えるか
全部自動で, よりはある程度人の手を加えた方がいい?
- ・ 動詞の性質を考慮 (e.g. 語彙概念構造など)
- ・ この資源を使って実際にどう処理するか
 - ・ 表現の同定 → 解釈
 - ・ 解釈候補の生成 → 選択