

物語生成システムのための動詞的／名詞的概念体系の構築に向けて Towards the Development of Systems for Verb/Noun Concepts in the Narrative Generation System

大石 颯祐[†], 小野寺 康[†], 小方 孝[‡]
Kensuke Oishi, Kou Onodera, Takashi Ogata

[†]岩手県立大学大学院, [‡]岩手県立大学
Graduate School of Iwate Prefectural University, Iwate Prefectural University
g231g010@s.iwate-pu.ac.jp

Abstract

In the narrative generation system that we have developed, generations of story and discourse are done in the level of conceptual representation, and conceptual systems such as noun concepts and verb concepts hierarchies gives the basis. This paper explains the current development phase. The description of constraint condition in each verb case frame that determines the range of possible or natural concepts in a narrative event relates to the classification of concepts in noun concepts hierarchy.

Keywords — Narrative Generation System, Verb Concepts, Noun Concepts, Concepts Hierarchy, Case Frames, Constraint Conditions

1. 研究の背景・位置付け・目的

物語生成システム(Ogata, 1991 ; 小方・堀・大須賀, 1996; 小方, 2003a, 2003b, 2007; 小方・金井, 2010)は, 大きく分けて概念生成機構と表層表現生成機構から成る(図 1). 概念生成機構では, 物語の意味内容と構造を, 物語木 (事象概念の木構造) の形式で生成する. 表層表現生成機構は, この事象概念の木構造を文・映像・音楽の各表現媒体に変換する.

物語木の終端ノードに用いる事象概念は物語中に現れる出来事を表す物語の基本的な表現単位であり, 物語生成システムでは動詞概念と名詞概念を深層格 (名詞概念の動詞概念に対する意味役割) で関連付けたフレーム形式で表現する. 例えば, 「学生がカツ丼を食べる」という意味の事象は, (event 食べる (agent 学生)(object カツ丼)) と表現する. この事象概念に対して, その構成要素である動詞概念や名詞概念の体系, 動詞概念における格フレーム及びそこに挿入する名詞概念の範囲を定義する制約条件の情報を提供するものが, 概念

体系である(図 2). また, 事象概念の木構造から言語表現を生成する場合, この概念体系と関連付けた言語辞書を使用する.

本研究は上述のように物語生成システムにおける必須要素として利用することが最大の目的であるが, この他に広告における単一事象の修辭的処理への応用にも利用出来る(阿部・小方・花田, 2008). この処理では, 概念体系から制約条件の範囲内の概念を取得することで広告風物語の場面における標準的な修辭 (日常的な情景に相当) の適用を, 制約条件の範囲外の概念を段階的に取得することで異化的な修辭 (非日常的な情景に相当) の適用を行う.

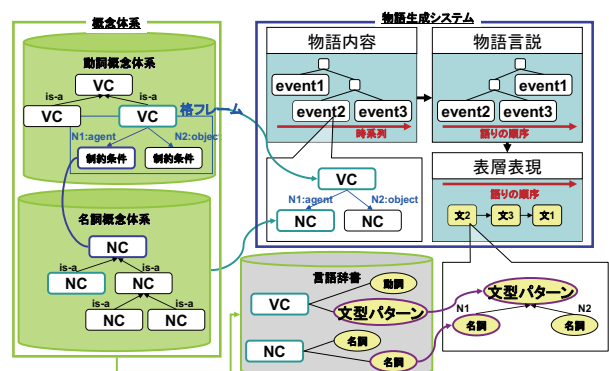


図 1 物語生成システムの全体像

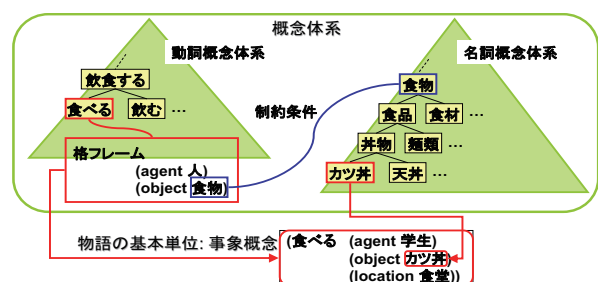


図 2 概念体系の機能

本稿では、現在開発中の、動詞概念体系とその格フレーム・制約条件、名詞概念体系を中心に、物語生成システムのための概念体系について紹介する。

概念体系構築の基本方針であるが、概念体系研究には蓄積があるので、既存のシステムを出来るだけ有効利用することを考えている。しかし物語生成システムの特徴や目的に沿ったカスタマイズを行わなければならない。特に本研究で利用した『日本語語彙大系』(池原 他, 1999)はもともと機械翻訳用の概念辞書であり、物語生成の観点からは動詞格の要素設定のための制約条件やそれと対応する名詞概念体系の範疇分類が粗過ぎるという欠点がある。そこで、それらを精緻化・充実させるという作業が研究の大きな部分を占める。

2. 成果と課題

これまで、詳細な制約条件の定義に必要な動詞概念の格フレームと名詞概念体系の構築を行って来た。具体的には、『日本語語彙大系』の用言意味属性体系 36 属性に関連付けられている動詞を述語とする文型パターン約 12000 個のうち、物理的な状態変化を表す 6 属性と関連付けられたものから自然現象や機械が行為の主体となるものを除く 4620 個に対して、5337 個の格フレームを定義した(図 3)。作成した格フレームは、『日本語語彙大系』における動詞概念と対になる文型パターンの名詞項を 8 種類の深層格のいずれかと対応付けたもので(図 4)、各深層格には制約条件が併記される(図 5)。

名詞概念体系については、『日本語語彙大系』の一般名詞意味属性体系 2710 属性と関連付けられた 145360 語を同音同義語ごとにまとめ(例:語「いぬ」「イヌ」「犬」「狗」を概念「犬」にまとめる)、139802 個の概念として登録した。このうち「人」「動物(個体)」「人工物」の下位概念については、『日本語 Word Net』(Bond et al., 2009)を参考に、705 個の中間概念を挿入することで詳細化を行った(図 6)。

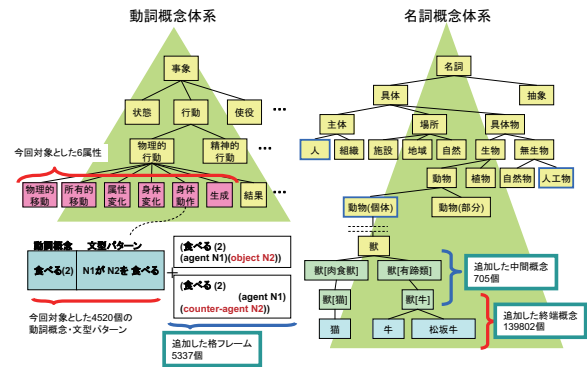


図 3 二種類の概念体系の構成

time	事象が発生する時間	動詞概念 文型パターン 制約条件 食べる(2) N1がN2を 食べる N1(4人 535動物) N2(838食料 534生物) 『日本語語彙大系』の情報 + (食べる (2)(agent N1)(object N2)) + (食べる (2)(agent N1)(counter-agent N2))
agent	行為者	
counter-agent	被行為者	
object	行為対象	
instrument	行為に用いられる道具	
location	事象が発生する場所	
from	事象の開始位置	
to	事象の終了位置	

図 4 深層格の種類 (左側) と各フレームの記述形式



図 5 動詞概念のプログラム記述

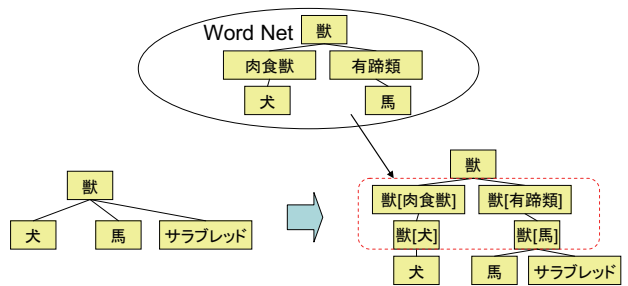


図 6 名詞概念体系の詳細化

問題点及び課題として、次のものが挙げられる。

- ① 概念に対応する深層格の検討：現在、作成した 5337 個の格フレームのうち、深層格との対応付けが困難なものが最終的に 174 個残っており、文型パターンのタイプ別に 7 種類に分類した(表 1)。

- ② 名詞概念体系の詳細化:作成した 5337 個の格フレームの制約条件を詳細に定義するため, 名詞概念体系の詳細化作業の範囲を拡大する必要がある.
- ③ 制約条件の詳細化: 試験的に動詞概念「食べる(2)」の制約条件を詳細化した. 具体には, ひとつの名詞項が複数の格に対応する場合, 対応する格ごとに制約条件を設定する. また元々の制約条件の範囲が広すぎる場合はそれをその下位概念に狭める. 事象概念をランダムに 300 パターン生成し, 生成結果を現実世界において日常的に起こり得る事象(自然な事象)/日常的には起こらないが物理的には実現可能な事象(不自然・可能な事象)/物理的に実現不可能な事象(不可能な事象)に分類した結果, 表 2 に示すように 83.3%の割合で自然な事象が生成された (この評価は筆者の常識に基づいて行っている). 今後はその他の動詞概念にも同様の詳細化を行う必要がある.

表 1 格フレーム未定義の文型パターンの 7 分類

問題点	文型パターン	コメント
1. 副詞的概念の扱い	N1が N2を N3に N4 締める	N4:副詞的概念
2. 所有格の扱い	N1が N2の N3にへ 潜る	N2:所有格
3. イベント(催し物)の扱い	N1が N2を N3で 破る	N3:イベント(催し物)
4. 行為目的の扱い	N1が N2から/より N3へ/に/まで N4に 出かける	N4:行為目的
5. 経験者格の必要な事象	N1が N2に 疲れる	N1:経験者(ある動作の影響を受けたり, 経験する実体)
6. 比較対象の扱い	N1が N2を N3の 倍にする	N3:比較対象
7. 慣用表現の扱い	N1は N2が N3から 遠のく	“足が遠のく”という慣用表現

表 2 事象概念のランダム生成実験とその評価

文型パターン	格フレーム	制約条件	自然	不自然/可能	不可能
N1がN2を食べる	食べる(2) (agent N1) (counter-agent N2)	N1:"人" "動物" N2:"生物" "食料"	28.7%	53.0%	18.3%
		N1:"人" "動物" N2:"生物(個体)"	83.3%	15.3%	1.3%

3. 課題の解決案

- ① 概念に対応する深層格の検討: 分類困難な深層格については, 今後, 各種文献を参考に改

訂を図って行く予定である. 例えば, (岡田, 1991)は動詞概念の深層格を 10 種類に, (竹内, 2011)は 71 種類に分類している. また, (小泉, 2007)は, 日本語の 10 種類の表層格に対する 40 種類の意味役割を分類している.

- ② 名詞概念体系の詳細化: これまでに作成した, 「人」「動物(個体)」「人工物」の下位概念以外についても, 試験的に行ったものと同様に名詞概念の詳細化を行う.
- ③ 制約条件の詳細化:作成した 5337 個の格フレームに対して試験的に行ったものと同様に制約条件の改善・詳細化を行う. この制約条件には, ②で作成した詳細化した名詞概念体系を使用する.

現在, ②については概念「場所」の下位概念を対象に, ③については「状態—事象変換システム」(小野寺・小方, 2011)で使用する 278 の制約条件を対象に作業を行っている.

参考文献

阿部 弘基・小方 孝・花田 健自 (2008). テレビ広告のショットにおけるブランド導入レトリックの分析—広告の物語生成システムの一機構のために—. 『日本認知科学会文学と認知・コンピュータ II 研究分科会(LCCII)第 16 回定例研究会予稿集』. 16G-05.

Bond, F., Isahara, H., Fujita, S., Uchimoto, K., Kuribayashi, T. & Kanzaki, K. (2009). Enhancing the Japanese WordNet. *The 7th Workshop on Asian Language Resources, ACL-IJCNLP 2009*.

Bringsjord, S. & Ferrucci, D. A. (2000). *Artificial Intelligence and Literary Creativity*. Lawrence Erlbaum.

小泉 保 (2007). 『日本語の格と文型—結合価値論に基づく新提案』. 大修館書店.

池原 悟・宮崎 正弘・白井 論・横尾 昭男・中岩 浩巳・小倉 健太郎・大山 芳史・林 良彦 (1999). 『日本語語彙大系 CD-ROM 版』. 岩

波書店.

Ogata, T. & Terano, T. (1991). Explanation Based Narrative Generation Using Semiotic Theory. *Proc. of National Language Processing Pacific Rim Symposium 9*. 321-328.

小方 孝・堀 浩一・大須賀 節雄 (1996). 物語のための技法と戦略に基づく物語の概念構造生成の基本的フレームワーク. 『人工知能学会誌』. 11(1). 148-159.

小方 孝 (2003a). 物語の多重性と拡張文学理論の概念-システムナラトロジーに向けて I-. In 吉田 雅明 編. 『複雑系社会理論の新地平』. 127-181. 専修大学出版局.

小方 孝 (2003b). 拡張文学理論の試み- システムナラトロジーに向けて II-. In 吉田 雅明 編. 『複雑系社会理論の新地平』. 309-356. 専修大学出版局.

小方 孝 (2007). プロップから物語内容の修辞学へ—解体と再構成の修辞を中心として—. 『認知科学』. 14(4). 532-558.

小方 孝・金井 明人 (2010). 『物語論の情報学序説—物語生成の思想と技術を巡って—』. 学文社.

小野寺 康・小方 孝 (2011). 概念体系と結びついたストーリーワールド—ストーリーライン変換機構. 『人工知能学会全国大会(第 25 回)論文集』. 1H2-OS1-14in.

岡田 直之 (1991). 『語の概念の表現と蓄積』. 電子情報通信学会.

竹内 孔一 (2011). 動詞項構造シソーラスの構築. 『人工知能学会全国大会(第 25 回)論文集』. 3H2-OS3-5.