

物語生成システムのための動詞概念格フレームと 名詞概念体系の開発と利用

Development and Use of Verb Concepts Case Frame and Noun Concepts System for the Narrative Generation System

大石 顕祐 *¹
Kensuke Oishi

小方 孝 *²
Takashi Ogata

小野寺 康 *¹
Kou Onodera

*¹ 岩手県立大学大学院
Graduate School of Iwate Prefectural University

*² 岩手県立大学
Iwate Prefectural University

Many narrative generation systems use small-scale knowledge-base/ontology, and deal with only narrow and limited narrative domain. On the other hands, we are attempting to give large-scale concepts ontology for verb concepts and noun concepts to the narrative generation system to generate various narratives in which the domain does not constrained. This paper introduces the framework and architecture of verb concepts and noun concepts ontologies utilizing Japanese lexicon by NTT. Verb concepts ontology contains case frames for 4620 concepts. After we explain on mechanisms of two kinds of ontologies, we evaluate the naturalness of constraint knowledge that case frame in each verb concept contains to search problems and the way and direction of the solution and application.

1. 研究の背景と目的

1.1 物語生成システムとは

物語生成システムとは、物語のテキスト(言語の他、映像も含む)を自動生成するシステムであり、また研究領域としては、人工知能、自然言語処理、認知科学、創造活動支援などと関連する。応用分野としては教育、ゲーム、娯楽、芸術、文学などが挙げられる。物語生成システムの例として、目標志向のプランニング(計画)によって行為系列を作る Tale-spin [Meehan 1980], 白昼夢のシミュレーションを実現した Daydreamer [Mueller 1990], イソップ童話風のストーリー生成システム[岡田 1991], 裏切りを含む物語に特化した BRUTUS [Bringsjord 2000]などが挙げられる。

1.2 筆者らの物語生成システム

筆者らの物語生成システム([Ogata 1991][小方 1996, 2003ab, 2007, 2010]等)は、事象概念の集合の生成を行う概念表現生成機構と、文・音楽・映像による表層表現の生成を行う表層表現生成機構からなる(図 1)。このうち概念生成は、木構造の生成と変換によって実現する。「何を語るか」を表す物語内容は事象概念の木構造を時系列順に並べたものとして、「如何に語るか」を表す物語言説は、事象概念の木構造を変形し、語りの順序に並べ替えたものとして表現する。筆者らは、この構想に基づいて各モジュール及び統合システムの開発を行っている。

1.3 物語生成システムにおける概念オントロジーの役割

オントロジーとは、一般的には共通概念を提供する体系化された辞書のことである[溝口 2006]。ここで概念オントロジーと言うのは、動詞的概念や名詞的概念を初めとする概念を組織化した辞書を意味する。物語生成システムにおいて、物語の基本単位である事象概念に対して、その構成要素である動詞概念や名詞概念の体系を提供する役割を持つ。事象概念は格フレー

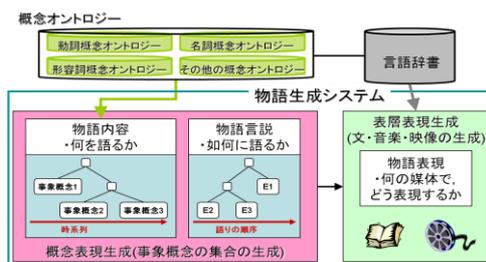


図 1 物語生成システムの概要

ムの形式で表現する。格には深層格と表層格が存在するが、このうち深層格は人工的な標識であるため定義者によってその分類は異なる。一例を挙げると、格文法をフレーム形式で記述することを提案したフィルモア[フィルモア 1976]が用いた深層格は、動作主格(Agent)、経験者格(Experiencer)、道具格(Instrument)、対象格(Object)、源泉格(Source)、目標格(Goal)、場所格(Location)、時間格(Time)の 8 つであり、本研究でも参考にしている。物語生成システムにおける格フレームとその表現形式を図 2 に示す。左側に示すのが格フレームの記述形式である。中央は、この形式を用いて記述した「学生がカツ丼を食べ、番茶を飲む。」という意味の事象概念の時系列順の連なりを表す。この時事象概念において、「学生」は agent(行為者)、「カツ丼」と「番茶」は object(行為対象)に相当する。

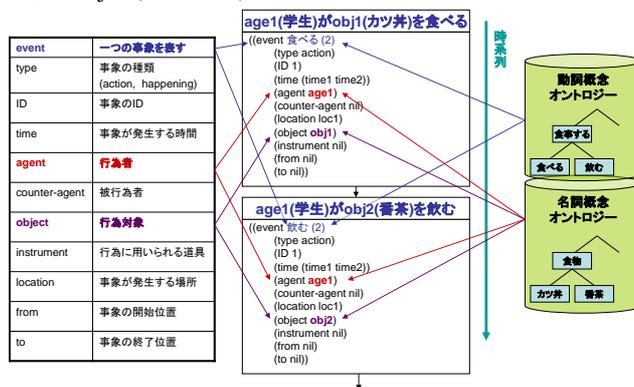


図 2 物語生成システムにおける格フレームとその表現形式

連絡先: 大石顕祐, 〒020-0193 岩手県立大学大学院ソフトウェア情報学
学研究科, 岩手県岩手郡滝沢村滝沢字菓子 152-52,
g231g010@s.iwate-pu.ac.jp

③では、より細かい概念体系を持つ『日本語 Word Net』[Bond 2009]を参考にした。表 1 に追加した新たな概念の例を示す。もともとは表中の「下位概念」に含まれる名詞すべてが「獣」という概念の下に含まれていたが、表中の「上位概念」を追加して細分化した。

現状で、①は同音同義語について約 15 万の名詞すべてに対して行った(類義語の統合は未着手)。②は未着手である。③は「人」、「動物」、「人工物」の属性以下の範囲に限定して試験的にを行い、新たに約 700 の概念を定義した。今後は範囲を拡大してこれらの作業を継続する予定である。

2.3 評価

構築した格フレームの制約条件の妥当性を確認するために、その制約条件を満たす範囲で生成した事象におかしなものが存在しないかどうかを実験する。この実験のために、格フレームの必須格の制約条件以下の名詞概念を無作為に取得して事象概念を生成するプログラムと、事象概念からその動詞概念の文型パターンを雛型として文を生成するプログラムを作成した(図 6)。



図 6 事象概念生成・事象概念-単文変換システムの構成

今回、「食べる(2)」(agent が object を食べる)を対象に実験を行った。これの必須格と制約条件は agent[人/動物]と object[食料/生物]である。object を「カレーライス」に固定して agent を無作為に決定した場合、agent を「学生」に固定して object を無作為に決定した場合、agent と object の両方を無作為に決定した場合のそれぞれについて 300、合計 900 の事象概念及びその文表現を生成した。生成された各事象の文を、自然な事象(例「老婦が漬け物を食べる」)、不自然だが可能な事象(例「女傑が三葉虫を食べる」)、不可能な事象(「日蝕が蛇を食べる」)の 3 種類に分類した。その結果自然な事象が生成される割合が低かったため、制約条件を agent[人/動物(個体)]と object[食料]に修正した上で同様の実験を行った。

以上の結果を表 2 に示す。制約条件を修正しない場合、自然な事象が生成された割合は最大でも 76.3%だったが、修正後はいずれのパターンでも 80%以上に上がった。今回制約条件の修正は評価対象のみに行ったが、今後すべての動詞概念に対して行う予定である。

表 2 動詞概念「食べる」に対する事象概念・単文生成実験

	自然な事象		不自然だが可能な事象		不可能な事象	
	修正前	修正後	修正前	修正後	修正前	修正後
N1可変	76.3%	88.3%	7.7%	11.0%	16.0%	0.7%
N2可変	43.0%	87.3%	56.3%	12.3%	0.7%	0.3%
両方可変	28.7%	83.3%	53.0%	15.3%	18.3%	1.3%

3. 課題

今後の課題として、格フレームの拡張と制約条件の修正が挙げられる。今回格フレームの構築に当って、文型パターンの名詞項に対応する格を確定出来ないものを 174 個発見した。これらを表 3 に示す 7 種類に分類した。例えば、問題点 1 は、「縮める(1)」(N1 が N2 を N3 に N4 縮める。例えば「選手が記録を 3 分に 2 秒縮める」)の名詞項 N4(副詞的概念)が問題となり、また問題点 2 は、「潜る(4)」(N1 が N2 の N3 に/へ潜る。例えば

「男が机の下に潜る」)の名詞項 N2(N3 に係る所有格)が問題となる。

表 3 格フレームの問題点の 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縮める	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

これらの名詞項に対応する格を整合的に定義可能な格フレームへの拡張に向けて、[小泉 2007]や[岡田 1991]など、フィルモア以降の格文法研究も参考にしながら考察する。表 4 にフィルモア、岡田、小泉のそれぞれが定義した格フレームを示す。表内の斜体字は本研究における格フレームに対応するものがない格を、太字は対応関係が一対一でない格を表す。例えば問題点 1 の「縮める(1)」の名詞項 N4 は岡田の格フレームにおける「OC:属性の補足」に対応する。このように、これら諸研究を参考に表 3 の問題点に対する改訂を行う予定である。

表 4 諸研究と本研究の格フレームの比較

小方ら	フィルモア	岡田	小泉	例文
agent	動作主格(A)	S:主体	<主格>	
	<i>経験者格(E)</i>			<i>観客の一部はその芝居を楽しんだ</i>
instrument	道具格(I)	I:道具または手段	<具格>	
object	対象格(O)	O:客体	<対格>	
counter-agent				
from	源泉格(S)	OF:出発点または源	<起点格>/<有界起点格>	<i>富士山は筑波山より高い</i>
to	目標格(G)	OT:目標	<着点格>/<有界着点格>	<i>金剛は五峰まで続いた</i>
location	場所格(L)	L:場所	<位置格>	
time	時間格(T)	T:時間	<時間的な起点・着点>	
		<i>OM:相互作用の相手</i>		<i>トラックがバスとぶつかる</i>
		<i>OS:拠り所</i>		<i>子が壁に似る</i>
		<i>OC:属性の補足</i>		<i>人がその風を遠く感じる</i>
			<主題格>	<i>象は鼻が長い</i>

また、自然な事象を生成するために、格フレームの制約条件の修正が必要となる。現在、図 7 に示す 2 種類の作業を予定している。ひとつは名詞概念オントロジーの概念体系の再構成/細分化(2.2 節で述べた)であり、もうひとつは「(一般的に)牛は食べるがネコは食べない」のような細かい制約条件の設定である。

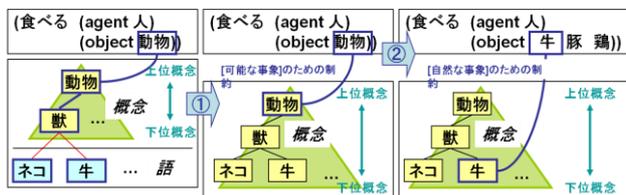


図 7 名詞概念オントロジーの問題点の解決方法

4. 応用可能性

動詞概念・名詞概念オントロジーが物語生成システムの各モジュールにおいてどのように利用できるかを考察する。まず、広告における単一事象の修辭的処理への応用について説明する。筆者らの広告生成システムの一連の研究では、広告の物語を分析し、それが商品もしくはブランドが組み込まれた物語になっている点に着目し、それを巨視的/微視的なレベル、ストーリー/

非ストーリーのそれぞれの水準でブランドが関連付けられていることを指摘している。このうち、[阿部 2008]による単一事象の修辭的処理では、テレビ広告を分析し、ストーリーの水準の微視的なレベルにおける商品導入の修辭を 12 種類抽出し、それらを単一事象の生成に用いたシステムを実装した。図 8 に商品導入の修辭(図右側の表)と、その応用システムの処理手順(図左側)を示す。

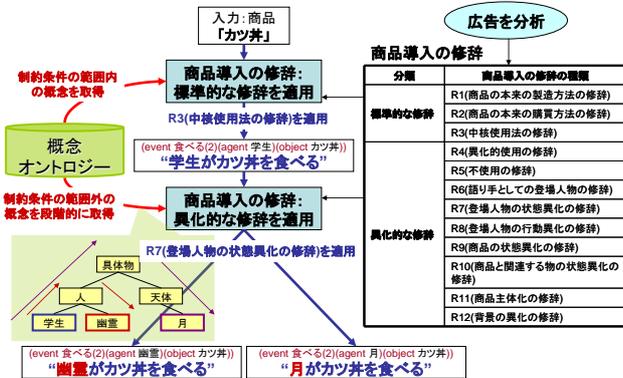


図 8 広告における単一事象の修辭的処理への応用

商品導入の修辭の中の R4~R5 は行為, R6~R8 は人物, R9~R11 は商品, R12 は場所に対して概念の異化を行う。システムでは、まず入力として商品名と商品の製造・購買・使用の各過程における標準的な修辭を選択することで、概念オントロジーから制約条件の範囲内の概念を取得し、「学生がカツ丼を食べる」のような標準的な事象を生成する。そして、この標準的な事象に対して異化的な修辭を適用する。図 8 では、例として「登場人物の状態異化の修辭」を適用し、「学生」を制約条件から段階的に逸脱させることで、低レベルの異化の場合には例えば「幽霊」、高レベルでは例えば「月」に置き換える。阿部のシステムで使用する概念はテレビ広告に登場する行為や商品、場面に限られていたが、概念オントロジーを用いればテレビ広告に限らず物語生成全般に応用出来る。

その他にも 2.3 節の評価実験で用いたように、事象概念生成及び文生成処理にも応用可能である。図 9 の左側に事象概念生成システム、右側に事象概念-単文変換システムの発展形をそれぞれ示す。これは、動詞概念「食べる(2)」をキーに事象概念「(event 食べる(2) (id 1) (agent 学生#1) (object カツ丼#1))」を生成し、これを「学生がカツ丼を喰う」という文に変換する過程を示した例である。学生#1 やカツ丼#1 は、2.3 節で用いたプログラムでは概念であったが、こちらではインスタンスである。また事象概念-単文変換システムでは、ひとつの概念に対してそれを表す複数の語彙を関連付けることで、事象概念の意味内容を変えない範囲での表現の多様化が可能である。

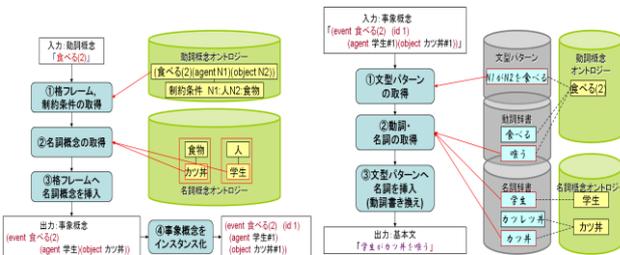


図 9 事象概念生成・事象概念-単文変換システムの発展形

5. おわりに

本研究では、物語生成システムの基盤としての動詞概念オントロジー及び名詞概念オントロジーの構築のため、動詞概念の

格フレームの構築、名詞概念オントロジーの構築、格フレームにおける制約条件の主要部分の確定の 3 つの作業を試みた。このうち動詞概念の格フレーム構築では、5337 個の格フレームを構築した。今後、格を確定出来ないパターン 174 個について、格フレーム化の規則を確立し修正する。名詞概念オントロジーは現在構築中で、約 15 万語の名詞の同音同義語を同一概念として登録する作業が終わり、同音同義語以外の概念抽出、概念のボトムアップな体系化、中間概念の挿入の各作業を現在行っている。今後、確定した格フレームと名詞概念オントロジーを用いて、格フレームにおける制約条件の主要部分の確定をする必要がある。

参考文献

[阿部 2008] 阿部弘基・小方孝・花田健自: 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, 16G-05, 2008.

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

[Bringsjord 2000] Bringsjord, S. & Ferrucci, D. A.: Artificial Intelligence and Literary Creativity, Lawrence Erlbaum Associates, Inc., Publishers, 2000.

[フィルモア 1976] フィルモア, J. C.: 格文法の原理—言語の意味と構造—, 田中春美・船城道雄 訳, 三省堂, 1976.

[池原 1999] 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: 日本語語彙大系 CD-ROM 版, 岩波書店, 1999.

[小泉 2007] 小泉保: 日本語の格と文型—結合価理論に基づく新提案, 大修館書店, 2007.

[Meehan 1980] Meehan, J. R.: The Meta-novel: Writing Stories by Computer, Garland Publishing, 1980.

[溝口 2006] 溝口理一郎 編: オントロジー構築入門, オーム社, 2006.

[Mueller 1990] Mueller, E. T.: Daydreaming in Humans and Machines, Ablex, 1990.

[Ogata 1991] Ogata, T. & Terano, T.: Explanation Based Narrative Generation Using Semiotic Theory, Proc. of National Language Processing Pacific Rim Symposium 9, pp.321-328, 1991.

[小方 1996] 小方孝・堀浩一・大須賀節雄: 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, Vol.11, No.1, pp.148-159, 1996.

[小方 2003a] 小方孝: 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, pp.127-181, 専修大学出版局, 2003.

[小方 2003b] 小方孝: 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, pp.309-356, 専修大学出版局, 2003.

[小方 2007] 小方孝: プロップから物語内容の修辭学へ—解体と再構成の修辭を中心として—, 認知科学, Vol.14, No.4, pp.532-558, 2007.

[小方 2010] 小方孝・金井明人: 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, 2010.

[岡田 1991] 岡田直之: 語の概念の表現と蓄積, 電子情報通信学会, 1991.