

依存構造を利用した言い換え文の自動生成

中野 滋 徳[†] 足 立 顕^{††} 牧 野 武 則[†]

自然言語文は、語順を変えることで多様な表現を作ることが可能である。語順を変更することで、ニュアンスの変更や強調表現をつくることができる。柔軟に語順を変えた文が生成可能であれば、タイトル文や埋め込み文の生成に役立てることができる。本稿では、語と語の直接の関係を定義した依存ネットワークを用いた言い換え文の自動生成について述べる。実験では、単文の依存ネットワークを用いて言い換え文を生成し、本手法の有効性を分析する。さらに、複文の依存ネットワークを用いた言い換え文生成の実験では、本手法がより複雑な構造を扱える手法であることを示す。

キーワード: 言い換え, 依存関係, ネットワーク

Automatic Paraphrase Generation based on Dependency Network

SHIGENORI NAKANO [†] AKIRA ADACHI ^{††} and TAKENORI MAKINO[†]

Sentence is able to generate various expressions by changing its phrase sequence. To recompose sentences by reordering a indicated phrase to end of sentence is expected to make its sentence more emphasetic and interesting. If this kind of operation can be achieved easily only by using surface information of sentences, it works usefully for title-making or emphasizing sentences. This paper describes about automatic generation of various and meaningful sentences so called paraphrase. The automatic paraphrase generation is driven from dependency network which is defined by phrase to phrase dependency relations. In the first experiment, simple sentences are tested to evaluate the usefullness of this generation metod, and the next experiment is applied for compound sentences. The result of experiment shows that automatic sentence generation based on dependency network is usefull.

Keywords: *paraphrase, dependency, network*

1. はじめに

自然言語文は、語順を変えることで強調表現やニュアンスの変更など多様な表現を作ることができる。タイトル文の生成や埋め込み文を生成する場合、語順を自由に変更した文を生成できるシステムが必須である。

本研究の目的は、語と語の直接の関係を定義した依存ネットワークを用いることで、文の本質的な意味を変えず、語を柔軟に配置した言い換え文の生成手法を発見することにある。文末表現とする語を指定することで自由な文の生成が可能な枠組みの構築を目指す。また、意味や概念などの情報を用いず、できる限り表層表現から得られる情報を用いることによって言い換え文を生成する。

言い換え技術の研究は、構文的言い換え、意味的言い換え、プラグマティック的言い換の 3 種類に分類される¹⁾。

この分類によると、構文的言い換えは統語的な情報により実現可能な言い換を指す。意味的言い換えは参照表現などに対し、同一の内容で置き換えたり、省略されているものを補完した言い換え文を生成することを指す。プラグマティック的言い換えはある状況に応じ同一の効果を持つ文への言い換を指す。

本研究はプラグマティック的言い換の研究に分類され、同様の研究には、文分割による連体修飾の言い換の研究²⁾、意味ネットワークからの文章生成の研究³⁾がある。前者の研究は、連体修飾節を含む文を 2 文に分割することで生じる、文脈上および構造上の問題点について事例分析を行った結果について述べている。後者の研究は、文や節に対して直接言い換え文を生成するのではなく、伝達すべき情報をすべて意味ネットワーク上に設定し、そこから好ましい伝達順序で、文を生成するための研究である。

本稿では、依存ネットワークを用いて言い換え文を生成した際、言い換え元である依存ネットワークと言い換え文との間で語と語の関係上の意味の保存ができるかについて、

[†] 東邦大学 大学院理学研究科 情報科学専攻
Department of Information Sciences Toho University
^{††} 富士通株式会社 情報メディアソリューション本部
Division of Media Solution, Fujitsu Co. Limited

事例分析を行う。語と語の関係上の意味を本稿では意味的關係と呼び説明で用いる。実験には、IPAL 基本動詞辞書⁴⁾を用いて意味的關係の保存の分析を行う。さらに、複文特有の構造について、人手で依存ネットワークを作成し言い換え文を生成し、複文の言い換えにおける本手法の問題点を述べる。

本稿の構成は、まず 2 節で依存ネットワークを定義し、3 節で依存ネットワークからの文生成手法を述べる。4 節で単文における言い換え文生成の実験結果を述べ、5 節で複文の言い換え実験について述べる。最後に、まとめを述べる。

2. 依存ネットワーク

語をノードとし、ノードとノードの直接の関係を示すアークを持つネットワーク型の構造を依存ネットワークと定義する。

アークには、ノードに対する関係の方向と関係情報を持つ。ノードには、文生成を行う際に用いる文生成順序情報を持つ。依存ネットワークには、構造をつくる元になった原文におけるノードの位置情報を持たず、ノード間の直接の関係のみ定義した構造である。図 1 に依存ネットワークの例を示す。

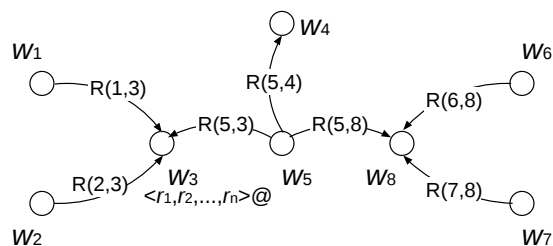


図 1 依存ネットワーク例

w_i はノードを表す。 w_i と w_j を結ぶアークは関係を意味し、矢印の向きは関係の方向を示す。 w_i と w_j を結ぶアークには関係情報を与え $R(w_i, w_j)$ と表現する。図 1 では、関係情報 $R(w_i, w_j)$ を $R(i, j)$ で表記している。 $\langle r_1, r_2, \dots, r_n \rangle @$ は、文生成順序情報を意味し、関係情報 $r_p = R(w_i, w_j)$ と $r_q = R(w_x, w_y)$ が与えられた場合に r_p, r_q のどちらを先に出力するかを決定する情報として用いる。 @ は文生成順序情報を持つノードを指す。 $\langle \dots, r_p, \dots, r_q, \dots \rangle @$ と指定されている場合、 r_p で関係付けられているノードを先に出力し、 r_q で関係付けられているノードを出力する。すべての関係情報に対して出力が終わると自ノードを後に出力し、文生成処理を終える。つまり、文生成順序情報は生成結果におけるノードと、関係するノードとの位置関係を指定したものである。

3. 文生成手法

図 2 に本実験で用いる依存ネットワークの具体例を示し、文生成の流れを説明する。

図 2 に示した依存ネットワークでは、ノードには語を設

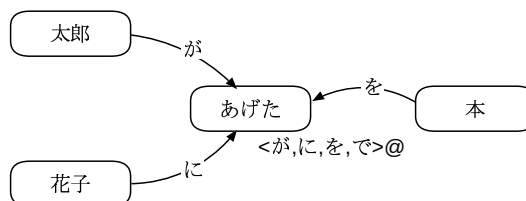


図 2 依存ネットワークの具体例

定し、関係情報には表層格をそのまま指定したものを用いる。文生成順序情報に設定する表層格は、依存ネットワークに關係が設定されて「いる／いない」に関わらずノードが文生成時において要求する生成文に対する好ましい順序を指定する。文生成情報に設定されていない関係情報は自ノードの直前に出力する関係として扱う。依存ネットワークにおけるアークの方向は、名詞が動詞に与える格関係の向きと一致する。

言い換え文の生成は、アークの方向と文生成順序情報によって行う。ノード「あげた」を文末表現とした文生成を行う場合、「あげた」に対して「が」「に」「を」を関係情報として持つアークが存在する。「あげた」を出力するために、関係情報と文生成順序情報とを比較し出力順序を決定する。生成結果として「太郎が花子に本をあげた」を生成結果に得る。

ここで、ノード「本」を文末表現とした文の生成について述べる。ノード「本」には、関係情報として $R(\text{“本”, “あげた”})$ が設定されている。これは、「本」の「あげた」に対する関係である。「本」を文末表現とするためには、 $R(\text{“あげた”, “本”})$ に関係の方向を変換する操作が必要となる。本実験における文生成システムは、アークの方向を変換する操作に対し表 1 に示す 2 つの制約を与える。

表 1 アークの方向変換操作における制約

$$\begin{aligned} R(w_i, w_j) = r_k &\Rightarrow R(w_j, w_i) = \phi \\ R(w_i, w_j) = \phi &\Rightarrow \text{関係の方向の変換不可} \end{aligned}$$

ϕ は関係のみ存在している状態を示す。

図 2 に示した構造に対し、ノード「本」を文末表現に指定して、アークの方向の変換を行った結果を図 3 に示す。なお、文末表現として指定したノードを図中の太線で示す。

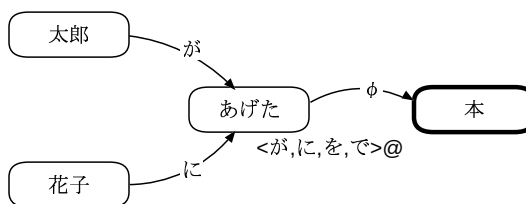


図 3 アークの方向の変換結果

「 $R(\text{“本”, “あげた”}) = \text{を}$ 」に対してアークの方向変換操作を行う。「 $R(\text{“あげた”, “本”}) = \phi$ 」となり関係のみが残され

る。これは、連体修飾節が名詞を修飾する場合の操作と一致する。

ノード「本」には、文生成情報を設定していないため関係するノードに対して文生成を行うが、関係するノード「あげた」に対する文生成が行われていないためアークをトラバースし、ノード「あげた」に文生成処理を移す。ノード「あげた」では、「が」「に」の関係と文生成順序情報から「太郎が花子にあげた」が生成され、ノード「本」に処理が戻ってくる。その結果として「太郎が花子にあげた本」を得る。

この操作は、より複雑な複文においても同様に扱うことができる。図 4 に依存ネットワークをあげ説明する。

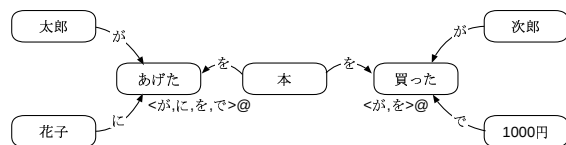


図 4 複文における依存ネットワーク例

図 4 は、「太郎が花子にあげた本を次郎が 1000 円で買った」に対応する依存ネットワークである。ノード「次郎」を文末表現に指定すると、図 5 に示すようにアークの方向の変換操作が行われる。なお、文末表現として指定した部分を図中の太線で示す。

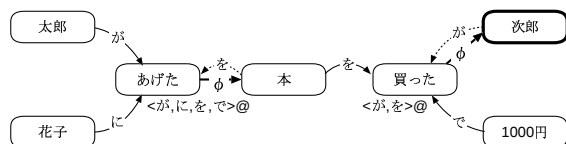


図 5 複文における依存ネットワークの変換例

アークを点線で表記した部分の関係が変換され、関係情報が ϕ となる。すでに述べたようにして、ノード「次郎」より文生成処理を行う。その結果、「太郎が花子にあげた本を 1000 円で買った次郎」を得る。各ノードを文末表現に指定し生成した言い換え文を表 2 に示す。

表 2 複文の言い換え例

文末表現	言い換え文
太郎	次郎が 1000 円で買った本を花子にあげた太郎
花子	次郎が 1000 円で買った本を太郎があげた花子
あげた	次郎が 1000 円で買った本を太郎が花子にあげた
本	太郎が花子にあげた次郎が 1000 円で買った本
1000 円	太郎が花子にあげた本を次郎が買った 1000 円
次郎	太郎が花子にあげた本を 1000 円で買った次郎
買った	太郎が花子にあげた本を次郎が 1000 円で買った

本文生成システムは、依存ネットワーク上のアークに関係情報をすべて設定している場合、すべてのノードを文末表現に指定して言い換え文を生成することができる。

4. 単文の言い換え実験

実験データには、IPAL 基本動詞辞書を用いる。この基本

動詞辞書にはサブエントリの動詞に対応して、述語素、文例 (3 つを上限)、格形式などが収録されている。述語素は動詞と名詞句の関係、名詞句同士の関係を同時に表そうとしたものである。場所、抽象的空間、原因、目的、変化、作用等に関する 57 種の述語素に分類されている。

“日本語動詞の諸相”⁵⁾によれば、この述語素は動詞と格形式をもった名詞句との間の意味的關係と定義されている。そして、「が」格、「を」格、「に」格のいずれかの名詞句が前提となつて、はじめて関係的な意味が取り出せると述べている。本稿では、この述語素が指し示す動詞と表層格との緊密な関係を意味的關係と呼ぶ。

実験では、依存ネットワークから言い換え文を生成するためにアークの方向の変更が必要な場合、その関係情報に対し表 1 で示した操作を行う。その結果、生成した言い換え文では、表層表現としての関係情報が失われ、修辭関係として表現した文を得る。語と語の関係を文中で示す表層表現を失うことによる言い換え文に与える意味的關係の影響を分析する。

アークの方向の変更を伴う言い換え文生成を行った際、依存ネットワークと言い換え文との間で、ノード間の意味的關係が保たれているかを検証することに目的を置く。なお、本稿ではノード間の意味的關係が保存されている状態を等価であると呼び、ノード間の意味的關係が保存された状態での言い換え操作に対して、等価性がある言い換えと呼ぶ。

4.1 依存ネットワークの生成

基本動詞辞書の表記 (サブエントリ) ごとの文例 5,284 文から複文を除去し、動詞が文末で使用されている単文を対象に、述語素が示す表層格と一致する文例を抽出し、3,227 件を得た。基本動詞辞書の文例は、「私の箸が」のように名詞部分は直接動詞と関係する単位で整理されたものである。

この辞書には文形式として語の関係を定義した情報が付与されている。 w_1, w_2 を名詞、 w_3 を動詞とする「 w_1 が w_2 を w_3 」の文形式を例に依存ネットワークの自動生成について述べる。文形式「 w_1 が w_2 を w_3 」は、 w_1 が関係情報「が」によって w_3 と関係していることを意味する。したがって、各ノード間の関係は、「 $R(w_1, w_3)=が$ 」「 $R(w_2, w_3)=を$ 」となり依存ネットワークを生成できる。

自動生成した依存ネットワークから言い換え文を自動生成し、アークの方向の変更により、表層表現としての関係情報を失った文を、表層格、意味的關係別に整理して原文と比較し、等価性のある言い換えができたかの分析を人手により行った。

4.2 単文言い換え結果の分析

表 3 に言い換え文と原文との間での等価性の分析結果を、表層格とそれに対応する意味的關係でまとめたものを示す。

表 3 の対象件数は、自動生成した言い換え文の件数である。() 内は疑問文、命令文、確言、概言等のムード表現を含んだ文の数 (内数) である。これらの文法的特徴は、基礎日本語文法⁶⁾を参考に人手で分類した。等価件数とは原文と比較した結果、等価性のある言い換え文を得た件数であ

る。比率は等価件数に対する対象件数に対する比率を示す。

表 3 には「へ」格に対応する項目がない。これは動詞辞書の中で「へ」格は「に」格と交替可能として定義されている。そのため、「へ」格は、交替可能な格を用いて分析した。また「が」格の抽出件数が少ない。基本動詞辞書での文例作成方針が「が」格に「は」が用いられたことによる。

表 3 格助詞単位の等価性評価の結果

格	対象件数	等価件数	比率
が	122(0)	122(0)	100%
を	1,597(28)	1,571(2)	98%
に	1,070(19)	994(4)	93%
から	201(0)	84(0)	42%
と	75(2)	53(0)	71%
より	11(0)	0(0)	0%
で	151(0)	113(0)	75%
合計	3,227(49)	2,937(6)	91%

対象件数全体の等価性は 91%で、等価性を失っている文は 3,227 文中 290(43) 文であった。() 内はムード表現を含んだ文の数である。表層格ごとに比較すると等価性に差がある。

全体に共通する傾向として、ムード表現を含んだ文は、一部の例外を除き等価性に欠けていた。以下に文例を示す。

「」内は基本動詞辞書の例文を示し、言い換え文は先頭に「→」をつけ、等価性を欠く文には文の先頭に「*」を付けた。言い換え文の語順が例文と異なっているのは、依存ネットワークの文生成順序情報を適用した結果である。

「結局 私が ばかを 見たのだ。」
→ *私が 結局 見たのだ ばか
「ビールを 上がりませんか。」
→ *上がりませんか ビール

ムード表現の中で等価性が認められたのは全部で 6 件で、以下に示す例文の願望表現や当為表現に等価性があった。

「私が あいつと 早く 縁を 切りたい。」
→ 私が 早く 縁を 切りたい あいつ
「私が 君に 不手際を 謝らねばならない。」
→ 私が 不手際を 謝らねばならない 君

ムード表現は分析を乱す要因となると判断し、本実験では対象件数の中からムード表現 49 件を除去した 3,178 件を分析対象とした。したがって、以降の説明で述べる分析対象からはムード表現を除去した。

分析対象の述語素の分類に一致した数は 57 種類中 42 種類であった。また、表層格ごとに述語素の意味的関係の表現が同じものをまとめ、32 種類とした。以下、表層格と意味的関係の単位に分析し集計した結果を示す。

「が」格,「を」格の分析結果

「が」格および「を」格の分析結果を表 4 に示す。

表 4 「が」格の等価性評価

格	意味的関係	対象件数	等価件数	比率
が	動作主	8	8	
	焦点	39	39	
	対象〔変化〕	58	58	
	対象〔消滅〕	2	2	
	対象〔出現〕	15	15	
	合 計	122	122	100%
を	対象〔受影〕	319	319	
	対象〔変化〕	321	321	
	対象〔消滅〕	41	41	
	対象〔出現〕	92	92	
	対象〔その他〕	747	747	
	空間	36	36	
	時間	13	13	
	合 計	1,569	1,569	100%

「が」格は 5 項目,「を」格は 7 項目の意味的関係がある。いづれの分類においても等価性がある言い換え文を得た。

「に」格の分析結果

「に」格の分析結果を表 5 に示す。

表 5 「に」格の等価性評価

格	意味的関係	対象件数	等価件数	比率
に	起因	27	27	100%
	態度の対象	17	17	100%
	関連	47	40	85%
	資格	32	28	88%
	動機	19	19	100%
	相手	71	69	97%
	範囲	60	60	100%
	空間的着点	227	202	89%
	空間的場所	283	280	99%
	非空間的着点	121	110	91%
	非空間的場所	147	138	94%
	合 計	1,051	990	94%

「に」格は 11 項目の意味的関係がある。このうち、4 項目はすべてのデータに対し等価性のある言い換え文を得た。「空間的場所」も次の 1 件を除いた以外のデータに対し等価性のある言い換え文を得た。

「牧草地在 一面に 広がる。」
→ *牧草地在 広がる 一面

「非空間的場所」は 94%の等価性のある言い換え文を得た。等価性に欠けた文は、以下に示す日時などの表現を含んだ文である。

「彼は 久しぶりに 布団を 日に 干した。」
→ *彼は 久しぶりに 布団を 干した 日
「試験が 三日後に 迫った。」
→ *試験が 迫った 三日後

「空間的着点」と「非空間的着点」は 90%前後の等価性

のある言い換え文を得た．等価性に欠けた文は、方向や程度を表す表現と結びついた文である．

「彼が 手を 左右に 振った。」
→ *彼が 手を 振った 左右
「会員が １０００名に 膨れた。」
→ *会員が 膨れた １０００名

「関連」「資格」「相手」に関しては、等価性の分析結果にばらつきがある．等価性に欠けた文は、「に」格を伴う名詞が動詞とが強く結び付いた文である．

「１マルクは 約１００円に 当る。」
→ *１マルクは 当る 約１００円
「彼が 人によく 絡む。」
→ *彼が よく 絡む 人
「豹が 人に 馴れない。」
→ *豹が 馴れない 人

「から」格,「と」格,「で」格の分析結果

「から」格,「と」格,「で」格の分析結果を表 6 に示す．

表 6 「から」格の等価性評価

格	意味的關係	対象件数	等価件数	比率
から	動機	3	0	0%
	相手	2	0	0%
	空間的起点	148	61	41%
	非空間的起点	48	23	48%
	合 計	201	84	42%
と	内容	21	9	43%
	対称	52	44	85%
	合 計	73	53	73%
で	手段	143	107	75%
	資格	8	6	75%
	合 計	151	113	75%

「から」格は 4 項目,「と」格と「で」格はそれぞれ 2 項目の意味的關係がある．これらの格により意味的關係を与える語が、起点ならびに対称を表すときに等価性が失われる傾向がある．

「で」格で等価性のない言い換えと判断する文のなかで、「ている」などのアスペクト表現が影響を与える傾向がみられた．下の文例で、「新幹線が 空いた 運賃の値上げ」の場合、等価性があると言えるが、「空いている運賃の値上げ」では等価性のある言い換えとはいえない．具体例を以下に示す．

「玄関に 彼が 奥の部屋から 現れた。」
→ *玄関に 彼が 現れた 奥の部屋
「富士山が 汽車の窓から 消えた。」
→ *富士山が 消えた 汽車の窓
「画商が 偽物の絵を 本物と 偽った。」
→ *画商が 偽物の絵を 偽った 本物
「小結が 大関と 取り組む。」
→ *小結が 取り組む 大関
「新幹線が 運賃の値上げで 空いていた。」
→ *新幹線が 空いていた 運賃の値上げ

[より格]

表 7 に示す「より」格では等価性がある言い換え文がなかった．

表 7 「より」格の等価性評価

格	意味的關係	対象件数	等価件数	比率
より	比較	11	0	0%

比較表現で使用するヨリ格の名詞句が文末に来ると、比較する対象が消滅するため等価性を失う．文例を示す．

「この商品が 見本より 落ちる。」
→ *この商品が 落ちる 見本
「彼は 名より 実を取った。」
→ *彼は 実を取った 名

以上、単文の言い換え実験における言い換え文をもとにした等価性の分析結果として、ムード表現が等価性に強い影響を与えていることがわかった．

「が」格,「を」格の等価性は安定している．「に」格も、数量、程度などを表す名詞のときなどの等価性が欠ける文を除けば、高い等価性を示した．

意味的關係で、比較を表す「より」格では等価が認められる言い換え文を見いだすことができなかった．また、起点を表す「から」格や、対称性を表す「と」格、着点を表す「に」格では、等価性を欠く傾向がある．

なお,「で」格で動詞に「ている」等のアスペクト表現を含むと等価性がなくなる傾向があった．このことから等価性の判断に、アスペクトの要素が影響を与えているといえる．

しかしアークの方向の変換を伴う言い換え文の生成により、表層表現としての関係情報を失う場合も、9 割以上の等価性のある言い換え文を得た．

5. 複文の言い換え実験

単文より複雑な文の構造をもつ文に対して、依存ネットワークによる言い換え文生成手法を適用して、全てのノードを文末表現とする言い換えについて実験する．

複文が単文と異なるその特徴として、全体の中心となる主節とそれに特定の関係で結び付く接続節で構成される．接続節は従属節と並列節に分かれ、従属節はさらに補足節、副詞節、連体節に分類⁶⁾できる．

本実験では連体節、副詞節、形式名詞をそれぞれ含む文を対象に言い換え実験を行い、依存ネットワークによる文生成手法が制限されるのはどういう条件のもとで起るかを検証する．

それぞれ依存ネットワークを人手で作成して言い換え文を出力し分析する．

図 6 の依存ネットワークには、「R(“盆地”,“入る”)＝に」を設定した．この例文では、表 8 に示す 5 つの文を生成した．原文の連体修飾の関係表現で「盆地」を文末にした言い換えでは「東京 2 3 区がすつぽり入る」と「氷点下 1 5

東京 2 3 区が すっぱり 入る 盆地は、しばしば
氷点下 1 5 度もの寒気で 覆われる。

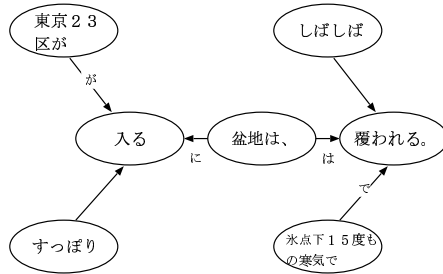


図 6 連体修飾を含んだ複文の依存ネットワーク図

表 8 連体修飾を含んだ複文の言い換え

文末表現	言い換え文
東京 2 3 区	しばしば 氷点下 1 5 度もの寒気で 覆われる 盆地に すっぱり 入る 東京 2 3 区
すっぱり	《関係表現がないため出力できなかった》
入る	しばしば 氷点下 1 5 度もの寒気で 覆われる 盆地に 東京 2 3 区が すっぱり 入る
盆地	東京 2 3 区が すっぱり 入る しばしば 氷点下 1 5 度もの寒気で 覆われる 盆地
しばしば	《関係表現がないため出力できなかった》
氷点下 1 5 度 もの寒気	東京 2 3 区が すっぱり 入る 盆地は しばしば 覆われる 氷点下 1 5 度もの寒気
覆われる	東京 2 3 区が すっぱり 入る 盆地は しばしば 氷点下 1 5 度もの寒気で 覆われる

度もの寒気で覆われる」の並列した連体修飾となっている。原文との等価性という点では問題ない。なお、「すっぱり」「しばしば」は関係情報が設定できないために、それらを文末表現にした言い換え文は生成しない。

さらに複雑な連体節をもつ文の言い換え文の生成実験を行う。

航空機に 積んだ レーダーで夜でも 海面の油汚染を 2 0 キロ幅で
キャッチする 装置が 1 0 日、電波研究所で 開発された。

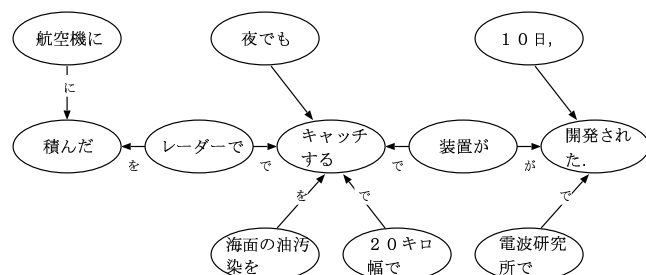


図 7 複雑な連体節をもった複文の依存ネットワーク図

図 7 の依存ネットワークには、「R(“レーダー”, “積む”)=を」「R(“装置”, “キャッチ”)=で」の関係表現を設定した。この例文では表 9 に示す 9 文を生成した。9 文いずれも原文との等価性をもっている。

表 9 複雑な連体節をもった複文の言い換え

文末表現	言い換え文
航空機	1 0 日、電波研究所で 開発された 装置で 2 0 キロ幅で夜でも 海面の油汚染を キャッチする レーダーを 積んだ 航空機
積んだ	1 0 日、電波研究所で 開発された装置で 2 0 キロ幅で 夜でも 海面の油汚染を キャッチする レーダーを 航空機に 積んだ
レーダー	1 0 日、電波研究所で 開発された 装置でも 2 0 キロ幅で 夜でも 海面の油汚染を キャッチする 航空機に 積んだ レーダー
夜でも	《関係表現がないため出力できなかった》
海面の油汚染	航空機に 積んだ レーダーで 1 0 日、 電波研究所で 開発された 装置で 2 0 キロ幅で 夜でも キャッチする 海面の油汚染
2 0 キロ幅	航空機に 積んだ レーダーで 1 0 日、電波研究所で 開発された 装置で 夜でも 海面の油汚染を キャッチする 2 0 キロ幅
キャッチする	航空機に 積んだ レーダーで 1 0 日、電波研究所で 開発された 装置で 2 0 キロ幅で 夜でも 海面の油汚染を キャッチする
装置	航空機に 積んだ レーダーで 2 0 キロ幅で 夜でも 海面の油汚染を キャッチする 1 0 日、 電波研究所で 開発された 装置
1 0 日、	《関係表現がないため出力できなかった》
電波研究所	航空機に 積んだ レーダーで 2 0 キロ幅で 夜でも 海面の油汚染を キャッチする 装置が 1 0 日、開発された 電波研究所
開発された	航空機に 積んだ レーダーで 2 0 キロ幅で 夜でも 海面の油汚染を キャッチする 装置が 1 0 日、電波研究所で 開発された

つぎに副詞節を含んだ複文について言い換え文の言い換え文の生成実験を行う。

採決が 終わった後で、大勢の人が 反対意見を 言いはじめた。

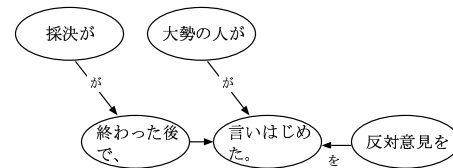


図 8 副詞節を含んだ複文の依存ネットワーク図

図 8 の依存ネットワークには、「R(“終わった”, “言いはじめた”)=φ」に対してアークの方向変換操作を行うための関係表現が存在していない。したがって「採決」「終わった」の副詞節内のノードを文末表現にした文生成ができない。

副詞節による接続は、時、原因、理由、条件、程度等の関係⁶⁾で構成される表層表現として、関係が示されていない。本文生成手法では表層表現を用いない関係を設定するため、表層表現を伴わない関係に対応した方向変換操作を行うことができない。したがって、主節部分は言い換え可能であるが、副詞節部分の言い換えができない。

表 10 副詞節を含んだ複文の言い換え

文末表現	言い換え文
採決	《関係表現がないため出力できなかった》
終わった	《関係表現がないため出力できなかった》
大勢の人	採決が終わった後で、反対意見を 言いはじめた 大勢の人
反対意見	採決が 終わった後で、大勢の人が 言いはじめた 反対意見
言いはじめた	採決が終わった後で、大勢の人が 反対意見を 言いはじめた

さらに形式名詞を含んだ文についても同様の言い換え文の生成実験を行う。

部長は 鈴木さんに 報告書を すぐに 提出する ことを 命じた。

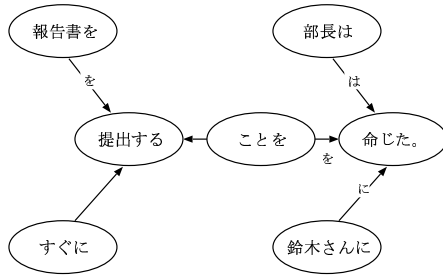


図 9 形式名詞を含んだ複文の依存ネットワーク図

表 11 形式名詞を含んだ複文の言い換え

文末表現	言い換え文
部長	報告書を すぐに 提出する ことを 鈴木さんに 命じた 部長
鈴木さん	報告書をすぐに 提出する ことを 部長は 命じた 鈴木さん
報告書	《関係表現がないため出力できなかった》
提出する	《関係表現がないため出力できなかった》
こと	報告書を すぐに 提出する 部長は 鈴木さんに 命じた こと
命じた	報告書を すぐに 提出する ことを 部長は 鈴木さんに 命じた

図 9 において、「R(“こと”, “提出する”)」の関係情報を与えることができない。つまり形式名詞を用いた構造においては、形式名詞を連体修飾する部分との逆方向の関係が定義できない。

以上、複文を対象とした実験からは、ノード間の関係情報を設定できるときは、複雑な文でもあらゆるノードを文末表現とする等価性のある言い換え文を生成できることが判明した。しかし、副詞節や形式名詞等のノード間の関係情報を設定できないときは、言い換え文の生成ができない。

6. ま と め

本稿では、依存ネットワークの定義を行い、それを用いた言い換え文生成手法について述べた。依存ネットワークを用いることで、柔軟に語順を変更した文が生成可能であることを示した。

単文の言い換え実験では、IPAL 基本動詞辞書を用い語

と語の意味的關係が、原文と言い換え文の間で保存されているかの分析を行った。その結果、およそ 90%程度の単文が本手法により言い換え可能であることが分かった。

複文での言い換え実験では、複文固有のいくつかの構造について人手で依存ネットワークを作成し、言い換え文が生成可能であるかの分析を行った。形式名詞を用いない連体修飾節のみ含む構造では、単文における言い換え文生成同様に、すべての語を文末表現とした言い換え文が生成可能であることが分かった。一方、形式名詞を用いた構造では、形式名詞を連体修飾する部分との逆方向の関係を定義できないなどの制約から、文末表現として指定できないノードが存在することが分かった。また、副詞節などを含む構造では、副詞節と関係する語との間にある理由、原因などの因果関係が表層表現として扱えないため、言い換え文を生成することができなかった。

今後の課題としては、複文から依存ネットワークを自動的に生成する仕組みおよび、語形変化を伴う言い換え文の自動生成を可能とする仕組みの研究について行っていく必要がある。

参 考 文 献

- 1) 佐藤理史: “論文表題を言い換える”. 情報処理学会, 論文誌, Vol.40, No.7, pp.2937-2945. (1998).
- 2) 野上優, 藤田篤, 乾健太郎: “文分割による連体修飾の言い換え”. 言語処理学会, 第 6 回年次大会発表論文集, pp.215-218, (2000)
- 3) 尾崎正太郎, 黒橋禎夫, 長尾真: “意味ネットワークからの文章生成”. 言語処理学会, 自然言語処理研究会, NL120-20, pp.133-140, (1997)
- 4) 情報処理振興事業協会技術センター: “計算機用日本語基本動詞辞書 IPAL(Basic Verbs) - 解説編 -”. (1987)
- 5) 村木新次郎: “日本語動詞の諸相”. ひつじ書房, (1991)
- 6) 益岡隆志, 田窪行則. “基礎日本語文法”. くろしお出版, (1992)