

中国語の文における照応生成アルゴリズム*

2Q-4

孫 土生 費 佳 唐澤 博†

山梨大学工学部‡

E-mail: son, karasawa@opal.esi.yamanashi.ac.jp

1 はじめに

自然言語対話システムにおいて、文生成部に位置して、f-構造から中国語の文を生成する。人とコンピュータが円滑なコミュニケーションを行なうため、照応表現の生成はたいへん重要である。本研究は、日本語における研究[1]を参考して、中国語の文法に従って前方照応制約などを記述し、前方照応制約に対して中国語で取り扱う単文、複文、重文のアルゴリズムを決定し、これを実装した。その結果を報告する。

2 中国語文生成に関する文法

2.1 中国語の文

中国語の文の定義：

文は一つの完全の意味を表す言語単位。文の種類は単文と複文と重文がある。

2.2 中国語の文の成分

文の成分は主に以下の七つである：主語、述語、目的語、修飾語、補語、状語、独立語。

2.3 中国語の詞の分類

詞類は実詞と虚詞二種類がある。実詞は単独に文の成分になれる。虚詞は実詞と一緒に合わせて文の成分になれる。実詞は名詞、代名詞、動詞、形容詞、副詞、数詞、量詞、嘆詞、擬音詞、区別詞を含む。虚詞は助詞、介詞、連詞、語気詞を含む[2]。

3 中国語の照応生成

3.1 中国語の代名詞

中国語の照応関係を構成する代名詞は人称代名詞と指示代名詞の二種類である[3]。人称代名詞と指示代名詞を表1、表2示す。

表1 人称代名詞

	単数	複数
第一人称	我	我们
第二人称	你	你们
第三人称(男)	他	他们
第三人称(女)	她	她们
疑问人称	谁	哪些人
其他	自己	大家

表2 指示代名詞

	近称		远称		疑问称	
	単数	複数	単数	複数	単数	複数
物	这个	这些	那个	那些	哪个	哪些
人	这位	这些人	那位	那些人	哪位	哪些人
场所	这里		那里		哪里	
方向	这边		那边		哪边	
指示	这样的		那样的		怎样的	
时间	这时		那时		几时	

3.2 中国語の照応生成の分類

中国語の照応は文内照応と文間照応である。

(1) 文内の照応

a. 単文の内部で照応する

小李 在自己的房间里 学习。

(李さんは自分の部屋で勉強する。)

先行詞：小李

照応詞：自己

b. 複文の内部で照応する

小李 想 自己 被老师 表扬 了。

(李さんは自分が先生に褒められたと思った。)

先行詞：小李

照応詞：自己

(2) 文間の照応

运动会之后，马克过去的大部分同学都去了查克的农场家用中餐。马克的父母亲也在那里，显然在等候我们。

(運動会の後、马克さんの過去の同級生は查克さんの農場家へ昼食を食べに行った。马克の両親もあそこに行って、私達を待っていた。)

先行詞：查克的农场家

照応詞：那里

*Anaphora Generation Algorithm for Chinese Text

†Son Dosei, Hi Ka, Hiroshi Karasawa

‡Yamanashi University, 4-3-11 Takeda, Kofu, Yamanashi 400-8511, Japan

先行詞：马克过去的大部分同学

照応詞：我们

3.3 照応表現に関するアルゴリズム

3.3.1 中国語の照応生成の制約

代用を行なう際には以下に述べるような制約を守って代用を行なわなければならない。この制約を守って代用を行なうことにより聞き手はこの制約を逆に使って代用詞を元の表現に戻すことができる。このため、この制約を守らないで代用を行なうと、聞き手は代用詞を元の表現に戻すことができなくなってしまうたり、間違った表現に戻してしまう恐れがある[4]。

制約 1：人称に矛盾を起こすような代用詞の生成はできない。

制約 2：数に矛盾を起こすような代用詞の生成はできない。

制約 3：因果関係に矛盾を起こすような代用詞の生成はできない。

制約 4：前文で視点になっていなければ代用詞の生成はできない。

3.3.2 文に関するアルゴリズム

(1) 単文

目的語や状語などの修飾語の属性と単文の主語の属性を一致する場合、目的語や状語などの修飾語の属性に対して代用を行なう。主語は人限りである。主語によって代用詞は「自己」や「第三人称の代名詞」や「第三人称の再帰代名詞」などを使う。また修飾語の属性と主語の属性は親族の関係を持てば、代用語として「兄弟」、「姉妹」、「親」を使う。

(2) 複文

場所や時間や方式や対象や目的語などに関する修飾文の主題と複文の主題が一致する時、代用を行なう。代用詞は、複文の主題によって「自己」や「第三人称の代名詞」や「第三人称の再帰代名詞」などを使う。また修飾語の属性と主語の属性は親族の関係を持てば、代用語として「兄弟」、「姉妹」、「親」を使う。

(3) 重文

重文の代用は前文と後文の関係で、前文で出ていた主題や場所や事や時間などを後文でもう一度出る場合照応を行なう。後文の各成分は前文で出ていた各成分と一致する場合に代用する。重文の代用は制約1から制約4まで従って代用を行なう。主題は人の場合、代用詞は人称代名詞を使う。その他の場合は指示代名詞を使う。

4 実装結果

入力文：

小李 在 小李 的 房间里 学习。

(李さんは李さんの部屋で勉強している。)

f- 構造にする記述：

```
([stype=decl,
endsym=[lexID=9201,pred='。'],
subj=[lexID=1938,pred='小李',
categ=pron,type=pers,
gender=masc,sp=sing,pers=3],
oblloc=[lexID=1985,pred='房间',
categ=noun,type=male,
poss=[lexID=1938,
pred='小李',
categ=noun,type=male],
lexID=6173,pred='学习',
categ=verb,type=intr,sent=3])).
```

出力文：

小李 在 自己 的 房间里 学习。

入力文：

小李 认为 小李 做得 对。

(李さんは李さんが正しくやっていると思う。)

出力文：

小李 认为 自己 做得 对。

5 終わりに

中国語の文に関しての照応生成を実現した。しかし、複文と重文に関しての照応生成は不十分であり、また、より自然会話を生成するため、省略の処理は重要である。よって、今後の課題として、よりさまざまな文の形に対応する複文と重文の照応生成ができること。さらに、中国語の文に関しての省略の処理は今後の一つのテーマになる。

参考文献

- [1] 大石,唐澤:日本語の複文・重文における照応表現の生成,情報処理学会第54回全国大会講演論文集(2),6B-4, pp.43-44 (1997.3)
- [2] 守屋 宏則:やさしく くいしい 中国語の文法の基礎,東方書店 (1995)
- [3] 朱德熙:语法讲义,商務印書館 (1997)
- [4] 山村毅,大西昇,杉江昇:日本語文の照応問題解決システム,電子情報通信学会論文誌,Vol.J73-D-2, No.6, pp.887-896,(1990)