

3K-8

省略語の回復処理と名詞文について

中嶋 正夫⁽¹⁾ , 池田 尚志⁽²⁾

(1) 明治大学 , (2) 電子技術総合研究所

【 1 】はじめに

日本語文章理解システム構築の1歩として、名詞文を中心とした省略の問題について考察する。これは、名詞句の同定処理の1部となる。ここで、名詞文とは「AはBだ」の形の名詞文をいう。名詞文は、A、Bにくる名詞句により解釈が異なり、「僕はウナギだ」のように省略文であることも多い。本稿では、名詞文における省略の問題を中心に考察し、その回復処理を組み込んだ質問応答システムについて述べる。

【 2 】名詞文と省略

(a) 名詞文の種類と省略

名詞文「AはBだ」の解釈としては、次の4種が考えられる。

①subclass関係

AがBに含まれるクラスである場合。

②instance関係

AがクラスBに所属するインスタンスである場合。

③eq関係

AとBが同じ者であることをいう場合。

④ellip

「僕はウナギだ」の類の省略文である場合。

このうち①～③が、純粹の名詞文である。

これらの解釈は、A、Bがclassであるか、instanceであるかによって次のように分布する。

(1) A:class→B:class

- subclass 猿は動物だ。
- ellip { ネズミは猫を恐れる。
 { コブラはマンガースだ。

(2) A:class→B:instance

- ellip { 寿司を食べたいのは太郎だ。
 { うなぎは二郎だ。

(3) A:instance→B:class

- instance 太郎は人間だ。
- ellip { 太郎は寿司が好きだ。
 { 二郎はうなぎだ。

(4) A:instance→B:instance

- eq 日本は首都は東京だ。
- ellip { 太郎は花子が好きだ。
 { 二郎は洋子だ。

本稿では④の型の省略と、次のような名詞句の一部についての省略について考察する。

ex. ネズミの天敵は猫だ。

{ a コブラはマンガースだ。

{ b 好物はチーズだ。

aでは主部に「コブラの天敵」が省略され、bでは「ネズミの好物」が省略されている。

(b) 名詞文の解釈

名詞文「AはBだ」が、前述の4種のうちのどれであるかを判断するには、次の2つが問題となる。

i. A、Bがそれぞれclassを指しているのか、instanceを指しているのか。

ii. 名詞文が省略文と解釈すべきものかどうか。

iについては、名詞句に「この」、「その」などの指示詞を伴って特定の対象を指すとき以外は、classと判断してよいように思われる。iiについては、次の2つの条件をたてることができる。

[条件1]・A、Bとも新情報である。すなわち、A、Bとも直前の文に現れない。

(但し、「AもBだ」の場合には、

Aが新情報、Bが旧情報である。)

[条件2]・直前の文にA、Bに対応する語があり、前文に重ね合わせることで解釈が可能なとき。

ex. 1) タールはタバコに含まれている。

タールは毒物だ。

2) タールはタバコに含まれている。

サリチル酸は酒だ。

1)では、タールが直前の文に現れており、省略文と解釈できない。一方、2)では“サリチル酸は酒に含まれている。”と解釈するのが妥当である。

ex. 1) 昨日は太郎が泳いだ。

今日は二郎だ。

2) 昨日は太郎が泳いだ。

今日は日曜日だ。

3) 太郎は花子が好きだ。

花子は美人だ。

1)は条件1を満たし、また二郎→太郎と対応させれば条件2も満たすから、省略文として解

Recovery of ellipsis and noun sentence in Japanese

Masao Nakajima⁽¹⁾ , Takashi Ikeda⁽²⁾

(1) Meiji University , (2) Electrotechnical Laboratory

積できる。2)は条件1を満たしているが、前文との重ね合わせでは解釈できないから、条件2に反しており、省略文としての解釈はできない。一方、3)は前文との重ね合わせによる解釈は可能であるが、条件1は満足せず、省略文と解釈できない例である。

(c)名詞文の解釈手順

以上の考察に基づいた、名詞文の解釈手順を図1に示す。

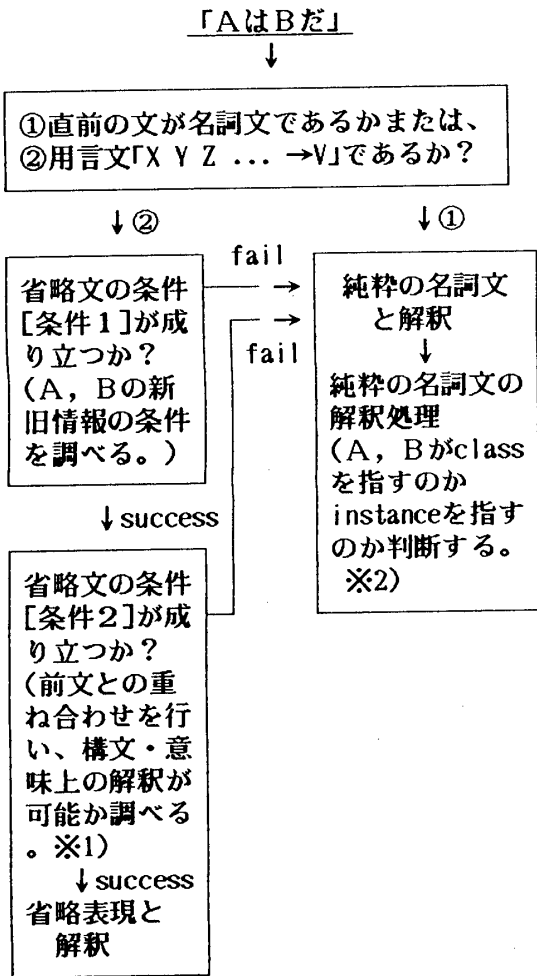


図1 名詞文の解釈手順

※1:名詞文「AはBだ」を前文と重ね合わせると言うことは、前文が用言文「X Y Z ... → V」のとき、X, Y, Z, ...のうちのどれか2つをA, Bで置き換えて解釈してみることである。但し、例えばXをAで置き換える場合、Xが[aのb]という構造をしているときは、X:=Aの置き換えだけでなく、X:=Aのb, X:=aのA.のような置き換えも行う。

※2:直前の文が名詞文「A'はB'だ」で、例えばA'が[aのb]の構造をしているときは、「aのAはBだ」、「AのbはBだ」の解釈も試みる。

【 3 】質問応答システム

(a) 質問応答システムの構成

現在、日本語文章を理解し、その内容についての質問に答える質問応答システムの作成を試みており、そこに、以上の省略語の処理を組み込んだ。システムは次の4つの部門から成る。

(1)構文・意味解析部門

入力文の構文・意味解析¹⁾を行い、依存構造木を出力する。

(2)翻訳部門

構文・意味解析結果を、その文が表現する述語式へ翻訳する。

(3)同定処理部門

代名詞や指示詞の解釈、普通名詞の同定、省略の回復等を行う。

(4)知識ベース操作部門

同定処理の終了した入力文を、平叙文なら知識ベースへ登録し、疑問文ならば知識ベースより答えを検索し、出力する。

構文解析部門についてはUTIIispで、その他の部門はprolog/KRで、M-380上にインプリメントしている。

(b) 対話例

入力文：太郎はアラブ人が好きだ。

花子はフランス人だ。

去年は、太郎はイランへ行った。

花子はフランスだ。

フランスはブドウ酒の産地だ。

イランは石油だ。

イランは石油の輸出国だ。

フランスは輸入国だ。

Q：花子は誰が好きか。

A：フランス人です。

Q：去年、花子はフランスへ行ったか。

A：はい、そうです。

Q：イランは石油の産地か。

A：はい、そうです。

Q：フランスは石油の輸出国か。

A：いいえ、違います。

【 4 】おわりに

名詞文の省略回復処理と、それを組み込んだ質問応答システムについて述べた。同定処理全般としては、談話文法の問題や知識を介在した文問の結末性に基づく同定処理などの大きな問題がある。今後はこれらについても、考察を進めていく予定である。

文献

1)池田：語法規則方式による日本語文の構文・意味解析

「情報処理学会論文誌」26-6