

小説における論理構造理解のための 因果・含意関係の抽出と分析

4 J-9

神山文子 西島恵介 藤田米春
大分大学 工学部

1 はじめに

我々は、短編推理小説を例にとり計算機による小説の自動理解の研究を進めている。計算機に小説を理解させる場合には、明示されている言語情報から作成される論理的構造を、読者が持っている常識、登場人物の心理等の様々な命題によって補完し、統合的な論理的体系を構築する必要がある。言語情報から作成される論理的構造は、文の因果・含意関係によって表現されることが多く、これらを文章間で連鎖することで、文章全体の論理的体系の構築が可能になると思われる。本報告では、文中から因果・含意を表わす語を手がかりとして、因果・含意関係（以下、このような関係を「因果的關係」と呼ぶ）にある文の抽出と分析について述べる。

2 文章中の因果・含意表現の分類

推理小説では、ある事象が原因となって次の事象を引き起こし、またその事象が次の事象を引き起こす原因・結果の関係を繰り返しながら結末に至るという特徴をもっている。この原因・結果の連鎖を抽出し、それを体系的に構築できれば計算機による小説理解に有効と考える。まず、文献[1](871文)より因果的關係を示す文を意味的な見地から探し分類した。この研究において因果的關係とは、原因とそれによって生じる結果（事実）との関係だけではなく、ある前提から推論によって導かれる結論、仮説から導かれるものなど因果的關係としては弱いものも含む。

Extraction and analysis of the cause and effect relations in a short detective story for understanding of the logical structure
Fumiko KAMIYAMA, Keisuke NISIJIMA,
Yoneharu FUJITA
Faculty of Engineering, Oita University

2.1 因果的關係を表わす手掛かり語

一文の中で、因果的關係が成立し、その際手掛かりとなる語として以下のものがあった。『たら』は助動詞でそれ以外は全て接続助詞である。手掛かり語の中でも、はっきり因果的關係がわかるものと判断が難しいものがある。

<主な意味>	<手掛かり語>
原因・理由	…『から』『ので』『で』
確定,仮定の順接	…『ば』『と』
確定,仮定の逆接	…『のに』『けれど』『が』 『ても』『でも』
過去完了,仮定形	…『たら』

2.2 文と文の因果的關係を表わす手掛かり語

一文の中ではなく、文と文を結びつけて因果的關係が成立するものがある。

2.2.1 接続詞によるもの

12種類の接続詞が因果關係を成立させるものとしてある。

<主な意味>	<手掛かり語>
順接	…『だから』『それで』『すると』 『したがって』『そうしたら』 『そこで』そして』
逆接	…『だが』『しかし』『ところが』 『でも』『それでも』

2.2.2 接続助詞によるもの

文献[1]では、先に結果を述べ、その後に原因を述べるという小説特有の表現による因果的關係の表現が多くあった。

[例1]私はまた赤くなった。美人教師といわれた
からである

3 因果的関係の抽出

以上の手掛かり語より計算機による因果的関係にある文の抽出を試みた。手掛かり語を探すために、形態素解析を行った。これは奈良先端大の開発した形態素解析システム[文献[5]]を利用した。

[表1]因果的関係の抽出実験（一部抜粋）

手掛かり語	抽出文	適合率	修正語抽出文	修正後適合率
ば	31文	58.1%	24文	75%
たら	20文	70%	18文	83.3%
と	93文	75%	87文	80.4%
が	53文	83%	52文	94%
ので	8文	100%	8文	100%
から	56文	91%	54文	94%
のに	4文	75%	4文	100%
けれど	12文	25%	9文	66.6%

これらより、因果的関係にある文(200文)は、全て抽出できるが、そうでない文(77文)も抽出してしまう。大きく分けて次の4つの要因があった。

①後の句を省略しているもの（結果、結論がなく読者の知識で補うもの）

[例2]腹をたてなければいけない相手なのに…

②後の句が省略されて、結論が前の文に書かれているもの。

[例3]間違いないと思ったのだ。あまりにも、英子に似ていたからである。

③語の用法が因果関係の機能をしないもの

[例4]教育問題だったら、お間違いですよ。

[例5]子供の認知をすればいいのだ。

④形態素解析の結果が間違っているもの

[例6]非常にスマートだと思う。

『と』が用言に接続されているため本来は格助詞なのに接続助詞と解析されてしまう。

①～④は以下の規則を用いて修正を行った。

I)①②も因果的関係の文として取り扱う。

II)会話文の中で手掛かり語の後に終助詞がつく場合は、念を押したり、判断、詠嘆などの用法なので含めない。終助詞『ね』などがある。

[例7]単なる憶測だからね。

III)手掛かり語の後に動詞、形容詞がある場合は、仮定の順接を表現しているのだが、因果的関係を認める事ができないので含めない。

[例8]犯罪ルポでも書けばいいのだ。

4 検討

I)～III)を規則として用い修正を行った結果、全体的に適合率は14.5%上がった。余分な文も抽出してくる問題として、③の要因が大きく占めこれは、表層の情報だけで区別することが困難であるので、論理構造を構築する際考える必要がある。この結果より、論理的関係にある文を利用して文章全体の論理的体系を構築するためには、手掛かり語の前後で原因と結果に分解する作業が次に必要である。また手掛かり語によって強い因果的関係にあるものと弱い因果的関係にあるものがあるので、強さを考慮した連鎖の張り方を考える必要がある。

5 おわりに

今回、因果的関係を示す手掛かり語を分析し、計算機による因果的関係を表わす文の抽出を行った。高い精度で抽出できたので論理構造を構築するために有効だと考えられる。

参考文献

- [1] 佐野洋：かわいそうなパパ，集英社文庫24「かわいい目撃者」より，集英社，(1979)
- [2] 藤田他：短編推理小説の論理構造の抽出．日本認知科学会第14回全国大会，June(1997)
- [3] 神山他：推理小説における暗黙情報獲得について．平成8年電関九支連大，(1996)
- [4] 西島他：短編推理小説における明示情報の論理的関係の抽出．人文コン研究会，Nov(1996)
- [5] 松本裕次他：日本語形態素解析システム ChaSenVer1.0使用説明書，奈良先端大学院