

スタック操作に基づく文脈処理について

67-5

木村 和広

(財) 新世代コンピュータ技術開発機構

1. はじめに

文の連なりである文章には、文構造があるのと同様にある種の構造＝談話構造があるように思える。Groszらの仮説[1]によれば、それは、木の形をした発話者（書き手）の意図の構造と、スタックで表現される聞き手（読者）の注意状態からなるとされる。発話者は、ある主張を聞き手に納得してもらおうという全体の目的のために、談話の各部分＝部分談話に目的をもたせ、それぞれの目的間に支配と被支配の関係、すなわち、プランーゴールというゴールとサブゴールのような関係が成立しているとする。また、聞き手は、談話に導入されたオブジェクトに注意を向けるが、それが談話の進行にともなって変化する様子は、プッシュポップ操作で代表されるスタック操作として考えることができる。そして、この注意の変化は、聞き手が認識した意図の構造によって制約されている。従って、このような談話構造が解析されるならば、照応／省略語の同定や、文の意味的曖昧性の除去、文章の要約などを行なう上で、有益な情報源となりうる。そこで、本稿では、この談話構造の解析を計算機上で実現するための検討を行なう。もちろん、各言語表現の解釈と談話構造とは、互いが互いを制約するものであり、どちらか一方が優先するというものではないので、この点に留意して、設計したい。

2. 談話構造解析処理



図1 談話構造

談話構造は、聞き手が認識するものであるし、また、発話者があらかじめ全体の構造を意識して発話しているとは限

らないから、実際の発話列から、談話の進行にともなってボトムアップに構築されていくことが電ましい。このとき、意図の構造は、支配関係で結ばれた木構造を仮定する。すなわち、ある談話の目的が複数の談話目的によって支配されることはない。そして、それぞれの意図を荷なった言語表現を部分談話と呼ぶことにする。部分談話の最大のものは、もちろん談話全体（文章全体）である。また、注意状態をあらわすスタックは、フォーカスの空間をこの木のノードに付随させることによって表現することにしよう。すると、構築すべき談話構造は、図1のような木構造になる。ここで、ノードFSiは、フォーカスの空間をあらわし、その中には、談話に導入されたオブジェクト、関係、および部分談話の目的（意図）が、保持されている。また、枝は、部分談話の目的間の支配－被支配の関係をあらわすと同時に注意状態の変化の履歴をあらわしている。ルートノードBSには、談話の背景にあるオブジェクト、関係、および談話全体の目的が付随する。

さて、この談話構造をどのようにして作っていくかを考える。まず、部分談話を構成する最小の単位を定める必要がある。ここでは、最小処理単位＝単文レベルと考え、以降この最小の部分談話を単に文と呼ぶことにする。したがって、複文・重文は、複数の文ということになる。そして、談話構造の解析戦略を次のような枠組みでとらえることにする。

- step 1. 隣接する文間の部分談話構造を解析する。
- step 2. 部分談話構造同士のマージをおこなう。
- step 3. デイフォルト処理（オプション）

step 1は、“まず”、“例えば”などのように部分談話の目的の切り換えを明示的に示す句＝会話句や、相、モダリティなどの情報を手掛りとして、解析する。このとき、ある判断を誤得するために使われる例示、対比、重複、理由、比喩などの典型的な修辭法の名称は、支配関係を分類するものとして用いる。この結果、隣接する文間の支配関係が部分木として表現される。ただし、これらの手掛り情報は、その解釈自体曖昧であることも多いので、隣接する文間の関係をstep 1で確定できなくてもよしとする。step 2は、

Context Analysis using Stack Operation

Kazuhiro KIMURA

Institute for New Generation Computer Technology

表現した次の岩と等価である。

$$\begin{array}{l} \Gamma \\ A \quad | 1 \rightarrow L1 \\ \quad | 2 \rightarrow L1+1 \\ \quad | 3 \rightarrow L1+1 \\ \quad | 4 \rightarrow L1+2 \\ \quad \vdots \\ B \quad | 5 \rightarrow L5+1 \\ \quad | 6 \rightarrow L6, L6 \leq L1+1 \\ \quad \vdots \end{array}$$

これは、図3のように、部分木A,Bがあり、かつ、Bのルートノードのぶらさがり枝を限定している様子を示している。

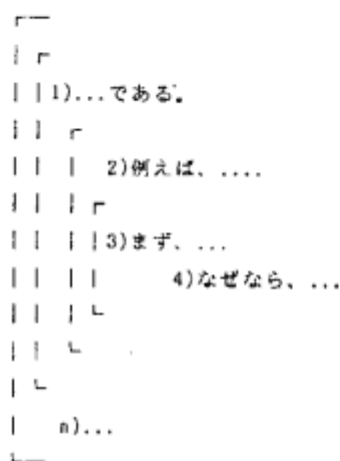


図2 被支配のレベル

図3 部分木の接続可能性

そこで、step 2 または 3 の解析とは、L1 と L6 の関係に焦点をあてて、L1 と L6 との間に成立する新たな制約を追加することになる。これにより、制約が強まり、全体の談話構造が、徐々に姿をあらわしていくことになる。

a) 先行する文に支配される。(L+1) 例) まず、...

談話構造を木構造と仮定し、これをボトムアップに構築していくための設計指針を提案した。なるべく一般性を失わないためには、制約の導入とその簡素化という方向が良いと考えた。今後は、実際に検証を行ない、照応／省略語の固定などの具体的な問題への適用を考えたい。

(1) Grosz, B. & Sidner, C.: Attention, Intentions, and the structure of discourse, *Computational Linguistics*, 12(3), 1986.