

TM-0956

ラムダ記法による文脈構造の表現と処理

小野 顕司、住田 一男、浮田 輝彦、
天野 貞家 (東芝)

September, 1990

© 1990, ICOT

ICOT

Mita Kokusai Bldg. 21F
4-28 Mita 1-Chome
Minato-ku Tokyo 108 Japan

(03)3456-3191~5
Telex ICOT J32964

Institute for New Generation Computer Technology

ラムダ記法による文脈構造の表現と処理^{*}

小野 顕司 住田一男 浮田 輝彦 天野 真家

(株) 東芝 総合研究所

本稿では、テキスト理解と文脈構造の関係について分析し、テキスト理解のモデルを示す。

我々は論説文等に於いて接続詞の系列が構成する文脈的構造を“思考の流れ”という観点で捉え、その表現方法及びそれを実テキストから抽出する方法を考察してきた。この構造はテキストの構造であるとともに、そのテキストを読む読者がテキストの内容を心象していく過程の構造でもある。本稿ではラムダ記法を用いてその表現方法を拡張し、話題の推移が捉えられるようにした。この表記方法を用いることにより、読み手の中に文脈構造がインクリメンタルに構成されていく過程をシミュレートすることができる。この過程はテキスト理解の認知心理的過程と対応するものである。

Expression and Processing of Discourse Structure based on λ Calculus^{**}

Kenji ono Kazuo Sumita Teruhiko Ukita Sin'ya Amano

R&D Center, Toshiba Corp.

Komukai-Toshiba-Cho 1, Saiwai-Ku, Kawasaki 210, Japan

e-mail address: ono@isl.rdc.toshiba.co.jp

In this paper we describe the relation between discourse structure and text understanding process and present a new psychological model of text understanding.

We thought that the contextual structure, which the sequence of conjunctions in a text forms of itself, reflects the "flow of thought" in comprehending the text. We have developed the way to express it and to extract it automatically from text. The structure reflects the way how readers figure out what is told in text. This paper presents the new way of expressing and processing discourse structure which are characterized by the composite sequence of conjunctive expressions and topic expressions. It incorporates lambda-calculus and can simulate the process in which readers psychologically enlarge discourse structure as he proceeds to read the text. It also corresponds to the cognitive process of text understanding.

^{*}本研究は、ICOTからの委託により
第五世代コンピュータプロジェクトの
一環として行っている。

^{**}This work was supported by ICOT
(Institute for New Generation
Computer Technology).

1. はじめに

我々は文脈構造を文と文、さらにそれらが結合されたものの間にある接続関係を記述したものが文脈構造であると捉え、「思考の流れ」という観点から、その表現方法及びテキストからの文脈構造の自動抽出の方法について考察してきた(小野[89], 木下[89])。

文脈構造は、文を最小単位とし、それらが2項接続関係(例示関係, 並列関係, 順接関係等)で互いにつながれて構成される2分木として表現した。接続関係には論旨展開の単位となる言明を構成するもの(言明のレベルの関係)と、それら言明を関連づけて論旨を展開するもの(思考のレベルの関係)とがある。図1にその概略を示す。これは、いわば文章中の各接続詞の“スコープ”を明確にしたものといえる。

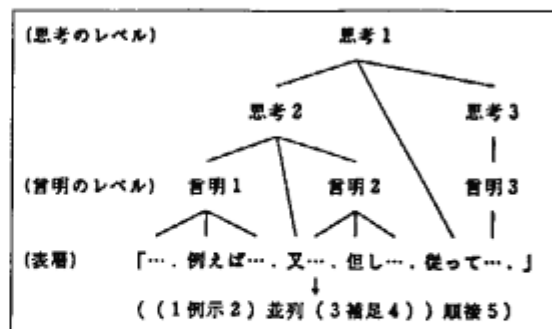


図1. 文脈構造の表現

この文脈構造を実テキストから抽出するため、従来は「思考制約規則」と称する、隣接する2接続詞の系列を見てその部分の構造を局所的に決定するような規則約400個を用いていた。しかし接続表現のみに着目するこのような構造抽出処理には限界がある。特に、文章中の話題表現が文脈構造に大きく影響しており、場合によっては思考制約規則にあてはまらないような構造を作る要因となっていることが観察された。

なぜ、またどのように話題表現はその文章の文脈構造に影響するのか、そもそも、接続表現を含めこれら表層的特徴に特徴づけられるような“思考の流れ”とは何なのか、これらの点を考察する過程で、読者がテキストを読む際の認知心理的過程と対応するテキスト理解のモデルを着想し、それをラムダ記法を用いて定式化した。本稿ではその概要を述べる。

2. 意図構造

いま、次のような文章を考えてみる。

“男は笑い、また女は騒いだ。”

この文章で書き手が言わんとしていることは、<男や女>(その場にいた特定の人の集合(特に他にコンテキストがなければ、その場にいた人全体)がく笑ったり騒いだりした>、すなわち、一時を楽しく過ごした、ということである。“…笑い、また…騒いだ”という形で、並列型の接続関係を介して2つの述語を併置することによって、この抽象的な述語“一時を愉快地過ごした”を表現している。

もう一つ別の文章を考えてみる。

“TTLは単一5Vで動作する、また(TTLは)大変高速である、従って…”

この文章では、表現“TTL”は<単一5Vで動作し、高速である>ようなものとして意図されている。以下に続く文章が、“TTL”の<単一5Vで動作し、高速である>という性質だけに依存するものである(すなわち、以下の文章で必要となる“TTL”の性質は、<単一5Vで動作し、高速である>ということだけである)ならば、この文章(コンテキスト)に於ては、“TTL”を他の<単一5Vで動作し、高速である>ようなもの(例えば、ある種のCMOS)に置き換えても、或はそれらを総称するような名詞に置き換えても一向に構わなはいはざである。つまり、この文章で書き手が表現“TTL”に意図している内容は<単一5Vで動作し、高速である>ようなもの、という内容であり、その点を押さえておくことがこの文章の内容理解(書き手の意図理解)の第一義である、ということである(もしその置換がうまくいかないなら、その文章は“TTL”の、<単一5Vで動作し、高速である>こと以外の性質を用いているのであり、その性質を満たすものとして、文章中で“TTL”は更に詳細に規定されていなければならない)。

したがって、この文章では(あるいは書き手の意識の元では)、表現“TTL”の意味は<単一5Vで動作し、高速である>で規定されているとみなすことができる。

また、筆者の意図したこの意味に於いて文章を理解すること、すなわち、“従って…”以降の議論が確かに“TTL”の、<単一5Vで動作し、高速である>という性質だけによっていることを理解することが、このテキストを“理解”するということなのである。

つまり、テキスト理解の過程とは、書き手が各表現に意図した意味とそれらの間に意図した相対関係とを読み手が正確に心象・再現していく過程に他ならないのである。

そのような“意図された意味”を扱う際、上記の説明の中で<>でくくって表現したようなものを単位とすることは有効である。これを以降心理オブジェクトと呼ぶ。また心理オブジェクト間の構造のことを、その文章の意図構造と呼ぶことにする。

3. 意図構造のλ表現

簡単の為、扱う文は“aはPである”、或は“aはPする”という形の文であるとし、それを“P(a)”と表記する。“aはPでない(しない)”という文章のときは、“ $\neg P(a)$ ”と表記する。その文の(文頭の)話題表現を‘a’、‘b’、…で表し、その文の述部を‘P’、‘Q’、…で表す。目的格等に位置する名詞句(項)は、すべて述部の一部とする。この記法により、2節で考えた2つの文章のうち、前者は“P(a)またQ(b)”，後者は“P(a)またQ(a)”と表記できる。意図構造はインフォーマルにそれぞれ

“<aやb>が<PやQ>である” あるいは “<PやQ>であるようなものが<aやb>である”

“<a>は<PやQ>である” あるいは “<PやQ>であるようなものが<a>である”

と表現できる。

数学やプログラム言語(LISP)、また高階の論理(タイプ理論)では、「…のようなもの」、「x such that x loves Mary.」、或は「 $\{x \mid f(x)=b\}$ 」といったように、ある文、条件式、論理式を満たすようなxという言い方で、特定のものや集合を定義する表記法がある。これは通常λ-expression (lambda, ラムダ表現)と呼ばれており、以下の書式で表記される。

$\lambda X [\dots (\text{変数 } X \text{ を含む式}) \dots]$ (ここでXはラムダ変数と呼ばれる)

以下では、この表記方法を利用して、<aやb>、<PやQである>というオブジェクトを表現することにする。尚、a, b, …は個体定数、x, y, …は個体変数、P, Q, …は述語定数、X, Y, …は述語変数、 λ は命題変数を表すとする。<aやb>は、 $\lambda X [X(a) \wedge X(b)]$ であらわし、<aやb>が<Pである>、ということ、 $\lambda X [X(a) \wedge X(b)] (P)$ と表現する。

タイプ理論では、表現 $\lambda X [X(a) \wedge X(b)]$ は、a及びbのどちらに対しても成立するような述語Xの集合を表しており、また、表現 $\lambda X [X(a) \wedge X(b)] (P)$ は、述語Pがその集合の元であることを主張する論理式となっている。この論理式は、 $P(a) \wedge P(b)$ と同値であるが、これは、表現 $\lambda X [X(a) \wedge X(b)]$ の中の大括弧[]のなかのXをPで置き換えたものとなっている。一般に、λ表現されたものの後ろに引数がきて構成されている論理式は、λ表現中の変数をその引数で置換して得られる論理式と同値になる。このような置換・書換えを通常λ-conversion (λ変換)と呼ぶ。

いまの場合、 $\lambda X [X(a) \wedge X(b)]$ と表現することにより、<aやb>を、aとbに共通する性質、つまりaとbがともに満たすような述語の集合という形で捉えているわけである。そして、述語Pは、確かにそのような述語の1つである、ということを主張する論理式によって、間接的に、Pで<aやb>を規定しているわけである。またこの表現は、談話理解の過程で読み手がこの心理的オブジェクトを心象する際、aやbに共通な部分にのみ注意を集中する(というよりも、aやbに共通な部分にのみ意識が集中される)、という心理的作用・機能の表現ともなっている。

さて、今の場合、表現“P(a)またP(b)”の“意図”として、 $\lambda X [X(a) \wedge X(b)] (P)$ を割り当てたわけであるが、これに通常のλ変換を行うと、 $P(a) \wedge P(b)$ となる。これはもちろん、表現“P(a)またP(b)”の意味として普通直観されるものに合致する。

すなわち、λ表現は、それが生成される段階、プロセス(手続き)が一部論理式の中に残った表現である、とみなすことができる。そして、そのプロセスの結果生成される“宣言的な”意味が、 $P(a) \wedge P(b)$ である、とみなすことができる。このように、本稿ではλオペレータに(心理的な)操作的意味を想定している。

<PやQである>という述語は、 $\lambda x [P(x) \wedge Q(x)]$ で表し、<a>が<PやQである>という文は、 $\lambda x [P(x) \wedge Q(x)] (a)$ で表す。この文は、述語P、Qのどちらをも満たすようなターム(個体、もの)の集合に、aが属している、ということと述べた論理式になっている。

<aやb>が<PやQである>という文は、上の2つを組み合わせることによって表現できる。すなわち、 $\lambda X [X(a) \wedge X(b)] (\lambda x [P(x) \wedge Q(x)])$ で表現される。

この分析は、“P(a)また $\neg P(b)$ ”(すなわち、“aはPだし、又bはPでない”)という文章の不自然性を説明してくれる。この文章の意図構造は、上述のQに $\neg P$ を代入したものになるが、これは

$\lambda X [X(a) \wedge X(b)] (\lambda x [P(x) \wedge \neg P(x)])$ となっており、書き手の心理中に $\lambda x [P(x) \wedge \neg P(x)]$ という、空集合のオブジェクトが存在することになる。このような、なにものをも規定しないような思考(無対象な思考)が談話理解のプロセスに存在するとは考えにくいし、このような表現に遭遇すると、読み手は書き手の“意図”がつかめず、混乱してしまうのである。

並列型以外の接続関係からなる2文の意図構造を、述部表現が否定されている/いない場合、及び2文間で話題表現・述部表現が推移する/しない場合の組み合わせ6通りで分類し、表1に示す。

尚、表1の論理式の中で $\lambda \lambda [\dots (\lambda \text{ を含んだ式}) \dots] (P(a))$ といった形式の表現を行うが、これは“P(a)という命題は、‘…(λを含んだ式)’中のλを満たすような命題である”という意味合いを持つ表現である。x、Xがそれぞれ個体、述語のタイプの変数であるのに対しλは命題のタイプ(文のタイプ)の変数である。また、論理式中の記号‘ $\supset$ ’、‘ $\equiv$ ’はそれぞれ、通常の論理における含意(AならばB)、同値(AならばBかつBならばA)を表す記号である。

表 1. 代表的な接続詞の意図構造

1. + ; また		4. = ; すなわち, つまり	
$P(a) + P(b)$	$\lambda X[X(a) \wedge X(b)] (\lambda x[P(x)])$ $= \lambda X[X(a) \wedge X(b)] (P)$	$P(a) = P(b)$	$P(a) \wedge \lambda x[P(a) \equiv P(x)] (b)$
$P(a) + \neg P(b)$	なし	$P(a) = \neg P(b)$	$P(a) \wedge \lambda x[P(a) \equiv \neg P(x)] (b)$
$P(a) + Q(a)$	$\lambda X[X(a)] (\lambda x[P(x) \wedge Q(x)])$ $= \lambda x[P(x) \wedge Q(x)] (a)$	$P(a) = Q(a)$	$P(a) \wedge \lambda X[P(a) \equiv X(a)] (Q)$
$P(a) + \neg Q(a)$	$\lambda x[P(x) \wedge \neg Q(x)] (a)$	$P(a) = \neg Q(a)$	$P(a) \wedge \lambda X[P(a) \equiv \neg X(a)] (Q)$
$P(a) + Q(b)$	$\lambda X[X(a) \wedge X(b)] (\lambda x[P(x) \wedge Q(x)])$	$P(a) = Q(b)$	$P(a) \wedge \lambda X[P(a) \equiv X(b)] (Q)$
$P(a) + \neg Q(b)$	$\lambda X[X(a) \wedge X(b)] (\lambda x[P(x) \wedge \neg Q(x)])$	$P(a) = \neg Q(b)$	$P(a) \wedge \lambda X[P(a) \equiv \neg X(b)] (Q)$
2. - ; 一方		5. $\rightarrow$; 従って	
$P(a) - P(b)$	$P(a) \wedge \neg \lambda X[X(a) \wedge \neg X(b)] (P)$	$P(a) \rightarrow P(b)$	$P(a) \wedge \lambda x[P(a) \supset P(x)] (b)$
$P(a) - \neg P(b)$	$P(a) \wedge \lambda X[X(a) \wedge \neg X(b)] (P)$	$P(a) \rightarrow \neg P(b)$	$P(a) \wedge \lambda x[P(a) \supset \neg P(x)] (b)$
$P(a) - Q(a)$	$P(a) \wedge \neg \lambda x[P(x) \wedge \neg Q(x)] (a)$	$P(a) \rightarrow Q(a)$	$P(a) \wedge \lambda X[P(a) \supset X(a)] (Q)$
$P(a) - \neg Q(a)$	$P(a) \wedge \lambda x[P(x) \wedge \neg Q(x)] (a)$	$P(a) \rightarrow \neg Q(a)$	$P(a) \wedge \lambda X[P(a) \supset \neg X(a)] (Q)$
$P(a) - Q(b)$	$\lambda X[X(a) \wedge \neg X(b)] (\lambda x[P(x) \wedge \neg Q(x)] \wedge \lambda X[\neg X(a) \wedge X(b)] (\lambda x[\neg P(x) \wedge Q(x)])$	$P(a) \rightarrow Q(b)$	$P(a) \wedge \lambda X[P(a) \supset X(b)] (Q)$
$P(a) - \neg Q(b)$	$\lambda X[X(a) \wedge \neg X(b)] (\lambda x[P(x) \wedge Q(x)] \wedge \lambda X[\neg X(a) \wedge X(b)] (\lambda x[\neg P(x) \wedge \neg Q(x)])$	$P(a) \rightarrow \neg Q(b)$	$P(a) \wedge \lambda X[P(a) \supset \neg X(b)] (Q)$
3. $\times$; しかし		6. $\leftarrow$; なぜなら	
$P(a) \times P(b)$	$P(a) \wedge \neg \lambda x[P(a) \supset P(x)] (b)$	$P(a) \leftarrow P(b)$	$P(a) \wedge \lambda x[P(b) \supset P(x)] (a)$
$P(a) \times \neg P(b)$	$P(a) \wedge \neg \lambda x[P(a) \supset P(x)] (b)$	$P(a) \leftarrow \neg P(b)$	$P(a) \wedge \lambda x[\neg P(b) \supset P(x)] (a)$
$P(a) \times Q(a)$	$P(a) \wedge \neg \lambda X[P(x) \supset X(a)] (Q)$	$P(a) \leftarrow Q(a)$	$P(a) \wedge \lambda X[Q(a) \supset X(a)] (P)$
$P(a) \times \neg Q(a)$	$P(a) \wedge \neg \lambda X[P(x) \supset X(a)] (Q)$	$P(a) \leftarrow \neg Q(a)$	$P(a) \wedge \lambda X[\neg Q(a) \supset X(a)] (P)$
$P(a) \times Q(b)$	$P(a) \wedge \neg \lambda X[P(a) \supset X(b)] (Q)$	$P(a) \leftarrow Q(b)$	$P(a) \wedge \lambda X[Q(b) \supset X(b)] (P)$
$P(a) \times \neg Q(b)$	$P(a) \wedge \neg \lambda X[P(a) \supset X(b)] (Q)$	$P(a) \leftarrow \neg Q(b)$	$P(a) \wedge \lambda X[\neg Q(b) \supset X(b)] (P)$

また、前述の如く、以下に列挙する論理式は、“ $P(a) - Q(b)$ ”, “ $P(a) - \neg Q(b)$ ” の 2 ケースを除いて、総て λ 変換すると対応する 2 文を連言した論理式と同値になる。

各式の意味合いについて、以下簡単に述べる。

○ - (一方) について:

まず、“ $P(a) - P(b)$ ” という文章を考えてみる。この文章の意図は、“ a と b は対立関係 (背反関係) にあり、 a が P だから、 b は $\neg P$ となりそうだが、そうではなく、 b も P だ。” というものである。すなわち、“ここで考えている P は、 a と b との対立関係に巻き込まれているような述語ではない” というものである。 a と b との対立関係に属している述語は、 a では成立するが b では成立しないような述語と、 a では成立せず b では成立するような述語との 2 つからなるが、その前者を $\lambda X[X(a) \wedge \neg X(b)]$ であらわすと、この文章の意図は $P(a) \wedge \neg \lambda X[X(a) \wedge \neg X(b)] (P)$ となる。(これは λ 変換すると $P(a) \wedge P(b)$ と同値になる。) $a =$ “日本人”, $b =$ “アメリカ人”, $P =$ “刺身が好きである” 等がその例である。

“ $P(a) - \neg P(b)$ ” の意図構造は反対に、 P が a と b との対立関係に部類する述語であることを主張するものとなっている。従ってこの文章の意図は $P(a) \wedge \lambda X[X(a) \wedge \neg X(b)] (P)$ となる。(これは λ 変換すると、 $P(a) \wedge \neg P(b)$ と同値になる。)

“ $P(a) - Q(a)$ ” の意図構造は、“ P と Q は対立する述語であり、 $P(a)$ だから、 $\neg Q(a)$ となりそうであるが、そうではなく、 $Q(a)$ である” というものである。すなわち、“ a は P と Q との対立関係に部類するような個体ではない” というものである。 P と Q との対立関係に属する個体は、 P を満たすが Q は満たさないような個体と、 P は満たさないが Q は満たすような個体との 2 つからなるが、前者の集合を $\lambda x[P(x) \wedge \neg Q(x)]$ で表すと、この文章の意図は $P(a) \wedge \neg \lambda x[P(x) \wedge \neg Q(x)] (a)$ と表現される。この式は λ 変換すると $P(a) \wedge Q(a)$ と同値になる。

“ $P(a) - \neg Q(a)$ ” の意図構造は、“ P と Q との対立関係に a は属している” というものであり、 $P(a) \wedge \lambda x[P(x) \wedge \neg Q(x)] (a)$ と表現できる。

“ $P(a) - Q(b)$ ” は、“ P と Q , a と b は、それぞれ対立関係にある述語、個体であり、この 2 つの対立

関係が一致している”すなわち，“a と b はそれぞれ、P であって Q でないような個体、Q であって P でないような個体であり、P、Q はそれぞれ、a に対して成立し b に対して成立しないような述語、b に対して成立し a に対して成立しないような述語となっている”という構造を持っている。

前述した λ 表現を用いてこれらの対立関係を表現すれば、この構造は

$$P(a) \wedge Q(b) \wedge \lambda X[X(a) \wedge \neg X(b)](\lambda x[P(x) \wedge \neg Q(x)]) \wedge \lambda X[\neg X(a) \wedge X(b)](\lambda x[\neg P(x) \wedge Q(x)])$$

と表現できる。この表現中の λ 表現された部分（2 行目）に λ 変換を施すと、

$P(a) \wedge Q(b) \wedge \neg P(b) \wedge \neg Q(a)$ となる。よって、最初の $P(a) \wedge Q(b)$ の部分は不必要となる。以下の議論ではその部分を取り除いて表記する。

これらの意図構造をよく調べると、 $\lambda X[X(a) \wedge \neg X(b)]$ や $\lambda X[\neg X(a) \wedge X(b)]$ という心理オブジェクトが存在している。これは、談話理解の過程で読み手がこれら心理オブジェクトを心象する際、a や b に共通の部分・性質は捨てられ、差異の部分だけに意識が集中されるという過程に対応している。

③ $\times$ （しかし）について：

$P(a) \times \neg Q(a)$ の意図構造が $P(a) \wedge \neg \lambda x[P(x) \supset Q(x)](a)$ となっているのは、“a が P を満たし、また P を満たすものは通常 Q も満たすので、Q(a) かと思ったが、a はそのような個体ではなかった”、つまり“a は $P(x) \supset Q(x)$ であるような x ではなかった”といった意味合いを反映してである。a = “ワープロ”、P = “高性能である”、Q = “値段が高い”等がその一例である。

ほかの 5 通りの構成についても同様である。

④ =（すなわち）について：

$P(a) = P(b)$ の意図構造が $P(a) \wedge \lambda x[P(a) \equiv P(x)](b)$ となっているのは、“b は、P に対しては a と同じである”という意味合いを反映してである。 $P(a) = Q(a)$ の意図構造が $P(a) \wedge \lambda x[P(a) \equiv X(a)](Q)$ となっているのは、“a にとっては P であることと Q であることは同じである”という意味合いを反映している。他にも同様である。

⑤ $\rightarrow$ （従って）、 $\leftarrow$ （なぜなら）について

論理式 $X(x) \supset Y(y)$ のどこに位置するようなものとして各述語・個体が想定されているか、によって各場合の構造が規定されている。

λ の前に $\neg$ があるような構造は、その構造が書き手にとっては何か“意外”なものであることを示している。通常はその後ろに、なぜそうであるかを説明した文が続くであろう。

λ 表現の前に羅列される命題は、後ろの λ 表現が表現する認識・意図の“前提”或は“枠”となるような、事実ないし宣言の意味を述べている。 λ 表現部分は“相対的”な関係・認識を表しているだけなので、この λ の外に外置された固定的な部分がないと、式全体の“絶対的”（宣言的）意味が確定しないのである。（この部分があることによって、式全体を λ 変換した結果が、対応する 2 文を連言した論理式と同値になる）

意図構造の λ 表現式は、それぞれ、対応する文章に於ける話題（焦点・関心の中心）の推移の構造をも表している。式中の λ 表現部分 ' $\lambda x [\dots]$ ' の ' λ ' の（ λ ）の中にくるものが、その文章全体の話題（書き手がその文章を用いて導入・規定しようとしたもの）を表し、 $[\dots]$ の部分は、その話題が“どのようなものであるか”を規定するための、その話題となるものが正しく解釈されるべきコンテキストを提供している。

さて、これら 2 文間で定義された意図構造を重層的に用いることによって、任意の文脈構造に対しその意図構造を計算することができる。

2 文からなる文脈構造には、その意図構造として表 1 に示したものをそのまま当てる。

文脈構造が 3 文以上からなる場合、すなわち 2 分木表現される文脈構造が部分構造を持つ場合、その部分構造に対して計算された意図構造中の、 λ 表現の引数部分（話題部分）を含むような命題部分をもって、親構造の意図構造計算に用いる。つまり、話題が 2 文目に推移するような接続関係（ $\rightarrow$, $\times$ など）を持つ部分構造については後半の文を、話題が推移しない接続関係（ $=$ など）をもつ部分構造については前半の文を、親構造の意図構造の計算に用いる。両方で話題を構成しているような接続関係（ $+$, $-$ 等）をもつ部分構造については、その部分構造の意図構造全体を用いる。

4. λ 表現による文脈理解モデル

談話理解においては、読解の過程で生成される中間的な文脈構造（と、それに対応する心理的オブジェクト）が次の文の内容の解釈（及び、その文を既存の文脈構造のどこに位置づけるか）を決定し、そして、そのように解釈された文が付加されて拡張された文脈構造（及びその心理的オブジェクト）が又次の文を解釈する、という具合に、構造化とその内容的解釈とは相互に影響しながら進行し、累積していく。

いま、以下の文章を読解する過程を考える。

「¹TTLは単一5Vで動作する素子である。²また(TTLは)IinやIoutが大きい。³従って(TTLは)LED等も直接点灯でき、扱いが便利である。⁴一方、CMOSも単一5Vで動作する素子である。⁵しかし、CMOSはIin・Ioutが小さく、逆電圧に弱い等扱いは不便である。⁶しかし、CMOSは省電力化に用いられるのに、TTLは用いられない。この差は…」

* () 内は省略されている話題を示す。

この文章を以下に定義する記号を用いて書き換えると、以下のようになる。

$$R(c) + S(c) \rightarrow T(c) - R(d) \times \neg T(d) \times (U(d) - \neg U(c))$$

ここで各記号はそれぞれ、 $c = \text{"TTL"}$ 、 $d = \text{"CMOS"}$ 、 $R = \text{"単一5Vで動作する素子である"}$ 、 $S = \text{"IinやIoutが大きい"}$ 、 $T = \text{"扱いが便利である"}$ 、 $U = \text{"省電力化に用いられる"}$ を表している。

第3文を読んだ段階で、2通りの読み方が可能になる。すなわち、第3文は第2文のみから導かれているのか、あるいは第1文と第2文の両方から導かれているのか、という解釈上の曖昧性である。

この2つの解釈はそれぞれ

$$1) \quad R(c) + (S(c) \rightarrow T(c))$$

$$2) \quad (R(c) + S(c)) \rightarrow T(c)$$

という、2つの文脈構造に対応している。

前者の読み方に対応する意図構造は $P(a) + Q(a)$ 型であり、

$$S(c) \wedge \lambda x[R(x) \wedge T(x)](c)$$

すなわち、「($S(c)$ であることを前提として) c は R であり、かつ、 $S(c)$ から導ける性質 T をも兼ね備えるようなものである」というものである。後者の読み方に対応する意図構造は $P(a) \rightarrow Q(a)$ 型であり、

$$R(c) \wedge S(c) \wedge \lambda x[(R(x) \wedge S(x)) \supset T(x)](T)$$

すなわち、「($R(c)$ 、 $S(c)$ であることを前提として) $R(c)$ と $S(c)$ から導ける性質 T を c はもつ」、すなわち、「 $T(c)$ である」というものである。

内容的には、 c に関する具体的な特徴・性質である R や S と、より一般的な言及である T を比較した場合、 R と T が並置される前者の構造よりも R と S が並置されてそこから T が導かれているような後者の構造の方が、相対的結束性の観点からも好ましい解釈となっている。しかし決定的な要因とはなっておらず、両者の解釈が可能である。

この状態で4文目が読み込まれ、ここで文章が終了すると考えると、以下の5通りの解釈が可能となる。

$$3) \quad ((R(c) + S(c)) \rightarrow T(c)) - R(d)$$

$$4) \quad (R(c) + (S(c) \rightarrow T(c))) - R(d)$$

$$5) \quad (R(c) + S(c)) \rightarrow (T(c) - R(d))$$

$$6) \quad R(c) + ((S(c) \rightarrow T(c)) - R(d))$$

$$7) \quad R(c) + (S(c) \rightarrow (T(c) - R(d)))$$

ところが、この5通りの解釈はすべて許容できない。3) の読み方は

$P(a) - Q(b)$ 型であり、

$$R(c) \wedge S(c) \wedge \lambda x[X(x) \wedge \neg X(d)](\lambda x[T(x) \wedge \neg R(x)]) \wedge \lambda x[\neg X(c) \wedge X(d)](\lambda x[\neg T(x) \wedge R(x)])$$

という構造に対応するが、 $R(c)$ が真であることが前提部にあり、 $\lambda x[X(x) \wedge \neg X(d)](\lambda x[T(x) \wedge \neg R(x)])$ (= $T(c) \wedge \neg R(c) \wedge \neg(T(d) \wedge \neg R(d))$) を心象することができない(この部分は、 c が持ち d が持たないどの様な性質が、述語 T を成立させ述語 R を成立させないようになっているのかを心象する構造であるが、 $R(c)$ が真なので、そのような性質を心象することが無意味になっている)。

同様に、 $P(a) - Q(b)$ 型である 4) の構造:

$$S(c) \wedge \lambda x[X(x) \wedge \neg X(d)](\lambda y[\lambda x[R(x) \wedge T(x)](y) \wedge \neg R(y)]) \wedge \lambda x[\neg X(c) \wedge X(d)](\lambda y[\neg \lambda x[R(x) \wedge T(x)](y) \wedge R(y)])$$

も、 $\lambda y[\lambda x[R(x) \wedge T(x)](y) \wedge \neg R(y)]$ (= $\lambda y[R(y) \wedge T(y) \wedge \neg R(y)]$) の部分で矛盾をおこしてしまう。(構造候補 5), 6), 7) もすべて、 λ 計算上このような矛盾を起こす。)

つまり、もしここで文章が終わっていれば、この文章に対する可能な解釈が存在しないということになる。これは、実際の読感とよく対応したものとなっている。

この状態で5文目が読み込まれると、これまでの読み方の不定部分・不整合部分が解決される。3), 4) の延長としての読み方には

$$8) \quad ((R(c) + S(c)) \rightarrow T(c)) - (R(d) \times \neg T(d))$$

$$9) \quad (R(c) + (S(c) \rightarrow T(c))) - (R(d) \times \neg T(d))$$

10) $((R(c) + S(c)) \rightarrow T(c)) \rightarrow R(d) \times \neg T(d)$

11) $((R(c) + (S(c) \rightarrow T(c))) \rightarrow R(d)) \times \neg T(d)$

の4通りがあるが、10), 11) は先同様、不適切な部分構造 $((R(c) + S(c)) \rightarrow T(c)) \rightarrow R(d)$ 及び $((R(c) + (S(c) \rightarrow T(c))) \rightarrow R(d))$ を含むので棄却される。9) の意図構造は $P(a) \rightarrow \neg Q(b)$ 型であり、

$S(c) \wedge R(d) \wedge$

$\lambda X[X(c) \wedge \neg X(d)] (\lambda Y[\lambda X[R(x) \wedge T(x)](y) \wedge T(y)]) \wedge$

$\lambda X[\neg X(c) \wedge X(d)] (\lambda Y[\neg \lambda X[R(x) \wedge T(x)](y) \wedge \neg T(y)])$

であるが、 $\lambda Y[\lambda X[R(x) \wedge T(x)](y) \wedge T(y)] (= \lambda Y[(R(y) \wedge T(y)) \wedge T(y)])$ の部分が冗長である。この構造は“c が持ち d が持たない性質のうちどのようなものが、述語 R かつ T を成立させ、かつ述語 T を成立させるのか”を心象する構造であるが、c が d で満たされている述語 T を満たすことが含まれているので、比較しても無駄な対象と比較を行っているような構造となっている。このような無益な心象過程を含むような意図構造は棄却される。

一方、 $P(a) \rightarrow \neg P(b)$ 型である 8) の構造は

$R(c) \wedge S(c) \wedge R(d) \wedge \lambda X[X(c) \wedge \neg X(d)](T)$

となり、矛盾を起こさない唯一の読み方となる。従って、ここまで読んだ時点で生成される文脈構造として、この構造が選択される。(5), 6), 7) を延長する読み方は、 λ 計算上矛盾を起こした部分構造がそのまま残るので、すべて却下される。

ここまでの過程を振り返ってみると、3文目の読了の時点で不定であった構造上の曖昧性(1)と2)が、5文目の読了の時点で、構造の相対的な規定作用によって解決されていることがわかる。また、4文目を読了した時点での、次の文への期待感(すなわち、4文目まででは、文章全体がまだまとまりを構成していないという読感)は、4文目までに対する部分構造が存在しないという形で反映されていることがわかる。

この後6文目が読み込まれ、8)の延長として

12) $((R(c) + S(c)) \rightarrow T(c)) \rightarrow (R(d) \times \neg T(d)) \times (U(d) \rightarrow \neg U(c))$

13) $((R(c) + S(c)) \rightarrow T(c)) \rightarrow ((R(d) \times \neg T(d)) \times (U(d) \rightarrow \neg U(c)))$

の2つの構造候補が生成される。13)の意図構造は

$\lambda V[V(T) \wedge \neg V(U)] (\lambda Y[\lambda X[X(c)](Y) \wedge \neg \lambda X[X(d) \wedge \neg X(c)](Y)]) \wedge$

$\lambda V[\neg V(T) \wedge V(U)] (\lambda Y[\neg \lambda X[X(c)](Y) \wedge \lambda X[X(d) \wedge \neg X(c)](Y)])$

となる。ここで V は、述語の集合を値域とするような変数である。

この構造も先と同様、 $\lambda Y[\neg \lambda X[X(c)](Y) \wedge \lambda X[X(d) \wedge \neg X(c)](Y)]$ の部分が

$\lambda Y[\neg Y(c) \wedge (Y(d) \wedge \neg Y(c))]$

という冗長な構造となっているので棄却される。

このようにして、12)のみが6文目までを妥当に構造化している構造となる。

さて、12)の意図構造

$R(c) \wedge S(c) \wedge R(d) \wedge T(c) \wedge \neg T(d) \wedge$

$\neg \lambda \mathcal{K}[\lambda X[X(c) \wedge \neg X(d)](T) \supset \neg \lambda \mathcal{K}[\lambda X[X(d) \wedge \neg X(c)](U)]$

がこの文章全体の意図となるが、これを翻訳すると次のようになる：

“(R(c), S(c), R(d), T(c), $\neg T(d)$ は前提として、) T(c) $\wedge \neg T(d)$ であれば、通常 $\mathcal{K}$ ということはないのだが、実際はそうではなかった。”ここで $\mathcal{K}$ とは、 $\lambda X[X(d) \wedge \neg X(c)](U)$ すなわち、 $U(d) \wedge \neg U(c)$ という内容のものである。これを更に翻訳すると、次のようになる：“5V単一で動作する素子の扱いに関しては、TTLは便利でCMOSは不便である、ということがある。だから、 $\mathcal{K}$ ということはないさうだが、実際はそうではなかった”

ここで $\mathcal{K}$ とは、省電力化についてはTTLを用いずCMOSを用いる、というものである。これは、この文章全体の論説の構造としては妥当なものであると思われる。

図2に以上の計算過程の全体像を示す。図より話題がローカルなレベルでは $c(TTL) \rightarrow T$ (扱いが便利である) $\rightarrow U$ (省電力化に用いられる) と、又グローバルなレベルでは $c(TTL) \rightarrow T$ (扱いが便利である) $\rightarrow \lambda X[X(d) \wedge \neg X(c)](U)$ (省電力化についてはTTLを用いずCMOSを用いる) と変化していく様子を読み取ることができる。

5. まとめ

ラムダ記法を文脈構造の記述に利用することによって、接続関係の構造と話題の構造とを融合した。この表現によって記述される構造は、接続詞や話題の推移の仕方に特徴づけられるような“思考の流れ”あるいは“意図”の構造である。これはまた、書き手の意識の下での各文の意味の相対規定関係を表したものであり、読み手が

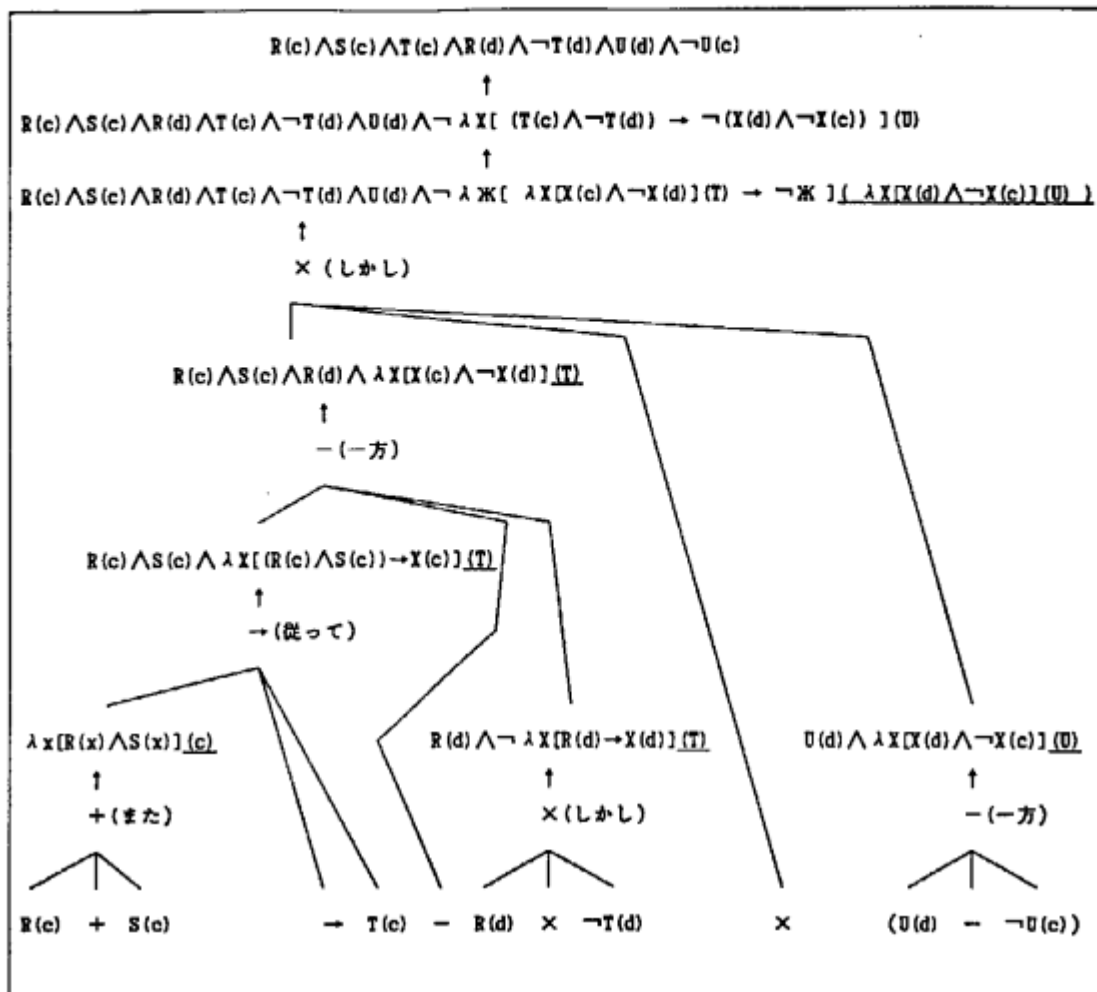


図2. λ 計算の模式図 (下線部は話題の推移を示す。)

文意を心象していく過程の構造でもある。

この記法によって、読解の過程で各文の意味が相対的に（及び累加的に）心象される部分（ラムダ表現中の大括弧の部分）を明確にした。そして、その部分の無矛盾性と無冗長性が談話に整合性（Coherency）を与えたと考え、この基準の下で文章から文脈構造を累加的に導出する方法を明確にした。

そして、このテキスト理解のモデルによって、もしここで文章が終わっていたらおかしい、と読み手が感じる現象が説明できることを示した。

ラムダ記法を用いた自然言語の意味論では、モンタギュー意味論（Dowty et al[81]等）がある。しかし本稿は以下の点でモンタギュー意味論と異なっている。

- ・対象範囲が1文ではなく文章全体である。
- ・モンタギュー意味論では、ラムダ計算は範疇文法で導出された構造（構文木）からその意味（論理式）を算出する過程で用いられるが、本稿ではラムダ計算は文の意味の相互作用性と対応するものとなっており、構造の形成過程そのものに関与する（何かの文脈文法を前提とするものではない）。
- ・宣言の意味論（モデル理論）ではなく、手続きの意味論である。
- ・読解の認知過程との対応がある。

今後、 λ 計算過程の精密化と、1文中に話題表現が複数ある場合の対応、省略された話題表現の復元への応用、また実テキスト処理のための理論のインプリメントを進めていく予定である。

6. 参考文献

- 小野, 浮田, 天野: 文脈構造の分析, 情報処理学会研究報告 89-NL-70.
 木下, 小野, 浮田, 天野: 日本語テキスト理解における文脈構造抽出法, 談話理解モデルとその応用シンポジウム論文集, pp. 125-136, 1989.
 D. R. Dowty, R. E. Wall and S. Peters: INTRODUCTION TO MONTAGUE SEMANTICS, D. Reidel Publishing Company, 1981.