

自然言語における程度概念の離散モデル

亀井 真一郎 村木 一至
NEC 情報メディア研究所

1 はじめに

実世界における物事の度合、程度は連続的に変化する。自然言語が有限の語彙を使用して程度概念をどのように表現しているかに関しては、従来から種々の興味深い言語現象が指摘され、修飾の選択制限、否定との関係、会話の含意、論理語との関係などが研究されてきた[1, 2, 3, 4, 5, 6]。我々は「自然言語には程度概念を離散的にとらえる理解の枠組が存在する」という観点にたったモデルを導入し、程度概念にまつわる諸現象を統一的に取り扱うことを試みている[7, 8, 9, 10]。本稿ではこのモデルを主として程度概念の否定操作に適用してその有効性を説明する。

2 程度概念の特徴とモデル化する際の困難

自然言語には物事の程度を表す表現が多数存在する。例えば以下に示すのがその例である。

- (1) a. always, often, sometimes, seldom, never
b. all, many, some, few, no
c. tall, short
- (2) a. 常に, しばしば, 時々, めったに.. ない, 決して..
ない
b. 全て, 多く, いくらか, ほとんど.. ない, 全く..
ない
c. 高い, 低い

これらの語が指し示す現実の量は状況や話者の判断によって大きく変化する。「しばしば」と「時々」とを分ける明確な基準があるわけではなく、これらの言葉が発話されたときに指し示す具体的な量は聞き手にはわからない。にもかかわらず我々は日常これらの表現を多用し、かつその意味が他の表現と比べて特に曖昧だとは感じていない。程度表現の意味理解とは、これらの言葉の指す現実の量がわかるということではなさそうである。

程度表現はどの言語にも存在するが、従来は異言語の表現を比較する枠組が欠如していた。通常、日本語の「しばしば」と英語の「often」は意味的に対応すると考えられるが、そう考えて良い理由は自明ではない。両者とも現実世界に具体的対象を持たないから意味的対応の根拠を現実世界の量に求めることはできない。「しばしば」と「often」が対応するというのはどういう意味にお

いてか、またその対応はどこまで厳密になりたつのか、それを記述するための枠組を定義する必要がある。

また程度表現は副詞、代名詞、限定詞、形容詞など品詞を越えて存在し、いくつかの共通した意味的性質を持っている。例えば程度強調詞類(Intensifier)に修飾されるときの制限を考えてみる。often, seldom, many, few, tall, short は very に修飾されるが、all, no, always, never は very には修飾されない。反対に前者は almost に修飾されないが、後者は almost による修飾を受ける。また一方 some, sometimes は very にも almost にも修飾されない。しかしながら従来はこのような品詞とは独立な意味的性質を理解する枠組が欠けていた。

表 1: very, almost と程度表現の修飾関係

	very	almost
often, seldom, many, few, tall, short	+	-
all, no, always, never	-	+
some, sometimes	-	-

我々はこれらの問題に答えるため程度表現の新しい意味モデルを提案した[7, 8, 9, 10]。次節以降ではそのモデルの基本概念を説明する。

3 程度の基本概念素とリスト表現

強調詞の修飾制限は、品詞に独立な、程度に関する概念素の存在を示唆している。そこで我々は程度表現の基本的概念素として、「all」「many」「some」「few」「no」の意味から抽出した5つの意味素「A」「M」「S」「F」「N」を仮定した。さらに、これらの意味素を並べたリストを導入し、具体的な程度表現の意味を表すのに用いる。

(3) {A, M, S, F, N}

上記が程度概念の意味を表すための基本リストである。個々の程度表現、例えば「高い」「低い」の意味を「高さ」という軸と「多い」「少ない」という程度に分け、後者の部分をこのリストを用いて表現すると、表3のようになる。すなわち、各々の程度表現は一つ一つで独立に意味が定まるのではなく、リストで表現されるような枠組全体の中の相対的位置によって意味が定まるのだと考える。現実世界の程度は連続量であるが、それを言

語世界ではこのリストの値のように離散的にとらえてい
ると考えて程度概念素を仮定すると、程度概念と程度
強調詞類の間の修飾の選択制限を表4のようにモデル化
できる。これは程度表現がもつ意味的性質の一つを統一
的に記述している。

表 2: 程度表現の意味のリスト表現 (1)

基本リスト	{A, M, S, F, N}
all, always	{A, -, -, -, -}
many, often	{-, M, -, -, -}
some, sometimes	{-, -, S, -, -}
few, seldom	{-, -, -, F, -}
no, never	{-, -, -, -, N}

表 3: 程度表現の意味のリスト表現 (2)

基本リスト	{M, S, F}
tall	{M, -, -}
not tall and not short	{-, S, -}
short	{-, -, F}

表 4: 程度強調詞類と程度概念素との修飾制限

程度強調詞	例	程度概念素					
		A	M	S	F	N	
a. 増幅語	very, extremely	-	+	-	+	-	
b. 緩和語	somewhat, pretty	-	+	-	+	-	
c. 弱化語	a little, slightly	-	+	-	+	-	
d. 近接語	almost, nearly	+	-	-	-	+	
e. 完結語	absolutely	+	-	-	-	+	

4 程度概念の2重リスト表現

4.1 数量を含む質問とそれに対する答え

一般の程度表現を取扱う前に問題点を単純化するた
め「数」を含む以下の例文を考える。

(4) I solved three of the problems.

この文は通常「問題をちょうど3問解いた」と解釈
される。しかし例えばテストの可否のボーダーラインが
3問であるような状況では、発話者が解いた問題数が実
際には4問以上であっても上記の文はごく自然に発せら
れる。この事実は発話の「数」部分の意味表現が単なる
「数」では不十分であることを示している。このことは
下記のような質問とその応答の考察でより明確になる。

(5) A: Did you solve three of the problems ?
B: - Yes, in fact I solved four.
- No, I solved four.

興味深いことに上記の問いには Yes/No 両方の答が
可能である。この現象を説明するために以下の5つの状
態を考える。(1) 問題が全て解けた状態、(2) 解けた問
題数が文中に表現されている数 (=この場合「3」) を越
えている状態、(3) 解けた問題数が文中に表現されてい
る数に一致している状態、(4) 解けた問題数が文中に表
現されている数に達していない状態、(5) 全く問題が解
けなかった状態。これらの状態を表現するために「A」
「>n」 「=n」 「<n」 「N」 5つの概念素を導入し、下
記のようなリスト表現でそれらの相対位置を表現する。

(6) {A, >n, =n, <n, N}

このリスト表現を用いると上記の5つの状態は表5
の様に表される。

表 5: 数の意味のリスト表現

基本リスト	{A, >n, =n, <n, N}
all	{A, -, -, -, -}
>three	{-, >n, -, -, -}
three	{-, -, =n, -, -}
<three	{-, -, -, <n, -}
none	{-, -, -, -, N}

さて上述の質問文に Yes/No 両方の答えが可能であ
ることを説明するため、文中の「数」部分の意味表現と
して以下のような2重リスト表現を導入する。

(7) $\left\{ \begin{array}{l} -, -, =n, -, - \\ A, >n, =n, -, - \end{array} \right\}$

この2重リストの上段は「直接解釈リスト」であり、
発話中にある数表現の値そのものを表している。下段は
「可能性解釈リスト」であり、発話中には直接現れない
解釈を表現している。この二つのリストの全体が発話さ
れた「数」部分の意味を表現するものとする。この2
重リストを用いると質問に対する答の可能性が図1のよ
うに単純な操作により求められる。図1の左の2重リス
トは質問文中にある数「3」を表現する。図の真中は実
際の状況を表しているリスト表現(実状況リスト)である。
ここでは実際に解いた問題数「4」を示している。実際
の状況は発話ではないので1重リストで表現される。さ
てこの二つのリストの共通部分(intersection)を計算す
る。問いの2重リストの上段(直接解釈リスト)と実状況

リストとを比較すると共通部分はない。したがってこの場合の答えは「No」になる。一方、質問の2重リストの下段(可能性解釈リスト)と実状況リストには共通部分(>n)が存在する。これは「Yes」の答えに対応する。2重リストの上下二つのリストの差の部分が、このモデルにおける「会話の含意」の可能性の明示的表現である。

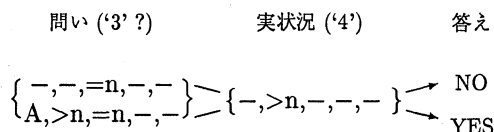


図 1: 数を含む問いに対する答えの可能性

4.2 数量を含む文に対する否定操作

本節では数を含む肯定文に対する否定文の意味を考察する。以下の文が文(4)に対応する否定文である。

(8) I didn't solve three of the problems.

この否定文にはいくつかの意味解釈が存在する。まずその一つは「解けなかった問題が3問ある」という解釈である。本論文のモデルを用いると、肯定文の意味表現から否定文のこの解釈の意味表現が図2のようにして導出できる。すなわち肯定文の2重リストの値の意味を「解けた問題の数」から「解けなかった問題の数」に変更するだけでよい。この操作は、この解釈で元の肯定文の「数」部分が否定語の作用領域の外にあることに対応している。

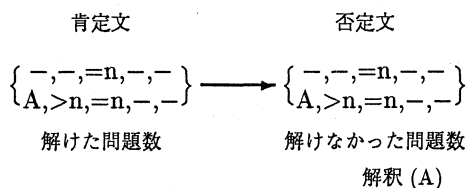


図 2: 数を含む否定文の解釈 (A)

この否定文の他の解釈としては下記の対話例が示すように「解いた問題の数が文中に表現されている数(=3)に達しなかった」という解釈がありうる。

(9) A: Did you solve three of the problems?
B: No, I didn't (get to) solve three of the problems.

——— 解釈 (B)

図3に肯定文の意味表現からこの解釈を導出する操作を示す。まず肯定文の2重リストの上下各リストの反転リストを作成する。次にその二つの反転リストの共通部分をとる。それが求めるべき解釈の下段「可能性解釈リスト」である。この可能性解釈リストから両端の値(=AとN)を削除して得られたリストが求める解釈の上段「直接解釈リスト」であり、この両者を合わせた2重リストが求める解釈の意味表現である。二つの反転リストの共通部をとったこの解釈は、肯定文の「直接解釈」の否定と「可能性解釈」の否定の両方に合致する否定解釈である。この否定解釈の「直接解釈リスト」を求める際に両端の値を削除する操作は、この否定解釈が通常、解いた問題の存在自体および解けなかった問題の存在自体は否定していないことに対応している。

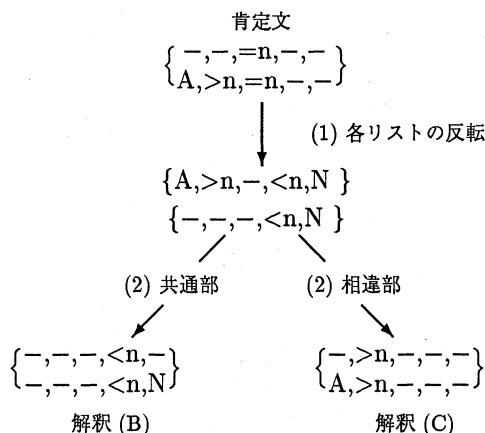


図 3: 数を含む否定文の解釈 (B,C)

この否定文は下記の例文(10)が示すように「解いた問題の数が文中に表現されている問題の数を越えている」という解釈も可能である。この解釈を肯定文の意味表現から導出するには、図3において2つの反転リストの相違部分をとればよい。すなわちこの解釈は元の肯定文の「直接解釈」の否定だけしか満たしていない否定解釈である。なおこの否定文がこうに解釈される場合には、文が発話される際に数部分に強勢(ストレス)がおかれる。この言語現象はこの否定操作の中で二つのリストの相違部分をとる操作に対応している。

(10) A: Did you solve three of the problems?
C: No, I didn't solve THREE of the problems: I solved ALL of them.

——— 解釈 (C)

5 程度概念の否定

5.1 'Many' 'A Few' 'Few'

本節では一般の程度表現にこのモデルを適用しそれらの否定の意味解釈を導出する。「many」「few」「a few」の意味を2重リストを用いて表現したのが図4である。通常「few」は否定的な少なさを表し「a few」は肯定的な少なさを表すと言われるが、そのことを明示的に表現したのはこのモデルが初めてである。

- (a) many $\begin{Bmatrix} -, M, -, -, - \\ A, M, -, -, - \end{Bmatrix}$
 (b) a few $\begin{Bmatrix} -, -, -, F, - \\ A, M, S, F, - \end{Bmatrix}$
 (c) few $\begin{Bmatrix} -, -, -, F, - \\ -, -, -, F, N \end{Bmatrix}$

図4: 'Many' 'A Few' 'Few' の2重リスト表現

5.2 'not many' の解釈の導出

「many」の意味を表す2重リストに対して前節で導入した否定操作を適用させたのが図5である。反転リストの共通部分をとった2重リストは「not many」が「単なる Some ではなくそれ以下をも含意する」ことを表す。反転リストの相違部分をとったもう一方の2重リストは「not many」が「全て」を意味しうることを表すが、上段の「直接解釈リスト」が空であることから、これが通常行なわれない解釈であることがわかる。このようにこのモデルは「not many」の vagueness も ambiguity も解釈の妥当性も同時に導出することができる。

6 おわりに

本論文では程度概念の意味を2重リスト内の離散的な相対位置で表現するモデルを説明した。この2重リスト表現上で定義した否定操作は肯定文の意味表現を元にして否定文の全ての解釈を導出することができる。

程度表現の意味理解とは指し示される現実量がわかることではなく、理解の枠組(リスト表現)内の相対的位置がわかるということである。また「often」と「しばしば」が対応するという直観は、日本語と英語の程度概念が類似した構造をもち、かつこの二語の相対的位置が対応しているという事実に基づいている。なお程度概念の素は本稿で導入したものだけではない。重要なのは自然言語における程度概念の離散的な把握の枠組である。

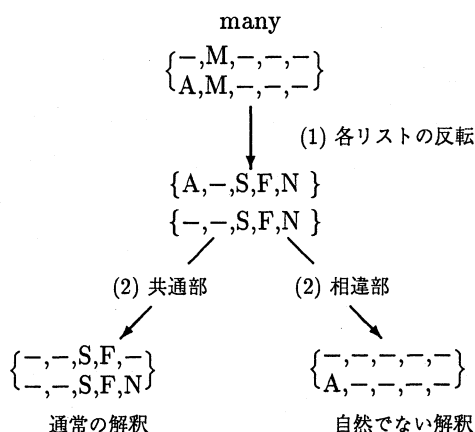


図5: 'Not many' の解釈の導出

本稿の程度表現の扱いを論理の観点から見ると量子子の一般化に相当する。現在はこのモデルの他の論理演算子への適用を試みている。例えば従来から程度表現と論理演算子「OR」との類似性が指摘されているが「inclusive OR」と「exclusive OR」とを2重リストに対応させられるものと考えている。「OR」「AND」「ならば」など一階の述語論理の演算子を拡張させることで自然言語の推論に適した論理の構築を目指したい。

参考文献

- [1] Grice, H. P. 'Logic and conversation' Syntax and Semantics 3. Academic Press. 1975.
- [2] Horn, L. R. 'On the semantic properties of logical operators in English' The Indiana University Linguistics Club 1976.
- [3] Bolinger, D. L. 'Degree words' Mouton. 1972.
- [4] Gazdar, G. 'Pragmatics: Implicature, Presupposition, and Logical Form' Academic Press. 1979.
- [5] 太田朗「否定の意味」大修館書店 1980.
- [6] Hirschberg, J. B. 'A theory of scalar implicature' ペンシルバニア大学博士論文. 1985.
- [7] 亀井、村木「程度表現のモデル化」電子情報通信学会 NLC 研究会 88-6. 1988.
- [8] Kamei and Muraki 「An Explanatory Model of Degree Concept」自然言語理解と人工知能国際シンポジウム ISKIT'92. 1992.
- [9] 亀井、村木「程度表現の意味モデル」情報処理学会第45回全国大会 予稿集 1E-10. 1992.
- [10] Kamei and Muraki 「A Discrete Model of Degree Concept in Natural Language」Coling-94, pp.775- 781. 1994.