

# 新知識を理解するための類推能力の育成

Development of Analogical Thinking Ability for Comprehending New Knowledge

石田 純太

Junta Ishida

砂山 渡

Wataru Sunayama

広島市立大学大学院情報科学研究科

Graduate School of Information Sciences, Hiroshima City University

When we try understanding of new knowledge, it performs an analogy to fit a features of the known knowledge to new knowledge, and to understand, but an opportunity and the method to bring up the ability for analogy do not exist generally. Analogy can be an effective means to help with deep understanding of new knowledge, as it can be seen well by comparing things do not know in known things. In this study, for the purpose of upbringing of the ability for analogy, we propose a system to encourage understanding of the analogy process and the practice due to repeated.

## 1. はじめに

以前に経験したことがらを現在直面していることがら、あるいは問題に当てはめることを類推という [鈴木 96]。我々が新たな知識の理解を試みる際には、既知の知識の特徴を新たな知識に当てはめて理解する類推を行っている。例えば、中学校の理科の授業では、目に見えない電圧と電流、抵抗について学ぶ際、電気回路を水流モデルに喩えた類推を用いることで生徒は理解を深めることができる。このように類推は、わからない物事を何か既知のもので喩えるとよくわかるように、新しい知識の深い理解を助けるために有効な手段になり得る。また、先行研究では、科学的発見においても、科学者は類推を頻繁に使用していることが明らかになっている [植田 00]。しかし、類推能力を育成する機会や方法は一般に存在しない。そこで本研究では、類推能力の育成を目的として、類推プロセスの理解と繰り返しによる練習を促すシステムを提案する。

本研究では、理解したい類推対象を「ターゲット」、類推対象と共通の特徴が多い既存知識を「ベース」と定義する。

次に、本研究でターゲットとして扱う事物について述べる。すべての事物は、明確な事物と不明確な事物に分けられる。明確な事物とは、個体差のない事物のことをいう。例えば、「酸素」のように、すべてのものが同じ性質を持つものをいう。不明確な事物とは、個体差のある事物のことをいう。例えば、「子供」のように、個々が違う性質を持つものをいう。本研究では、具体名詞を明確な事物、抽象名詞を不明確な事物とし、明確な事物をターゲットとして扱う。

## 2. 関連研究

類推は一般的に以下のようなプロセスからなると考えられている [鈴木 96]。

類推プロセス 1: ターゲット問題の表象

類推プロセス 2: ベースの検索

類推プロセス 3: 写像

類推プロセス 4: 正当化

連絡先: 石田純太, 広島市立大学大学院情報科学研究科, 731-3194 広島市安佐南区大塚東 3-4-1

E-mail:junta@sys.info.hiroshima-cu.ac.jp

このプロセスに従えば、類推が可能になると考えられる。このプロセス通りに類推をするためには、類推プロセスを理解するとともに繰り返し練習する必要がある。しかし、類推プロセスの理解支援や繰り返し練習する機会が一般に存在していない。そこで、類推プロセスを理解するために、プロセスに沿った手順を明確にし、繰り返し練習ができる環境を設けることが望まれる。

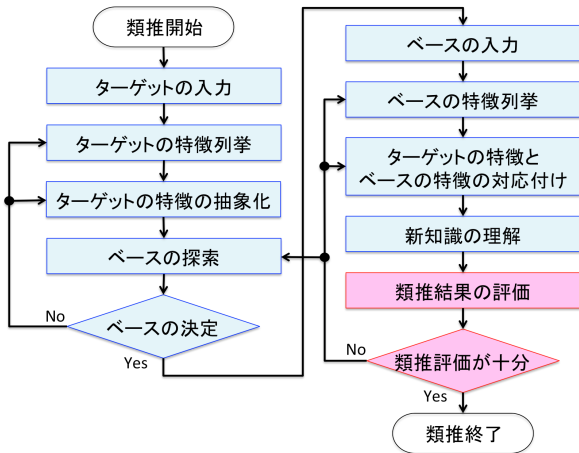
解決方法がわからない問題に対して創造的発想を行って解決しようとする思考者の着想段階を、糸口となる知識を豊富に含んだフルテキスト情報データベースの検索結果からの類推を用いて支援する方式を提案した研究がある [田中 92]。この研究では、フルテキスト情報データベースの検索結果から類推を行うことが、問題の解決策の着想に有効なことが明らかになった。しかし、この研究では、類推による着想支援として用いるフルテキスト情報データベースが極めて小規模なものである問題がある。そこで、類推を支援するための大規模なデータベースとして、検索エンジンの利用が考えられる。

類推を用いたスキル獲得支援としては、従来のルール発想推論に、類推公理を加えることによって、ルールの生成および新述語の発見、類比の補完のすべてを統一的に実現させ、スキル獲得に有効に機能するかを検証した研究がある [金城 14]。この研究では、実際に新しいスキルが発見され、類推はスキル獲得に有効なことが明らかになった。しかし、この研究は、ターゲットと類似したベースが与えられており、本来困難であるベースを探す過程を省略している。

また、類推的思考に着目した教材開発として、日常的な課題を技術的な見方や考え方を用いて解決する力を育成する e-ラーニング教材を開発した研究がある [佐藤 10]。この研究では、類推の練習回数が多いければ、未知の課題に類推を適用できるようになることが明らかになっている。しかし、この開発教材だけでは類推が上手くできない生徒に対して、システムの支援が不十分であり、教員による支援が必要になっている。

そこで本研究では、ベースを探す過程を含んだ類推手順を設定することで、類推プロセスを理解し、検索エンジン<sup>\*1</sup>での検索結果を用いて支援を行うことで、指導者を必要とせず、繰り返し練習を促すシステムを提案する。

\*1 本研究では、検索エンジンは Bing を使用。



□ ユーザ    □ システム

図 1: 類推による新知識の理解の枠組み

### 3. 類推能力の育成の枠組み

本システムでは、新知識を理解するための類推能力を育成する要素として「類推手順」「類推結果の自動評価」を実装した(図1)。類推手順は、類推プロセスを理解するために、類推プロセスに沿って設定した。類推結果の自動評価は、新知識の理解度が十分かを判定するために用意した。本システムでは、理解したい事物をターゲットとして名詞で入力する。

類推手順は6つの手順から構成され、すべての手順を終了後、システムが類推結果を自動評価する。類推結果は、8つの評価指標から評価する。類推結果の評価を上げるために手順を繰り返すことで、繰り返し練習が可能となる。

### 4. 類推手順

本研究において、[鈴木 96]の類推プロセスを理解するために、設定した類推手順を以下に示す。

- 類推手順 1: ターゲットの特徴列挙
- 類推手順 2: ターゲットの特徴の抽象化
- 類推手順 3: ベースの探索
- 類推手順 4: ベースの特徴列挙
- 類推手順 5: ターゲットの特徴とベースの特徴の対応付け
- 類推手順 6: 新知識の理解

この類推手順に沿って類推を繰り返し練習するシステムの画面を図2に示す。

最初に、理解したい事物をターゲットとして名詞で入力する。ターゲットとしては扱った名詞は、具体名詞とする。

類推手順1では、ターゲットの特徴を列挙する。入力フォームにターゲットの特徴を入力する。列挙した特徴は、図2左上のように、追加ボタンを押すことで縦に追加されていく。ターゲットの特徴は、単語で入力する。

類推手順2では、ターゲットの特徴を抽象化する。類推手順1で列挙した特徴を1つ以上選択し、選択した特徴を集合とみなし、その集合に名前を付ける。抽象化した集合一覧は、図2中央下の緑の枠内に表示する。抽象化した集合一覧は、名前を付けた集合名とその集合を作成する時に選択した特徴の通し番号をセットで表示する。特徴の通し番号は、上から1,2,...と

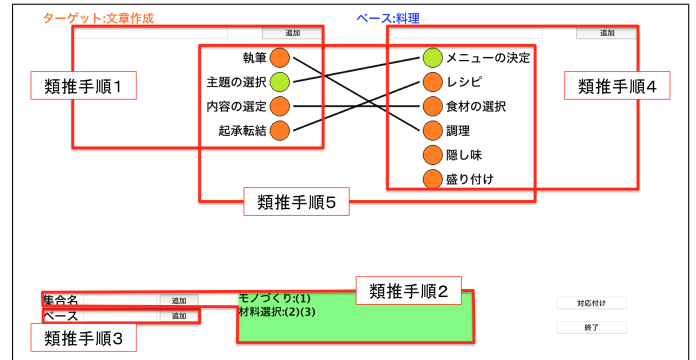


図 2: 類推手順に沿った繰り返し練習システム画面

表 1: 評価指標一覧

	評価指標	評価
ターゲット	特徴の数	5段階評価
	特徴の重要性	
	特徴の必要性	
ベース	特徴の数	
	特徴の重要性	
	特徴の必要性	
ターゲットとベース	シソーラス距離	
	対応付けの数	

する。特徴を抽象化することによって、ベースの探索とベースの特徴列挙を容易にする。

類推手順3では、ベースを探索する。類推手順2で抽象化してできた集合一覧から連想される事物をターゲットを類推するためのベースとして決定する。図2左下でベースを入力すると、図2右上に類推手順4のインターフェースが出現する。

類推手順4では、ベースの特徴を列挙する。入力フォームにベースの特徴を入力する。列挙した特徴は、図2右上のように、追加ボタンを押すことで縦に追加されていく。ターゲットと同様、ベースの特徴も単語で入力する。

類推手順5では、ターゲットの特徴とベースの特徴で共通のものを対応付けする。ターゲットの特徴とベースの特徴で共通のものを1つずつ選択すると、線で結ばれ対応付けを行うことができる(図2中央)。

類推手順6では、類推手順5の対応付けの結果をもとに、新知識の理解を深める。ベースの特徴からターゲットの新たな特徴を列挙することができ、理解を深めることが可能となる。

### 5. 類推結果の自動評価

本研究では、新知識の理解度が十分かを判定するため、類推結果を評価する。評価が一定基準を超えていなければ、新知識についての理解が不十分な可能性があることを示す。評価が特に不十分な指標に関しては、ユーザに評価を良くすることを促す。

評価指標はターゲットの特徴について3つ、ベースの特徴について3つ、ターゲットとベースの組み合わせについて2つの合計8つの指標を用意した(表1)。8つの指標はそれぞれ5段階で評価する。

## 5.1 ターゲットとベースの特徴の評価

### 5.1.1 特徴の数

ターゲットとベースともに特徴が列挙できなければ、類推プロセスが成立しない。ターゲットの特徴が挙がらなければ、ベースを探せず、新知識の理解ができない。また、列挙したターゲットやベースの特徴の数が少ない場合、対応付けが上手くできず、新知識の理解が不十分と考えられる。したがって、新知識の理解に対して、特徴の数は多いことが望まれる。

特徴の数の評価は、列挙した特徴の数に基づいて、5段階で評価を行う。特徴の数が不十分と評価された場合、特徴の列挙を促す。

### 5.1.2 特徴の重要性

ターゲットとベースに関して、多くの特徴が列挙されることが望まれるが、列挙する特徴が何でも良いというわけではなく、特徴の重要性が高いことが望まれる。本研究では、特徴の重要性を「ターゲット（あるいはベース）の主要な特徴」と定義する。重要性を評価する方法は、検索エンジンにおける検索ヒット件数を用いて、条件付き確率を求める。ターゲットの重要性を評価する場合は、「ターゲット」、「特徴かつターゲット」についての検索ヒット件数を取得し、以下の式から条件付き確率を求める。式内の  $T$  はターゲット、 $F$  は特徴を表す。

$$Importance = \frac{Hit(F \wedge T)}{Hit(T)} \quad (1)$$

上記で求めた条件付き確率が高い場合、その特徴はターゲットの主要な特徴となる。これは、特徴がターゲットの主要な特徴の場合、AND 検索をしても検索ヒット件数は多いと考えたことによる。ベースの特徴の重要性を評価する場合は、式 (1) 内のターゲットをベースに置き換え、条件付き確率を求める。

特徴の重要性の評価は、式 (1) で求めた値に基づいて、5段階で評価を行う。特徴の重要性が不十分と評価された場合、主要な特徴の列挙を促す。

### 5.1.3 特徴の必要性

ターゲットとベースに関して、多くの特徴が列挙されることが望まれるが、列挙する特徴が何でも良いというわけではなく、特徴の必要性が高いことが望まれる。本研究では、特徴の必要性を「ターゲット（あるいはベース）を想起可能な特徴」と定義する。必要性を評価する方法は、検索エンジンにおける2つの場合の検索ヒット件数の差を求める。ターゲットの必要性を評価する場合は、「特徴」、「特徴かつターゲット」についての検索ヒット件数を取得し、以下の式から検索ヒット件数の差を求める。式内の  $T$  はターゲット、 $F$  は特徴を表す。

$$Necessity = Hit(F) - Hit(F \wedge T) \quad (2)$$

上記で求めた差が小さい場合、その特徴はターゲットを想起可能な特徴となる。これは、ターゲットを想起可能な特徴の場合、特徴のみの検索ヒット件数と特徴かつターゲットの AND 検索の検索ヒット件数の差は小さいと考えたことによる。ベースの特徴の必要性を評価する場合は、式 (2) 内のターゲットをベースに置き換え、検索ヒット件数の差を求める。

特徴の必要性の評価は、式 (2) で求めた値に基づいて、5段階で評価を行う。特徴の必要性が不十分と評価された場合、ターゲット（あるいはベース）を想起可能な特徴の列挙を促す。

## 5.2 ターゲットとベースの組み合わせの評価

### 5.2.1 ターゲットとベースの距離

決定するベースは、ターゲットと共通の特徴が多いことが望まれる。ターゲットとベースが階層的シソーラスで距離が近い

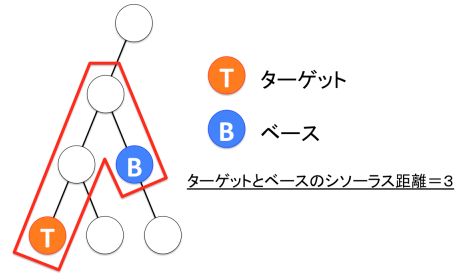


図 3: シソーラス距離

場合は、共通の特徴が多いため、ターゲットとベースのシソーラス距離が近い場合は、特徴の対応付けが容易で、新知識の理解も容易になる。しかし、ターゲットとベースのシソーラス距離が遠い場合は、シソーラス距離が近い場合では知りえなかった新知識の理解ができる可能性が高い。したがって、決定するベースは、ターゲットと共通の特徴が多いことはもちろん、ターゲットとシソーラス距離が遠いことが望まれる。本研究では、ターゲットとベースの距離をシソーラス距離によって決定する。シソーラス距離は、日本語 WordNet[Bond 09] の階層的シソーラスを用いて、ターゲットのノードからベースのノードまで辿ったノード数とする (図 3)。

ターゲットとベースの距離の評価は、シソーラス距離に基づいて、5段階で評価を行う。ターゲットとベースの距離の評価が不十分と評価された場合、ターゲットとシソーラス距離が遠いベースの再決定を促す。

### 5.2.2 対応付けの数

ターゲットの特徴とベースの特徴の対応付けを行った結果から、新知識を理解する。対応付けの数が少ない場合は、ターゲットまたはベースの列挙した特徴の数が不十分である可能性がある。また、決定したベースよりも新知識の理解を促すことができるベースがある可能性がある。したがって、新知識の理解に対して、対応付けの数が多いことが望まれる。つまり、対応付けの数が新知識の理解度に繋がる。

対応付けの数の評価は、対応付けの数に基づいて、5段階で評価を行う。対応付けの数が不十分と評価された場合、ターゲットとベースの特徴の列挙を促す。また、ターゲットとシソーラス距離が遠いベースの再決定を促す。

## 6. 今後の計画

### 6.1 各類推手順の支援

#### 6.1.1 特徴列挙の支援

類推手順 1 において、ターゲットの特徴の列挙が困難な場合は、検索エンジンのサジェスト機能を用いて特徴の候補を提示する。具体的には、ターゲットを検索エンジンにかけて、出力された検索候補を特徴の候補として提示する。類推手順 4 において、ベースの特徴の列挙が困難な場合も同様の方法で特徴の候補を提示する。

#### 6.1.2 集合作成の支援

類推手順 2 において、集合の作成が困難な場合は、ターゲットの各特徴間の類似度を可視化することで、集合として作成する際に選択する特徴を選択しやすくする。具体的には、ターゲットの各特徴間のレーベンシュタイン距離<sup>\*2</sup>やシソーラス距離を計算し、各特徴間の類似度を可視化する。

\*2 文字列同士の類似度として、文字単位の挿入・変更・削除という基本操作の最低回数で求める編集距離に基づく手法をレーベンシュタイン距離と呼ぶ。

### 6.1.3 ベース決定の支援

類推手順 3 において、ベースの決定が困難な場合は、検索エンジンのサジェスト機能を用いてベースの候補を提示する。具体的には、集合一覧を検索エンジンにかけて、出力された検索候補をベース候補として提示する。また、ターゲットとベース候補のシソーラス距離を計算し、ベース候補とともにそのシソーラス距離も提示する。

### 6.1.4 対応付けの支援

類推手順 5 において、対応付けが困難な場合は、ターゲットとベースのそれぞれの特徴間の類似度を提示する。ターゲットとベースのそれぞれの特徴間のレーベンシュタイン距離やシソーラス距離を計算し、類似度が高い組み合わせを提示する。

## 6.2 システムを利用して新知識を理解した例

本システムを利用して、文章作成について理解した例を示す。今回は「文章」をターゲットとした。

まず、ターゲットの特徴を列挙する(類推手順 1)。列挙した「文章」の特徴とその特徴の重要性和必要性を表 2 に示す。重要性<sup>\*3</sup>と必要性はそれぞれ式(1)と式(2)に基づいて計算した。

次に、列挙した特徴を抽象化するために集合を作成する(類推手順 2)。「執筆」は「モノづくり」と抽象化し、「主題」「内容」をまとめて「材料」と抽象化した。これにより、「モノづくり」と「材料」という集合を作成した。

次に、作成した集合一覧から連想される事物をベースとして決定する(類推手順 3)。今回は、「モノづくり」と「材料」から連想される「料理」をベースとして決定した。「文章」と「料理」のシソーラス距離は 5 となった。

次に、ベースの特徴を列挙する(類推手順 4)。列挙した「料理」の特徴とその特徴の重要性和必要性を表 3 に示す。「料理」の特徴は作る手順に沿って列挙した。

次に、ターゲットとベースのそれぞれの特徴を対応付ける(類推手順 5)。「料理」の特徴を「文章」の特徴に以下のように対応付けを行った。対応付けの際、「隠し味」「盛り付け」にそれぞれ対応する「文章」の特徴を新しく追加した。

- メニュー - 主題
- レシピ - 起承転結
- 食材 - 内容
- 調理 - 執筆
- 隠し味 - コーモア
- 盛り付け - 体裁

最後に、ターゲットとベースの特徴の対応付けの結果から、新知識を理解する(類推手順 6)。上記の対応付けの結果より、以下が得られた。

主題を選択し、起承転結に沿って内容を選定し、コーモアを加えながら執筆し、最後に体裁を整えることで上手く文章が作成できる。

## 6.3 システムの活用方法

類推能力を身に付けたい人に練習する環境として本システムを提供する。また、教育現場での活用や未知事項を理解する研究での活用も考えられる。

\*3 式(1)に基づいて重要性を計算する際に、ターゲットのみの検索ヒット件数よりも、ターゲットと特徴の AND 検索の検索ヒット件数が多い場合は、条件付き確率を 100%とした。

表 2: 「文章」の特徴一覧

特徴	重要性	必要性
執筆	8.69	4980000
主題	13.98	10020000
内容	28.98	32330000
起承転結	0.59	384000

表 3: 「料理」の特徴一覧

特徴	重要性	必要性
メニュー	39.55	14490000
レシピ	100	20600000
食材	100	36600000
調理	96.18	1800000
隠し味	6.43	169000
盛り付け	11.72	480000

## 6.4 評価実験の計画

本研究で開発した類推による新知識の理解の枠組みを実装したシステムを利用することで、類推プロセスの理解と繰り返しによる練習を促し、類推能力を育成できるのかについて検証するため、評価実験を行う。実験では、明確な事物をターゲットとして扱う。被験者は数日間、提案システムを利用して、用意する様々なジャンルのターゲットについて類推を用いて理解する。

## 7. まとめ

本稿では、新知識を理解するための類推能力の育成方法として、類推プロセスを理解し、繰り返し練習を促すシステムの枠組みを提案した。

## 参考文献

- [Bond 09] Francis Bond, Hitoshi Isahara, Sanae Fujita, Kiyotaka Uchimoto, Takayuki Kuribayashi and Kyoko Kanzaki : Enhancing the Japanese WordNet, The 7th Workshop on Asian Language Resources, in Conjunction with ACL-IJCNLP 2009.
- [植田 00] 植田一博: 科学者の類推による発見, 人工知能学会論文誌 15 巻 4 号, pp608-617(2000).
- [金城 14] 金城敬太, 尾崎知伸, 古川康一, 原口誠: アナロジーを組み込んだルール発想推論によるスキル獲得支援, 人工知能学会論文誌 29 巻 1 号, pp188-193(2014).
- [佐藤 10] 佐藤治樹, 松田稔樹: 類推的思考に着目した問題解決力育成のための技術教育教材の開発, 日本教育工学会研究報告集, pp129-136(2010).
- [鈴木 96] 鈴木宏昭: 説明と類推による学習, 波多野誼余夫(編)『認知心理学 5 学習と発達』pp. 149-179, 東京大学出版会(1996).
- [田中 92] 田中一男: アナロジーを用いた着想支援方式, 情報処理学会研究報告知能と複雑系(ICS)1992 巻 70 号, pp41-50(1992).