

2022年度（令和4年度）卒業論文

属性の頻度で選出した類似単語において
引き算を用いた類推

指導教員

村田真樹
村上仁一

鳥取大学工学部 電気情報系学科

自然言語処理研究室

B19T2059C 田代 寛治

概要

発想支援とは、既存のアイデアや物事を様々な方向から発展させ新たな発想を生み出すことである。発想支援の手法の一つにチェックリスト法 [1][2][3] というものがある。チェックリスト法とはアレックス・F・オズボーン氏が作った発想法で、いくつかの項目を利用して発想支援を行う手法である。

小野の研究 [5] と須ヶ牟田の研究 [6] では、概念属性データベースを用いて発想支援が行われた。概念属性データベースは、ある単語を表す概念を属性としたデータベースである。小野の研究では属性の拡大と縮小による類推、須ヶ牟田の研究では属性の結合による類推が行われた。

本研究では、概念属性データベースを用いて発想支援を行う。概念属性データベースは、ある単語を表す概念を属性としたデータベースである。先行研究でも用いられた McRae のデータセット [4] を概念属性データベースとする。概念属性データベースから似た単語同士において属性の類推を行って、「単語 A」 - 「単語 B」 → 「属性 C」, 「単語 A」 - 「属性 C」 → 「単語 B」, 「単語 B」 + 「属性 C」 → 「単語 A」 の 3 種類の類推に 4 種類の手法を用いて類推し、類推方法の追加を目指す。

手法 1 は類似単語組の選出に属性票数を使い、属性の選出に票数比率を使う。手法 2 は類似単語組の選出には属性票数を使うが、属性の選出は乱数で行う。手法 3 は類似単語組の選出に属性票数を使わないで、属性の選出には票数比率を使う。手法 4 は類似単語組の選出に属性票数を使わないで、属性の選出には乱数を使う。これらの 4 種類の手法を 3 種類の類推に用いた結果から「類推可能」, 「大方類推可能」, 「見方によっては類推可能」, 「類推不可能」の 4 段階で評価した。被験者実験を 5 人で行い、上記の評価内容と同様に評価を行った。類推全体において「類推可能」と「大方類推可能」を正解とする場合の平均正解率を比較すると精度の良さは手法 3 が 0.34, 手法 1 が 0.31, 手法 4 が 0.19, 手法 2 が 0.18 の順番になった。また、「類推可能」と「大方類推可能」と「見方によっては類推可能」を正解とする場合の平均正解率を比較すると、精度の良さは手法 3 が 0.56, 手法 1 が 0.51, 手法 4 が 0.40, 手法 2 が 0.34 の順番になった。類推全体では手法 3 が最も精度が高かったが、類推「単語 B」 + 「属性 C」 → 「単語 A」では手

法1の方が精度が高かった。これにより属性の選出には票数比率を使った方が良いことが分かった。この手法3の類推「単語A」-「単語B」→「属性C」が「類推可能」, 「大方類推可能」を正解とする場合の正解率が0.43, 「類推可能」, 「大方類推可能」, 「見方によっては類推可能」を正解とする場合の正解率が0.69の精度で本研究では比較的良い精度だった。

小野の研究[5]では属性の拡大と縮小による類推, 須ヶ牟田の研究[6]では属性の結合による類推が行われて, さらに本研究では似た単語同士における属性の差を用いて間接的に単語から属性が類推されるという先行研究では行えなかった類推が可能になった。類推「単語B」+「属性C」→「単語A」の手法4は小野の研究[5]の手法とほぼ同じである。手法1と手法3はこの手法4に対して属性選出に票数比率を使うという改善を加えた手法である。そのため先行手法と提案手法は比較することができる。先行研究では「類推可能」と「大方類推可能」を正解とする場合の正解率は0.21であった。本研究での類推「単語B」+「属性C」→「単語A」の手法1と手法3の「類推可能」と「大方類推可能」を正解とする場合の正解率は0.31であった。これは先行研究よりも0.1大きい値である。これにより提案手法の属性票数は精度向上の役に立つことが分かる。また提案手法が先行研究の手法よりも優れていることが分かる。

目次

第1章	はじめに	1
第2章	先行手法	4
2.1	McRae のデータセット	4
2.2	小野の先行研究	4
2.3	須ヶ牟田の先行研究	5
第3章	提案手法	6
3.1	概念属性データベース	6
3.2	手法の概要	7
3.3	手法1: 票数ありの属性の \cos 類似度による単語組選出と票数比率による属性選出	7
3.4	手法2: 票数ありの属性の \cos 類似度による単語組選出と乱数による属性選出	7
3.5	手法3: 票数なしの属性の \cos 類似度による単語組選出と票数比率による属性選出	8
3.6	手法4: 票数なしの属性の \cos 類似度による単語組選出と乱数による属性選出	8
第4章	提案手法の実験	9
4.1	評価方法	9
4.2	著者の実験結果	10
4.3	著者の評価例	11
第5章	被験者実験	12
5.1	実験結果	12
5.2	手法1の評価例	14

5.3	手法 2 の評価例	16
5.4	手法 3 の評価例	18
5.5	手法 4 の評価例	20
第 6 章	考察	22
6.1	類推全体における手法ごとの精度比較	22
6.2	類推ごとの比較	22
6.2.1	類推「単語 A」 - 「単語 B」 → 「属性 C」 の考察	22
6.2.2	類推「単語 A」 - 「属性 C」 → 「単語 B」 の考察	23
6.2.3	類推「単語 B」 + 「属性 C」 → 「単語 A」 の考察	23
6.3	人手評価の正解例の考察	24
6.4	人手評価の不正解例の考察	24
6.5	先行研究との比較	24
第 7 章	今後の課題	26
第 8 章	おわりに	27

表 目 次

2.1	概念属性データベースの例	4
2.2	a_cat を持つものを取り出す	5
2.3	a_cat と is_large の両方を持つものを取り出す	5
3.1	概念属性データベースの例	6
4.1	4段階評価の内容	9
4.2	4段階評価の例	9
4.3	手法1(票数ありの属性の cos 類似度による単語組選出と票数比率で属性選出)の正解率	10
4.4	手法2(票数ありの属性の cos 類似度による単語組選出と乱数による属性選出)の正解率	10
4.5	手法3(票数なしの属性の cos 類似度による単語組選出と票数比率による属性選出)の正解率	10
4.6	手法4(票数なしの属性の cos 類似度による単語組選出と乱数による属性選出)の正解率	10
4.7	手法1(票数ありの属性の cos 類似度による単語組選出と票数比率による属性選出)の例	11
4.8	手法2(票数ありの属性の cos 類似度による単語組選出と乱数による属性選出)の例	11
4.9	手法3(票数なしの属性の cos 類似度による単語組選出と票数比率による属性選出)の例	11
4.10	手法4(票数なしの属性の cos 類似度による単語組選出と乱数による属性選出)の例	11
5.1	手法1の平均正解率	12
5.2	手法2の平均正解率	12

5.3	手法 3 の平均正解率	13
5.4	手法 4 の平均正解率	13
5.5	一人目の手法 1 での実験結果の例	14
5.6	二人目の手法 1 での実験結果の例	14
5.7	三人目の手法 1 での実験結果の例	15
5.8	四人目の手法 1 での実験結果の例	15
5.9	五人目の手法 1 での実験結果の例	15
5.10	一人目の手法 2 での実験結果の例	16
5.11	二人目の手法 2 での実験結果の例	16
5.12	三人目の手法 2 での実験結果の例	17
5.13	四人目の手法 2 での実験結果の例	17
5.14	五人目の手法 2 での実験結果の例	17
5.15	一人目の手法 3 での実験結果の例	18
5.16	二人目の手法 3 での実験結果の例	18
5.17	三人目の手法 3 での実験結果の例	19
5.18	四人目の手法 3 での実験結果の例	19
5.19	五人目の手法 3 での実験結果の例	19
5.20	一人目の手法 4 での実験結果の例	20
5.21	二人目の手法 4 での実験結果の例	20
5.22	三人目の手法 4 での実験結果の例	21
5.23	四人目の手法 4 での実験結果の例	21
5.24	五人目の手法 4 での実験結果の例	21

第1章 はじめに

発想支援とは、既存のアイデアや物事を様々な方向から発展させ新たな発想を生み出すことである。発想支援の手法の一つにチェックリスト法 [1][2][3] というものがある。チェックリスト法とはアレックス・F・オズボーン氏が作った発想法で、いくつかの項目を利用して発想支援を行う手法である。

発想支援の研究を、以前の小野の研究 [5] では、「猫」という単語の属性を考えて、その属性に「大きい」という属性を追加することで「虎」という単語を類推して、「猫を大きくすると虎が導き出される」といった拡大や縮小の類推を行った。単語およびそれに付随する概念属性の集合データである McRae のデータセット [4] を基に研究を進めた。須ヶ牟田の研究 [6] は小野の研究 [5] を発展させようとし、小野の研究 [5] で行われなかった属性を結合して単語を類推する研究を行うことによる類推方法の増加を目指した。この研究も McRae のデータセットを基に進めた。

本研究では、似た単語同士において類推を行って、類推方法の追加を目指す。手法1は類似単語組の選出に属性票数を使い、属性の選出に票数比率を使う。手法2は類似単語組の選出には属性票数を使うが、属性の選出は乱数で行う。手法3は類似単語組の選出に属性票数を使わないで、属性の選出には票数比率を使う。手法4は類似単語組の選出に属性票数を使わないで、属性の選出には乱数を使う。この4種類の手法を「単語A」－「単語B」→「属性C」、「単語A」－「属性C」→「単語B」、「単語B」＋「属性C」→「単語A」全ての類推に用いて類推した結果を「類推可能」、「大方類推可能」、「見方によっては類推可能」、「類推不可能」の4段階で評価する。

本研究でも小野 [5]、須ヶ牟田 [6] の研究と同じく McRae のデータセット [4] を基に研究を進める。いずれは既存データ内に囚われずに新たな発想を生み出す発想および類推の補助に繋がっていきたい。

本研究の主な主張点は以下の4点である。

- 本研究の目的は似た単語同士から概念属性データベースを基に発想支援を行えるようにする。(例: 「チーター」引く「トラ」から「速い」を類推)

- 4種類の手法を用いて類似単語組と属性を選出し、3種類の類推を評価した。3種類の類推全体において「類推可能」と「大方類推可能」を正解とする場合の平均正解率を比較すると精度の良さは手法3が0.34, 手法1が0.31, 手法4が0.19, 手法2が0.18の順番になった。また、「類推可能」と「大方類推可能」と「見方によっては類推できる」を正解とする場合の平均正解率を比較すると、精度の良さは手法3が0.56, 手法1が0.51, 手法4が0.40, 手法2が0.34の順番になった。類推全体では手法3が最も精度が高かったが、類推「単語B」+「属性C」→「単語A」では手法1の方が精度が高かった。これにより手法1と手法3で用いている方法である属性の選出に票数比率を使う方法が良いことが分かる。しかし、手法1と手法3は優劣がつかないため手法1と手法3で用い方の異なる方法である「類似単語組の選出に属性票数を使う方法」を使うか否かは状況次第であると分かった。
- 小野の研究[5]では属性の拡大と縮小による類推、須ヶ牟田の研究[6]では属性の結合による類推が行われて、さらに本研究では似た単語同士における属性の差を用いて間接的に単語から属性が類推されるという先行研究では行えなかった類推が可能になった。類推「単語A」-「単語B」→「属性C」の類推精度は「類推可能」、「大方類推可能」を正解とする場合の正解率が0.43で、「類推可能」、「大方類推可能」、「見方によっては類推可能」を正解とする場合の正解率が0.69となり、本研究では比較的良い類推精度だった。
- 類推「単語B」+「属性C」→「単語A」の手法4は小野の研究[5]の手法とほぼ同じである。手法1と手法3はこの手法4に対して属性の選出に票数比率を使うという改善を加えた手法である。そのため先行手法と提案手法は比較することができる。先行研究では「類推可能」と「大方類推可能」を正解とする場合の正解率は0.21であった。本研究での類推「単語B」+「属性C」→「単語A」の手法1と手法3の「類推可能」と「大方類推可能」を正解とする場合の正解率は0.31であった。これは先行研究よりも0.1大きい値である。これにより提案手法の属性票数は精度向上の役に立つことが分かる。また提案手法が先行研究の手法よりも優れていることが分かる。

本論文の構成は以下のとおりである。第2章は、本研究に関連する研究としてどのような研究が行われてきたか記述する。第3章は、本研究での4つの提案手法を記述する。第4章は、本研究の4つの手法についての被験者実験の実験結果を記述する。第5章は、

本研究の被験者実験について記述する。第6章は、考察を記述する。第7章は、今後の課題について記述する。

第2章 先行手法

本章では、先行研究について記述する。2.1節ではMcRaeのデータセット [4] について記述する。2.2節では小野が行った単語を複数の属性で表現したことについて記述する。

2.1 McRaeのデータセット

McRaeのデータセット [4] は、基本レベルに属すると考えられる 541 種類の英単語概念に対し、725 名の評定者が与えた総計 7,529 種類の `is_large`, `a_cat` のような意味属性データを収集したものである。McRaeのデータセットの一部を表 2.1 に示す。

表 2.1: 概念属性データベースの例

単語	概念属性
cat	<code>an_animal has_4_leg a_pet has_fur ...</code>
tiger	<code>a_cat an_animal has_4leg has_fur is_large ...</code>
gun	<code>a_weapon made_of_metal is_dangerous ...</code>
pistol	<code>a_gun a_weapon made_of_metal is_small ...</code>
rifle	<code>a_gun a_weapon made_of_metal is_long ...</code>

2.2 小野の先行研究

小野は表 2.1 のような概念属性データベースから属性の拡大と縮小による類推を行う。例えばネコ (`cat`) を大きくしたいならネコ (`cat`) の属性に `is_large` という属性を加える。そして加えた属性を持つ項目を取り出す。そしてトラなどが抽出されることを期待する。以下に手順を示す。

- 手順1 a_単語 (a_cat など) の属性を持っている単語を取り出し, 属性として扱う.
- 手順2 取り出した単語に加えたい属性 (is_large など) を探す.
- 手順3 手順1 で取り出した属性と 手順2 で取り出した属性の両方を持つ単語を取り出す.

表 2.2 と表 2.3 に手順 1 から手順 3 の具体例を示す。

表 2.2: a_cat を持つものを取り出す

単語	概念属性
tiger	a_cat an_animal, has_4leg, has_fur, is_large …
lion	a_cat an_animal, has_4leg, has_fur, is_large …
cheetah	a_cat an_animal, has_4leg, has_fur, is_large …

表 2.3: a_cat と is_large の両方を持つものを取り出す

単語	概念属性
tiger	a_cat an_animal, has_4leg, has_fur, is_large …
lion	a_cat an_animal, has_4leg, has_fur, is_large …

2.3 須ヶ牟田の先行研究

須ヶ牟田は表 2.1 のような概念属性データベースから属性の結合による類推を行う。例えば概念属性データベースから「食器洗いに使われる」と「水が必要」から「食器洗浄機」を類推するといった類推を行う。

第3章 提案手法

本章の構成は次の通りである。3.1節では概念属性データベースについて、3.2節では手法の概要について、3.3節から3.6節では手法1~4について記述している。

3.1 概念属性データベース

本研究ではMcRaeのデータセットを概念属性データベースとして用いて類推を行う。概念属性データベースとは、ある単語を表す概念を属性としたデータベースである。McRaeのデータセットの意味属性データを概念属性とする。概念属性データベースの一部を表3.1に示す。

表 3.1: 概念属性データベースの例

単語	概念属性	属性票数
cat	has_fur	22
cat	an_animal	21
cat	a_pet	17
cat
tiger	has_stripes	22
tiger	a_carnivore	17
tiger	an_animal	15
tiger
pistol	inbeh_fires	26
pistol	a_weapon	17
pistol	a_gun	16
pistol
rifle	used_for_hunting	23
rifle	inbeh_fires	18
rifle	a_gun	14
rifle

3.2 手法の概要

属性票数を割合に換算した票数比率と正規分布と内積を利用して類似しているかを判断する \cos 類似度を用いて, 4 通りの手法で類推する. まず

- (単語 A) - (単語 B) \rightarrow (属性 C)

の引き算を行い, 次に矢印をイコールとして見て

- (単語 A) - (属性 C) \rightarrow (単語 B)
- (単語 B) + (属性 C) \rightarrow (単語 A)

と引き算を変形した類推を行う. 具体的な説明では最初の引き算のみを例示する.

3.3 手法 1: 票数ありの属性の \cos 類似度による単語組選出と票数比率による属性選出

単語 A と単語 B の \cos 類似度と単語 A での属性 C の票数比率の積が大きい 3 つ組を抽出する. 単語 A と単語 B のベクトルは属性の票数の値をベクトルの要素の値として表現し, \cos 類似度を算出する. 属性 C は単語 A の属性のうち単語 B の属性でないものとする. 属性 C の票数比率は, 単語 A での属性 C の票数を属性の単語 A の属性の票数の和で割ったものである.

- 【例 1】 (単語 urn の持つ属性) - (単語 bucket の持つ属性) \rightarrow (urn のみ持つ属性 used_for_ashes)

3.4 手法 2: 票数ありの属性の \cos 類似度による単語組選出と乱数による属性選出

単語 A と単語 B の \cos 類似度が大きい 3 つ組を抽出する. 単語 A と単語 B のベクトルは属性の票数の値をベクトルの要素の値として表現し, \cos 類似度を算出する. 属性 C は単語 A の属性のうち単語 B の属性でないものとする.

- 【例 2】 (単語 strainer の持つ属性) - (単語 colander の持つ属性) \rightarrow (strainer のみ持つ属性 used_for_vegetables)

3.5 手法 3: 票数なしの属性の cos 類似度による単語組選出と票数比率による属性選出

単語 A と単語 B の cos 類似度と単語 A での属性 C の票数比率の積が大きい 3 つ組を抽出する。単語 A と単語 B のベクトルは属性をベクトルの要素とし、要素の値は全て 1 に固定して表現して cos 類似度を算出する。属性 C は単語 A の属性のうち単語 B の属性でないものとする。属性 C の票数比率は、単語 A での属性 C の票数を属性の単語 A の属性の票数の和で割ったものである。

- 【例 3】 (単語 shield の持つ属性) - (単語 catapult の持つ属性) → (shield のみ持つ属性 used_for_protection)

3.6 手法 4: 票数なしの属性の cos 類似度による単語組選出と乱数による属性選出

単語 A と単語 B の cos 類似度が大きい 3 つ組を抽出する。単語 A と単語 B のベクトルは属性をベクトルの要素とし、要素の値は全て 1 に固定して表現し、cos 類似度を算出する。属性 C は単語 A の属性のうち単語 B の属性でないものとする。

- 【例 4】 (単語 partridge の持つ属性) - (単語 pheasant の持つ属性) → (partridge のみ持つ属性 lives_in_pear_trees)

第4章 提案手法の実験

本実験は4手法それぞれを類似した単語同士で30組ずつ3種類の類推を実行して、それぞれを◎, ○, △, ×の4段階で評価する.

4.1 評価方法

本実験の評価方法について, 出力された単語と属性に対して, ◎, ○, △, ×の4段階で評価し, 評価した結果から正解率を求める. ◎のみを正解とした場合, ◎と○を正解とした場合, ◎, ○, △を正解とした場合の3種の評価基準群における正解率を求める. この評価基準は先行研究でも同様に行われたため, 比較することも可能になると考えている. 表 4.1 と 4.2 に 4段階評価の内容とその例をそれぞれ示す.

表 4.1: 4段階評価の内容

	評価の内容
◎	類推できる
○	大方類推できる, 少し変だと思う
△	見方によっては類推できる
×	類推できない

表 4.2: 4段階評価の例

引かれる単語	引く単語	類推された属性	評価
カミソリ	斧	剃るのに使う	◎
ナマズ	サバ	湖に住む	○
ポット	バケツ	料理に使う	△
さば	サーモン	人が狩る	×

4.2 著者の実験結果

本実験で、著者の行った人手評価の正解率を表 4.3 から表 4.6 に示す。

表 4.3: 手法 1(票数ありの属性の \cos 類似度による単語組選出と票数比率で属性選出)の正解率

	◎の割合	◎と○合計の割合	◎と○と△合計の割合
A-B	0.10	0.67	0.73
A-C	0.03	0.40	0.60
B+C	0.03	0.43	0.60
平均	0.53	0.50	0.64

表 4.4: 手法 2(票数ありの属性の \cos 類似度による単語組選出と乱数による属性選出)の正解率

	◎の割合	◎と○合計の割合	◎と○と△合計の割合
A-B	0.13	0.50	0.67
A-C	0.00	0.50	0.67
B+C	0.00	0.53	0.63
平均	0.04	0.51	0.66

表 4.5: 手法 3(票数なしの属性の \cos 類似度による単語組選出と票数比率による属性選出)の正解率

	◎の割合	◎と○合計の割合	◎と○と△合計の割合
A-B	0.57	0.83	0.90
A-C	0.07	0.50	0.77
B+C	0.10	0.53	0.77
平均	0.25	0.62	0.81

表 4.6: 手法 4(票数なしの属性の \cos 類似度による単語組選出と乱数による属性選出)の正解率

	◎の割合	◎と○合計の割合	◎と○と△合計の割合
A-B	0.23	0.50	0.60
A-C	0.00	0.43	0.60
B+C	0.03	0.50	0.67
平均	0.09	0.48	0.62

4.3 著者の評価例

本実験で、著者の行った人手評価の例を表 4.7 から表 4.10 に示す。

表 4.7: 手法 1(票数ありの属性の cos 類似度による単語組選出と票数比率による属性選出)の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
octopus	squid	has_8_tentacles	◎	◎	◎
clam	squid	has_a_shell	○	△	△

表 4.8: 手法 2(票数ありの属性の cos 類似度による単語組選出と乱数による属性選出)の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
tongs	colander(水切りボウル)	used_for_picking_up_items	○	×	×
raven	oriole(ムクドリモドキ)	is_large	○	○	○

表 4.9: 手法 3(票数なしの属性の cos 類似度による単語組選出と票数比率による属性選出)の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
porcupine(ヤマアラシ)	gopher(ホリネズミ)	has_quills(トゲトゲ)	◎	○	○
finch(アトリ科)	nightingale	is_small	×	×	×

表 4.10: 手法 4(票数なしの属性の cos 類似度による単語組選出と乱数による属性選出)の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
mackerel(さば)	salmon	hunted_by_people	×	×	×
catfish(ナマズ)	mackerel	lives_in_lakes	○	△	△

第5章 被験者実験

4手法それぞれを類似した単語同士で30組ずつ3種類の類推を実行して、それぞれを◎, ○, △, ×の4段階で評価する評価実験を5人の被験者に対して行う。

5.1 実験結果

被験者実験の手法ごとの平均正解率を表5.1から表5.4に示す。

表 5.1: 手法1の平均正解率

	◎の割合	◎と○合計の割合	◎と○と△合計の割合
A-B	0.18	0.37	0.55
A-C	0.10	0.25	0.45
B+C	0.15	0.31	0.54
平均	0.14	0.31	0.51

表 5.2: 手法2の平均正解率

	◎の割合	◎と○合計の割合	◎と○と△合計の割合
A-B	0.07	0.26	0.47
A-C	0.02	0.09	0.10
B+C	0.05	0.19	0.46
平均	0.05	0.18	0.34

表 5.3: 手法 3 の平均正解率

	◎の割合	◎と○合計の割合	◎と○と△合計の割合
A-B	0.20	0.43	0.69
A-C	0.11	0.28	0.48
B+C	0.12	0.31	0.49
平均	0.14	0.34	0.56

表 5.4: 手法 4 の平均正解率

	◎の割合	◎と○合計の割合	◎と○と△合計の割合
A-B	0.05	0.29	0.43
A-C	0.02	0.10	0.38
B+C	0.03	0.17	0.40
平均	0.03	0.19	0.40

5.2 手法1の評価例

手法1(票数ありの属性の cos 類似度で単語組と票数比率で属性選出)の被験者実験の評価例を表 5.5 から表 5.9 に示す.

表 5.5: 一人目の手法1での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
urn	bucket	used_for_ashes	△	×	×
tongs	colander(水切りボウル)	used_for_picking_up_items	×	×	×
peg	bucket	made_of_wood	×	×	×
peg	cap_(bottle)	made_of_wood	×	×	×
bin_(waste)(ごみ置き場)	bucket	used_for_holding_garbage	○	△	△
pot	bucket	used_for_cooking	◎	○	◎
colander(水切りボウル)	tongs(トング)	used_for_straining(細かい物の除去)	×	×	×
bluejay(アオカケス)	nightingale(サヨナキドリ)	is_blue	◎	○	○
shotgun	bazooka	requires_bullets	×	×	×
shield	bazooka	used_for_protection	△	×	×

表 5.6: 二人目の手法1での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
urn	bucket	used_for_ashes	×	×	×
tongs	colander(水切りボウル)	used_for_picking_up_items	×	×	×
peg	bucket	made_of_wood	×	×	×
peg	cap_(bottle)	made_of_wood	×	×	×
bin_(waste)(ごみ置き場)	bucket	used_for_holding_garbage	×	×	×
pot	bucket	used_for_cooking	×	×	×
colander(水切りボウル)	tongs(トング)	used_for_straining(細かい物の除去)	×	×	×
bluejay(アオカケス)	nightingale(サヨナキドリ)	is_blue	◎	○	○
shotgun	bazooka	requires_bullets	×	×	×
shield	bazooka	used_for_protection	×	×	×

表 5.7: 三人目の手法 1 での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
urn	bucket	used_for_ashes	×	△	△
tongs	colander(水切りボウル)	used_for_picking_up_items	×	×	×
peg	bucket	made_of_wood	×	×	×
peg	cap_(bottle)	made_of_wood	×	×	×
bin_(waste)(ごみ置き場)	bucket	used_for_holding_garbage	◎	◎	◎
pot	bucket	used_for_cooking	△	◎	◎
colander(水切りボウル)	tongs(トング)	used_for_straining(細かい物の除去)	×	△	○
bluejay(アオカケス)	nightingale(サヨナキドリ)	is_blue	○	○	△
shotgun	bazooka	requires_bullets	×	×	×
shield	bazooka	used_for_protection	△	○	△

表 5.8: 四人目の手法 1 での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
urn	bucket	used_for_ashes	×	○	×
tongs	colander(水切りボウル)	used_for_picking_up_items	○	×	△
peg	bucket	made_of_wood	×	×	×
peg	cap_(bottle)	made_of_wood	×	×	×
bin_(waste)(ごみ置き場)	bucket	used_for_holding_garbage	×	×	◎
pot	bucket	used_for_cooking	○	○	○
colander(水切りボウル)	tongs(トング)	used_for_straining(細かい物の除去)	×	×	×
bluejay(アオカケス)	nightingale(サヨナキドリ)	is_blue	◎	◎	◎
shotgun	bazooka	requires_bullets	×	×	×
shield	bazooka	used_for_protection	△	×	×

表 5.9: 五人目の手法 1 での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
urn	bucket	used_for_ashes	×	×	×
tongs	colander(水切りボウル)	used_for_picking_up_items	△	×	×
peg	bucket	made_of_wood	×	×	×
peg	cap_(bottle)	made_of_wood	×	×	×
bin_(waste)(ごみ置き場)	bucket	used_for_holding_garbage	△	×	○
pot	bucket	used_for_cooking	◎	△	△
colander(水切りボウル)	tongs(トング)	used_for_straining(細かい物の除去)	×	×	×
bluejay(アオカケス)	nightingale(サヨナキドリ)	is_blue	◎	○	○
shotgun	bazooka	requires_bullets	×	×	×
shield	bazooka	used_for_protection	×	×	×

5.3 手法2の評価例

手法2(票数ありの属性の cos 類似度で単語組と乱数で属性選出)の被験者実験の評価例を表5.10から表5.13に示す.

表 5.10: 一人目の手法2での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
strainer(ザル)	colander	used_for_vegetables	×	△	×
gun	bazooka	used_for_shooting	△	×	△
knife	machete	used_with_forks	△	×	×
sword	machete	used_for_war	△	○	◎
sweater	coat	is_knitted	○	△	△
pants(男もの)	trousers(男女もの)	worn_by_women	×	×	×
beatle(甲虫類)	cockroach(ゴキブリ)	has_wings	×	×	×
perch(スズキ)	mackerrel(サバ)	tastes_good	×	×	×
dagger(短剣)	machete	used_in_MacBeth	×	×	×
squid(イカ)	octopus(タコ)	lives_in_saltwater	×	×	×

表 5.11: 二人目の手法2での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
strainer(ザル)	colander	used_for_vegetables	×	×	×
gun	bazooka	used_for_shooting	×	×	×
knife	machete	used_with_forks	×	×	×
sword	machete	used_for_war	◎	△	◎
sweater	coat	is_knitted	◎	△	△
pants(男もの)	trousers(男女もの)	worn_by_women	×	×	×
beatle(甲虫類)	cockroach(ゴキブリ)	has_wings	×	×	×
perch(スズキ)	mackerrel(サバ)	tastes_good	×	×	×
dagger(短剣)	machete	used_in_MacBeth	×	×	×
squid(イカ)	octopus(タコ)	lives_in_saltwater	×	×	×

表 5.12: 三人目の手法 2 での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
strainer(ザル)	colander	used_for_vegetables	×	×	×
gun	bazooka	used_for_shooting	△	△	△
knife	machete	used_with_forks	△	○	○
sword	machete	used_for_war	◎	○	○
sweater	coat	is_knitted	◎	◎	◎
pants(男もの)	trousers(男女もの)	worn_by_women	×	×	×
beatle(甲虫類)	cockroach(ゴキブリ)	has_wings	△	△	△
perch(スズキ)	mackerrel(サバ)	tastes_good	△	△	△
dagger(短剣)	machete	used_in_MacBeth	×	×	×
squid(イカ)	octopus(タコ)	lives_in_saltwater	×	×	×

表 5.13: 四人目の手法 2 での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
strainer(ザル)	colander	used_for_vegetables	×	×	×
gun	bazooka	used_for_shooting	×	×	×
knife	machete	used_with_forks	×	△	◎
sword	machete	used_for_war	×	×	○
sweater	coat	is_knitted	×	△	○
pants(男もの)	trousers(男女もの)	worn_by_women	○	×	×
beatle(甲虫類)	cockroach(ゴキブリ)	has_wings	×	×	×
perch(スズキ)	mackerrel(サバ)	tastes_good	×	×	×
dagger(短剣)	machete	used_in_MacBeth	×	×	×
squid(イカ)	octopus(タコ)	lives_in_saltwater	×	×	×

表 5.14: 五人目の手法 2 での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
strainer(ザル)	colander	used_for_vegetables	△	×	△
gun	bazooka	used_for_shooting	×	×	△
knife	machete	used_with_forks	○	△	○
sword	machete	used_for_war	×	×	△
sweater	coat	is_knitted	○	△	△
pants(男もの)	trousers(男女もの)	worn_by_women	◎	△	△
beatle(甲虫類)	cockroach(ゴキブリ)	has_wings	×	×	×
perch(スズキ)	mackerrel(サバ)	tastes_good	×	×	×
dagger(短剣)	machete	used_in_MacBeth	×	×	×
squid(イカ)	octopus(タコ)	lives_in_saltwater	△	×	△

5.4 手法3の評価例

手法3(票数なしの属性の cos 類似度で単語組と票数比率で属性選出)の被験者実験の評価例を表 5.14 から表 5.19 に示す.

表 5.15: 一人目の手法3での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
urn	bucket	used_for_ashes	△	△	△
pajamas	nightgown	worn_to_bed	×	×	×
shield	catapult	used_for_protection	△	×	×
paintbrush	broom	used_for_painting	○	×	△
bluejay	sparrow	is_blue	◎	○	○
broccoli	califlower	is_green	◎	◎	◎
catapult	shield	used_for_throwing	△	×	×
peg	cap(ボトルの)	made_of_wood	×	×	×
peg	bin_(waste)(ごみ置き場)	made_of_wood	×	×	×
bluejay	starling(ムクドリ)	is_blue	△	×	×

表 5.16: 二人目の手法3での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
urn	bucket	used_for_ashes	×	×	×
pajamas	nightgown	worn_to_bed	△	△	△
shield	catapult	used_for_protection	×	×	×
paintbrush	broom	used_for_painting	○	△	○
bluejay	sparrow	is_blue	◎	◎	◎
broccoli	califlower	is_green	◎	◎	◎
catapult	shield	used_for_throwing	×	×	×
peg	cap(ボトルの)	made_of_wood	×	×	×
peg	bin_(waste)(ごみ置き場)	made_of_wood	×	×	×
bluejay	starling(ムクドリ)	is_blue	△	△	△

表 5.17: 三人目の手法 3 での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
urn	bucket	used_for_ashes	×	△	△
pajamas	nightgown	worn_to_bed	○	○	○
shield	catapult	used_for_protection	△	△	△
paintbrush	broom	used_for_painting	○	○	○
bluejay	sparrow	is_blue	○	○	○
broccoli	califlower	is_green	◎	◎	◎
catapult	shield	used_for_throwing	△	△	△
peg	cap(ボトルの)	made_of_wood	×	×	×
peg	bin_(waste)(ごみ置き場)	made_of_wood	×	×	×
bluejay	starling(ムクドリ)	is_blue	○	△	△

表 5.18: 四人目の手法 3 での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
urn	bucket	used_for_ashes	×	○	×
pajamas	nightgown	worn_to_bed	×	×	×
shield	catapult	used_for_protection	×	×	×
paintbrush	broom	used_for_painting	△	○	○
bluejay	sparrow	is_blue	◎	○	○
broccoli	califlower	is_green	◎	◎	◎
catapult	shield	used_for_throwing	×	×	×
peg	cap(ボトルの)	made_of_wood	×	×	×
peg	bin_(waste)(ごみ置き場)	made_of_wood	×	×	×
bluejay	starling(ムクドリ)	is_blue	◎	○	○

表 5.19: 五人目の手法 3 での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
urn	bucket	used_for_ashes	×	×	×
pajamas	nightgown	worn_to_bed	○	○	◎
shield	catapult	used_for_protection	△	×	×
paintbrush	broom	used_for_painting	◎	○	◎
bluejay	sparrow	is_blue	◎	○	○
broccoli	califlower	is_green	◎	◎	◎
catapult	shield	used_for_throwing	△	×	×
peg	cap(ボトルの)	made_of_wood	×	×	×
peg	bin_(waste)(ごみ置き場)	made_of_wood	×	×	×
bluejay	starling(ムクドリ)	is_blue	◎	○	○

5.5 手法4の評価例

手法4(票数なしの属性の cos 類似度で単語組と乱数で属性選出)の被験者実験の評価例を表5.20から表5.24に示す.

表 5.20: 一人目の手法4での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
partridge	pheasant	live_in_pear_trees	×	×	×
budgie	parakeet	is_green	○	○	△
parakeet	budgie	is_tropical	×	×	×
starling	blackbird	is_small	×	×	×
blackbird	starling	good_chirps	×	×	×
perch(スズキ)	trout(マス)	is_small	×	×	×
trout	perch	is_smelly(川魚特有の臭み)	×	×	×
starling	finch(アトリ科)	is_black	×	×	×
finch	starling	is_yellow	△	×	×
drapes(ヒラヒラした仕切り布)	curtains	used_for_insulation	×	×	×

表 5.21: 二人目の手法4での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
partridge	pheasant	live_in_pear_trees	×	×	×
budgie	parakeet	is_green	×	×	×
parakeet	budgie	is_tropical	×	×	×
starling	blackbird	is_small	×	×	×
blackbird	starling	good_chirps	×	×	×
perch(スズキ)	trout(マス)	is_small	×	×	×
trout	perch	is_smelly(川魚特有の臭み)	○	○	△
starling	finch(アトリ科)	is_black	×	×	×
finch	starling	is_yellow	△	×	×
drapes(ヒラヒラした仕切り布)	curtains	used_for_insulation	×	×	×

表 5.22: 三人目の手法 4 での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
partridge	pheasant	live_in_pear_trees	○	×	×
budgie	parakeet	is_green	○	△	△
parakeet	budgie	is_tropical	×	○	×
starling	blackbird	is_small	△	×	×
blackbird	starling	good_chirps	○	△	△
perch(スズキ)	trout(マス)	is_small	×	×	×
trout	perch	is_smelly(川魚特有の臭み)	△	×	×
starling	finch(アトリ科)	is_black	◎	△	△
finch	starling	is_yellow	◎	△	△
drapes(ヒラヒラした仕切り布)	curtains	used_for_insulation	×	×	×

表 5.23: 四人目の手法 4 での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
partridge	pheasant	live_in_pear_trees	△	△	△
budgie	parakeet	is_green	◎	◎	◎
parakeet	budgie	is_tropical	△	△	△
starling	blackbird	is_small	○	○	○
blackbird	starling	good_chirps	○	○	○
perch(スズキ)	trout(マス)	is_small	○	○	○
trout	perch	is_smelly(川魚特有の臭み)	○	○	○
starling	finch(アトリ科)	is_black	◎	◎	◎
finch	starling	is_yellow	△	△	△
drapes(ヒラヒラした仕切り布)	curtains	used_for_insulation	○	△	△

表 5.24: 五人目の手法 4 での実験結果の例

引かれる単語 A	引く単語 B	類推される属性 C	A-B	A-C	B+C
partridge	pheasant	live_in_pear_trees	△	△	△
budgie	parakeet	is_green	○	○	○
parakeet	budgie	is_tropical	△	△	△
starling	blackbird	is_small	○	△	△
blackbird	starling	good_chirps	△	△	△
perch(スズキ)	trout(マス)	is_small	×	×	×
trout	perch	is_smelly(川魚特有の臭み)	○	○	○
starling	finch(アトリ科)	is_black	○	△	△
finch	starling	is_yellow	○	△	△
drapes(ヒラヒラした仕切り布)	curtains	used_for_insulation	○	○	○

第6章 考察

6.1 類推全体における手法ごとの精度比較

4種類の手法を用いて類似単語組と属性を選出し、3種類の類推を評価した。3種類の類推全体において「類推可能」と「大方類推可能」を正解とする場合の平均正解率を比較すると精度の良さは手法3が0.34, 手法1が0.31, 手法4が0.19, 手法2が0.18の順番になった。また、「類推可能」と「大方類推可能」と「見方によっては類推可能」を正解とする場合の平均正解率を比較すると、精度の良さは手法3が0.56, 手法1が0.51, 手法4が0.40, 手法2が0.34の順番になった。これによって類似単語組と属性の選出する時の属性票数の使い方によって精度に違いが発生することが確認できた。

類推全体では手法3が最も精度が高かったが、類推「単語B」+「属性C」→「単語A」では手法1の方が精度が高かった。これにより属性の選出には票数比率を使った方が良いことが分かった。手法1と手法3は優劣がつかないため手法1と手法3で用い方の異なる方法である「類似単語組の選出に属性票数を使う方法」を使うか否かは状況次第であると分かった。

6.2 類推ごとの比較

6.2.1 類推「単語A」-「単語B」→「属性C」の考察

類推「単語A」-「単語B」→「属性C」の類推精度は手法3の「類推可能」と「大方類推可能」を正解とする場合の正解率が0.43, 「類推可能」と「大方類推可能」と「見方によっては類推可能」を正解とする場合の正解率が0.69, 手法1の「類推可能」と「大方類推可能」を正解とする場合の正解率が0.37, 「類推可能」と「大方類推可能」と「見方によっては類推可能」を正解とする場合の正解率が0.55, 手法4の「類推可能」と「大方類推可能」を正解とする場合の正解率が0.29, 「類推可能」と「大方類推可能」と「見方によっては類推可能」を正解とする場合の正解率が0.43, 手法2の「類

推可能」と「大方類推可能」を正解とする場合の正解率が0.26, 「類推可能」と「大方類推可能」と「見方によっては類推可能」を正解とする場合の正解率が0.47となった.

つまり精度の良さは手法3, 手法1の順番は確定で, 手法2と手法4は評価基準によって順番が入れ替わった.

6.2.2 類推「単語 A」 - 「属性 C」 → 「単語 B」の考察

類推「単語 A」 - 「属性 C」 → 「単語 B」の類推精度は手法3の「類推可能」と「大方類推可能」を正解とする場合の正解率が0.28, 「類推可能」と「大方類推可能」と「見方によっては類推可能」を正解とする場合の正解率が0.48, 手法1の「類推可能」と「大方類推可能」を正解とする場合の正解率が0.25, 「類推可能」と「大方類推可能」と「見方によっては類推可能」を正解とする場合の正解率が0.45, 手法4の「類推可能」と「大方類推可能」を正解とする場合の正解率が0.10, 「類推可能」と「大方類推可能」と「見方によっては類推可能」を正解とする場合の正解率が0.38, 手法2の「類推可能」と「大方類推可能」を正解とする場合の正解率が0.09, 「類推可能」と「大方類推可能」と「見方によっては類推可能」を正解とする場合の正解率が0.10となった.

つまり精度の良さは手法3, 手法1, 手法4, 手法2となった.

6.2.3 類推「単語 B」 + 「属性 C」 → 「単語 A」の考察

類推「単語 B」 + 「属性 C」 → 「単語 A」の類推精度は手法1の「類推可能」と「大方類推可能」を正解とする場合の正解率が0.31, 「類推可能」と「大方類推可能」と「見方によっては類推可能」を正解とする場合の正解率が0.54, 手法3の「類推可能」と「大方類推可能」を正解とする場合の正解率が0.31, 「類推可能」と「大方類推可能」と「見方によっては類推可能」を正解とする場合の正解率が0.49, 手法4の「類推可能」と「大方類推可能」を正解とする場合の正解率が0.17, 「類推可能」と「大方類推可能」と「見方によっては類推可能」を正解とする場合の正解率が0.40, 手法2の「類推可能」と「大方類推可能」を正解とする場合の正解率が0.19, 「類推可能」と「大方類推可能」と「見方によっては類推可能」を正解とする場合の正解率が0.46となった.

つまり精度の良さは手法1, 手法3の順番は確定で, 手法2と手法4は評価基準によって順番が入れ替わった.

6.3 人手評価の正解例の考察

表 5.15 から表 5.19 を見ると人によって評価はかなりバラつきがあった。例えば類似単語組から類推された属性が色であるときに「ブロッコリー」から「カリフラワー」を引いて「緑色」が類推された時には全員◎の評価をしているが「アオカケス」から「スズメ」を引いて「青色」が類推された時には◎で評価している人も○で評価をしている人もいた。

「アオカケス」から「スズメ」を引いて「青色」が類推された時には類推 A-C と類推 B+C は類推 A-B よりも評価が良くなることは少なく、A-B は他の類推でも比較的他の手法よりも評価が良かった。

6.4 人手評価の不正解例の考察

正解例と同様に表 5.15 から表 5.19 を見ると不正解の例でもやはりバラつきがあった。例えば「パジャマ」から「ナイトガウン」を引いて「寝る時に着る」が類推された時には◎から×まで全部使われて評価されていた。これは、「ナイトガウン」が日本人にあまり馴染みのない単語であるため評価が分かれたと考える。

しかし、「ペグ」から「ボトルのキャップ」を引いて「木製」のような明らかに違う類推は、全員一致で×になった。この類推のような、そもそも類似していない単語組が選出されることも精度低下の要因の一つと考える。

以上より、類推の評価は容易ではない。

6.5 先行研究との比較

小野の研究 [5] では属性の拡大と縮小による類推、須ヶ牟田の研究 [6] では属性の結合による類推が行われて、さらに本研究では似た単語同士における属性の差を用いて間接的に単語から属性が類推されるという先行研究では行えなかった類推が可能になった。

表 5.3 の B+C の手法 4 は小野の研究 [5] の手法とほぼ同じである。手法 1 と手法 3 はこの手法 4 に対して属性の選出に票数比率を使うという改善を加えた手法である。そのため先行手法と提案手法は比較することができる。先行研究では「類推可能」と「大方類推可能」を正解とする場合の正解率は 0.21 であった。本研究での B+C の手法 1 と手法 3 の「類推可能」と「大方類推可能」を正解とする場合の正解率は 0.31 であった。これは先行研究よりも 0.1 大きい値である。これにより提案手法の属性票数は精度向

上の役に立つことが分かる．また提案手法が先行研究の手法よりも優れていることが分かる．

第7章 今後の課題

本研究では、概念属性データベースからなる McRae のデータセットから引き算を用いた類推を行ったが、いくつかの問題が残っている。以下に今後の課題を示す。

- 類推精度の向上

手法3の類推「単語A」－「単語B」→「属性C」の「類推可能」と「大方類推可能」と「見方によっては類推可能」を正解とする場合の正解率は0.69と比較的高くなった。しかし、「単語A」－「単語B」→「属性C」の「類推可能」と「大方類推可能」を正解とする場合の正解率は0.43でとどまっている。

また類推「単語A」－「単語B」→「属性C」, 「単語A」－「属性C」→「単語B」, 「単語B」＋「属性C」→「単語A」の3種類では「類推可能」と「大方類推可能」と「見方によっては類推可能」を正解とする場合の平均正解率は0.56でとどまっている。そのため、これらの類推の精度向上を目指したい。

- McRae のデータセット以外のデータセットの活用

本研究では、McRae のデータセット [4] を概念属性データベースとして研究を行ったが、このデータセット以外のデータベースを使用しなかった。McRae のデータセット [4] 以外のデータベースの使用など、様々な条件を今後検討したい。

- 類推の考え方の追加

先行研究や本研究で行われた類推はデータセットに無い発想を行うのは難しい。これを解決できる新しい類推方法を追加できればより新規性のある発想が可能になるかもしれない。

- 評価について

発想や類推を評価するのは明確な基準がないため人手評価しか評価する術がない。評価基準が今よりも明確なものがあれば評価が適切に行えるかもしれない。

第8章 おわりに

本研究では、概念属性データベースを用いて発想支援を行った。概念属性データベースは、ある単語を表す概念を属性としたデータベースである。先行研究でも用いられた McRae のデータセット [4] を概念属性データベースとする。概念属性データベースから似た単語同士において属性の類推を行って、「単語 A」 - 「単語 B」 → 「属性 C」, 「単語 A」 - 「属性 C」 → 「単語 B」, 「単語 B」 + 「属性 C」 → 「単語 A」 の 3 種類の類推に 4 種類の手法を用いて類推し、類推方法の追加を目指した。

手法 1 は類似単語組の選出に属性票数を使い、属性の選出に票数比率を使う。手法 2 は類似単語組の選出には属性票数を使うが、属性の選出は乱数で行う。手法 3 は類似単語組の選出に属性票数を使わないで、属性の選出には票数比率を使う。手法 4 は類似単語組の選出に属性票数を使わないで、属性の選出には乱数を使う。これらの 4 種類の手法を用いて類似単語組と属性を選出し、3 種類の類推を評価した。「単語 A」 - 「単語 B」 → 「属性 C」, 「単語 A」 - 「属性 C」 → 「単語 B」, 「単語 B」 + 「属性 C」 → 「単語 A」 の 3 種類の類推全体において「類推可能」と「大方類推可能」を正解とする場合の平均正解率を比較すると精度の良さは手法 3 が 0.34, 手法 1 が 0.31, 手法 4 が 0.19, 手法 2 が 0.18 の順番になった。また、「類推可能」と「大方類推可能」と「見方によっては類推可能」を正解とする場合の平均正解率を比較すると、精度の良さは手法 3 が 0.56, 手法 1 が 0.51, 手法 4 が 0.40, 手法 2 が 0.34 の順番になった。これによって類似単語組と属性の選出する時の属性票数の使い方によって精度に違いが発生することが確認できた。

類推全体では手法 3 が最も精度が高かったが、類推「単語 B」 + 「属性 C」 → 「単語 A」では手法 1の方が精度が高かった。これにより属性の選出には票数比率を使った方法が良いことが分かった。しかし、手法 1 と手法 3 は優劣がつかないため手法 1 と手法 3 で使い方の異なる方法である「類似単語組の選出に属性票数を使う方法」を使うか否かは状況次第であると分かった。

小野の研究 [5] では属性の拡大と縮小による類推、須ヶ牟田の研究 [6] では属性の結合による類推が行われて、さらに本研究では似た単語同士における属性の差を用いて間接

的に単語から属性が類推されるという先行研究では行えなかった類推が可能になった。

類推「単語 A」－「単語 B」→「属性 C」の類推では「類推可能」と「大方類推可能」を正解とする場合の正解率が 0.43, 「類推可能」と「大方類推可能」と「見方によっては類推できる」を正解とする場合の正解率が 0.69 で比較的精度の良い類推となった。

類推「単語 B」＋「属性 C」→「単語 A」の手法 4 は小野の研究 [5] の手法とほぼ同じである。手法 1 と手法 3 はこの手法 4 に対して属性の選出に票数比率を使うという改善を加えた手法である。そのため先行手法と提案手法は比較することができる。先行研究では「類推可能」と「大方類推可能」を正解とする場合の正解率は 0.21 であった。本研究での類推「単語 B」＋「属性 C」→「単語 A」の手法 1 と手法 3 の「類推可能」と「大方類推可能」を正解とする場合の正解率は 0.31 であった。これは先行研究よりも 0.1 大きい値である。これにより提案手法の属性票数は精度向上の役に立つことが分かる。また提案手法が先行研究の手法よりも優れていることが分かる。先行研究も本研究も McRae のデータセット [4] に依存しているため、このデータセット以外のデータベースの使用など、様々な条件を今後検討したい。

謝辞

研究の進め方や本論文の書き方など、細部にわたる御指導を頂きました、鳥取大学工学部知能情報工学科自然言語処理研究室の村田真樹教授に心から御礼申し上げます。また、本研究を進めるにあたり、御指導、御助言を頂きました、村上仁一准教授に心から御礼申し上げます。その他様々な場面で御助言を頂いた自然言語処理研究室の皆様には感謝の意を表します。

参考文献

- [1] Osborn, Alex F. (1957) Applied imagination : principles and procedures of creative problem-solving. New York : Scribner
- [2] 梅村雄貴, アイデア連想のつながりを明示化するブレインライティング支援システム, 修士論文, 北陸先端科学技術大学院大学先端科学技術研究科, 2019.
- [3] 國藤進, 発想支援システムの研究開発動向とその課題, 北陸先端科学技術大学院大学, 人工知能学会誌, 8 巻, pp.552-559, 1993.
- [4] Ken McRae, George S Cree, Mark S Seidenberg, and Chris McNorgan, Semantic feature production norms for a large set of living and nonliving things. Behavior Research Methods, 37, pp.547-559, 2005.
- [5] 小野正稔, 概念属性データベースを利用した発想支援の研究, 卒業論文, 鳥取大学工学部電気情報系学科, 2020.
- [6] 須ヶ牟田稔, 概念の結合に基づく類推を利用した発想支援, 卒業論文, 鳥取大学工学部電気情報系学科, 2021.