

# シソーラスと大規模テキストを用いた汎化推論規則の導出

伊東 直弘<sup>†</sup> 吉永 直樹<sup>†,††</sup> 鍛冶 伸裕<sup>†,††</sup> 豊田 正史<sup>††</sup>

<sup>†</sup> 東京大学大学院情報理工学系研究科 〒113-8654 東京都文京区本郷 7-3-1

<sup>††</sup> 東京大学生産研究所 〒153-8505 東京都目黒区駒場 4-6-1

<sup>†††</sup> 独立行政法人情報通信研究機構 〒184-8795 東京都小金井市貫井北町 4-2-1

E-mail: †{naohiro\_ito,ynaga,kaji,toyoda}@tkl.iis.u-tokyo.ac.jp

あらまし 計算機による言語理解には多様な世界知識が必要であり、ある事象が成立するときに別の事象が成立しうるのかを推論規則の形で獲得し、知識として蓄積しておくことは重要である。現在、推論規則獲得に関する研究の多くでは、テキスト中に出現する具体的な事例に基づいて推論規則が獲得されている。しかし、このアプローチで獲得できる推論規則は、テキストに実際に書かれている事象に限定されるため、実世界の多種多様な事象の組み合わせをカバーすることは困難である。そこで本研究では、この問題を解決するために、大規模テキストから獲得した事例に基づく推論規則をシソーラスを用いて汎化する手法を提案する。提案手法では、推論規則中に含まれる名詞をより抽象的な概念に汎化するが、この際、名詞をどの概念まで汎化するかを適切に決定する必要がある。そこで、汎化先となる上位概念の下位にある概念の網羅性と多様性に基づく素性を導入し、これらを教師あり学習を用いて適切に組み合わせることで汎化の程度を自動的に決定する。実験では、大規模ブログ記事から抽出した事例に基づく推論規則に対して提案手法を適用し、汎化推論規則を得た。この汎化推論規則からシソーラスを介して単語に展開して得られる推論規則の数は2,774,510となり、汎化前の推論規則の約122倍であった。15件の動詞対の汎化推論規則をランダムサンプルして適合率を計算したところ、その適合率は、約84%であった。

キーワード 知識獲得, 推論規則, 汎化, シソーラス

## 1. はじめに

我々は、実生活において、ある事象が成立するときに、別の事象が成立しうるかを推論しながら思考を行っている。例えば、「大臣が賄賂をもらう」という事象が成立するとき、「その大臣は逮捕され」さらに「辞任する」という事象が成立すると推論し、そこから「次の大臣は誰になるだろうか」と思考を展開する。しかし、我々にとっては容易と思われるこの推論も、計算機にとっては容易でない。計算機上でこの推論を行うことを可能にするためには、こうした2つの事象間の関係を知識として蓄積しておくことが重要である。

この「ある事象が成立するときに、(蓋然的に)別の事象が成立する」という知識は、推論規則と呼ばれる。本研究で扱う推論規則の具体例としては、「インフルエンザに罹る ⇒ 咳が出る」のような因果関係や「彼を殺す ⇒ 彼が死ぬ」のような含意関係などがある。

推論規則のような莫大に存在する世界知識は、人手で整備することが現実的でないため、実世界に関する記述である大規模なウェブテキストから自動獲得する研究が広く行われている[1][2]。しかし、そのようなアプローチでは、テキスト中に実際に書かれた事象を対象とするため、得られた推論規則の網羅性が低く、固有名詞を含む事象や出現頻度が低い語句を含む事象など実世界で起こり得る多種多様な事象を網羅することが難しい。

そこで本研究では、テキストから獲得した推論規則の網羅性

を向上させることを目的とし、シソーラスを用いて推論規則中に含まれる語を上位概念に置き換えることで、従来の事例に基づく推論規則を汎化した汎化推論規則を導出することを目指す。このように各語を汎化することにより、シソーラスを介して単語に展開することで、(事例に基づく)推論規則を獲得し、推論規則の網羅性を向上させることができる。

次に、事例に基づく推論規則のどの部分を汎化するかを述べる。推論規則とは、大きく述語とそれに係る項から形成される。上の「インフルエンザに罹る ⇒ 咳が出る」の例では、項が「インフルエンザに」「咳が」であり、述語が「罹る」「出る」である。推論規則を汎化するには、項と述語の両方が汎化の対象となりうる。本研究では、項と述語のうち項の方が固有名詞を含むなど推論規則を適用する文脈を強く制限すると考え、項のみを汎化の対象とする。

推論規則に含まれる項を汎化する際に、シソーラス中の任意の上位概念が汎化の結果の候補となり得るが、そのどの段階の概念まで汎化すべきかを適切に決定する必要がある。この課題を解決するために、本研究では、教師あり学習に基づく分類器を利用し、各項をその上位概念に汎化できるかどうかを順次判定する。分類器の素性としては、汎化先となる上位概念の下位にある概念の網羅性と多様性に基づく素性を用いる。

実験では、大規模ブログ記事コーパスから取得された事例に基づく推論規則の集合に対して、提案手法を適用して汎化処理を行い、汎化推論規則を導出する。その結果を人手で評価し、汎化推論規則の精度と汎化推論規則としてカバーできる推論

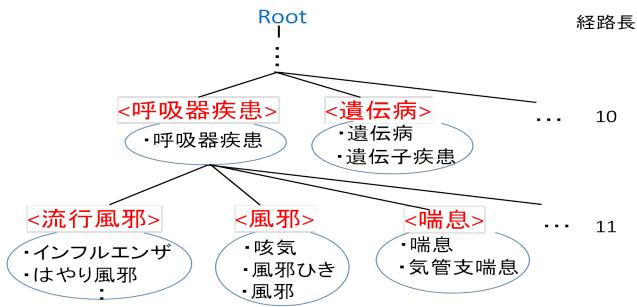


図1 シソーラスの構造

規則数の増加量と精度を測ることで、提案手法の有効性を検証する。

本稿の構成は以下の通りである。まず2章では、本研究で用いる用語の定義や知識について確認する。3章で提案手法について述べ、4章ではその評価を行う。次に、5章では関連研究を紹介し本研究との違いを述べる。最後に6章でまとめと今後の課題について述べる。

## 2. 前提知識

本研究で扱う推論規則とは、「ある事象が成立するとき別の事象が成立しうる」という規則であり、「事象 ⇒ 事象」の形で表す。ここで各事象は、動詞（述語）と、その述語に格助詞を介し直接係る名詞（項）から成るとする。このとき、述語を修飾する項の順番は区別しない。

本研究で扱う「事例に基づく推論規則」とは、テキストから直接獲得され、テキスト中で実際に書かれた事象を対象とする推論規則のことである。その「事例に基づく推論規則」に含まれる項の名詞を上位概念に汎化した推論規則を「汎化推論規則」と呼ぶ。

次に、本研究で推論規則を汎化する際の手がかりとして用いるシソーラスの構造について、特に代表的なシソーラスである WordNet を例にして説明する。シソーラスとは概念の上位、下位関係をまとめた辞書のことであり、図1のように類義語句の集合として概念が定義され、それらの概念間には、上位下位関係に基づく木構造が定義されている。WordNet では本来概念のノードが synset の id で表されるが、本論文では分かりやすさのため、類義語句の中から代表して一語を選び、<概念>と表記する。例えば、<インフルエンザ>などと表記する。また、シソーラスの最上位のノードのことを便宜上 Root と呼ぶ。本稿では、ある概念が Root からどのくらい離れているかという経路長を定義し、手法に利用する(図1中の10, 11)。

## 3. 提案手法

### STEP0: 事例に基づく推論規則の獲得

乾[1]らの手法を参考にして、手がかり表現をもとにテキストから事例に基づく推論規則の集合を獲得する。

得られた事例に基づく推論規則に対して、2つのSTEPからなる推論規則の汎化手法を提案する。なお提案手法の入力とする事例に基づく推論規則は動詞の対が共通する推論規則（罹る

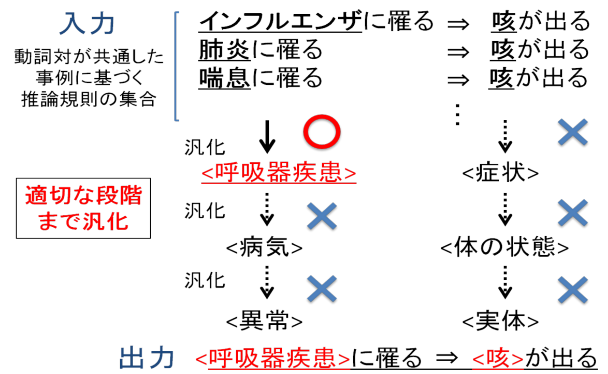


図2 手法の概要

⇒ 出る、など) ごとにまとめて処理を行う。

### STEP1: シソーラスへの項の紐づけ

事例に基づく推論規則に含まれる項の名詞をシソーラス上で対応する概念に紐づける。

### STEP2: 分類器を用いた項の汎化

次に、シソーラスを利用して、その概念が上位概念に汎化できるかどうかを機械学習の分類器を用いて判定しながら1段階ずつ上の概念に汎化する。

STEP2では、分類器により、ある概念がそれ以上汎化ができないと判定されるまで汎化を繰り返す。事例に基づく推論規則中の全ての項に関して、それ以上汎化ができないと判断された時点で処理を終了し、汎化推論規則を出力する。分類器の素性としては、汎化先となる上位概念の下位にある概念の網羅性と多様性に基づく素性を用いる。

例を挙げると、図2のように、入力として、テキストから獲得した「インフルエンザに罹る ⇒ 咳が出る」「肺炎に罹る ⇒ 咳が出る」「喘息にかかる ⇒ 咳が出る」のような事例に基づく推論規則の集合をとる。STEP1として、その事例に基づく推論規則の中の項の名詞であるインフルエンザ、肺炎、喘息や咳をそれぞれ、シソーラス中の対応する概念に紐づける。

次に、STEP2でシソーラスを用いて上位の概念に汎化する。その際に、<インフルエンザ>、<肺炎>、<喘息>の概念を<呼吸器疾患>という概念に汎化できるか、さらには<病気>という概念まで汎化できるか、<異常>まで汎化できるか等を順に分類器を用いて判定していき、汎化できないと判断された時点で処理を終了する。

今回の例の場合は、インフルエンザ、肺炎、喘息を<呼吸器疾患>まで汎化でき、咳は、<咳>という概念より上位概念に汎化することができないと分類器が正しく判断できればこれにより、「<呼吸器疾患>に罹る ⇒ <咳>が出る」という汎化推論規則が出力される。

以下、手法について詳しく説明をする。

#### 3.1 STEP0: 事例に基づく推論規則の獲得

本研究では、乾ら[1]の手法を参考にして事例に基づく推論規則の獲得を行う。具体的には、手がかり表現として接続標識「の」「ば」を利用してテキストから推論規則を取り出す。利用する手がかり表現は、接続される前後の事象の間で、ある程

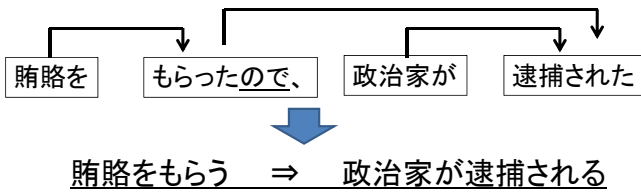


図3 コーパスからの事例に基づく推論規則の獲得

度の必然性を保持していることが必要となる。「ので」や「ば」は、多くの場合、接続される事象の間に必然性を持つ関係が成り立っているため、手がかり表現として相応しいと考えられる。

上記の手がかり表現により結ばれる2つの述語を抽出し、その述語に格助詞を介して係る文節(項)を抽出することによって事例に基づく推論規則を得る。本研究では、代表的な格助詞である「が」、「を」、「に」を抽出対象とする。

具体例として、図3のようにテキストから事例に基づく推論規則を獲得する。

### 3.2 STEP1: シソーラスへの項の紐づけ

事例に基づく推論規則中に現れる項の名詞の中から汎化対象とするものを選別し、それらをシソーラス上の対応する概念と紐づける。ここでは、項の名詞がシソーラスに収録されていない場合は、その項は汎化の対象とせず、その項はないものとして扱う。例を挙げると、「太郎がインフルエンザに罹る ⇒ 咳が出る」という事例に基づく規則に対しては、「太郎」はシソーラスに登録されていないため、その項は汎化の対象としない。一方で、インフルエンザと咳は、それぞれ<流行風邪>、<咳>に紐づけ可能であるため、これらは汎化の対象とする。

### 3.3 STEP2: 分類器を用いた項の汎化

本節では、STEP1により項が紐づいたシソーラス上のノードからボトムアップで上の概念に汎化を行う際に、具体的にどのような汎化処理を行うのかを説明する。推論規則中に含まれる汎化対象となる項の名詞が1つの場合と複数の場合で汎化の処理が異なるため、それぞれ分けて説明する。

含まれる項の名詞が1つである場合は、シソーラス上でRootからの経路長が最も大きいノードから始めて、その汎化レベルにあるノードを1段階ずつ上の概念に汎化を行う。その際に、汎化できるかを順次判定していき、ある概念からその上位の概念に汎化できなくなったときに、この汎化の処理を終了し、その時点までに汎化できた汎化推論規則を出力する。

一方、含まれる項の名詞が複数である場合については、まず2章で定義をしたRootからの経路長が一番長い概念に着目して、汎化処理を行い、汎化できなくなるまで汎化を行ったのち、次にシソーラス中での経路長が長い項に着目して同様の処理を繰り返す。Rootからの経路長が最も長く、最も深い項の概念から汎化を始める理由は、できるだけ深いノードの概念の方が汎化できる可能性が高く、安全に汎化できると考えられるためである。

経路長が一番長い概念から始めて1つずつ上の概念に汎化するが、その際にそれ以外の項は固定しておく。汎化処理の対象となっている名詞がそれ以上、上の概念に汎化できなくなった

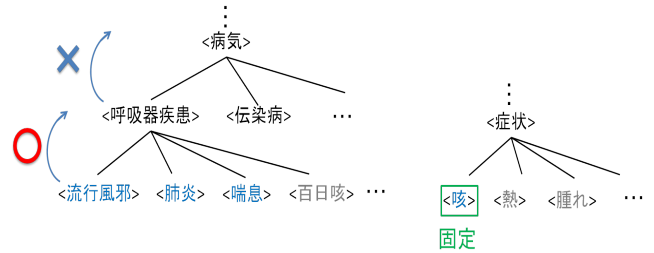


図4 項が2つのときの汎化処理1

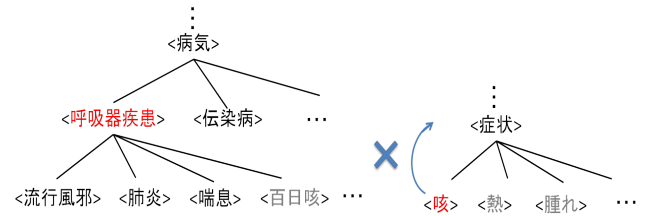


図5 項が2つのときの汎化処理2

ときに、その次に経路長が長い項の名詞に対する汎化処理を行う。同様に、全ての名詞に対して順番に行う。含まれる全ての項で汎化がそれ以上できないと判定された時点で汎化推論規則を出力する。導出された汎化推論規則にシソーラス上で含まれない(事例に基づく)推論規則に関しては、STEP1でシソーラスに紐づけられている全ての項の名詞がなくなるまで同様の処理を再帰的に繰り返す。これにより、特定の動詞に対して、複数の汎化推論規則が導出され得ることに注意されたい。

具体的に、項の名詞が2つのときの汎化の処理の流れを下図4および5に示す。図4のように、一番深い項の概念(<流行風邪>、<肺炎>、<喘息>)から汎化処理を始める。その際、他の項の概念(<咳>)は固定する。<流行風邪>、<肺炎>、<喘息>が<呼吸器疾患>に汎化できるかどうかを検討する。この場合、<呼吸器疾患>に汎化できるため、次に、<呼吸器疾患>と<伝染病>が<病気>に汎化できるかどうかを検討する。この場合、<病気>に汎化すると、「<病気>に罹る ⇒ <咳>が出る」という汎化推論規則が得られるが、全ての病気が咳が出るとは限らないため、<病気>まで汎化するのは不適切である。<病気>に汎化できないと判定されたときに、その時点までに汎化できていた概念である<呼吸器疾患>に固定をして、図5のように残りの項の概念を汎化する。<咳>という概念から<症状>に汎化できないと判定されたときに、「<呼吸器疾患>に罹る ⇒ <咳>が出る」という汎化推論規則を出力する。

#### 3.3.1 汎化の可否の判定に用いる4つの素性

我々は、シソーラス上である概念からその上位概念に汎化できるかを自動的に決定するために、汎化の可否を判定する素性を導入し、これらを教師あり学習に基づく分類器で組み合わせることで汎化の可否の判定を行う。ある概念(汎化元の概念)をその上位概念(汎化先の概念)に汎化できるかを判定する場合を考える。以下、導入した4つの素性についてそれぞれ詳しく説明する。

## 汎化元の概念の網羅率

汎化元の概念を汎化先の概念に汎化できるかどうかを判定するために、その汎化先の概念の直下にある汎化元の概念がより多く観測されるほど汎化しやすいと考えられる。その汎化元の概念に着目をして、次の式のように指標を定義する。

$$score_{網羅率} = \frac{\langle \text{実際に推論規則に出現した下位概念の数} \rangle}{\langle \text{シソーラスに収録されている下位概念の数} \rangle}$$

この式は、コーパス中の推論規則が汎化元の概念をどのくらい網羅できているかを数値化している。例えば、〈流行風邪〉、〈肺炎〉、〈喘息〉を〈呼吸器疾患〉に汎化できるかを判定する際に、汎化先の概念である〈呼吸器疾患〉がシソーラス上でその下位概念として〈流行風邪〉、〈肺炎〉、〈喘息〉の他に、〈百日咳〉、〈気管支炎〉などの合計 14 種類の概念を収録しているとすると、推論規則に実際に現れた汎化元の概念は、〈流行風邪〉、〈肺炎〉、〈喘息〉の 3 種類であるとする。このとき、指標は  $score_{網羅率} = \frac{3}{14}$  となる。

## 汎化元の概念の頻度の偏り

汎化元の概念から汎化先の概念に汎化を行う際に、出現する汎化元の概念の頻度が偏り過ぎている場合、汎化を行うことが適切でない可能性があると考えられる。そのため、コーパスに現れた汎化元の概念の頻度の偏りを反映した指標を提案する。ある概念の汎化元の概念が何回出現したかを観測し、そのエントロピーを以下の式で計算する。

$$score_{偏り} = \sum_{i=1}^n -p_i \log_2 p_i$$
$$p_i = \frac{\langle \text{その概念の出現数} \rangle}{\langle \text{汎化元の概念の総出現数} \rangle}$$

具体的には、〈流行風邪〉、〈肺炎〉、〈喘息〉が〈呼吸器疾患〉に汎化できるかを判定する際に、「〈流行風邪〉に罹る ⇒ 〈咳〉が出る」という推論規則がテキスト中に 4 回現れていたとする。同様に、〈肺炎〉〈喘息〉を含む推論規則もそれぞれ、6 回、2 回現れているとすると、その際の出現確率は、〈流行風邪〉は  $\frac{1}{3}$ 、〈肺炎〉は  $\frac{1}{2}$ 、〈喘息〉は  $\frac{1}{6}$  となる。これをもとにエントロピーを計算すると、

$$score_{偏り} = -\frac{1}{3} \log_2 \left( \frac{1}{3} \right) - \frac{1}{2} \log_2 \left( \frac{1}{2} \right) - \frac{1}{6} \log_2 \left( \frac{1}{6} \right)$$

となる。

## 汎化先の概念の Root からの経路長

シソーラスの構造上の特徴から、最上位の概念である Root からその概念のノードが離れていけば離れているほど、汎化が起きやすく、最上位の Root に近い概念ほど汎化が起きにくいことが期待される。例えば、〈流行風邪〉や〈肺炎〉を〈呼吸器疾患〉にするなど、Root から離れた概念は汎化されやすいと考えられる。

経路長はシソーラスの Root ノードからその概念までの距離で定義した。経路長が大きいほど汎化しやすいように指標を設定するため、具体的に経路長が 7 以上の場合は、指標の値を

1.0 にし、それより小さい場合は指標の値を 0 とした。その境界に経路長 7 の値を採用する理由としては、実際に、獲得した事例に基づく推論規則を分析すると、経路長が 7 を基準にして汎化の可否に違いがあることが確認されたからである。

## 汎化先の概念の類義語を含む推論規則の有無

汎化先の概念の類義語を含む推論規則が既にテキストから獲得されている場合、その汎化先の概念への汎化が起りやすいことが期待される。

例えば、「〈流行風邪〉に罹る ⇒ 〈咳〉が出る」という推論規則中の〈流行風邪〉という汎化元の概念が、〈呼吸器疾患〉という汎化先の概念に汎化できるかどうか、すなわち「〈呼吸器疾患〉に罹る ⇒ 〈咳〉が出る」という汎化推論規則が得られるかどうかを検討する際に、既に汎化先の概念である〈呼吸器疾患〉という概念の類義語を含む推論規則である「呼吸器疾患に罹る ⇒ 咳が出る」のような（事例に基づく）推論規則が既にテキストから獲得できているならば、〈流行風邪〉から〈呼吸器疾患〉へ汎化できる可能性が大きいと判定できる。具体的には、実際のコーパスで名詞が現れている場合に指標の値を 1.0 に、現れていない場合は指標を 0 と定義した。

これらの 4 つの指標を素性として組み合わせることで、項の汎化の可否を判定する分類器 (SVM) を学習する。

## 4. 評価実験

本章では、前章で述べた提案手法を実際にテキストから獲得した事例に基づく推論規則の集合に対して適用することで、汎化推論規則の導出を行う。さらに、得られた汎化推論規則について人手で評価を行い、その結果を考察する。

推論規則を獲得する知識源としては、我々の研究室で 2006 年 2 月から 2012 年 12 月まで収集した 7 年分のブログ記事コーパス、約 2 億記事、約 20 億文を対象にした。係り受け解析には、Yoshinaga らが開発した J.DepP [5] を利用した。

3.1 節の手順で手がかり表現を利用して、事例に基づく推論規則の集合を獲得した。その際に、ノイズを除くために、一定頻度以上現れている規則を扱うことにした。具体的には、2 回以上コーパス中に出現した事例に基づく推論規則を獲得した。事例に基づく推論規則の出現頻度の計測には、FREQT を利用した<sup>(注1)</sup> [4]。

実際に実験では、軽動詞（ある、なる、する、など補助的な意味しか持たない動詞）を含む動詞の対を除いて出現頻度が高い上位 200 対の動詞の対に提案手法を利用した。そしてその 200 対からランダムに選んだ 15 対に対してラベル付けをし、評価を行った。

事例に基づく推論規則を汎化する際に利用するシソーラスとして、約 9 万語句を収録した日本語 WordNet [6] を利用した。

### 4.1 分類器の学習データの作成

項の汎化の可否を判定する分類器の学習データとテストデータを以下の手順で作成した。ちょうど適切な汎化ができてい

(注1) : <http://chasen.org/~taku/software/freqt/>

汎化推論規則を手で作成し、その規則を基準とすることで自動的に正例と負例を作成した。具体的には、その適切な汎化ができていない汎化推論規則よりも汎化し過ぎている汎化推論規則を負例、その汎化推論規則の概念より下の概念を持つ汎化推論規則を正例となるように自動的に分類した。

高頻度で出現している 200 対の動詞の対からランダムに選んだ 15 対を対象にデータを作成した。具体的に選んだ推論規則の動詞対は、「飲む ⇒ 治る」、「使う ⇒ 探す」、「乗る ⇒ 行く」、「使う ⇒ 疲れる」、「乗る ⇒ 着く」、「ひねる ⇒ 出る」、「残る ⇒ 食べる」、「届く ⇒ 作る」、「載る ⇒ 読む」、「使う ⇒ 買う」、「来る ⇒ 遊ぶ」、「録画する ⇒ 見る」、「届く ⇒ 読む」、「当たる ⇒ 行く」、「行く ⇒ 寄る」の 15 対である。

これにより 655 個の正例と 247 個の負例を得た。その評価用データの 7 割を学習データにし、3 割をテストデータとした。

#### 4.2 予備実験：汎化の可否を判定する分類器の精度

提案手法の STEP0 で獲得した事例に基づく推論規則のうち、高頻度で出現している前節で記した 15 対の動詞の対を対象に実験を行った。分類器により判定した汎化の可否に関しての精度を確かめる。

分類のための学習器としては、SVM を実装した LIBLINEAR [7] を用いた。その汎化が正しい（正例）か、不適切（負例）かの 2 値に分類をした。

前節の方法で、作成した学習データとテストデータに関して評価を行い、その結果、分類精度は、81.7%であった。一方でマジョリティベースラインでの分類精度は、61.7%であったため、マジョリティベースラインを上回ることができた。正例に関して、適合率は、89.6%となり、再現率は、79.5%となった。この結果から汎化の可否を判定する分類器は、一定以上の信頼性があることが確認された。

#### 4.3 実験結果

以下に提案手法の STEP1, STEP2 を適用する実験を以下の手順で行った。

3 章の STEP0 で獲得した事例に基づく推論規則に対して、各項に関して適切な汎化ができていない汎化推論規則を手で作成した。その人手で作成した正解の汎化推論規則をシソーラスを介してより下にある単語に展開することで、得られる（事例に基づく）推論規則を正解データとした。

その適切な汎化がされている汎化推論規則が包含する事例に基づく推論規則の集合を手法の入力として提案手法を用い、正解とする汎化推論規則を再現できるかどうかの実験を行った。

その事例に基づく推論規則に提案手法を用いて汎化推論規則を導出した。上記の方法で獲得した正解データの（事例に基づく）推論規則と導出した汎化推論規則を同様にシソーラスを介して単語に展開し得られる（事例に基づく）推論規則とを比較することで適合率、再現率を求めた。

##### 4.3.1 汎化により、獲得した推論規則の数の調査

200 対の動詞対の事例に基づく推論規則に対して提案手法を用いて 913 の汎化推論規則を導出した。その汎化推論規則をシソーラスを介して単語に展開することでカバーできる（事例に基づく）推論規則数を見積もった。その数を汎化処理を行う前

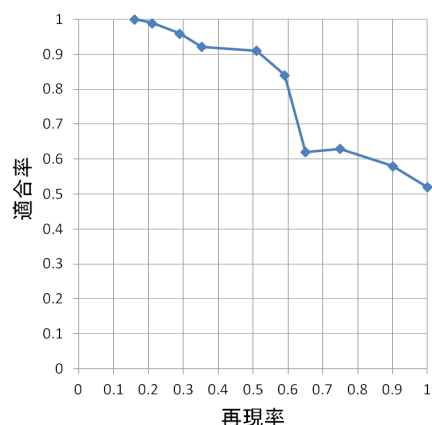


図 6 適合率-再現率曲線

の事例に基づく推論規則数と比較した。

その結果は、汎化前の（事例に基づく）推論規則数が 22,738 であるのに対し、提案手法で獲得した汎化推論規則から、2,774,510 の（事例に基づく）推論規則を獲得できた。約 122 倍の推論規則を獲得することができた。

汎化推論規則がカバーする推論規則の数の見積もり方を具体例を交えて説明する。例えば、「＜抗菌剤＞を飲む ⇒ ＜呼吸器疾患＞が治る」という汎化推論規則の場合であれば、＜抗菌剤＞の下位の概念にある語句の数を全て足し合わせた数と、＜呼吸器疾患＞の下位の概念にある語句の数を全て足し合わせた数を掛けるものが、この汎化推論規則がカバーする推論規則の数となる。

一方、汎化前の事例に基づく推論規則の数を計算する際に、事例に基づく推論規則に対しても、シソーラス上でその項の名詞が紐づく概念の類義語に関しても推論規則の数に含めた。具体例を挙げると、事例に基づく推論規則中に、「インフルエンザに罹る ⇒ 咳が出る」がある場合に、インフルエンザが属する概念である＜流行風邪＞に含まれる類義語の集合である、はやり風邪や流行性感冒などの推論規則も獲得できていることとした。

##### 4.3.2 汎化により、獲得した推論規則の精度の調査

評価データの適合率は、0.84 であり、再現率は、0.59 となった。前節で得られた約 277 万の推論規則のうち、正解は、約 233 万程度であると見積もることができる ( $277 \times 0.84$ )。

質の高い知識ベースを維持するためには、正例の適合率が高くなるような識別をすることが好ましい。そこでパラメータを調整することで、適合率と再現率のトレードオフを調整できるかどうか調査した。適合率と再現率のバランスを決定するために、パラメータを設定したときの適合率-再現率曲線は、図 6 のようになる。

この図から、適合率を 90%程度に保ったまま、約 55%程度まで再現率を上げることができることがわかった。一方、適合率を約 95%まで上げて保つ場合でも、再現率を約 30%と高い値に保つことができた。

適合率が約 85%程度から約 60%に急激に落ちた前後に得られた汎化推論規則を確認したところ、適合率が 60%に落ちたときの汎化推論規則は、正解の汎化推論規則に対して 2 段階汎

表 1 素性の有効性の検証

手法	適合率	再現率
提案手法	0.84	0.57
–下位概念の網羅率 ( $score_{網羅率}$ )	0.71	0.61
–下位概念の頻度の偏り ( $score_{偏り}$ )	0.80	0.59
–概念の Root からの経路長 ( $score_{経路長}$ )	0.63	0.78
–上位概念へ属する語句の有無 ( $score_{規則の有無}$ )	0.71	0.55

汎化し過ぎている項が複数あった。このため、大幅に適合率が下がったと考えられる。

#### 4.3.3 素性の有効性の検証

どの素性が有効かを確かめるため、分類器の素性をそれぞれ 1 つ抜いたベースラインと適合率、再現率を比較した。導出する汎化推論規則の精度に関して、3 章で提案をした 4 つある分類器の素性を 1 つずつ減らした手法をベースラインとして、提案手法の精度と比較を行った。その結果は、表 1 のようになった。

本研究では、適合率が特に重要な素性であり、その適合率の変化に着目した。汎化先の概念の Root からの経路長の素性を除いたときに、最も適合率が大きく減少したため、概念の Root からの経路長の素性が有効であると確認できた。一方で、汎化元の概念の頻度の偏りの素性を除いたときに、適合率の変化がほとんど見られなかったため、汎化先の概念の頻度の偏りの素性はあまり重要でない素性であることがわかった。

#### 4.4 分析

##### 4.4.1 汎化の適切性の評価

獲得した汎化推論規則に対して、その汎化が適切な汎化であるかどうかを評価する。提案手法により得られた汎化推論規則に対して、人手で適切な段階の汎化が行われているかどうかを分析した。

前述の 15 対の動詞対に対して得られた汎化推論規則をランダムに 50 個選び、5 つのラベルに分類をした。

まず不適切な汎化と適切な汎化を分けた。さらに不適切な汎化に関して、

(1) 汎化が不十分である

得られた汎化推論規則に、汎化が不十分な項が含まれている

(2) 汎化をし過ぎている

得られた汎化推論規則に、汎化をし過ぎている項が含まれている

(3) 汎化が不十分な項と汎化をし過ぎている項がある

得られた汎化推論規則に、汎化が不十分な項と汎化をし過ぎている項の両方が含まれている

という (1), (2), (3) の 3 つに分類した。

そして、適切な汎化に関して、

(4) 汎化が行われ、適切である

項の汎化が行われていて、推論規則として意味が成立している

(5) どの項も汎化が行われていないが適切である

どの項も汎化が行われていないが、推論規則として成立している

という (4), (5) の 2 つに分類した。

その結果を表 2 にまとめる。

表 2 汎化推論規則中の項の汎化の適切性と頻度

分類	頻度
汎化が不十分である	8
汎化をし過ぎている	2
汎化が不十分な項と汎化をし過ぎている項がある	4
汎化が行われ適切である	22
どの項についても汎化が行われていないが適切である	14

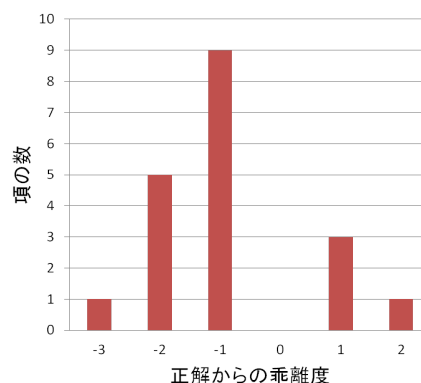


図 7 項の正解からの乖離度と項の数

項が複数ある場合、汎化が行われ、全ての項に関して適切な汎化が行われたときのみを「汎化が行われ適切である」に分類した。

さらに、「汎化が不十分である」、「汎化をし過ぎている」、もしくは「汎化が不十分な項と汎化をし過ぎている項がある」に分類された場合は、その項が適切な汎化が行われた場合の正解の概念からの乖離度を記録した。ここで乖離度が負のときは、その項の汎化が不十分であることを表し、乖離度が正のときは、その項を汎化し過ぎていることを表す (図 7)。

汎化が不適切である汎化推論規則の項の概念の正解の概念からの乖離度を図 7 で確認すると、乖離度の絶対値が 1 か 2 である項の数は、18 であり、それに対して、乖離度の絶対値が 3 以上である規則は、1 であった。乖離度の最大値が 3 であったことから大きく正解から外れている規則は少ないことが確認できた。

また、汎化が不十分な項の数が 15 であるのに対して、汎化をし過ぎた項の数は、3 であった。汎化が不十分な項の数が、汎化をし過ぎた項の数よりも多いことから、提案手法では、誤った汎化を行い汎化し過ぎることが少ないことがわかった。誤った汎化をしてしまう可能性がある場合は汎化を行わず、安全な汎化のみを行うという我々の方針に合っていることが確認できた。

##### 4.4.2 実際に得られた汎化推論規則の例

得られた汎化推論規則の一部を動詞対ごとにまとめて表 3 に載せる。表中の太字で示した汎化推論規則に関しては表 4 で、その汎化推論規則をシソーラスを介してその概念より下位にある単語に展開することで、得られた (事例に基づく) 推論規則の例を載せる。この例から、人間の直感にある程度近い推論規則が得られていたことが確認できる。

4.3.1 節で示したように、テキストからは直接獲得できない推論規則を多く獲得できていた。2 つ例を挙げると表 4 で太文

流行する ⇒ つける
<伝染病>が流行する ⇒ 気をつける
飲む ⇒ 治る
<抗菌剤>を飲む ⇒ <呼吸器疾患>が治る
水を飲む ⇒ <風邪>が治る
当たる ⇒ 行く
<ギャンブル>が当たる ⇒ <交換所>に行く
<予報>が当たる ⇒ <買い物>に行く
残る ⇒ 食べる
<料理>が残る ⇒ <食事>に食べる
<料理>が残る ⇒ 食べる
<料理>が<食器>に残る ⇒ 食べる
乗る ⇒ 行く
<輸送手段>に乗る ⇒ <町>に行く
<輸送手段>に乗る ⇒ <教育機関>に行く
<輸送手段>に乗る ⇒ <買い物>に行く
届く ⇒ 作る
<果物>が届く ⇒ <ジュース>を作る
<野菜>が届く ⇒ <カレー>を作る
<材料>が届く ⇒ テーブルを作る
行く ⇒ 寄る
<買い物>に行く ⇒ <小売店>に寄る
<気晴らし>に行く ⇒ <海>に寄る
<教育機関>に行く ⇒ 寄る
使う ⇒ 疲れる
<活力>を使う ⇒ 疲れる
<目上>に気を使う ⇒ 疲れる
<筋肉>を使う ⇒ 疲れる

<伝染病>が流行する ⇒ 気をつける
ハンセン病が流行する ⇒ 気をつける
腺熱が流行する ⇒ 気をつける
回帰熱が流行する ⇒ 気をつける
パラチフスが流行する ⇒ 気をつける
赤痢が流行する ⇒ 気をつける
リウマチ熱が流行する ⇒ 気をつける
<抗菌剤>を飲む ⇒ <呼吸器疾患>が治る
抗生物質を飲む ⇒ インフルエンザが治る
抗生物質を飲む ⇒ 気管支炎が治る
ペニシリンを飲む ⇒ 肺炎が治る
抗生物質を飲む ⇒ 百日咳が治る
抗生物質を飲む ⇒ 風邪が治る
イソニアジドを飲む ⇒ インフルエンザが治る
<料理>が残る ⇒ <食事>に食べる
ビーフストロガノフが残る ⇒ 夕飯に食べる
リゾットが残る ⇒ ブランチに食べる
パテが残る ⇒ 朝食に食べる
スナック菓子が残る ⇒ おやつに食べる
テトラツィーニが残る ⇒ 昼食に食べる
サンドイッチが残る ⇒ ピクニックに食べる
<輸送手段>に乗る ⇒ <町>に行く
2階建てバスに乗る ⇒ 町に行く
電車に乗る ⇒ 郷里に行く
おんぼろ車に乗る ⇒ ゴースタウンに行く
地下鉄に乗る ⇒ 町に行く
車に乗る ⇒ 集落に行く
急行に乗る ⇒ 町に行く

字で記した「ハンセン病が流行する ⇒ 気をつける」という推論規則や「ビーフストロガノフが残る ⇒ 夕飯に食べる」という推論規則は、我々の研究室が7年間収集したウェブテキストからは獲得することができなかった。研究室のブログテキストからは直接獲得できないが今回の提案手法により獲得できた例を見ると、ハンセン病やビーフストロガノフなどのあまりテキストに出現しないような事例に関する知識が多いことが分かった。このように推論規則を汎化することで直接テキストに書かれていない知識を獲得でき、網羅性が高い推論規則を獲得できることを確認した。

一方で、汎化推論規則をソーラスを介して適用することで、誤った知識を獲得してしまう可能性もある。例えば、表4中の「<抗菌剤>を飲む ⇒ <呼吸器疾患>が治る」という汎化推論規則の<抗菌剤>という概念をその下位の概念の語句であるイソニアジドに展開し、<呼吸器疾患>という概念を下位の概念の類義語であるインフルエンザに展開することで、「イソニアジドを飲む ⇒ インフルエンザが治る」という規則を獲得することができた。しかし、イソニアジドは、インフルエンザの治療に使われないため、この推論規則は成立していない。このように、汎化推論規則をソーラスの下位の語句に適用する際に注意が必要であることがわかる。

## 5. 関連研究

推論規則の自動獲得の研究としては、2001年のDIRT [8]の研究が先駆けとなり、その後、DIRTを改良する研究 [9] [10] が行われている。

推論規則の獲得に関する研究の多くは、テキストから直接取り出した事例に基づく推論規則を対象にしている [2] [3]。しかし、本研究のように推論規則の網羅性の向上を目指す研究もいくつかある。本章ではそのうち2つの研究を紹介し、本研究との違いについて述べる。

乾ら [11] は、潜在変数モデルを使って、因果関係知識の事象の汎化を行い、推論規則の網羅性を向上させることを目指している。潜在変数モデルに、事象間に因果関係が成り立っているかどうかというクラスラベルを導入することで、他の先行研究より、精度の高い因果関係知識を得ることができると報告している。

Melamud ら [12] は、推論規則の項をベクトルで表すことで、推論規則の汎化を行っている。彼らの手法では、ある述語の対の推論規則の項にどのような語句が来やすいかを語句ベクトルで表し、さらに項と述語の部分にLDAモデルを用いてトピックベクトルを導入した。そのようにして、その項が、その述語の前後のどのようなトピックで現れやすいかの情報を考慮することで、精度の高い推論規則を獲得している。

これらの2つの研究では、汎化できる推論規則の型が限定されている。具体的には、1つの推論規則中に項が3つ以上含まれる推論規則の汎化を扱うことができない。それに対して本研究では、項が3つ以上含まれる推論規則も扱うことができるという利点がある。例えば、本手法では、「<ペット>を<公園>に連れてく⇒<排泄物>をする」などの項が3つの汎化推論規則も獲得ができる。

また、提案手法では、シソーラスを利用して汎化を行うため、汎化の過程や汎化の結果を人が見て解釈しやすいという利点があると考える。

## 6. まとめと今後の課題

本研究では、推論規則の網羅性の向上を目的として、シソーラスを用いて、事例に基づく推論規則を項を汎化する手法を提案した。実験では、推論規則の網羅性が向上したかどうかを調べるため、汎化推論規則の項をシソーラスを介して単語に展開することで、カバーできる推論規則の増加数を分析した。汎化前の事例に基づく推論規則の数に対して汎化推論規則を拡張してカバーできるようになる(事例に基づく)推論規則の数は、約122倍に増えた。評価データの適合率は0.84と高い値を示した。また、汎化推論規則の拡張によりカバーできた推論規則の例を確認することで、テキストからは直接獲得できない推論規則を獲得できていることを確認した。

今後の課題としては、以下の2点が挙げられる。

1つ目は、シソーラスに収録されていない語句も汎化の対象とすることが挙げられる。今回、項の名詞がシソーラスに収録されていない場合、その語句を汎化の対象としなかった。今後は、シソーラスの自動拡張などを行うことによって、そうした語句に対しても汎化を行うことを検討したい。

2つ目として、他の推論規則の網羅性の向上を目指す手法との詳細な比較が挙げられる。利用する知識ベースや獲得する推論規則の形式が異なることなどから、本研究では、関連研究で紹介した。他の推論規則の網羅性を向上させる研究との比較を行わなかった。

## 文 献

- [1] 乾孝司, 乾健太郎, 松本裕治, 接続標識「ため」に基づく文書集合からの因果関係知識の自動獲得, 情報処理学会論文誌, vol.45, no.3, 2004.
- [2] T. Shibata, and S. Kurohashi, Acquiring strongly-related events using predicate-argument co-occurring statistics and case frames, Proceedings of the 5th International Joint Conference on Natural Language Processing (IJCNLP), 2011.
- [3] B. Jans, S. Bethard, I. Vulic, and M. F. Moens. Skip n-grams and ranking functions for predicting script events, Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics (EACL), 2012.

- [4] K. Abe, S. Kawasoe, T. Asai, H. Arimura and S. Arikawa, Optimized Substructure Discovery for Semi-structured Data, Proceedings of the 6th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD), 2002.
- [5] N. Yoshinaga, and M. Kitsuregawa, Kernel slicing: Scalable online training with conjunctive features, Proceedings of the 23rd International Conference on Computational Linguistics (Coling), 2010.
- [6] H. Isahara, F. Bond, K. Uchimoto, M. Utiyama and K. Kanzaki, Development of Japanese WordNet, Proceedings of the 6th edition of the Language Resources and Evaluation Conference (LREC), 2008.
- [7] R. Fan, K. Chang, C. Hsieh, X. Wang, and C. Lin, Liblinear: A library for large linear classification, Journal of Machine Learning Research, vol.9, 2008.
- [8] D. Lin, and P. Pantel, DIRT - Discovery of Inference Rules from Text, Proceedings of the seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2001
- [9] G. Dinu, and R. Wang. Inference Rules and their Application to Recognizing Textual Entailment, Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL), 2009.
- [10] J. Berant, I. Dagan, and J. Goldberger, Global Learning of Typed Entailment Rules, Proceedings of the 49th Association for Computational Linguistics (ACL), 2011.
- [11] T. Inui, H. Takamura, and M. Okumura. Latent Variable Models for Causal Knowledge Acquisition, Proceedings of the 8th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing), 2007.
- [12] O. Melamud, J. Berant, I. Dagan, J. Goldberger, and I. Szpektor. A two level model for context sensitive inference rules, Proceedings of the 51st Association for Computational Linguistics (ACL), 2013.