

# 時期依存性を有するイベント連鎖の獲得

中島 直哉<sup>†</sup> 吉永 直樹<sup>††</sup> 鍛冶 伸裕<sup>††</sup> 豊田 正史<sup>††</sup> 喜連川 優<sup>††</sup>

<sup>†</sup> 東京大学 大学院情報理工学系研究科 〒113-8656 東京都文京区本郷 7-3-1

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

E-mail: <sup>†</sup>{nakashima, ynaga, kaji, toyoda, kitsure}@tkl.iis.u-tokyo.ac.jp

あらまし 大量に蓄積されたウェブテキストは、人間の行動に関する記述を多量に保有している。これらを知識として収集し、様々な状況下で成り立つイベント連鎖を獲得することが出来れば、文脈に即した行動提示などへの応用に役立つ。例えば「窓を閉める」と「クーラーをつける」は特に夏に成立する連鎖である。このような連鎖は時期を指定した上で獲得することが望ましい。本稿では、ウェブから収集した時系列テキストを利用し、接続標識などの手がかり表現に加え、各季節におけるイベントの出現情報やイベント間の共起情報を特徴量とした機械学習により、それぞれの季節の下で成立しやすい連鎖の獲得を目指す。さらに、獲得された連鎖に対して人手で評価を行い、その精度と獲得可能な知識の量について見積もることで、提案手法の有効性を検証する。

キーワード 知識獲得, 推論規則, 時系列

Naoya NAKASHIMA<sup>†</sup>, Naoki YOSHINAGA<sup>††</sup>, Nobuhiro KAJI<sup>††</sup>,

Masashi TOYODA<sup>††</sup>, and Masaru KITSUREGAWA<sup>††</sup>

<sup>†</sup> Graduate School of Information Science and Technology, The University of Tokyo

7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656 Japan

<sup>††</sup> Institute of Industrial Science, The University of Tokyo

4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505 Japan

E-mail: <sup>†</sup>{nakashima, ynaga, kaji, toyoda, kitsure}@tkl.iis.u-tokyo.ac.jp

**Key words** knowledge acquisition, inference, time series

## 1. はじめに

ウェブに大量の情報が流れるようになって久しい。さらに、ウェブ上に存在する情報を長期にわたってアーカイブすることによって、情報の利用価値はますます高くなってきている。このような状況は、ウェブテキストを利用した知識の自動獲得などの研究にとって大きな追い風である。例えば、人間はブログにおいて、自身の行動や見聞きした出来事を順を追って書き連ねる。このようなブログテキストが大量に存在することで、行動や出来事（イベント）の繋がり（連鎖）を獲得することが可能となってきている。

イベント連鎖は、人間への行動提示や発話生成などへの利用することが可能な知識の一つである。人間は、様々なイベントに対して、次に起こりうる出来事や取るべき行動などを想定する事が可能である。例えば、「朝起きて熱があれ」ば「風邪ではないか」と疑うだろうし、「風邪である」ことが分かれば「風邪薬を飲む」だろう。人間にとって、ほぼ無意識のうちに行われるこのような行動や出来事の連鎖を理解することは、計算機

にとって困難である。そこで、このような連続する出来事をイベント連鎖として取り出し、知識として蓄えることが研究課題 [1]~[3] となっている。

さて、このような知識を人間への行動提示や、発話生成などへ応用することを考えた時、よりもっともらしい出力を得るために、対象の人間や発話がどのような状況下にあるかを考慮に入れることが望ましい。例えば「窓を閉める」と「クーラーをつける」というイベントの連鎖は、夏という季節の下でのみ利用されるべき知識である。しかし、現在までになされている研究のほとんどは、文脈によらず成立する規則を獲得することに主眼が置かれており、それらの知識がどのような文脈で成立するかは意識されていない。本研究では、特に、知識が成立する時期に注目し、ウェブから収集した大規模時系列テキストから、時期依存性を考慮に入れたイベント連鎖知識の獲得を目指す。

イベント連鎖の獲得には大規模テキストを用いるが、全体を一度に処理するだけではイベント連鎖の時期依存性の有無を判定することは難しい。そこで、我々は、月毎に分けて収集された時系列ブログテキストを利用することで時期依存性を考慮し

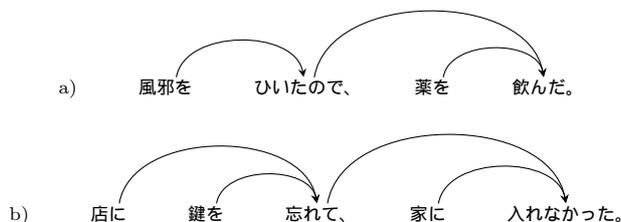


図 1 イベントの取得

たイベント連鎖の獲得を試みる．具体的には，まず，季節ごとに分割した時系列ブログテキストから，イベントの出現情報やイベント間の共起情報を特徴量として取得する．これに接続標識に代表される手がかり表現に基づいた特徴量を組み合わせた機械学習により，イベント連鎖の獲得と同時に，それぞれの連鎖がどの季節の下で成立しやすいかの分類を目指す．

実験では，実際に時系列ブログテキストから取得されたイベント連鎖に対して，人手で分類結果の評価を行い，その精度と時期依存性を有するイベント連鎖として獲得可能な知識の量について見積もることで，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 章で本研究で獲得するイベント連鎖とその時期依存性について述べる．次に，3 章で提案手法について述べ，4 章ではその評価を行う．そして，5 章で関連研究について述べる．最後に 6 章でまとめと今後の課題について述べる．

## 2. イベント連鎖と時期依存性

本章では，本研究で扱うイベントについて定義を与え，その連鎖と時期依存性について説明する．

### 2.1 イベントとイベント連鎖

本研究では，イベントを係り受け関係にある名詞と格助詞，動詞の 3 つの要素で構成されるものと定義する．また，格助詞は取得可能な場合にのみ取得し，動詞は肯定か否定の情報を保持するものとする．イベントの抽出には，係り受け解析を利用する．例えば，「風邪をひいたので、薬を飲んだ。」という文があったとき，構文解析を行うことで，係り受け関係にある名詞と動詞のペアを図 1 の a) ように見つけることができる．よって，この文からは“風邪をひく”と，“薬を飲む”という 2 つのイベントを取得する．また，「店に鍵を忘れたから、家に入れなかった。」という文の場合は，構文解析により図 1 の b) ような係り受け構造を見つけることができる．したがって，この文からは“店に忘れる”，“鍵を忘れる”，“家に入れない”という 3 つのイベントが取得される．この例のように，本研究では 1 つの動詞に複数の名詞がかかる場合にも，それぞれをイベントとして取得する．

次に，イベント連鎖とは，イベント対  $\langle X, Y \rangle$  があったときに，イベント  $X$  が起きた際に，それに続いてイベント  $Y$  が生じることに不自然性のないようなイベント対と定義する．以下にイベント連鎖の示す．これ以降，2 つのイベントを矢印で結んだ表現をイベント連鎖の表現とする．

- (1) 風邪をひく → 薬を飲む
- (2) 外に出る → 公園を散歩する
- (3) 風呂に入る → 体を洗う
- (4) 風呂に入る ⇨ 山に登る

特に (3) の例では，外に出た際に毎回公園を散歩するわけではないが，外に出た後の行動として，公園を散歩するというのは不自然ではない．このような関係についても，本研究では獲得の対象とする．ただし，人間が見て連続して起きるのはおかしいと感じるイベント対，また，めったに連続して起こることのないイベント対は獲得の対象ではない．例えば (4) の例は，風呂に入った後に山に登るといったイベントが起きるのは不自然であるため，連鎖しないイベント対と判断する．

### 2.2 連鎖の時期依存性

イベント連鎖の内，ある季節に，特に連続して生じるようなことが多いイベント対を時期依存性を有するイベント連鎖とする．例を以下に示す．例において，右肩に季節名のついた矢印は，記述された季節に起きる連鎖を示す．

- (5) 散歩に行く →<sup>春</sup> 桜を見る
- (6) 窓を閉める →<sup>夏</sup> 冷房をつける
- (7) 山に登る →<sup>秋</sup> 紅葉を見る
- (8) 公園に行く →<sup>冬</sup> 雪だるまを作る
- (9) 公園に行く → 遊具で遊ぶ
- (10) 暖房を入れる ⇨<sup>冬</sup> 雪だるまを作る

例えば，散歩に行って桜を見るという連鎖が不自然でないのは春のみである．したがって (5) の例のように春に成立する連鎖として獲得する (6) ~ (8) の例も，それぞれ矢印に付記した季節にのみ成立するような連鎖である．また，時期依存性のないイベント連鎖としては (9) のようなものが，時期依存性はあるが連鎖しないイベント対としては (10) のようなものが考えられる．

例 (5) ~ (8) で示したようなイベント連鎖は，文脈として季節を指定しなければ獲得できない知識である．すなわち，時期依存性を考慮に入れたイベント連鎖の獲得は，文脈を考慮に入れない場合に比べて，より多くの知識の獲得につながるという．

## 3. 提案手法

本章では，時系列ウェブテキストから時期依存性を有するイベント連鎖を獲得する手法について述べる．まず，基本となるアイデアについて説明し，次に，連鎖の獲得の対象となるイベントの取得と機械学習のための特徴量抽出について述べる．その後，学習データを得るために行ったラベル付けなどについて説明する．

### 3.1 アプローチ

本稿では，イベント連鎖の候補としてウェブテキストからイベント対を抽出し，抽出された候補を季節に関係なく生じるイベント連鎖と春夏秋冬それぞれに起こるイベント連鎖，それに連鎖しないイベント対を含めた 6 つのクラスに分類することを目指す．

表 1 イベントの接続に基づく手がかり表現

連用形, ため, が, なら, から, ながら  
と, れば (ば, らば), たら, ので, のに

具体的には、まず、時系列ウェブテキストから分類を行う際に利用する特徴量を獲得する。今回は、言語的な手掛かりと各季節における単語やイベントの出現頻度を特徴量として利用する。特に、各季節における頻度情報は、時系列ウェブテキストをそれぞれの季節に書かれたものに分割し、それぞれに対して、処理を行うことで得ることができる。その後、抽出した特徴量を組み合わせることで、イベント対の特徴ベクトルを作成し、人手で作成した正解データを利用して分類器の学習を行う。

さらに、学習された分類器で未知のイベント対を分類した結果に対して、人手でラベルの修正を行い、これを学習データに加えることでより精度の高い分類を行う。これは、時期依存性を有するイベント連鎖は、一般的に成り立つイベント連鎖に比べて数が少なく、その学習データを十分量確保するのが困難であるためである。

以下では、本節で述べた内容を、より具体的に説明していく。

### 3.2 イベントの抽出とイベント連鎖候補の獲得

まず、イベントの抽出について述べる。最初に、係り受け解析により、全ウェブテキストからイベントを獲得する。係り受け解析には、Yoshinaga らが開発した J.DepP [4] を利用した。また、その際ストップワードを含むイベントのフィルタリングを行う。具体的には、時相名詞や形式名詞、数詞などに属する名詞を含むイベントや、“する”、“ある”などの軽動詞を含むイベントは、具体的な意味を成さないことが多いため削除する。さらに、フィルタリングされたイベントを spacesaving アルゴリズム [5] によりカウントしていく。また、その際、動詞は全てひらがなに正規化する。これにより「風邪をひく」と「風邪を引く」のような動詞の表記のゆらぎを吸収したカウントが可能である。最終的には、低頻度なイベントを切り捨て、残ったイベントを連鎖獲得の対象として利用する。

次に、上記の手順で得られたイベントの間の組み合わせをイベント連鎖として獲得する。ただし、抽出された全てのイベントの組み合わせを考えると膨大な量となるため、本研究では連用形による接続を手がかりとしてイベント連鎖候補を獲得した。具体的には、イベントがテキスト中で図 2 の a) のように連用形で接続されていた場合、その 2 つのイベントによるイベント対をイベント連鎖の候補として獲得している。

### 3.3 特徴量抽出

本節では、3.2 節で得られた各イベント連鎖候補を、特徴ベクトルで表現する方法について述べる。

#### 3.3.1 手掛かり表現による特徴量

連鎖するイベントはテキスト中で、特定の接続標識や動詞の連用形によって結ばれていることが多い。そこで、手がかり表現に基づく特徴量として、イベント対が特定の接続標識や連用形で結ばれ、複文としてテキスト中に出現した頻度を用いる。利用する接続標識は乾らの研究 [1] を参考にした。表 1 に特徴量として利用する手がかりを示す。

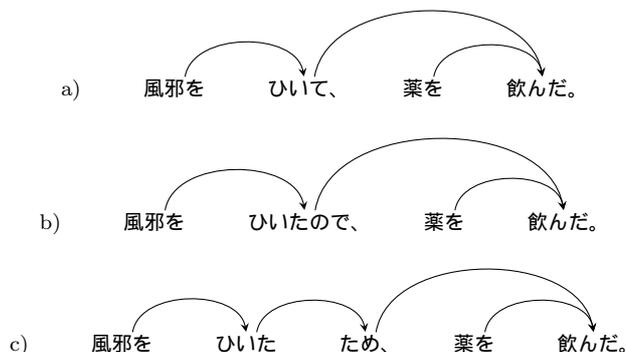


図 2 手がかり表現の取得

表 2 動詞のモダリティなどに基づく手掛かり表現

動詞の過去形, 推量「だろう」「らしい」  
受け身「れる、られる」、使役「せる、させる」

取得対象となるようなテキストのパターンを図 2 に示す。テキスト中に a) のようなパターンがあれば連用形接続手がかりを b) と c) のようなパターンがあれば接続標識手がかりを、それぞれ取得することになる。

また、特定の動詞のモダリティや過去形がイベント対の中に存在した回数も特徴量として取得する。例えば、“～らしい”のような推量の表現が多用されるイベント対は、実際に、連鎖が起こっている可能性が減少すると考えられる。また、動詞の過去形は、連鎖する順序に影響する。例としては、「薬を買うため、薬局に行った」というテキストがあったとき、実際のイベントは、薬局に行く(後項)→薬を買う(前項)という順で連鎖する。そこで、動詞の過去形の情報を特徴量として保持することで、このような例に対処する。表 2 に特徴量として利用する表現を示す。動詞の過去形については、前項のみ過去の場合と後項のみ過去の場合、そして、双方過去形の場合に分けて頻度を取得する。

以上で取得される特徴量は、イベント対の連鎖性の判定に用いられる。

#### 3.3.2 頻度ベースの特徴量

ここでは、頻度ベースの特徴量について述べる。頻度ベースの特徴量には、イベント対の連鎖性判定に寄与する情報として利用するウィンドウ共起に基づく特徴量と時期依存性の判定に用いる単語やイベントの出現数などの時系列ヒストグラムを利用した特徴量が存在する。

##### a) ウィンドウ共起

連鎖するイベントは、テキスト中の近い位置に記述されると考えられる。そこで、本手法では、イベントの共起を 1 文内、3 文内、同一のウェブページ内という 3 つのウィンドウに分けて頻度を取得し、特徴量として利用する。ただし、共起の際、テキスト中でのイベントの出現順が異なる場合は別々に数えることとする。ウィンドウを複数用意するのは、イベントが共起した際の位置関係で関連度の強さが変化するためである。一般的に、近い位置で共起したイベントのほうが強い関連性を有すると考えられる。3 つのウィンドウを設定することで、関連度の

表 3 特徴ベクトルの構成
言語的手掛かり (36 個)
ウィンドウ共起 (3 個)
時系列ヒストグラム (28 個)

強さを学習結果に反映させることができる。

#### b) 時系列ヒストグラム

季節によって連鎖しやすさの異なるようなイベント対は、特定の季節の共起頻度が特に高くなるはずである。したがって、ブログテキストを取得された季節に分け、各季節のもとでの共起頻度を取得する。具体的には、時系列ウェブテキストに対して、3~5月に取得されたものを春、6~8月を夏、9~11月を秋、12~2月を冬と分割し、それぞれブログ記事内でのイベントの共起頻度を取得した。この特徴量は、対象のイベント連鎖の季節性の有無を判定する一助となることが期待される。

また、時期依存性を有するイベント連鎖では、イベントそれぞれ自身が季節性を有していたり、イベントに含まれる単語（名詞や動詞）に季節性があるものも多い。したがって、単語やイベントの各季節における出現頻度も特徴量として取得する。これに関しても、上記で行ったように時系列ウェブテキストを分割し、前項の名詞、前項の動詞、後項の名詞、後項の動詞のそれぞれについて、各季節で出現頻度を数える。この特徴量も、イベント連鎖の季節性の有無を判定するために用いられる。

#### 3.4 特徴量の正規化と特徴ベクトルの作成

上記までで取得された特徴量は頻度そのものであり、各特徴量の出現頻度によって大きな偏りが生まれてしまう。そのため、このまま特徴ベクトルを作成すると、出現頻度が高い特徴量の影響が大きくなり学習が上手くいかない可能性がある。そこで、各特徴量の正規化を行う。(a)(b)(c)の特徴量については、イベント対に関して横断的に正規化を行う。具体的には、特定の特徴量を取り出して全てのイベント対で比較してみたとき、その最大値で対象の特徴量の値を割ることで正規化する。(d)(e)については、春夏秋冬の特徴量の比を保ったまま、各特徴量毎に値の範囲が0から1の間となるように正規化する。

正規化した特徴量を組合せることで特徴ベクトルを作成する。(a)(b)については、イベント対の前項と後項を逆にした場合の特徴量も加える。これは、前項、後項を入れ替えた場合に、連鎖の可否が変化するイベント対の分類に有効である。結果として67種類の特徴量を有する特徴ベクトルが作成される。特徴ベクトルの内訳を表3にまとめた。

#### 3.5 学習データの作成方法

連鎖候補であるイベント対に対して人手でラベル付けを行うことによって、学習データを作成する。ラベル付けは、特徴量の連用形接続の頻度が上位のデータに対して行う。これは、頻度が上位のイベント対の方が、イベント連鎖である可能性が高いためである。また、頻度が上位のデータから季節性の高いものを抽出し、それらについてもラベル付けを行う。これは、季節性を有するイベント連鎖は特定の時期にのみ出現するので、連用形接続の頻度も季節性のないイベント連鎖に比べて少なく、単純に連用形接続の頻度の上位のみを見ても、あまり学習

データを得ることができないためである。季節性の高いものを抽出は、季節ごとのイベントの共起頻度の偏りを最大頻度と次点との差の大きさに基づいて行った。ラベル付に際しては、人間でも判断に迷うようなイベント対や、具体的な意味を成さないイベントを持つイベント対に関しては学習データから除いている。

しかし、単純に人手でラベル付けをする作業は、大きなコストが掛かり、十分なデータ量を集めるのが難しい。そこで、得られた学習データを利用して、適合性フィードバックにより学習データを拡張する。まず、既に得ている学習データから分類器を学習する。その分類器に連用形接続の頻度が上位100万件以内の未分類のイベント対の特徴ベクトルを入力し、分類を行う。ここで、上位100万件を利用するのは、下位のイベント対の特徴ベクトルが非常にスパースであり、分類問題を解くために十分な情報を保持できていないと判断できるためである。次に、分類結果から、各ラベルに関して確度が高いイベント対を抽出する。今回利用した分類器ではロジスティック回帰モデルを用いているため、それぞれのイベント対に対して各ラベルへの分類確率が求められる。よって、分類確率が高いイベント対を抽出する。このようにして得られたイベント対に対して、人手でラベルの修正を行う。そして、そのラベル付けされたデータを元の学習データに加えることで、学習データの量を増加させ、より精度の高い分類を実現する。

## 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データに適用することで、時期依存性を有するイベント連鎖の獲得を行う。さらに、得られたイベント連鎖について人手で評価を行い、その結果を考察する。

### 4.1 実験設定

今回、イベント連鎖を獲得するために用いたデータには、2006年から2012年11月まで収集したブログ記事を用いた。これらは、約24億文からなっている。さらに、3.2節で述べた方法を用い、全ブログテキスト中でおおよそ1000回以上出現するイベントを連鎖の獲得対象とした。具体的には17万個のイベントが獲得の対象となった。このブログテキストから前節で述べた手法によってイベント対と対応する特徴ベクトルを作成する。データには、予め獲得月がタグ付けされているため、この情報を用いて各季節ごとの単語やイベントの出現頻度、そしてイベントの共起頻度を取得することが可能である。結果として、イベント連鎖の候補として約700万件のイベント対が得られた。また、分類のための学習器としては、ロジスティック回帰モデルを実装したLIBLINER[6]を用いた。

### 4.2 学習データの作成

3.5節で説明した手法に従い、学習データの作成を行った。まず、人手によるラベル付けは、特徴量の連用形接続の頻度が上位の300件のイベント対に対して施した。また、同上位10000件のイベント対から季節性の高いものを700件抽出し、それらについてもラベル付けをした。結果として、連鎖の有無と、その季節性について、3.1節に記述した6つのラベルを人手でつ

表 4 人手による学習データの作成

連鎖しない	季節なし	春	夏	秋	冬	計
282	52	20	26	9	47	436

表 5 フィードバック後の学習データの内訳

連鎖しない	季節なし	春	夏	秋	冬	計
393	149	72	95	53	121	883

表 6 各ラベルへの分類数

	分類された数
季節なし	78763
春	7165
夏	8998
秋	810
冬	7295

けることで、表 4 に示したように 436 件の学習データを作成した。

さらに、436 件の学習データにより分類器を学習し、適合性フィードバックを行った。結果として、データの総量としては、フィードバック前の約 2 倍となる 883 件の学習データを得ることができた。特に、人手によるラベル付け作業であまり多くの学習データを得られなかった秋のラベルに関しては、フィードバックによりおよそ 6 倍の量とすることができた。最終的に得られた学習データにおける各ラベルの内訳を表 5 に示す。「連鎖しない」というラベルが多くなるのは、実際にデータを眺めてみると、連鎖しないイベント対が大半であり、学習データを集めやすいためである。また、時期依存性を有する連鎖を見てみると、夏と冬の学習データが多くなった。これは、夏や冬という極端な気候の季節のほうが、その時期特有のイベントが起こりやすいためであると考えられる。

#### 4.3 獲得できた知識量の概算

上記までで得られた 883 件の学習データを用いて学習した分類器に、連用形接続の頻度が上位 100 万件以内の未分類のイベント対の特徴ベクトルを入力し、分類を行った。上位 100 万件を利用する理由は、前節のとおりである。

今回利用した分類器ではロジスティック回帰モデルを用いているため、それぞれのイベント対に対して各ラベルへの分類確率が求められる。

結果として、個々のイベント対において分類確率が最も高かったラベルが、そのイベント対の分類結果となる。上記の設定で分類を行った結果、各ラベルへ分類されたイベント対の総量を表 6 に示した。

次に、分類されたイベント対の中にどれだけ量の正しいイベント連鎖が存在するかを見積もるため、分類の精度を算出する。まず、各ラベルに対して分類結果を無作為に 500 件抽出し、人手で正誤判定を行った。そして、分類の際の分類確率が一定以上のイベントのみをとりだし、その条件下での適合率を算出する。例えば、分類確率を  $p$  とし、 $p > 0.7$  以上の分類結果の適合率は、式 1 のような式によって求めた。

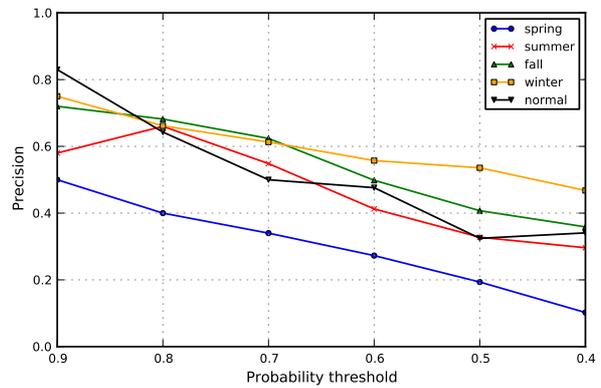


図 3 各ラベルにおける一定の分類確率以上の適合率

表 7 獲得可能なイベント連鎖の概算

	$p > 0.7$	$p > 0.5$
季節なし	2742	8677
春	20	507
夏	330	2162
秋	206	272
冬	1302	3712

$$\text{適合率} (p > 0.7) = \frac{p > 0.7 \text{ のイベント対中の正解数}}{p > 0.7 \text{ のイベント対の数}} \quad (1)$$

季節なしと春夏秋冬の 5 つのラベルについて分類確率 0.4 ~ 0.9 まで 0.1 刻みで分類の適合率を算出した結果を図 3 に示す。結果として、分類確率が 0.7 以上あれば分類された数のおよそ半数以上は正しいイベント連鎖を獲得できることがわかった。また、春の分類精度のみ特に適合率がよくないという結果が出た。これについては、4.4 節で誤分類の原因について触れながら考察を示す。さらに、提案手法を利用することで獲得されるイベント連鎖の数を式 2 に従って見積もる。

$$\text{獲得できるイベント連鎖の数} = \text{適合率} \times \text{分類数} \quad (2)$$

まず、分類確率が 0.7 以上の分類結果より、2000 件程度の時期依存性を有するイベント連鎖がおよそ 6 割の適合率で獲得可能だと推定される。また、分類確率が 0.5 以上の分類結果について見てみると、分類精度はあまり良いとは言えないが、特に夏や冬については数千件単位で時期依存性のあるイベント連鎖の存在を確認することができる。これは、実際に本手法によって、時期依存性のあるイベント連鎖を獲得できることを示している。

実際に未分類のイベント対を分類することで獲得されたイベント連鎖の例を表 8 に示す。特に、春夏秋冬の各ラベルに分類されたイベント対は、本手法の特徴である時期依存性のあるイベント連鎖として獲得されたイベント対である。それぞれの季節の例を見てみると、直感的に妥当な例であるといえる。また、表中の太字で示したイベント対は、どれも前項が「公園を訪れる」という同じ行為を表しているが、連鎖として起きるイベントは、季節ごとに異なるものとなっている。このように、前項に季節性は無いが、他のイベントとの連鎖を考えたときに季節性を帯びるものについても獲得できていることが確認すること

表 8 獲得されたイベント連鎖の例

連鎖しない	季節なし
話題になる → 気持ちになる 話を聞く → 大人になる シャワーをあびる → 汗をかく 水を飲む → 喉が乾く 目がさめる → 状態になる 話題になる → 機会がある リーダーにつける → 捨札に送る	おなかが空く → 飯を食べる 鍵を忘れる → 家に入れない 熱を出す → 病院に行く 事故を起こす → 警察に捕まる 喉が乾く → 水を飲む 風邪をひく → 咳が止まらない ばい菌が入る → 炎症を起こす
春	夏
公園で遊ぶ → 桜を見る 雨が降る → 花粉が飛ばない 弁当を作る → 花見に出かける 満開を過ぎる → 桜が散る 桜が終わる → ツツジが咲く 桜が咲く → 花見に行く 風が吹く → 桜が散る	公園に行く → 花火を見る 浴衣に着替える → 花火大会に行く 縁側に座る → スイカを食べる 水着に着替える → 海で遊ぶ 窓を閉める → 冷房をいれる 甲子園に行く → 高校野球を観る 海に行く → 花火をみる
秋	冬
公園へ行く → どんぐりを拾う 落ち葉を集める → 火をつける 京都に行く → 紅葉をみる 木枯らしが吹く → 葉っぱが落ちる 旅に行く → 紅葉をみる 紅葉が終わる → 葉を落とす 散歩に行く → 紅葉をみる	公園に行く → 雪だるまを作る 雪が残る → 路面が凍結する 寒波がくる → 雪が積もる 気温が下がる → 氷がはる 天気が崩れる → 雪が降る 手袋を忘れる → 手がかじかむ 気温が下がる → 道路が凍る

表 9 各ラベルごとの分類結果の誤り分析

	季節なし	春	夏	秋	冬	合計
無関係のイベント対	36	26	33	42	37	174
イベントの情報不足	30	19	14	20	27	110
時期依存性の誤認	1	32	25	9	10	77
イベントの連鎖順	1	6	4	24	4	39
特定の出来事に由来	5	8	11	2	2	28
逆説的な連鎖	14	1	4	0	4	23
慣用的な表現	5	1	1	1	0	8
その他	8	7	8	2	16	41

ができた。

#### 4.4 誤分類考察

前節の実験において、評価者によって誤分類だと判定されたイベント対を季節なしと春夏秋冬の5つのラベルからそれぞれ100件ずつランダムに抽出し、誤分類の原因としていくつかのカテゴリを考え分類を行った。分類した結果を表9に示す。分類の結果、全体を通してみたとき、誤分類の原因として最も多かったのは、無関係のイベント同士が連鎖するとして分類されているものと、イベント自体が具体的な状況を表現できていないものが多く、ついで、時期依存性の誤認など、幾つかの特徴的な要因が見られた。以下では、それぞれの要因について例を挙げながら原因を分析していく。

##### 無関係のイベント対

このカテゴリには「カメラを持つ → 雨が降る」のような、直感的に互いに無関係とわかるような2つのイベントを連鎖すると分類してしまっているものが入る。このような誤認識が生じる原因としては、なかなかありえないようなこと、珍しいことがあったときに、その内容をブログなどに書いておきたいというモチベーションがあがることが多いため、なかなかありえ

ないような連鎖に関する記述も多数存在することがあげられるのではないかと考える。

##### イベントの情報不足

これについては、連鎖の判定ミスというよりは、イベント自体を獲得する際に生じる問題である。このカテゴリには「気持ちになる → 手を出す」のような、不明瞭なイベントを含むイベント対を分類した。このようなものに関しては、人間にとっても具体的な状況をイベントから想起することができないため、連鎖性の判定は難しい。

##### 時期依存性の誤認

時期依存性に関する誤分類として、一年を通して連鎖するイベント対を特定の季節に分類してしまう例が挙げられる。例えば「風邪をひく → 病院に行く」という連鎖は季節に関係なく生じるが、分類結果は冬となっている。これは、「風邪をひく」というイベントそのものが冬に起こりやすいことに起因していると考えられる。また、春の連鎖における誤分類に関して言えば、この要因が最も多いことがわかった。したがって、春における季節性の判定が、他の季節と比較して難しいことが、春の連鎖の分類における精度低下の原因であると考えられる。

この誤認識を回避するためには、イベント単体が各季節に出現する頻度などを利用して、ある程度出現の偏りをならすなどの工夫が必要であると考えられる。しかし、あまり偏りをならしすぎてしまうと、実際に時期依存で連鎖するイベント対の分類に支障が出てしまうため、うまくバランスを取る必要がある。

##### イベントの連鎖順

誤分類例として、特に多く見られたのが、「雪が積もる → 雪が降る」のように、前項と後項を入れ替えると成立するような例である。これについては、実際に前項と後項を入れ替えた成立すると特徴量のパターンが類似することが多いということが原因であると考えられる。

##### 特定の出来事に由来する連鎖

例えば、東日本大震災が3月にあった関係上「地震がくる → 津波が発生する」のような地震関連の自体連鎖が春の連鎖であると誤分類してしまう例がよく見られた。また、同様にオリンピックなどが原因で「アメリカに勝つ → 金メダルをとる」のように本来は連鎖しないイベント対を連鎖すると誤って分類してしまう例も見られた。

##### 逆説的な連鎖

例えば「お腹が空く → ご飯を食べない」のように、一方のイベントが肯定の表現に変わるとイベント連鎖として成り立つような誤分類が、このカテゴリに含まれるイベント対である。これについては「お腹がすいたけど、ご飯を食べられなかった」のような逆説的な記述が多くのブログで散見されることが原因ではないかと考えられる。

##### 慣用的に用いられるイベント対

慣用的に用いられるイベント対が連鎖するイベント対として分類される例も散見された。例えば「背中を見る → 子供が育つ」が連鎖するイベント対として誤分類されてしまったのは、「(親

の) 背中を見て子供が育つ」というフレーズが、多くのブログで書かれていたためであると考えられる。

#### 4.5 適合性フィードバックの有効性の検証

ここでは、学習データの作成の際に利用した適合性フィードバックの効果を検証する。まず、各ラベルへの分類結果から、分類確率が高い順番に 100 件のイベント対を抽出し、評価用のデータとする。そして、2 人の評価者に協力してもらい、評価用データの正誤を判定してもらい、分類結果の適合率を算出する。評価にあたっては、以下のような評価基準を与えて、分類器によってイベント対に付けられたラベルが正しいか否かを評価してもらった。

(1) 前項のイベントから、後項のイベントが連続して起こることが妥当か

(2) 時間的な順序関係は正しいか

(3) 季節性を有する場合は、特定の 1 つの季節のみに依存すると言えるか

また、適合性フィードバックをかける前の学習データのみで学習した分類器に対しても、同様の評価を行い、上記で得られた結果と比較することで、適合性フィードバックの効果についても検証する。すなわち、フィードバックをかけることによって、分類精度の向上がみられるかを確認する。

##### 4.5.1 評価のための準備

人手による評価においては、より正確に評価をしてもらうために、評価者に余計な先入観を与えないことが重要である。本評価では 2 人の評価者に先入観を与えないため、適合性フィードバックをする前の学習データを利用して学習した際の分類結果と、適合性フィードバックをかけた後の学習データを利用して学習した際の分類結果を混合して、評価用のデータとした。これは、適合性フィードバックをかけた後の方が、良い結果になるはずだという先入観を排除する目的がある。また、混合した際、イベント対の順番をランダムに変更している。これも、評価用のデータが確度の高い順に並んでしまい、それが評価者の判定にノイズを与えることを防ぐためである。

また、正確な評価のためには、評価者の判定基準が一致している必要がある。ここでは、2 人の評価者の基準が一致しているか否かをカッパ値によって評価する。本評価においては、6 つのラベルに関して各 100 件ずつ、それをフィードバック有りと無しの分類結果に対して評価するため、合計 1200 件のイベント対に関して評価を行う必要がある。しかし、フィードバックの有無にかかわらず、どちらの分類結果にも重複して現れるイベント対に関しては 1 度の評価をすれば良い。評価用のデータを作成したところ、合計 977 件のイベント対に対して評価を行えばよいことがわかった。2 人の評価者にこの 977 件のイベント対を評価してもらい、カッパ値を計算したところ  $\kappa = 0.558$  という値を示した。これは、非常に高い一致率というわけではないが、2 人の評価者の評価が良好な一致を示していると言える値である。

##### 4.5.2 評価結果と考察

評価には、分類確率の上位から順に  $x$  個サンプルを取り出

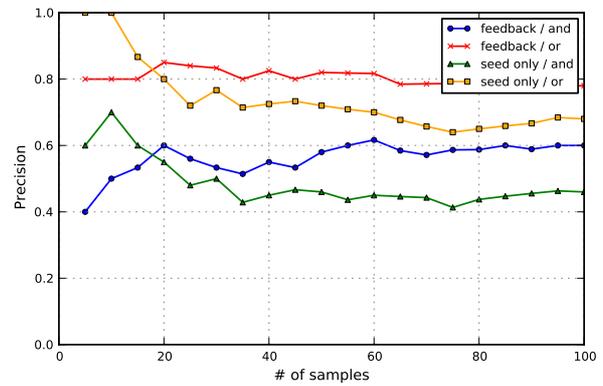


図 4 秋に成立する連鎖として分類されたイベント対の評価

した際の適合率を利用する。具体的には、サンプル数を 5 個刻みで増加させながら適合率を算出しグラフを作成した。各ラベルに分類されたイベント対の評価の結果を図 4 に示す。全ての季節のラベルに対して同様の傾向が見られたため、今回は、フィードバックにより最もデータ数が増加した秋のラベルの評価結果を示す。

グラフにおいては縦軸が適合率を、横軸がサンプル数を示している。また、凡例において feedback と書かれているものが適合性フィードバックを利用した分類器による分類結果の評価であり、seed only と書かれているものがフィードバックを利用していない分類器による分類結果の評価である。また、and と書かれているものは、2 人の評価者双方が正しいと判定した時のみ正解と判断した場合の適合率であり、or と書かれているものは、2 人の評価者の片方でも正しいと判定したときに正解と判断した場合の適合率を示す。適合性フィードバックの効果については、and のついた凡例同士、または or のついた凡例同士を比較することで確認できる。実際にグラフを見てみると、seed only のグラフより feedback のグラフの方が高い適合率で推移しており、適合性フィードバックによって学習データを増やしたことで、分類の精度が向上したとすることができる。

## 5. 関連研究

本稿で取り扱ったイベント連鎖は、関連性を強いイベントを獲得する研究と類似している。このような知識は一般に含意関係や推論規則、因果関係と呼ばれ、質疑応答や文書要約、機械翻訳、談話理解などへの応用を見越して古くから研究対象となってきた。

含意関係は、言語表現対  $\langle T, H \rangle$  があったとき、 $T$  を読んだ人間が  $H$  も真実だと判断できるような関係と定義されている [7]。特に、含意関係の判定に関する研究は、2005 年から毎年作成されている RTE Challenge [8] というベンチマークの存在などが後押しとなり、最も研究が盛んとなっている。最近の研究例としては、解析木を用いた構文情報と単語類似度を利用した意味情報を組み合わせることで、入力として与えられた 2 つの文が含意の関係にあるかを判定を行い、その精度の向上を示した Mehdad らの研究などがある [9]。また、含意の判定以外にも、含意関係にある表現対の獲得 [10], [11] や、含意関係に

あるのテキストの生成 [12], [13] などの研究が行われている。

また、推論規則や因果関係とは、事態対  $\langle X, Y \rangle$  があったとき、 $X$  が生じた際に  $Y$  も蓋然的に起きるような関係のことをいう。乾ら [1] は、ある程度の必然性を有する因果関係を表現する“ため”という接続標識に注目し、これによって結ばれる複文に注目して、因果関係知識の獲得を試みている。最終的に、1年分の新聞記事より 27,000 件以上の因果知識を獲得できることを示している。また、Shibata ら [2] はイベントの共起に注目して推論規則の獲得を目指している。彼らの手法では、係り受け構造を利用して推論規則の候補となる事態対を取得し、その後、事態の共起に着目して、実際に推論規則になる事態対を選出する。最後に、格フレーム [14] を適用し省略されている表現を補うことによって、より詳細な推論規則を獲得し、およそ 1 億のウェブページ上のテキストから約 2 万件の推論規則が獲得したと報告している。

しかし、これらの研究において、獲得される関係が成立する文脈まで考慮しているものはなかった。したがって、本稿で述べたように文脈を意識した関係知識の獲得に向けた議論が必要であると考えられる。

## 6. おわりに

本論文では、推論規則のような知識の獲得において、規則が成立する時の状況を意識することが、文脈に沿った行動提示や発話生成などへの応用を考えた際に重要な要素となると考えた。そして、その第一歩として、各季節における単語やイベントの出現回数、また、イベント間の共起情報を言語の手掛かりとともに特徴量として利用した機械学習により、時期依存性を意識しながら、イベント連鎖を獲得する手法を提案した。さらに、提案手法の評価実験により、一定量の時期依存性のあるイベント連鎖が獲得可能であることを明らかにし、本手法の有効性を示した。

今後の課題としては、精度向上があげられる。特にイベント対が連鎖するか否かに関しては誤分類が多くさらなる考察が必要である。また、本稿では、イベント連鎖を個々の季節へと分類したが、実際には、複数の季節において連鎖するイベントというものも存在する。例えば、本稿で夏の連鎖として紹介した「外に出る → 汗をかく」というイベントは、冬以外であればある程度妥当だということもできる。その点を考えて、提案手法にマルチラベル分類を取り入れることも考えている。さらに、本手法を応用することで、年齢や地域、性別など様々な文脈におけるイベント連鎖の獲得が可能ではないかと思われる。

## 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 (自然言語)” 情報処理学会論文誌, vol.45, no.3, pp.919-933, mar 2004.
- [2] T. Shibata, and S. Kurohashi, “Acquiring strongly-related events using predicate-argument co-occurring statistics and case frames,” Proceedings of 5th International Joint Conference on Natural Language Processing, 2011.
- [3] Q. Do, Y. Chan, and D. Roth, “Minimally supervised event causality identification,” EMNLP, Edinburgh, Scotland, 7

- 2011.
- [4] N. Yoshinaga, and M. Kitsuregawa, “Kernel slicing: Scalable online training with conjunctive features,” COLING, pp.1245-1253, 2010.
- [5] G. Cormode, and M. Hadjieleftheriou, “Methods for finding frequent items in data streams,” The VLDB Journal, vol.19, no.1, pp.3-20, Feb. 2010.
- [6] R.E. Fan, K.W. Chang, C.J. Hsieh, X.R. Wang, and C.J. Lin, “Liblinear: A library for large linear classification,” J. Mach. Learn. Res., vol.9, pp.1871-1874, June 2008.
- [7] I. Androutsopoulos, and P. Malakasiotis, “A survey of paraphrasing and textual entailment methods,” CoRR, vol.abs/0912.3747, 2009.
- [8] I. Dagan, G. Oren, and B. Magnini, “The pascal recognising textual entailment challenge,” Proceedings of the PASCAL Challenges Workshop on Recognising Textual Entailment, 2005.
- [9] Y. Mehdad, A. Moschitti, and F.M. Zanzotto, “Syntactic/semantic structures for textual entailment recognition,” Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, pp.1020-1028, Stroudsburg, PA, USA, 2010, Association for Computational Linguistics.
- [10] J. Berant, I. Dagan, and J. Goldberger, “Global learning of typed entailment rules,” Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1, pp.610-619, Stroudsburg, PA, USA, 2011, Association for Computational Linguistics.
- [11] J. Berant, I. Dagan, and J. Goldberger, “Global learning of focused entailment graphs,” Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, pp.1220-1229, Stroudsburg, PA, USA, 2010, Association for Computational Linguistics.
- [12] P. Malakasiotis, and I. Androutsopoulos, “A generate and rank approach to sentence paraphrasing,” Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, pp.96-106, 2011.
- [13] S. Zhao, H. Wang, X. Lan, and T. Liu, “Leveraging multiple mt engines for paraphrase generation,” Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010), pp.1326-1334, Beijing, China, August 2010, Coling 2010 Organizing Committee.
- [14] 河原大輔, 黒橋禎夫, “格フレーム辞書の漸次的自動構築,” Journal of natural language processing, vol.12, no.2, pp.109-131, 2005.