

## 2種類のアンカー情報と共起パターンの組み合わせによる事態間関係獲得

阿部 修也<sup>†</sup> 乾 健太郎<sup>†</sup> 松本 裕治<sup>†</sup>

<sup>†</sup> 奈良先端科学技術大学院大学 情報科学研究科  
奈良県生駒市高山町 8916-5  
E-mail: †{shuya-a,inui,matsu}@is.naist.jp

あらまし 行為-効果, 行為-目的のような事態間の関係を大規模コーパスから自動的に獲得する。文内共起パターンを利用する手法では, 事態間でどの項が共有されるかの知識を獲得することが難しい。そこで事態間で共有される名詞(アンカー)を用いて項共有情報を獲得し, 文内共起パターンによる事態間関係と組み合わせることで項を必要とする事態間関係を獲得する。このとき2種類の異なるアンカーを用いることで, 精度を保ったまま再現率を向上できることを確認した。

キーワード 関係獲得

## Combining Pattern-Based and Anchor-Based Approaches for Event Relation Acquisition

Shuya ABE<sup>†</sup>, Kentaro INUI<sup>†</sup>, and Yuji MATSUMOTO<sup>†</sup>

<sup>†</sup> Nara Institute of Science and Technology Graduate School of Information Science  
Takayama, Ikoma, Nara 630-0192, Japan  
E-mail: †{shuya-a,inui,matsu}@is.naist.jp

**Abstract** Addressing the task of acquiring semantic relations between events from a large corpus, we first argue the complementarity between the pattern-based relation-oriented approach and the anchor-based argument-oriented approach. We then propose a two-phased approach, which first uses lexico-syntactic patterns to acquire predicate pairs and then uses two types of anchors to identify shared arguments. The present results of our empirical evaluation on a large-scale Japanese Web corpus have shown that (a) the anchor-based filtering extensively improves the precision of predicate pair acquisition, (b) the two types of anchors are almost equally contributive and combining them improves recall without losing precision, and (c) the anchor-based method achieves high precision also in shared argument identification.

**Key words** relation extraction

### 1. はじめに

含意関係や因果関係などの事態間関係は, 人間に近い高度な言語処理能力を工学的に実現する上で欠かせない知識のひとつであり, 質問応答, 情報抽出, 対話, 要約など, 幅広い言語処理アプリケーションに役立つことが期待できる。

本稿では大規模コーパスから事態間関係を自動獲得する手法について論じる。本稿で対象とする事態間関係とは, 「X=電話;Xをかける→Xが通じる(行為-効果の関係)」のように, (a) 事態表現間に成り立つ関係が「行為-効果関係」や「行為-前提関係」のように分類されており, さらに (b) 「X (=電話)」のように2つの事態間で同一の要素からなる項(「かける」のラ

格と「通じる」のガ格)があればその情報が特定されているという条件を満たす関係を云う。したがって, 事態間関係を自動獲得するには, (a) 事態表現対の関係の識別および (b) 項の共有関係の同定が必要である。事態間関係獲得についてはすでにいくつかの手法が提案されている [1]~[7] が, いずれの手法も上の (a), (b) の問題を同時に解決するものにはなっていない。以上の背景を踏まえ, 本稿では事態表現の文内共起情報と文章内共起情報を組み合わせることによって, 関係の識別と項の共有情報の同定をともに実現する手法を提案する。

### 2. 関連研究

既存の手法を項に注目した手法と共起パターンを利用する手

法に分けて紹介する。

事態間の項に注目した手法を2つに分けて紹介する。ひとつめは項の要素集合の分布が似ている事態対は類義/同義/含意関係にあるという仮説に基づく手法である(DIRT [3], TE/ASE [8])。もうひとつは同一談話関係にあり項を共有する事態対は一連の状況で一緒に起りやすい事態対になっているという仮説に基づいている[5]。これらの手法は項の共有を伴う含意関係を獲得できるという利点はあるが、関係を推定する証拠が少ないために前提/結果/手段関係等のより詳細な区分が必要な関係を識別することが困難である。

類義/同義/含意関係をより区別した前提/結果/手段関係等を獲得するためにはより多くの関係を表わす証拠が必要である。そこで文内で述語対が共起したときの共起パターンを関係を表わす証拠とすることで、前提/結果/手段関係のような事態間関係を認識することができる[1], [2], [4]。例えば、Inuiらは因果関係を表わす共起パターンとして接続助詞「ため」を利用した。しかし、文内で述語対が共起する場合はしばしば片方の事態の格が省略されるため、事態間の項の共有がわからなくなる欠点がある。例えば、「お茶を淹れて飲む」では、「飲む」のヲ格の「お茶」が省略されている。

ここまでの手法を要約すると、項に注目した手法は項の共有情報を認識できるが、類義/同義/含意関係よりも詳細な区別が必要な関係を認識することが困難である。一方、共起パターンに基づく手法では前提/結果/手段関係のような関係を認識できるが、項の共有情報を認識できない。

この問題に対し、Torisawa [7]は項に注目した手法と共起パターン手法を組み合わせることで、項の共有情報を認識しつつ「時間的な前後関係のある推論規則」を獲得する手法を提案している。この手法は動詞テ形接続や連用中止接続のように頻度が高く一般的な手がかりと、別途収集した格関係の統計を巧妙に組み合わせることによって、時間的な前後関係を獲得することを狙うもので、「時間的な前後関係のある推論規則」の獲得で成果を上げている。ただし、この手法を広く他の事態間関係に適用できるかどうかは今のところ明らかでない。

こうした背景を踏まえ、本稿では広く他の事態間関係に適用可能でありながら、項の共有を認識可能な手法を提案する。本手法は、Pekarの談話関係解析による手法を参考にして、文章内共起を用いて一連の状況で一緒に起りやすい述語対を項の共有情報と共に獲得する。また項の要素集合の類似を見て、こちらでも述語対を項の共有情報と共に獲得する。この項の共有情報と、共起パターンを用いて認識した事態間関係を組み合わせると、項の共有情報を伴う事態間関係を獲得する。言い換えると、共起パターンで認識した事態間関係に、文章内共起を用いて認識した項の共有情報を割り当てる手法である。この手法は、関係認識手法に共起パターンを用いているため、広く他の事態間関係を認識することが可能であり、Torisawaの手法の欠点を克服している。同時に、文章内共起を用いて項の共有情報を認識しているため、項の共有情報を付与することが難しいという共起パターンの欠点を克服している。詳細は3.節で説明する。

### 3. 事態関係獲得

#### 3.1 概要

2つの相補的な手法を組み合わせることで述語間の関係を獲得する。ひとつ目の手法は事態間の関係を表す共起パターンに着目する手法で、ふたつ目の手法は述語対それぞれの項の間で共通する名詞に着目する手法である。提案手法の概要をFigure 1に示す。

提案手法は大きく2つの手続きに分かれる。「行為-効果」関係や「行為-手段」関係など、獲得したい述語対の関係のクラスが与えられたとき、最初に共起パターンを用いて候補となる述語対を獲得する。次に、ここで獲得した述語対の項と項の間で共有されている名詞(アンカー)を探し、アンカーが見付かれれば述語対は何らかの含意関係にあると見なす。見付からなければ述語対を候補から除く。このとき我々はインスタンス・アンカーとタイプ・アンカーの2種類のアンカーを用いた。

節4.の実験で、アンカーを用いて述語対候補をフィルタリングすることで精度が向上すること、2種類のアンカーを用いることで再現率が向上することを示す。

#### 3.2 述語対獲得

述語対獲得手法として共起パターンに基づくAbe [9]の手法を用いた。この手法はPantelら [4]が提案したEspressoと呼ばれる実体間関係獲得手法を事態間関係獲得に拡張した手法である。Espressoは信頼性の高い実体間関係の少数のシードインスタンスを入力として共起パターンを獲得し、この共起パターンを使って新たな実体間関係インスタンスを獲得する、ということを実トラップ的に繰り返し学習する関係獲得手法である。

##### 3.2.1 事態を表わす表現

コーパス中では様々な形式で事態が表現されている。本研究では動詞句の他に事態を含意する名詞句からも事態間関係を獲得する。事態を含意する名詞句は様々なものが、本実験においてはサ変名詞と接尾辞の組み合わせだけを対象とする。例えば、「掃除機できれいにする」から「掃除する→きれい」という述語対(行為-結果の関係)を獲得できる可能性がある。

##### 3.2.2 共起パターンの表現

Espressoは実体間のテキストを一般化したものを共起パターンに用いたが、事態間関係の場合は前述したように項が存在するため、係り受け関係に基づく共起パターンを用いた。

述語対が直接係り受け関係になっている場合と、述語対が任意の文節要素を介して係り受け関係になっている場合だけを事態間関係獲得の対象にした。共起パターンは事態の接尾辞(～者、～機、～中、…)や助詞(～が、～を、～に、～ために、…)や事態の間の任意の文節を含み、これを「日本語機能表現一覧」[10]を用いて一般化した。さらに、事態が行為(走る、食事をする、…)か出来事(風邪をひく、事故がある、…)かの区別も共起パターンに含ませた。この区別は後述の意志性辞書を利用した。

例えば「電話をかけたけれども通じない」では、「電話をかける」は動詞で行為、「通じる」は動詞で出来事なので、

「<verb;action>だけれども<verb;effect>ない」という共起パターンを獲得する。

### 3.3 共有項同定

前節(3.2)の手法で獲得した述語対が、どの項とどの項が共有関係にあるかを次に述べるアンカー情報を参照して同定する。

#### 3.3.1 インスタンス・アンカー

インスタンス・アンカーを利用した手法はアンカーを用いて動詞の含意関係を獲得する Pekar の手法を参考にしている。この手法は、同じ登場人物や対象物や場所を項とする動詞対は広く含意関係にあると仮定し、同じ談話文脈にある動詞を見付け、動詞の項の間で同じ名詞を共有するならばその動詞対は含意関係にあるとした。例えば Figure 1 の(2a)では、名詞「パン」は2回出現し、「焦げる」のガ格に「パン」があり、「焼く」のヲ格に「パン」がある。このとき2つの「パン」は同じものを指している。このように同じ物を指す語をアンカーと呼び、「焦げる」のガ格と「焼く」のヲ格は同じものを指している。

我々はこれをインスタンス・アンカーと呼ぶ。一方のタイプに基づくアンカーについては3.3.2で述べる。

インスタンス・アンカーを次の手順で利用する。

述語対  $Pred_1$  と  $Pred_2$  が与えられたとき、 $(Pred_1, Arg_1; Pred_2, Arg_2; Anc)$  をコーパスから次の条件を満たすように探す。

(a) Web ページに出現する  $Pred_1$  の項  $Arg_1$  となるようなアンカー  $Anc$  が名詞句の主辞

(b)  $Anc$  が先程と同じ Web ページに表われる  $Pred_2$  の項  $Arg_2$  となる

(c)  $Anc$  がストップリストに含まれていない。

(d)  $pmi(Pred_i, Arg_i) \geq -1.0$  for  $i \in \{1, 2\}$

実験では人手で作成した219語ストップリストを用いた。これは代名詞、数詞、「こと」、「もの」、「とき」のような一般的な名詞を含んでいる。

条件(d)の  $pmi(Pred_i, Arg_i)$  は  $Pred_i$  と  $Arg_i$  の間の pointwise mutual information である。これは係り受け解析器の誤りに起因する誤ったアンカーを除くためである。

Pekar は談話文脈の境界を同一段落内ではかつ談話解析の結果を用いて注意深く定義したが、我々は同じ Web ページにあるアンカーを共有している述語対は何らかの関係にあると仮定した。Pekar と比べて我々の談話文脈の境界が広過ぎるにもかかわらず、4.2で示すように我々の仮定は十分正確である。その理由は、3.2の手法による述語対と組み合わせることで、誤ったインスタンス・アンカーによる結果がフィルタリングされるためである。

最終的にそれぞれの項の対にあるアンカー集合を計算する  $Pred_1-Arg_1$  と  $Pred_2-Arg_2$  で得られた組を集める。

$$\text{AnchorSet}(Pred_1-Arg_1, Pred_2-Arg_2) = \{Arg | (Pred_1, Arg_1; Pred_2, Arg_2; Anc)\}.$$

#### 3.3.2 タイプ・アンカー

Figure 1 の文(3a)と文(3b)を例に説明する。2文とも述語「焼く」と述語「焦げる」を含む。(3a)では名詞「パン」が「焼く」のヲ格にあり、(3b)では同じ名詞「パン」が「焦げる」のガ格にある。この場合、名詞「パン」は焼くのヲ格と焦げるの

ガ格互いに共有していることを示すアンカーであると仮定する。このようなアンカーをタイプ・アンカーと呼ぶ。このように呼称する理由は、(3a)のパンと(3b)のパンは同じオブジェクトを指しているわけではないが、同じタイプに言及しているからである。

述語対  $Pred_1$  と  $Pred_2$  が与えられたとき、 $Pred_1$  と  $Pred_2$  が共起する文をコーパスから探し、項に  $Anc$  が出現する回数を計算する。

•  $Pred_1$  の項  $Arg_1$  が名詞  $Anc$  であれば、 $(Pred_1-Arg_1; Pred_2; Anc)$  の出現回数を増やす。

•  $Pred_2$  の項  $Arg_2$  が名詞  $Anc$  であれば  $(Pred_1; Pred_2-Arg_2; Anc)$  の出現回数を増やす。

項の対のアンカー集合  $Pred_1-Arg_1$  と  $Pred_2-Arg_2$  の項の集合の共通集合、すなわち

$$\text{AnchSet}(Pred_1-Arg_1, Pred_2-Arg_2) = S_1 \cap S_2,$$

$$S_1 = \{Arg | (Pred_1, Arg_1; Pred_2; Anc)\},$$

$$S_2 = \{Arg | (Pred_1; Pred_2, Arg_2; Anc)\}.$$

#### 3.3.3 アンカー集合の適用

以上の方法で収集したアンカー集合について、ある述語対がアンカーによってカバーされているかどうかはある述語対  $Pred_1$  と  $Pred_2$  が任意の  $Arg_1$  と  $Arg_2$  について、 $\text{AnchSet}(Pred_1-Arg_1, Pred_2-Arg_2)$  に含まれているかどうかで判断する。本節で説明している共有項同定の処理では、前節で獲得した述語対候補それぞれに次の手続きを行う。

(1) アンカーによってカバーされていない述語対を捨てる

(2) 頻度順に最大  $k$  個の項対を選ぶ(実験では  $k = 3$ )。

(3) それぞれの項対に対して、頻度順に最大  $l$  個のアンカーを選ぶ(実験では  $l = 3$ )。

## 4. 実験

### 4.1 設定

河原ら[11]がWebから収集したコーパス約5億文を利用した。これらの文をCaboCha[12]で係り受け解析して、述語対が共起する事例を抽出した。計算コストの問題から、述語対と共起パターンの頻度が20回未満の事例を除いた。

実験ではInui et al.[13]の4つの因果関係のうち行為-効果関係と行為-手段関係の2つを用いた。行為-効果関係は事象  $x$  と事象  $y$  の間で意志性のない事象  $y$  が起るように直接または間接的な意志的な事象  $x$  を行う。例えば、行為「Xが運動する」と出来事「Xが汗をかく」はこの関係になる。必然性がなくてもよい。例えば、「飲む」の結果として「二日酔いになる」は必然性がなく、評価者はこれを正解と評価する。一方で行為-手段関係は意志性のある事象  $x$  を行うためにしばしば意志性のある事象  $y$  を行う。例えば、「Xが走る」は「Xが運動する」ためにしばしば行なうことである。

実験では、意志性のラベルを伴う12000以上の動詞、これには8968の意志性動詞、3597の非意志性動詞、547の曖昧なものを含む、を人手で構築した。意志性動詞には「食べる」、「研究する」、非意志性動詞には「温まる」、「壊れる」、「悲しむ」などがある。意志性情報はパターンを用いたの述語対獲得の述語

表 2 関係分類の精度

	LSPs		covered by anchors		
	all	top-N	インスタンス	タイプ	組み合わせ
行為-効果	400	254	175	169	254
	304	202	167	154	238
(precision)	(0.76)	(0.79)	(0.95)	(0.91)	(0.93)
(recall)	(1.00)	(0.66)	(0.54)	(0.50)	(0.78)
行為-手段	400	254	176	178	254
	324	206	165	170	242
(precision)	(0.81)	(0.81)	(0.93)	(0.95)	(0.95)
(recall)	(1.00)	(0.63)	(0.50)	(0.52)	(0.74)

の素性として使う。

#### 4.2 結果と議論

##### 4.2.1 述語対獲得

拡張 Espresso アルゴリズムを、行為-効果関係で正例 25, 負例 4 のシード関係インスタンス, 行為-手段関係で正例 174, 負例 131 のシード関係インスタンスで走らせた。ランキングとセレクションを 40 回繰り返した後、行為-効果では 9,511 共起パターンと 22,489 関係インスタンス, 行為-手段では 14,119 共起パターンと 13,121 関係インスタンスを獲得した。獲得した事例を Table 1 に示す

4 つの部分 (1-500, 501-1500, 1501-3500, 3500-7500) から 100 個の述語対をランダムに選び、2 人の評価でそれぞれの述語対が正しいかを評価した。適切な共有項を付与することで適切な関係になる場合を正解とした (合計で 800 述語対)。例えば、「かける」と「つながる」を正解とした。kappa 値は行為-効果で 0.68, 行為-手段で 0.61。以降の評価では、片方の評価者の結果をゴールドスタンダードと見なす。3.2 の共起パターン手法を適用した結果を LSPs の all に示す。

3.3 節で説明したアンカーをもちた手法を先程の 800 述語対に適用した。Table 2 の “covered by anchors” に結果を示した。傾向としてどちらのクラスも大体同じなので行為-効果の結果に的をしばって説明する。

インスタンス・アンカー手法を適用した結果をテーブルの「インスタンス」に示した。インスタンス・アンカー手法は先程サンプリングした 400 述語対のうち 175 事態対をカバーした。そのうち 167 述語対が正しい関係であった。167 / 175 で精度を、167 / 304 で再現率を計算した。その図はインスタンス・アンカーを用いることで再現率を犠牲にして高い精度で正しい述語対を選んだことを示している。しかし、インスタンス・アンカーとタイプ・アンカーを組み合わせることで精度を犠牲にせず再現率を改善した。“covered by anchors” の「組み合わせ」にその結果を示す。インスタンス・アンカーを用いた場合もタイプ・アンカーを用いた場合も同じ傾向 (精度は高いが再現率が低い)、適度に異なる事例をカバーしているので組み合わせることで再現率の改善に寄与した。

まとめると、共起パターン手法で生成した述語対に対してアンカーを用いることで高い精度を得て、インスタンス・アンカーとタイプ・アンカーを用いることで適用することで再現率

を保つことができた。

##### 4.2.2 共有項同定

次に共有項同定の accuracy を示す。アンカーでカバーした前述の述語対 (254 対が行為-結果, 254 対が行為-手段) に対して、評価者 2 人に共有している項は正しいか尋ねた。結果を Table 3 に示す (インスタンス・アンカー手法とタイプ・アンカー手法とこの 2 つの組み合わせを比較している)。

“arg-strict” は事態間の関係が正しく、最大 3 つの共有項が全て正しい場合を正解とした。“arg-lenient” は事態間の関係が正しく、最大 3 つの共有項のうち 1 つでも正しい場合を正解とした。“anc-strict” は項対で共有されている最大 3 つの名詞の候補が全て正しい場合を正解とし、“anc-lenient” は項対で共有されている名詞の候補が 1 つでも場合を正解とし、“anc-any” は項対で共有されている名詞の候補を評価の対処としなかった (名詞候補が全て誤りでも正解とした)。評価者間の一致率は高く、“arg-strict” と “anc-any” の場合の kappa 値は行為-効果関係で 0.70, 行為-手段関係で 0.71 である。

行為-効果では再現率の分母を 202, 行為-手段では 206 として、近似的な再現率を Table 2 に示す。

本実験の設定では、述語対の関係が間違っていればどの評価においても誤りとしている。精度と再現率の上限は Table 2 で与えられる。“arg-” と “combined” の上限は精度 0.93, 再現率 0.78 である。“arg-lenient” と “combined” の上限は精度 0.91, 再現率 0.76 である。上限を考慮すれば、“arg-lenient” と “combined” と “anc-lenient” は精度 0.91, 再現率 0.76 で、我々の手法は高い精度を達成した。一方で “arg-strict” と “anc-strict” の結果はまだ問題があり、これは将来の課題である。

#### 5. まとめ

補補的な関係にある共起パターンを用いた手法とアンカーを用いた手法を用いて事態間関係を獲得した。最初に共起パターンを用いて所与の関係にある述語対候補を獲得する。次にインスタンス・アンカーとタイプ・アンカーの 2 種類のアンカーを用いて述語対候補をフィルタリングする。アンカーを用いてフィルタリングすることで精度が向上し、2 種類のアンカーを用いることで再現率を保つことができた。

今後の課題は大きく 2 つある。ひとつはより明確で意味のある評価を行うことである。例えばこの事態間関係知識をタスクに適用してその結果を評価する。もうひとつの課題は今回はアンカーを用いた手法を用いたが、この部分に照応解析やゼロ照応解析の手法を取り入れることである [14], [15, etc.]。

謝辞 「Web 上の 5 億文の日本語テキスト」の使用許可を下された情報通信研究機構の河原大輔氏と京都大学の黒橋禎夫氏に感謝いたします。

#### 文 献

- [1] S. Abe, K. Inui and Y. Matsumoto: “Acquiring event relation knowledge by learning cooccurrence patterns and fertilizing cooccurrence samples with verbal nouns”, Proceedings of the 3rd International Joint Conference on Natural Language Processing, Hyderabad, India, pp. 497-504 (2008).
- [2] T. Inui, K. Inui and Y. Matsumoto: “Acquiring causal

表 1 獲得した例

	Pred1	Arg1	Pred2	Arg2	Anc
行為-効果	開始する	ヲ	終了する	ガ	インストール, トランザクション
行為-効果	デザインする	ヲ	かわいい	ガ	ロゴ
行為-効果	寝る	デ	眠れる	デ	ベット, 布団
行為-手段	治療する	デ	処方する	ヲ	薬
行為-手段	治療する	ヲ	処方する	ニ	患者
行為-手段	帰宅する	デ	運転する	デ	車, 自動車
行為-手段	利用する	ヲ	コピーする	ヲ	ファイル, データ

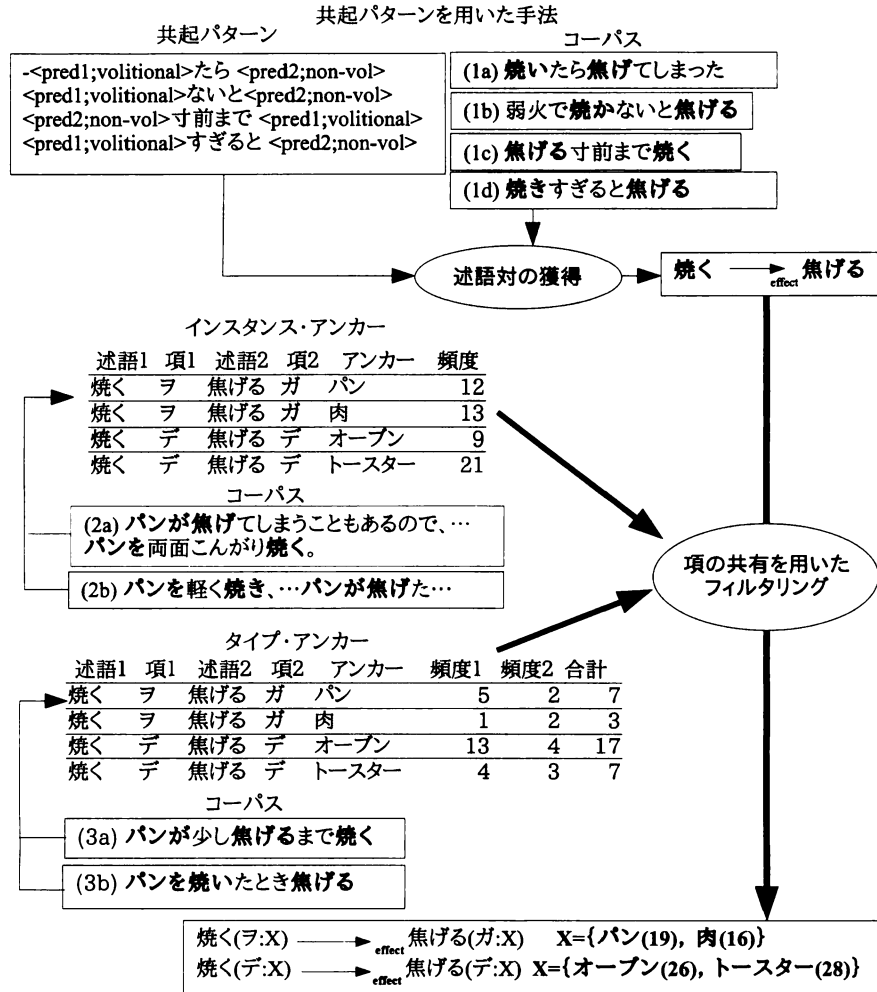


図 1 Two-phased event relation acquisition

表 3 共有項同定の精度

		行為-効果				行為-手段							
		anc-strict		anc-lenient		anc-strict		anc-lenient					
		prec	recall	prec	recall	prec	recall	prec	recall				
arg-strict	token	0.80	0.46	0.86	0.50	0.88	0.50	0.82	0.48	0.88	0.51	0.88	0.51
	type	0.76	0.42	0.80	0.44	0.82	0.45	0.86	0.50	0.88	0.51	0.88	0.51
	combined	0.76	0.64	0.82	0.69	0.83	0.70	0.83	0.70	0.87	0.73	0.87	0.73
arg-lenient	token	0.92	0.53	0.93	0.53	0.94	0.54	0.92	0.54	0.94	0.55	0.94	0.55
	type	0.84	0.47	0.86	0.48	0.87	0.48	0.90	0.52	0.92	0.53	0.92	0.53
	combined	0.89	0.74	0.91	0.76	0.92	0.76	0.92	0.78	0.94	0.78	0.94	0.79

- knowledge from text using the connective marker tame”, *ACM Transactions on Asian Language Information Processing (TALIP)*, 4, 4, pp. 435–474 (2005).
- [3] 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*, New York, NY, USA, ACM, pp. 323–328 (2001).
  - [4] P. Pantel and M. Pennacchiotti: “Espresso: Leveraging generic patterns for automatically harvesting semantic relations”, *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL*, pp. 113–120 (2006).
  - [5] V. Pekar: “Acquisition of verb entailment from text”, *Proceedings of the Human Language Technology Conference of the NAACL, Main Conference*, New York City, USA, Association for Computational Linguistics, pp. 49–56 (2006).
  - [6] I. Szpektor, E. Shnarch and I. Dagan: “Instance-based evaluation of entailment rule acquisition”, *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, Prague, Czech Republic, Association for Computational Linguistics, pp. 456–463 (2007).
  - [7] K. Torisawa: “Acquiring inference rules with temporal constraints by using japanese coordinated sentences and noun-verb co-occurrences”, *Proceedings of Human Language Technology Conference/North American chapter of the Association for Computational Linguistics annual meeting (HLT-NAACL06)*, pp. 57–64 (2006).
  - [8] I. Szpektor, H. Tanev, I. Dagan and B. Coppola: “Scaling web-based acquisition of entailment relations”, *Proceedings of EMNLP 2004* (Eds. by D. Lin and D. Wu), Barcelona, Spain, Association for Computational Linguistics, pp. 41–48 (2004).
  - [9] S. Abe, K. Inui and Y. Matsumoto: “Acquiring event relation knowledge by learning cooccurrence patterns and fertilizing cooccurrence samples with verbal nouns”, *Proceedings of the 3rd International Joint Conference on Natural Language Processing*, pp. 497–504 (2008).
  - [10] S. Matsuyoshi, S. Sato and T. Utsuro: “Compilation of a dictionary of japanese functional expressions with hierarchical organization”, *ICCPOL*, pp. 395–402 (2006).
  - [11] D. Kawahara and S. Kurohashi: “A fully-lexicalized probabilistic model for japanese syntactic and case structure analysis”, *Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 176–183 (2006).
  - [12] T. Kudo and Y. Matsumoto: “Japanese dependency analysis using cascaded chunking”, *Proceedings of the 6th Conference on Natural Language Learning 2002 (COLING 2002 Post-Conference Workshops)*, pp. 63–69 (2002).
  - [13] T. Inui, K. Inui and Y. Matsumoto: “What kinds and amounts of causal knowledge can be acquired from text by using connective markers as clues?”, *Proceedings of the 6th International Conference on Discovery Science*, pp. 180–193 (2003).
  - [14] R. Iida, K. Inui and Y. Matsumoto: “Exploiting syntactic patterns as clues in zero-anaphora resolution”, *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the ACL*, Association for Computational Linguistics, pp. 625–632 (2006).
  - [15] M. Komachi, R. Iida, K. Inui and Y. Matsumoto: “Learning based argument structure analysis of event-nouns in japanese”, *Proceedings of the Conference of the Pacific Association for Computational Linguistics*, pp. 120–128 (2007).