

文脈的アナロジーによるストーリー中の事象間の類似

山内 玲未[†]秋元 泰介[‡][†]九州工業大学情報工学部知能情報工学科[‡]九州工業大学大学院情報工学研究院知能情報工学研究系
akimoto@ai.kyutech.ac.jp

1. はじめに

物語の基本的な性質の一つとして、部分の意味が前後の文脈、すなわち他の部分ないし全体との相関のもとに定まることが挙げられる。例えば、「親が子を殺す」という事象でも、それが親による子の虐待という文脈に置かれた場合と、子から親への家庭内暴力に対する行為として置かれた場合とでは、多少なりとも異なる意味を帯びるだろう。

このような性質を適切に扱うことは、計算機上で物語を処理する上での基本的な問題である。このような観点から、本研究では、二つのストーリーそれぞれの中の一つの事象に注目したときに、その事象間の類似度を、前後の文脈を加味して計算するモデルを提案する。

この問題に対しては、いくつかの異なるアプローチが考えられる。物語理解の認知モデル(Zwaan & Radvansky, 1998)のように、物語の読解過程で生じる表象や記憶に基づいて新しい文の意味を理解するような仕組みも必要だろう。スキーマ的な構造(Schank & Abelson, 1977; Rumelhart, 1975)に基づいて部分の意味が定まるという考え方もある。それから、表面的に異なる事物の間の類似には、アナロジーの認知も深く関与するだろう(Gentner & Markman, 1995)。物語の認知は総合的にとらえる必要があるが、本研究では特にアナロジー的認知に焦点を合わせる。

2. 文脈的アナロジー

アナロジーの代表的な理論の一つとして、Gentner (1983)の structure mapping 理論やその計算モデル(Falkenhainer, Forbus & Gentner, 1989)が挙げられる。この理論は、アナロジーを、ストーリーや状況、物理的・機械的なシステム等の構造的な表象間に、要素レベルではなく、要素間の「関係」の水準で整合的な対応関係を構成することとして説明している。本研究が提案する文脈的アナロジーは、主としてGentnerらの理論に依拠しているが、いくつか異なる点がある。

まず、従来のアナロジーのモデルは、二つの構造的な表象の全体と全体の間に対応関係を構成する仕組みを扱っている。それに対して文脈的アナロジーは、各ストーリーの「部分」に注目した時に生じる前後の文脈的な構造間に対応関係を構成する。

それから、従来のアナロジーの計算モデルの多くが、予め関係構造が付与された表象を扱っているのに対して、本研究で計算モデルに与えるストーリーは、事象の列のみとする。事象間の関係に相当する構造は、計算過程でプログラムが自ら構成する。また、部分間の対応や類似の計算に、単語の分散表現モデルを用いることも、提案モデルの一つの特徴である。

分散表現に基づく言語モデルや、構造化されていない表象を用いるという点では、Turney (2008)が Latent Relational Analysis に基づくアナロジーモデルを提案している。しかし、それが扱うのは何らかの対象(例えば太陽系と原子)を表す単語集合の間に対応関係を構成するという問題であり、ストーリーの文脈的な構造を扱うものではない。

なお、提案モデルでは、二つのストーリーの間に base と target のような非対称的關係は設けずに、ストーリー間に対応関係を構成することをアナロジーあるいはマッピングと呼ぶ。

3. 計算モデル

類似度の計算対象となる事象(焦点事象と呼ぶ)が指定された二つのストーリーを入力とし、焦点事象間の類似度を0~1の数値として算出するモデルを構築した。プログラムはC++言語により実装されている。

3.1 ストーリーの表現形式

このモデルが扱うストーリーは、事象の時系列順の並びであり、 $S = (E_1, \dots, E_n)$ という形で表現される。この中の焦点事象を E_c と表す。後述する計算の中で二つのストーリーを比較する際は、一方にプライム記号を付け、 S' や E'_c のように表記する。

事象は格フレーム形式で表現する。すなわち、動詞(行為・動作を表すサ変名詞も含むが、以下では単に動詞と呼ぶ)を中心要素とし、一つ以上の名詞が動詞に対する意味役割(格)を明記した形で記述される。例えば、「親が子を殴る」という意味の事象は、[殴る (agent 親) (counter-agent 子)]と表現される。このプログラムが扱える格は agent (動作主), counter-agent (被動作主), object (対象), instrument 道具), goal (目標地点), from (出発地点), location (場所)の7種類である。

3.2 文脈構造

提案モデルは、Gentnerらの理論を踏襲して、アナロジーを、要素の上位に位置する関係レベルの対応関係に基づく認知であると仮定する。ストーリーにおける関係レベルの構造を階層的に捉えると、事象中の格要素(人物や物を表す名詞)にかかる関係に相当するのは動詞(動作・行為)であり、動詞にかかる関係はそれと他の動詞との関係であるという見方ができる。言い換えると、ストーリー中の人物や物の位置付けはそれが関与する動作・行為によって定まり、各動作・行為の意味は、前後の動作・行為との関係によって定まるということである。

以上の考え方に基づいて、提案モデルはまず、焦点事象を中心として、それに関係する前後の事象の動詞部分を結び付ける。こうして形成される構造を「文脈構造」と呼び、

$$C = \langle E_c, P, F, L \rangle \quad (1)$$

と定義する．ここで， P は焦点事象よりも前から引き寄せられた動詞の集合であり，これを「過去空間」と呼ぶ． F は焦点事象よりも後から引き寄せられた動詞の集合であり，これを「未来空間」と呼ぶ．ここで引き寄せられる動詞は，焦点事象の格要素と同一の名詞が含まれる事象の動詞部に限られる． L は，焦点事象中の格要素と過去空間及び未来空間の動詞を結ぶリンクの集合である．リンクには，格要素と動詞の関係を表すラベルとして，agent または not-agent，すなわち，当の人物や物が動詞に対して agent として関与するか，agent 以外の格として関与するかを表す記号が付与される．なお， P と F それぞれにおいて，重複する動詞は一つにまとめられ，リンクのラベルは先に現れた事象における関係のみが反映される．

図1に文脈構造の例を示す．焦点事象は[殺す (agent 親) (counter-agent 子)]という事象であり，過去空間の「蹴る」及び「殴る」という動詞が，「親」に対しては not-agent，「子」に対しては agent の関係で結び付けられている．また，未来空間においては「自首」が「親」に対して agent 関係で結合されている．

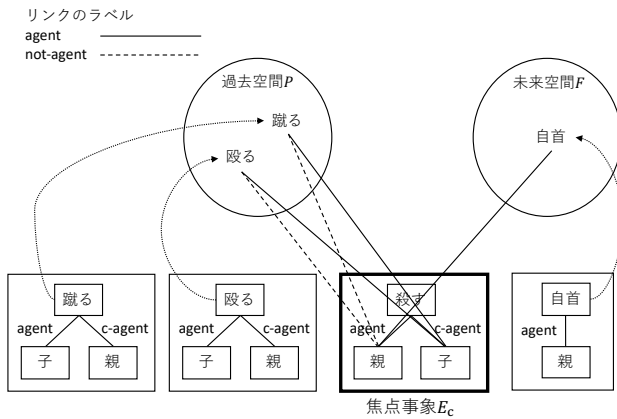


図1: 焦点事象を中心とする文脈構造

3.3 類似度の計算

焦点事象間の類似度を，文脈構造間のマッピングに基づいて計算する．計算の過程で，単語間の類似度を計算する場面があるが，そこでは単語の分散表現モデルの一つである word2vec (Mikolov et al., 2013) の skip-gram モデル (300次元, window 10, negative sampling) と，現代日本語書き言葉均衡コーパス(Maekawa et al., 2014)により構築した，175,801の単語ベクトルを使用する．

3.3.1 焦点事象間のマッピングと類似度

まず，焦点事象 E_c と E'_c の間に，格フレームの構造に基づく要素間のマッピングを形成する．事象間のマッピングは，動詞対及び格要素対の集合とし， $M^E = \{m_i^E\}$ と表す．一つの事象は一つの動詞と一つ以上の格要素(名詞)から構成されるため， m_1^E は動詞対， m_2^E 以降は格の種類が一致する格要素の対となる．焦点事象間で格の対応が取れない部分はマッピングに含まれない．なお，agent を必須格と仮定し， m_2^E を agent 間の対とする．

そして，焦点事象間の類似度を， M^E に基づく動詞対及び個々の格要素対の類似度の平均とし，

$$\text{sim}_E(E_c, E'_c) = \frac{1}{|M^E|} \left(\text{sim}_V(m_1^E) + \sum_{i=2}^{|M^E|} \text{sim}_N(m_i^E) \right) \quad (2)$$

と定義する．ここで sim_V は動詞対の類似度， sim_N は格要素対の類似度である．これらは，後述する文脈構造間のマッピングに基づいて計算される．

3.3.2 文脈構造間のマッピング

文脈構造間のマッピングにおいては， M^E の各対に注目したときに，それぞれ異なる対応関係が構成される．

まず，焦点事象間の格要素対 m_i^E に注目したときに生じる過去空間及び未来空間のマッピングをそれぞれ $M^{PN_i} = \{m_j^{PN_i}\}$ ， $M^{FN_i} = \{m_j^{FN_i}\}$ と表す． M^{PN_i} は， $P - P'$ 間における，注目格要素とリンクで結ばれた動詞の集合間の1対1マッピングであり， M^{FN_i} は $F - F'$ 間におけるそれである．これらのマッピングは，動詞間の類似度(単語ベクトル間のコサイン類似度)に基づいて，次のように構成される．

- i. 動詞集合間の動詞対を全通り探索し，類似度が最大でありかつ類似度が閾値 θ 以上の動詞対を抽出し，マッピングの要素とする．
- ii. 類似度が θ 以上の動詞対がなくなるまで処理iを繰り返す．

これにより， M^{PN_i} 及び M^{FN_i} はそれぞれ，0個以上の動詞対が類似度の高い順に並んだものとなる．図2に文脈構造間のマッピングの例を示す．焦点事象の agent 格の対に注目した時に生じる M^{PN_2} と M^{FN_2} を表している．

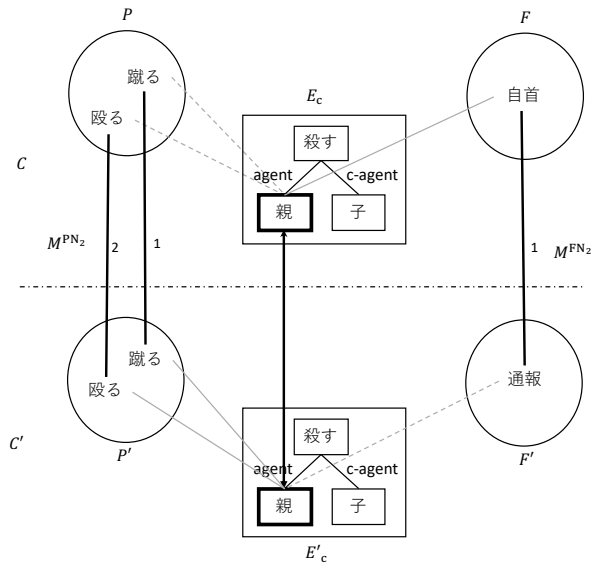


図2: agent 格の対に注目した文脈構造間マッピング

一方，動詞対 m_1^E に注目したときに生じる過去空間及び未来空間のマッピングをそれぞれ $M^{PV} = \{m_j^{PV}\}$ ， $M^{FV} = \{m_j^{FV}\}$ と表す．これらは，動詞間ではなく，焦点事象の動詞と過去空間または未来空間の動詞との間に生じる，「動詞間関係」のマッピングである．これらのマッピングは，

焦点事象の agent 格が関与する動詞に注目して構成する。すなわち、agent 対に注目した時に生じる M^{PN_2} と M^{FN_2} の各動詞を、焦点事象の動詞との関係に置き換えたものとする。例えば、 $M^{PN_2} = \{(v_1, v'_2), (v_2, v'_1)\}$ ならば、 $M^{PV} = \{(v_1: v_c, v'_2: v'_c), (v_2: v_c, v'_1: v'_c)\}$ となる（動詞間関係をコロンにより表す）。図3は、動詞対に注目した時に生じる文脈構造間のマッピングの例である。

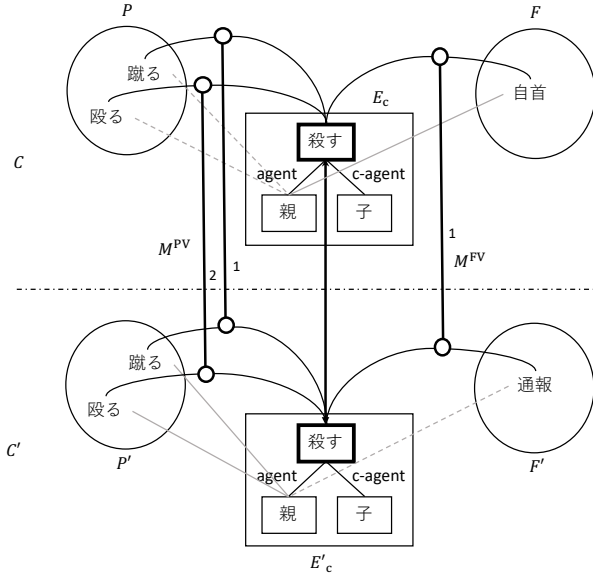


図3: 動詞対に注目した文脈構造間マッピング

3.3.3 事象要素間の類似度

焦点事象間のマッピング M^E における各対の類似度は、その対に注目した時の過去空間と未来空間それぞれの類似度の重み付き平均とし、格要素対と動詞対について、それぞれ以下の式により計算する。

$$\text{sim}_N(m_i^E) = \frac{\text{sim}_A(M^{PN_i})\alpha_P + \text{sim}_A(M^{FN_i})\alpha_F}{\alpha_P + \alpha_F} \quad (3)$$

$$\text{sim}_V(m_1^E) = \frac{\text{sim}_A(M^{PV})\alpha_P + \text{sim}_A(M^{FV})\alpha_F}{\alpha_P + \alpha_F} \quad (4)$$

ここで、 $\text{sim}_A(M^x)$ は、過去空間または未来空間のマッピングに基づく類似度であり、 α_P と α_F は各類似度にかかる重みである。

過去空間及び未来空間の類似度は、マッピングにおいて対になった動詞間または動詞間関係間の類似度をもとに、次の式により計算する。

$$\text{sim}_A(M^x) = \zeta \left(\sum_{j=1}^{|M^x|} \cos(m_j^x) r_j \omega^j \right) \quad (5)$$

ここで、 $\cos(m_j^x)$ は、 m_j^x が動詞対の場合は動詞のベクトル間のコサイン類似度、 m_j^x が動詞間関係対の場合は、動詞間関係をベクトルの差とするコサイン類似度、すなわち $\cos(v_p - v_c, v'_q - v'_c)$ である。それにかかる r_j は、過去空

間または未来空間の動詞対 m_j^x と、焦点事象の格要素対 m_i^E ($\text{sim}_V(m_i^E)$ の計算時は agent 対 m_1^E) を結ぶリンクのラベル (agent か not-agent か) が、ストーリー間で異なる場合は -1、同じ場合は 1 となる。これは、「(子)を殴った親」と「(子)に殴られた親」のように、ストーリー上での動詞と名詞の関係の相違が対照的な意味を生じさせるという仮定に基づく。それから、 ω は 1 未満の正数であり、これにより、マッピングの順位 j が下がるほど類似度に対する重みが減少する。これは、マッピングが多数生じるほど類似度が高まるのに対して、それがストーリーの長さ強く依存することを避けるための仕組みである。最後に、符号と重みのかかった各対の類似度の総和に対して、シグモイド関数

$$\zeta(x) = \frac{1}{1 + e^{-ax}} \quad (6)$$

が適用されて、0~1 の値が得られる。式6中の a は変数であるが、後述する実験を通じた経験的な値として、 sim_N の計算時は $a = 3$ 、 sim_V の計算時は $a = 2$ とする。

4. 実験

焦点事象が定められた複数のストーリー対を用意し、提案モデルによる類似度と、人の評定による類似度との比較を行った。

4.1 ストーリー対の作成

実験用のストーリー対を作成するために、ストーリー中の事象間の類似の種類として、次の2つのカテゴリを設けた。

- A) 文脈類似かつ単体非類似: 焦点事象を単体として比較しても似ていないが、前後の文脈の中での意味合いは似ている。
- B) 文脈非類似かつ単体類似: 焦点事象を単体として比較すると似ているが、前後の文脈の中での意味合いは似ていない。

上記の各カテゴリに当てはまるストーリーを作るための指針として、ブロップ (1987) の昔話の形態学の考え方を参考にし、焦点事象のストーリー中での機能的なカテゴリが共通するものを A、対照的なものを B とした。本実験では「発端」「結末」「加害(問題)」「解消」「妨害」「補助」という6種類の機能的カテゴリを定めた。

以上の枠組みに基づいて、A と B に該当するストーリー対をそれぞれ6つ作成した。その内訳を表1に示す。一つのストーリーの事象数は7~10程度とした。カテゴリAとBそれぞれの例(文表現)を表2と表3に示す。

表1: 実験に使用したストーリー対

A		B			
No.	E_c	E'_c	No.	E_c	E'_c
2	発端	発端	1	加害	解消
4	解消	解消	3	補助	妨害
6	妨害	妨害	5	発端	結末
8	解消	解消	7	加害	解消
10	妨害	妨害	9	妨害	補助
12	発端	発端	11	発端	結末

表 2: カテゴリ A の例 (No. 2)

S	S'
父親が弟にお菓子を与える。	鹿がタヌキに木の実を貰う。
兄が弟を羨む。	キツネがタヌキに木の実を強請る。
兄が弟からお菓子を奪う。	タヌキが断る。
兄が弟と喧嘩する。	キツネがタヌキに噛み付く。
母親が兄にお菓子を与える。	鹿がキツネを止める。
兄が弟にお菓子を返す。	鹿がキツネに木の実を分ける。
兄がお菓子を食べる。	キツネがタヌキに謝る。
弟がお菓子を食べる。	キツネが鹿と木の実を食べる。
母が微笑む。	

表 3: カテゴリ B の例 (No. 1)

S	S'
親が子を蹴る。	子が親を蹴る。
親が子を殴る。	子が親を殴る。
親が子を殺す。	親が子を殺す。
医者が親を通報する。	親が後悔する。
警察が親を逮捕する。	親が自首する。
親戚が親を罵る。	警察が親を逮捕する。
祖父が困る。	親戚が親に同情する。

4.2 人の評定データの収集

人による類似度の評定データを収集するために、アンケート調査を行った。アンケートは E メールにより個別に配布し、Web 上のフォームから回答するようになっている。情報工学系の大学生 23 名にアンケートを送付し、全員から回答が得られた。

回答者には、表 2 や表 3 と同様の形で、12 のストーリー対が一覧できる形で提示される。焦点事象は下線により示される。回答者は、それらを見ながら、12 のストーリー対を a) 似ている、b) 中間、c) 似ていない、という 3 ランクに均等に (4 つずつ) 振り分ける。その際、回答者には「下線部に注目して、その文のストーリー中での (前後の文脈中での) 意味合いがどの程度似ているか」を評定するよう指示が与えられる。なお、最後に、どのような基準で評定を行ったかを自由記述形式で記入する欄も設けたが、この回答は本稿では扱わないこととする。

4.3 結果

各ストーリー対に対する類似度の値と順位、及び人-モデル間における値と順位それぞれの相関係数を表 4 にまとめる。人の評定値は、ランク a を 3、b を 2、c を 1 に換算した平均値である。提案モデルのパラメータ設定は、文脈構造間マッピングにおける閾値 $\theta = 0.2$ 、類似度計算 (式 3, 4) における過去と未来の重み $\alpha_P = 1.2$ 、 $\alpha_F = 0.8$ 、式 5 における重み $\omega = 0.8$ とした。参考までに、文脈構造を用いずに、焦点事象単体で計算した類似度も比較対象に加えた (単体比較)。この値は、焦点事象間のマッピング M^E における各単語対のコサイン類似度を平均したものである。

結果を見ると、人の評定値は予想通りカテゴリ A の方が高くなる傾向にあり、文脈構造を用いない単体比較の計算値は B の方が高くなる傾向があった。それに対して、提案モデルと人の間には、値と順位の両方において、やや強い正の相関が見られた。よって、ストーリー中の事象間の類似の計算において、文脈的アナロジーが重要な働きをすることができると言うことができる。

表 4: 人の評定値とモデルの計算値の比較

ストーリー No.	人		提案モデル		単体比較	
	値	順位	値	順位	値	順位
1 [B]	1.783	10	0.335	12	1.000	1
2 [A]	2.043	5	0.701	1	0.302	7
3 [B]	1.174	12	0.417	10	0.242	10
4 [A]	2.043	5	0.687	2	0.606	2
5 [B]	1.870	8	0.613	5	0.471	4
6 [A]	2.565	2	0.615	4	0.290	9
7 [B]	1.913	7	0.539	8	0.334	6
8 [A]	2.739	1	0.599	6	0.232	11
9 [B]	1.609	11	0.399	11	0.339	5
10 [A]	2.348	3	0.573	7	0.231	12
11 [B]	2.087	4	0.631	3	0.487	3
12 [A]	1.870	8	0.467	9	0.300	8
相関係数	-	-	0.590	0.649	-0.197	-0.361

5. むすび

本稿では、ストーリー中の事象間の類似度を、当の事象そのものではなく、その前後の文脈構造間のアナロジー的なマッピングに基づいて計算するモデルを提案した。実験用に作成したストーリー対を用いて、提案モデルの計算値と人の評定値の比較を行った結果、ストーリー中の事象間の類似の計算における文脈的アナロジーの有効性が示された。

今後の課題は、モデルの振る舞いをより詳細に分析し、仕組みの問題点や、このモデルが上手く扱うことができないストーリー対の特徴を明らかにすることである。

謝辞 本研究は JSPS 科研費 JP18K18344 の支援を受けた。

参考文献

- Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence*, 41, 1-63.
- Gentner, D. (1983). Structural-mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2), 155-170.
- Gentner, D., & Markman, A. B. (1995). Similarity is like analogy: Structural alignment in comparison. In C. Cacciari (Ed.), *Similarity in Language, Thought, and Perception* (pp. 111-147). Brepols.
- Maekawa, K., Yamazaki, M., Ogiso, T., Maruyama, T., Ogura, H., Kashino, W., Koiso, H., Yamaguchi, M., Tanaka, M., & Den, Y. (2014). Balanced corpus of contemporary written Japanese. *Language Resources and Evaluation*, 48(2), 345-371.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Proc. 26th International Conference on Neural Information Processing Systems* (pp. 3111-3119).
- ブロップ, ウラジーミル (著), 北岡誠司, 福田美智代 (訳) (1987). 昔話の形態学. 水声社. (原著 1928 年)
- Rumelhart, D. E. (1975). Notes on a schema for stories. In D. G. Bobrow & A. Collins (Eds.), *Representation and Understanding: Studies in Cognitive Science*. Academic Press.
- Schank, R. C., & Abelson, R. P. (1977). *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures*. Lawrence Erlbaum.
- Turney, P. D. (2008). The latent relation mapping engine: Algorithm and experiments. *Journal of Artificial Intelligence Research*, 33, 615-655.
- Zwaan, R. A., & Radvansky, G. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2), 162-185.