

文内最適化による時間的順序関係推定

吉川 克正

日本アイ・ビー・エム(株) 人間文化研究機構 国立国語研究所
東京基礎研究所
katsuy@jp.ibm.com

浅原 正幸

コーパス開発センター
masayu-a@ninjal.ac.jp

飯田 龍

東京工業大学
大学院情報理工学研究科
ryu-i@cl.cs.titech.ac.jp

1 はじめに

時間情報解析は、事象・時間表現の抽出、事象・時間表現の属性同定、時間情報正規化、時間的順序関係推定など、いくつかの部分問題に分割される。中でも時間的順序関係推定は、近年、特に盛んに研究が行われている問題である。時間的順序関係は、因果関係との関連が密接であり、それを同定することは、深い意味理解を行うために不可欠な問題だと考えられている。時間的順序関係推定は、TempEval-1,2 [13, 14] などの評価型ワークショップを中心にして、様々な手法が提案されてきた。

TempEval-2 では時間的順序関係推定問題を種類別に以下の4タスクに分割する。同一文内にある事象と時間表現間の順序関係推定 (Task C)、文書作成日時 (DCT) と事象との順序関係推定 (Task D)、隣接2文にある主事象間の順序関係推定 (Task E)、同一文内における事象間の順序関係推定で、1つの事象がもう1つの事象に対して直接係り受け関係になる場合 (Task F)。図1はTempEval-2のデータにおける時間的順序関係の例である。ここには5つの事象と1つの時間表現があり、先に述べた4種類の時間的順序関係が示されている。この図にある時間的順序関係は、表1にまとめられており、この表にあるような時間関係ラベル (BEFORE, OVERLAP, AFTER など) を推定するのが時間的順序関係推定問題である。

表 1: Temporal Relations (TLINKs) in Figure 1

task	relation
Task C	e53 (change) OVERLAP t10 (a couple of years)
Task D	e50 (think) OVERLAP t0 (DCT)
Task D	e52 (think) OVERLAP t0 (DCT)
Task D	e53 (change) AFTER t0 (DCT)
Task E	e50 (think) OVERLAP e57 (reposition)
Task F	e50 (think) OVERLAP e51 (gloomy)
Task F	e52 (think) BEFORE e53 (change)

初期の研究では、時間的順序関係推定は局所的な分類問題として扱われていたが [1, 7], 近年の研究では、この問題を全体最適化問題として捉え、文書全体とし

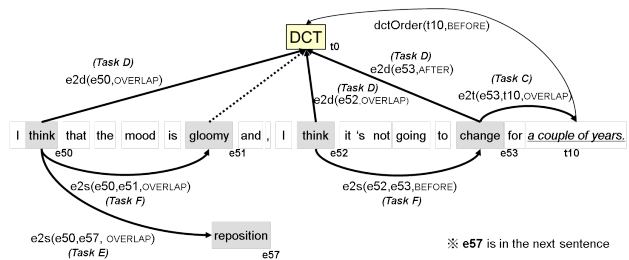


図 1: TempEval-2 における時間的順序関係

て最適な関係ラベルの配置を見つける手法が提案されている。以下、本稿では、前者のモデルを局所モデル、後者のモデルを大域モデルと呼ぶことにする。全体最適化手法には整数線形計画法を利用した手法 [2, 3] や、Markov Logic を利用した手法がある [15, 5]。

本研究では、文書全体ではなく、文単位で順序関係を最適化する全体最適化手法を新たに提案する。このような文内最適化モデルを提案する動機は2つある。まず、文内最適化モデルが文書最適化モデルと比較して単純であり、最適化問題が縮小されることで、様々な素性や制約を扱いやすくなることは大きな長所である。次に種類が異なる順序関係には、それぞれ異なった最適化方法が本来必要と考えられるからである。つまり、順序関係の種類によっては、文書全体よりも小さな範囲での最適化の方が望ましいものと仮定できる。例えば次のような仮説を考えることができる。

- Task C と F では係り受け関係などの統語的な情報が利用できるため、文内最適化が十分に機能する
- Task D の場合には時制の情報が直接効くため局所モデルでもよい
- Task E には統語的な情報が少ないため、文書最適化が必要となる

このような動機のもと、本研究の文内最適化モデルは、TempEval-2 のタスクのうち、Task C,D,F を最適化して解くことができる。実験では、この3つのタスクに対し、文内最適化モデルがどのように働くかを検証する。

2 Markov Logic による文内最適化モデル

本研究では、最適化モデルの実装に Markov Logic [10] を利用する。Markov Logic は一階述語論理と Markov Network を組み合わせた手法である。本来矛盾を許容しない一階述語論理に対し、確率的な重み付けを行うことで、ある程度の矛盾を許容する枠組みと言える。既にいくつもの自然言語処理タスクにおいて応用されている [11, 9, 8].

Markov Logic モデルの構築では、まず述語の定義を行う。ここでは4つのタスクそれぞれに対応するよう、表2の4種類の**潜在述語**を定義する。潜在述語はモデルによって推定したい情報を表現しており、正解を与えるのは学習時だけである。

次に**観測述語**を定義するが、これは学習時及び推定時に素性となる情報を表現する。例えば、表2以外の“推定されない”順序関係も観測述語となる。本研究では $dctOrder(t, R)$ がそれにあたり、文書作成日時と、時間表現 t の関係が R であることを表現する。図1にはこれらの時間的順序関係と述語の対応も図示されている。また、他の観測述語については、表3にまとめている。

以降の節では、本研究で提案する局所モデル及び、大域モデル（文内最適化モデル）について述べる。

2.1 局所モデル

局所モデルは局所的素性のみを利用し、個々の順序関係を独立に局所分類問題として解くモデルである。Markov Logic において局所的素性は**局所論理式**で表される。局所論理式は、ただ1つの潜在述語を持つ論理式で、ゆえに同時に推定する時間的順序は1つだけである。それに対し、2つ以上の潜在述語を持ち、同時に複数の時間的順序を推定するのが**大域論理式**になり、こちらについては次の2.2節で述べることにする。

局所的素性は先行研究 [12, 6] を参考にし、表3の観測述語を各タスクに対応した潜在述語と組み合わせて表現する。例えば、時制を表現する $tense(e, te)$ を利用して、Task F の素性を実装する場合には、次のような局所論理式を導入する：

$$tense(e1, +te1) \wedge tense(e2, +te2) \Rightarrow e2s(e1, e2, +R) \quad (1)$$

この論理式は、 $(e1 \times e2)$ の2つの事象に対する時制の組み合わせを素性とする。ここで、“+”記号が表すのは、この式(1)を素性テンプレートとして、定数を割り当てて素性を展開した時、展開される**基底論理式**は、各々異なる重みを持つことを表している。式(1)は、あらゆる

表 2: 潜在述語と対応する時間的順序関係

タスク	述語	説明
C	$e2t(e, t, R)$	事象 e と時間表現 t の関係は R
D	$e2d(e, R)$	事象 e と文書作成日時との関係は R
E	$e2e(e1, e2, R)$	隣接2文の主事象 $e1$ と $e2$ の関係は R
F	$e2s(e1, e2, R)$	同一文内の事象 $e1$ と $e2$ の関係は R

表 3: 各タスクに対する観測述語

素性	述語	C	D	E	F
EVENT-class	$class(e, c)$	Y	Y	$e1 \times e2$	$e1 \times e2$
EVENT-tense	$tense(e, te)$	Y	Y	$e1 \times e2$	$e1 \times e2$
EVENT-aspect	$aspect(e, a)$	Y	Y	$e1 \times e2$	$e1 \times e2$
EVENT-polarity	$polarity(e, p)$	Y	Y	$e1 \times e2$	$e1 \times e2$
EVENT-stem	$stem(e, s)$	Y	Y	$e1 \times e2$	$e1 \times e2$
EVENT-word	$wordEvent(e, w)$	Y	Y	Y	Y
EVENT-POS	$eventPos(e, p)$	Y	Y	$e1 \times e2$	$e1 \times e2$
TIME-type	$type(t, ty)$	Y	Y	Y	Y
TIME-value	$value(t, ty)$	Y	Y	Y	Y
TIME-word	$wordTime(e, w)$	Y			
TIME-POS	$posTime(t, p)$	Y			
TIME-DCtOrder	$dctOrder(t, r)$	Y	Y	Y	Y
Dep-Word	$depWord(e, w)$	Y	Y	Y	
Dep-POS	$depPos(e, p)$	Y	Y	Y	
Dep-Label	$dep(e1, e2, l)$	Y			Y
SR-Word	$srlWord(e, w)$	Y	Y	Y	Y
SR-POS	$srlPOS(e, p)$	Y	Y	Y	Y
SR-Label	$srl(e1, e2, l)$	Y			Y

時制と順序関係の組み合わせを展開する。例えば、

$$tense(e1, PRESENT) \wedge tense(e2, FUTURE) \Rightarrow e2s(e1, e2, BEFORE) \quad (2a)$$

$$tense(e1, PRESENT) \wedge tense(e2, PRESENT) \Rightarrow e2s(e1, e2, BEFORE). \quad (2b)$$

これら2つの論理式は式(1)から自動で展開されるもので、学習の結果、それぞれ異なる重みが付与されることになる。図1の事例から学習するならば、式(2a)が事例と一致するので、相対的に高い重みを獲得することになる。

2.2 文内最適化モデル

全体最適化を利用する時間的順序関係手法は、文書全体での最適化が行われてきたが、局所モデルと比較して必ずしもうまくいくとは限らず、また問題のサイズが大きくなるため、適切な素性や制約を見つけ出すことが困難になる傾向がある。そこで本研究では、文内最適化モデルという、限定的な全体最適化を行う手法を提案する。これにより最適化すべき問題が小さくなり、素性や制約を扱い易くなる。文内最適化モデルが扱うことのできるタスクは TempEval-2 の Task C, D, F である。よって、潜在述語 $e2t$, $e2d$, $e2s$ の3種類に対し、個々の局所論理式と、潜在述語間の整合性を保つための大域論理式を加えたモデルが文内最適化モデルになる。ここでは、大域論理式について述べることにする。

まず, Task C と D の間には,

$$dctOrder(t, +R1) \wedge e2t(e, t, +R2) \Rightarrow e2d(e, +R3) \quad (3)$$

$$dctOrder(t, +R1) \wedge e2d(e, +R2) \Rightarrow e2t(e, t, +R3) \quad (4)$$

の2つの大域論理式が成立する. これは $dctOrder$, $e2t$, $e2d$ の3つの関係の間にある推移関係を利用して, その論理的整合性を維持する式である. しかし, これらの推移関係は必ず成立するものではないため, その強さは学習によって決められるよう, *soft* 制約を利用する. 強い制約に違反する順序関係の配置は禁止されていないが, 違反による罰則は大きく, 解として選ばれる確率は低くなる. 逆に弱い制約は違反しても大きな罰則にならないため, 解に対する影響も小さくなるのである.

Task D と F についての大域的制約は,

$$e2d(e1, +R1) \wedge e2d(e2, +R2) \Rightarrow e2s(e1, e2, +R3), \quad (5)$$

$$e2d(e1, +R1) \wedge e2s(e1, e2, +R2) \Rightarrow e2d(e2, +R3) \quad (6)$$

これは $e2d$ と $e2s$ の間の整合性を確保する. Task D と F の推定精度には大きな開きがあるため, 相対的に精度の低い Task F にとって効果的な制約である.

Task C と F については, 次の大域的制約が定義できる.

$$e2t(e1, t, +R1) \wedge e2t(e2, t, +R2) \Rightarrow e2s(e1, e2, +R3), \quad (7)$$

$$e2t(e1, t, +R1) \wedge e2s(e1, e2, +R2) \Rightarrow e2t(e2, t, +R3) \quad (8)$$

これは, 事象が時間表現と結びつくことが多く, その時間表現を複数の事象が共有している場合に意味がある.

さらに, 意味役割を利用した大域的制約についても紹介しよう. ある事象が別の事象と同じ時間を共有している時, 文書作成日時との関係 $e2d$ も同じになり易いと考えられるので, 次のような制約を実装することができる.

$$\begin{aligned} & srl(e1, t, AM-TMP) \wedge srl(e2, t, AM-TMP) \\ & \Rightarrow e2d(e1, R1) \wedge e2d(e2, R2) \wedge R1 = R2. \end{aligned} \quad (9)$$

同様に, $e2t$ についても, 2つの事象が同じ時間表現を“AM-TMP”に持つ時, ある事象と時間表現の関係はもう一方の事象との関係に影響を与えると仮定できる.

$$\begin{aligned} & srl(e1, t, AM-TMP) \wedge srl(e2, t, AM-TMP) \\ & \wedge e2t(e1, t, +R1) \Rightarrow e2t(e2, t, +R2). \end{aligned} \quad (10)$$

このように, 文内最適化であれば, 他のタスクで提案されているような, 複雑な制約や素性を加えることもできるものと考えられる.

表 4: 実験結果

モデル / チーム	C	D	E	F
局所モデル	0.65	0.75	0.55	0.52
文内最適化モデル	0.67	0.74	-	0.55
文書最適化モデル	(0.65)	(0.76)	(0.57)	(0.56)
TRIOS*	0.65	0.79	0.56	0.60
TIPSem	0.55	0.82	0.55	0.59
TRIPS*	0.63	0.76	0.58	0.59
NCSU-indi	0.63	0.68	0.48	0.66

3 実験と結果

3.1 実験設定

本研究では次の2点に焦点を当て, 実験を行った. (1) 提案手法が時間的順序関係推定の性能を向上させられるか. (2) 提案手法の性能は TempEval-2 の参加チームと比較してどの水準にあるか.

まず実験設定について概説する. 本研究では TempEval-2 で提供された training と test のデータセットを利用する. TempEval-2 は複数言語のデータセットを持つが, 本研究では英語のみを利用する. training データは学習セットと開発セットの2つに分割し, パラメータの調整を行うものとする. その結果, 学習, 開発, テストの各セットが含む文書数は, 各々147, 15, 20 文書となった.

素性の抽出のために以下のツールを利用した. 品詞タグ付けには stanford-POS-tagger¹を, 構文解析及び意味役割割付与には LTH semantic parser²を利用し, そして Markov Logic エンジンとして, *Markov thebeast* を利用している. これは自然言語処理のタスク向けに調整された Markov Logic エンジンである. 評価には, TempEval-2 で用いられた精度 (正解/出力数) による評価法を用いるが, 考察時には再現率の誤りも考慮する.

3.2 実験結果

表4には, テストセットにおける局所モデルと文内最適化モデルの結果及び, TempEval-2 の参加チームのうち, 各タスクで最高性能の結果を示してある. まず, 文内最適化の効果を確認しよう. 文内最適化により, Task C と F においては一定の精度向上が見られる. Task D においては, 若干の性能低下が見られたが, これは制約式 (3) や (6) の導入によって, Task C や F の誤りを含んだ結果が悪影響を与えているものと考えられる. またこの結果は, Task D に対しては局所モデルが効果的である, という1節で挙げた仮説とも一致する. これに, 参考値として示してある文書最適化を行った場合 [15] の結果も見ると, Task C については文内最適化の性能

¹<http://nlp.stanford.edu/software/tagger.shtml>

²<http://nlp.cs.lth.se/>

が文書最適化を上回っている。逆に Task E は文内最適化では扱えないため、文書最適化の結果が本研究の最高値となる。

次に、TempEval-2 の参加チーム [14] との性能比較を行うと、やはり Task C と E では、本研究の文内最適化モデルと文書最適化モデルがそれぞれ高い性能を達成していることが分かる。Task E では、TRIPS [12] の性能が上回るように見えるが、TRIPS は、全ての関係について結果を出力するわけではなく、再現率エラーがあるため、TempEval-1 [13] と同様に F 値評価にすると、2-3pt の低下がある。ゆえに、Task E の文書最適化モデルの性能は TRIPS の性能と同等以上であると言える。

また、Task D と F ではともに最高値へは届かなかった。特に Task F では、オントロジー素性を利用した NCSU [4] が非常に高い性能を達成しており、同様の素性や制約を導入する手法を考案する必要があると考えられる。

4 おわりに

本稿では、時間的順序関係推定に関する新たな最適化モデルを提案した。これは文書全体で最適化するのではなく、文内での最適化を行うことで、モデルを単純かつ頑健にしたものである。その結果、TempEval-2 の Task C (同一文内の事象-時間表現間の順序関係推定) においては、従来の文書全体最適化よりも高い性能を達成した。

今後の展望として考えていることは2つあり、まず、今回十分に性能を改善できなかった Task D と Task F について、必要な素性や制約を見直すこと。もう1つは、本研究の自然な拡張を考案することで、これは先に文内最適化を行った上で、その素性や制約を固定してやれば、文書最適化への拡張も比較的容易に行えることが期待できるし、さらには、全体としての処理の高速化を求める場合には、文内最適化モデルの出力結果をそのまま既知として、文間の時間的順序関係のみを同定する手法も有望である。

参考文献

- [1] Branimir Boguraev and Rie Kubota Ando. Timeml-compliant text analysis for temporal reasoning. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence*, pp. 997–1003, 2005.
- [2] Nathanael Chambers and Daniel Jurafsky. Jointly combining implicit constraints improves temporal ordering. In *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, pp. 698–706, Honolulu, Hawaii, October 2008. Association for Computational Linguistics.
- [3] Quang Do, Wei Lu, and Dan Roth. Joint inference for event timeline construction. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pp. 677–687, Jeju Island, Korea, July 2012. Association for Computational Linguistics.
- [4] Eun Ha, Alok Baikadi, Carlyle Licata, and James Lester. Ncsu: Modeling temporal relations with markov logic and lexical ontology. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, pp. 341–344, Uppsala, Sweden, July 2010. Association for Computational Linguistics.
- [5] Xiao Ling and Daniel S. Weld. Temporal information extraction. In *Proceedings of the Twenty-Fifth National Conference on Artificial Intelligence*, 2010.
- [6] Hector Llorens, Estela Saquete, and Borja Navarro. Tipsem (english and spanish): Evaluating crfs and semantic roles in tempeval-2. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, pp. 284–291, Uppsala, Sweden, July 2010. Association for Computational Linguistics.
- [7] Inderjeet Mani, Marc Verhagen, Ben Wellner, Chong Min Lee, and James Pustejovsky. Machine learning of temporal relations. In *ACL-44: Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*, pp. 753–760, Morristown, NJ, USA, 2006. Association for Computational Linguistics.
- [8] Ivan Meza-Ruiz and Sebastian Riedel. Jointly identifying predicates, arguments and senses using markov logic. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 155–163, Boulder, CO, USA, June 2009. Association for Computational Linguistics.
- [9] Hoifung Poon and Pedro Domingos. Joint inference in information extraction. In *Proceedings of the Twenty-Second National Conference on Artificial Intelligence*, pp. 913–918, Vancouver, Canada, 2007. AAAI Press.
- [10] Matthew Richardson and Pedro Domingos. Markov logic networks. *Machine Learning*, Vol. 62, No. 1-2, pp. 107–136, 2006.
- [11] Parag Singla and Pedro Domingos. Entity resolution with markov logic. In *Proceedings of the Sixth International Conference on Data Mining (ICDM)*, pp. 572–582, Washington, DC, USA, 2006. IEEE Computer Society.
- [12] Naushad UzZaman and James Allen. Trips and trios system for tempeval-2: Extracting temporal information from text. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, pp. 276–283, Uppsala, Sweden, July 2010. Association for Computational Linguistics.
- [13] Marc Verhagen, Robert Gaizaukas, Frank Schilder, Mark Hepple, Graham Katz, and James Pustejovsky. Semeval-2007 task 15: Tempeval temporal relation identification. In *Proceedings of the 4th International Workshop on SemEval-2007.*, pp. 75–80, 2007.
- [14] Marc Verhagen, Roser Sauri, Tommaso Caselli, and James Pustejovsky. Semeval-2010 task 13: Tempeval-2. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, pp. 57–62, Uppsala, Sweden, July 2010. Association for Computational Linguistics.
- [15] Katsumasa Yoshikawa, Sebastian Riedel, Masayuki Asahara, and Yuji Matsumoto. Jointly identifying temporal relations with markov logic. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pp. 405–413, Suntec, Singapore, August 2009. Association for Computational Linguistics.