

情報構造抽出に向けた 深層学習・表現学習

豊田工業大学 知能数理研究室

准教授 三輪誠

makoto-miwa@toyota-ti.ac.jp

自己紹介 (略歴)

- 2008

- 東京大学大学院 新領域創成科学研究科基盤情報学専攻
近山研究室 博士課程修了

2008 - 2011

- 東京大学大学院 情報理工学系研究科 コンピュータ科学専攻
辻井研究室 特任研究員

2011 - 2014

- イギリスマンチェスター大学コンピュータ科学研究科
National Centre for Text Mining (NaCTeM) Research Associate
- 東京大学大学院 工学系研究科 鶴岡研究室 客員研究員

2014 -

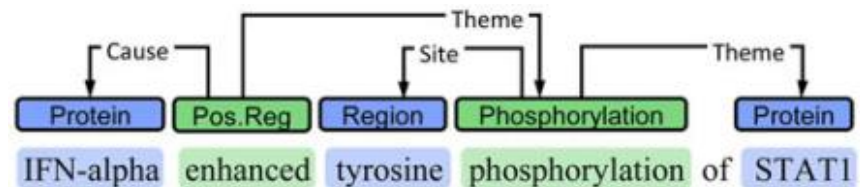
- 豊田工業大学 知能数理研究室 准教授
- 産業総合研究所人工知能研究センター 客員研究員 (2015-)

研究テーマ

- 機械学習によるデータに埋もれた情報・知識の構造化・形式知化
 - ゲームの特徴表現の自動抽出 (近山研)
 - 生命医学文献からの情報抽出 (辻井研・NaCTeM)
 - 深層学習・表現学習 (鶴岡研・豊田工大) など



将棋レボリューション 激指13
(©毎日コミュニケーションズ)



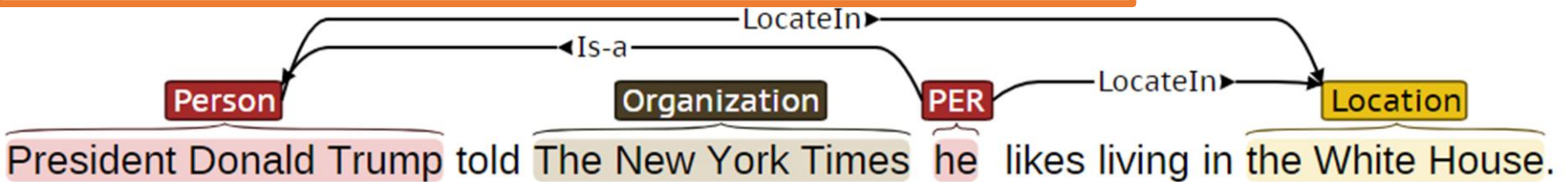
情報構造の例

注意: 適当に
作った例です

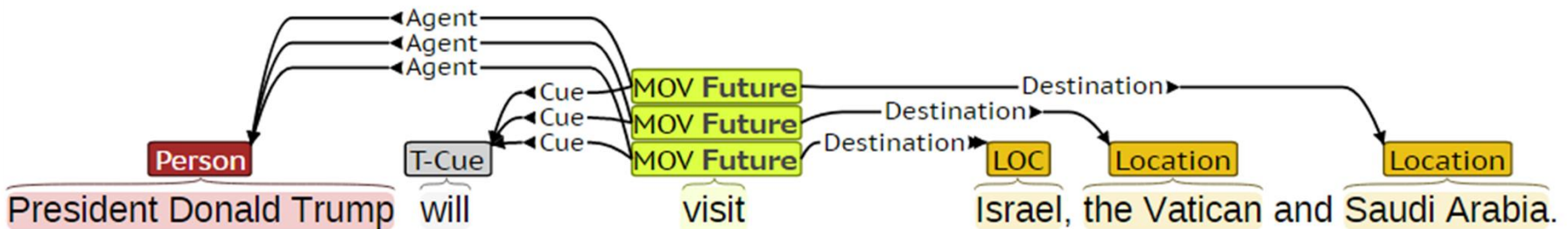
固有表現：文の領域



関係（主に静的な関係）：2つの（固有）表現間の関係



事象構造（主に動的な関係）：トリガと複数の項としての(固有)表現



情報構造抽出

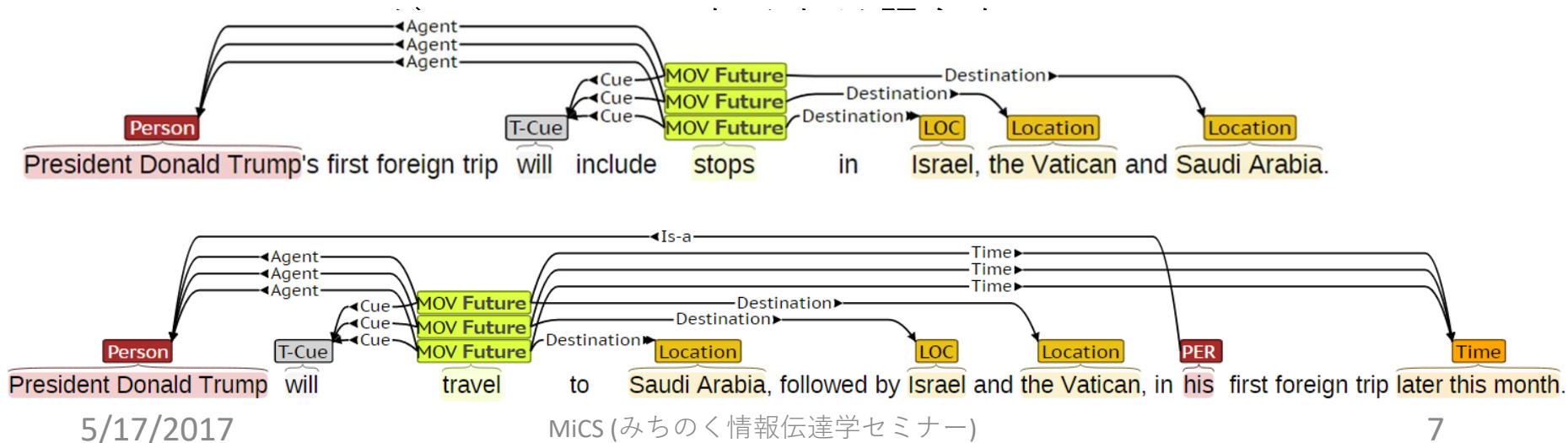
- 自然言語文に記述された情報を構造化・汎化して抽出する一連の技術
 - 固有表現抽出
 - 人・場所など文中の(固有)表現を抽出・分類
 - 関係分類
 - 文中に記述された2つの固有表現間の関係を分類
 - 事象(イベント)構造抽出
 - 文中に記述された固有表現を含む事象をN-aryの関係として抽出
- 本発表では**タグ付きコーパスを用いた機械学習**による情報抽出を対象
 - タグ種類は定義済み

表層や構文の違いを情報構造として汎化



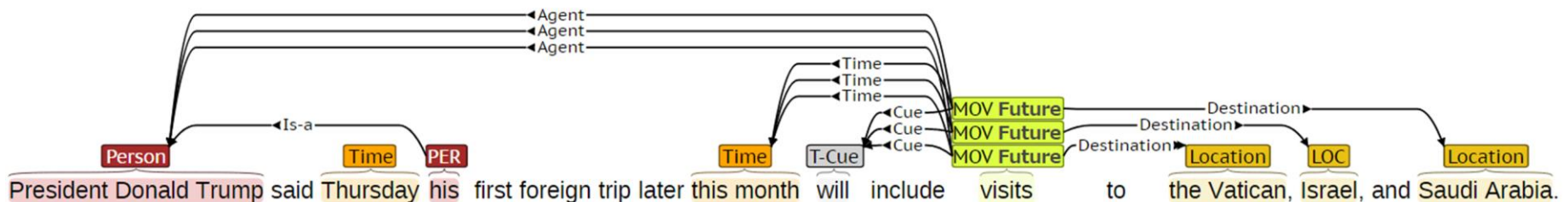
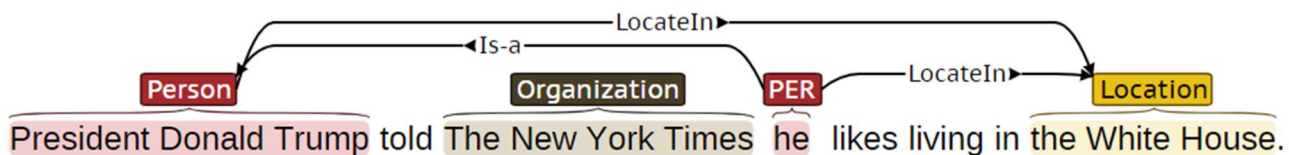
情報構造抽出の難しさ (1)

- 表現の多様性・曖昧性
 - 同じMOVEMENTのキーとなる表現（トリガー）の表層も品詞も違う
 - visits（名詞）, stops（名詞）, travel（動詞）
 - 統語構造的にも一致しない
 - 同じ表層でも同じ型になるとは限らない



情報構造抽出の難しさ (2)

- 構造の複雑さ（関係・事象構造）
 - 文から範囲の**重複**した非連結グラフを抽出
 - 様々なレベルでの汎化（型付け）
 - 事象では
 - 事象ごとに異なる多項関係 (全てが出現するとも限らない)
 - 知識ベースなどの3つ組では表現できない
- 照応解析など他の問題との関連



情報構造抽出の応用例: システム生物学への利用

反応の種類: Inhibition

修飾物質: Dorsomorphin

反応物: SMAD 1/5/8

生成物: pSMAD 1/5/8 (phosphorylated: リン酸化)



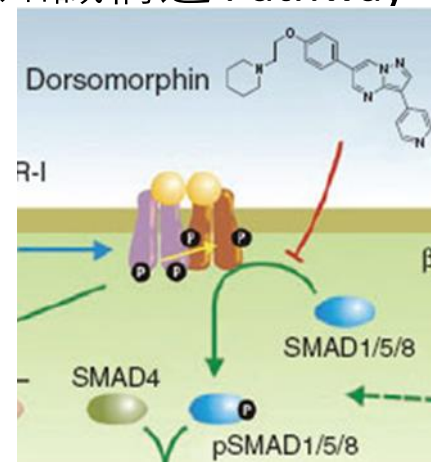
対応

(実際には対応がつかない
こともある)

事象構造



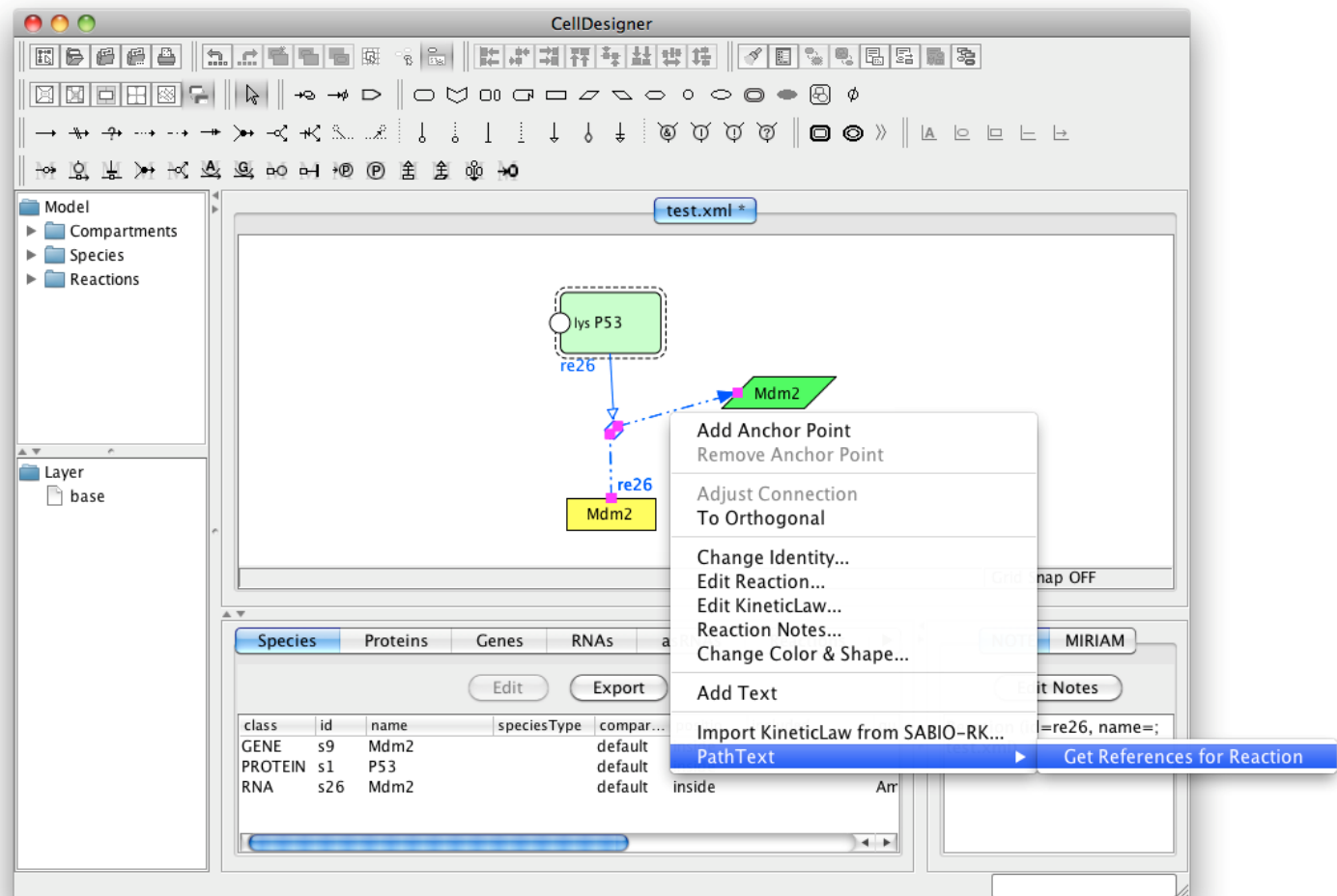
システム生物学における
知識構造 Pathway

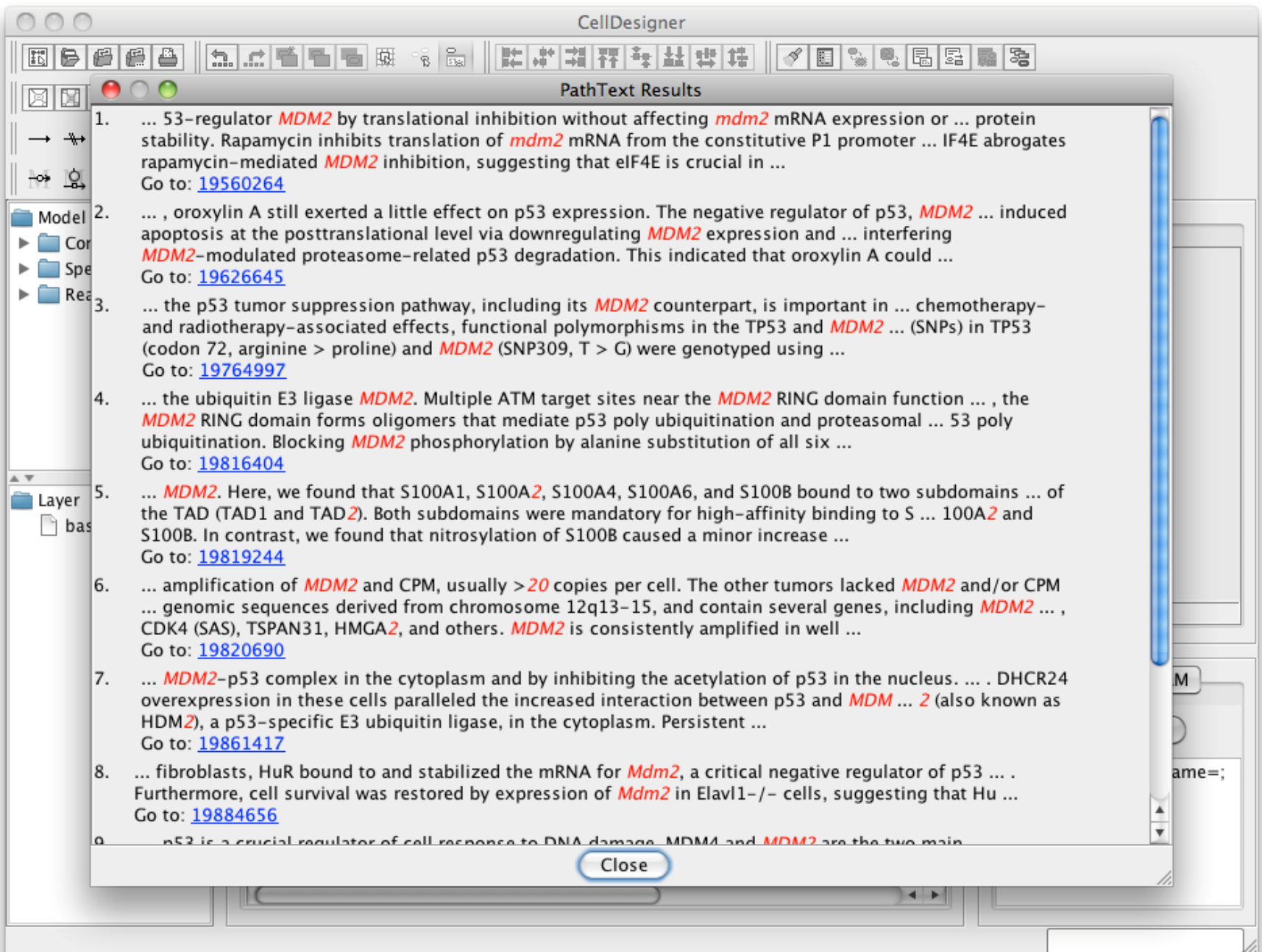


※分野オントロジと連携した事象・固有表現の型の定義

PathText [Miwa et al., 2013]

PathwayエディタCellDesigner との連携





The image shows a computer screen with two overlapping windows. The background window is 'CellDesigner', a software for biological modeling, showing a list of pathways on the left and a text area on the right. The foreground window is a web browser displaying a PubMed article titled 'Rapamycin increases the p53/MDM2 protein ratio and p53-dependent apoptosis by translational inhibition of mdm2 in cancer cells'.

CellDesigner Window:

- 1. ... 53-regulator **MDM2** by translational inhibition of MDM2 protein stability. Rapamycin inhibits translation of MDM2 protein. Rapamycin-mediated **MDM2** inhibition, suggesting that MDM2 is a crucial regulator of cell response to DNA damage. Go to: [19560264](#)
- 2. ... , oroxylin A still exerted a little effect on cell apoptosis at the posttranslational level via inhibiting **MDM2**-modulated proteasome-related p53 pathway. Go to: [19626645](#)
- 3. ... the p53 tumor suppression pathway, in which p53 is phosphorylated (codon 72, arginine > proline) and **MDM2** is phosphorylated. Go to: [19764997](#)
- 4. ... the ubiquitin E3 ligase **MDM2**. Multiple **MDM2** RING domain forms oligomers that mediate ubiquitination. Blocking **MDM2** phosphorylation inhibits ubiquitination. Go to: [19816404](#)
- 5. ... **MDM2**. Here, we found that S100A1, S100B, and S100P are involved in the TAD (TAD1 and TAD2). Both subdomains of S100A1 and S100B are involved in the TAD. In contrast, we found that nitrosylation of S100A1 and S100B inhibits their interaction with **MDM2**. Go to: [19819244](#)
- 6. ... amplification of **MDM2** and CPM, usual in many human cancers. Genomic sequences derived from chromosomes 12, 17, and 21 (CDK4 (SAS), TSPAN31, HMGA2, and others) are amplified in many human cancers. Go to: [19820690](#)
- 7. ... **MDM2**-p53 complex in the cytoplasm. Overexpression of **MDM2** in these cells paralleled the overexpression of **MDM2**, a p53-specific E3 ubiquitin ligase. Go to: [19861417](#)
- 8. ... fibroblasts, HuR bound to and stabilized MDM2 mRNA. Furthermore, cell survival was restored by overexpression of MDM2. Go to: [19884656](#)
- 9. ... p53 is a crucial regulator of cell response to DNA damage.

Web Browser Window:

PathText Results

Rapamycin increases the p53

www.ncbi.nlm.nih.gov/pubmed/19560264?dopt=Abst...

NCBI Resources How To

PubMed.gov

US National Library of Medicine
National Institutes of Health

PubMed

Advanced

Display Settings: ☒ Abstract [Send to:](#) ☐

Cancer Lett. 2009 Dec 28;286(2):250-9. Epub 2009 Jun 26.

Rapamycin increases the p53/MDM2 protein ratio and p53-dependent apoptosis by translational inhibition of mdm2 in cancer cells.

Kao CL, Hsu HS, Chen HW, Cheng TH.

Institute of Biochemistry and Molecular Biology, National Yang-Ming University, Taipei, Taiwan, Republic of China.

Abstract

Rapamycin, a potential anti-cancer agent, modulates activity of various factors functioning in translation, including eIF4E, an initiation factor selectively regulating expression of a subset of cellular transcripts. We show here that rapamycin suppresses levels of the p53-regulator MDM2 by translational inhibition without affecting mdm2 mRNA expression or protein stability. Rapamycin inhibits translation of mdm2 mRNA from the constitutive P1 promoter, which contains two upstream ORFs (uORFs) in the 5'UTR. Suppression is accompanied by increased hypo-phosphorylation of 4EBP-1, an inhibitory eIF4E binding protein. Ectopic expression of eIF4E abrogates rapamycin-mediated MDM2 inhibition, suggesting that eIF4E is crucial in modulating MDM2 expression in rapamycin-treated cells. Rapamycin administration also results in elevated PUMA expression and PARP cleavage, which is reproduced by siRNA knockdown of eIF4E or MDM2, suggesting that MDM2 suppression by rapamycin stimulates p53-mediated apoptosis. Together, our results define translational regulation of MDM2 expression by eIF4E and provide a molecular mechanism underlying rapamycin-induced p53-dependent apoptosis.

深層学習以前の関係抽出

JointER [Miwa et al., 2014]

- 固有表現とその関係を表の穴埋め問題に定式化
 - 表全体を埋めて構造化パーセプトロン等でスコア付
 - セルの素性 + これまで見つけた構造の素性

<https://github.com/tticoin/JointER>

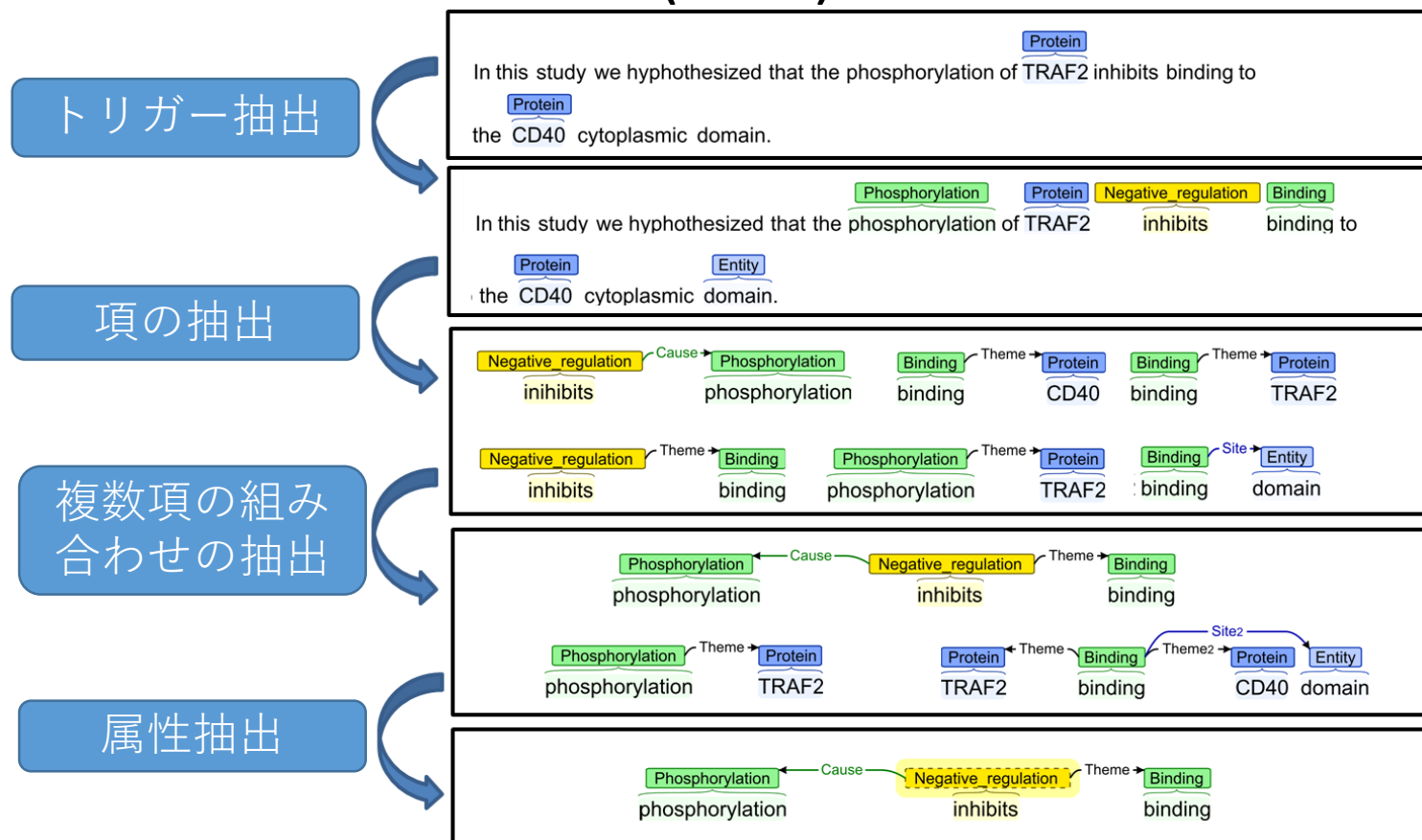
The diagram illustrates the JointER model's state for the sentence "US forces in Somalia, Haiti, Bosnia, and Kosovo". It consists of a table representing the current state of the parse and a graph showing the relationships between entities.

US	GPE							
forces	EMP-ORG	ORG						
in	⊥	⊥	⊥					
Somalia	⊥	<PHYS	⊥	GPE				
Haiti	⊥	<PHYS	⊥	⊥	GPE			
Bosnia	⊥	<PHYS	⊥	⊥	⊥	GPE		
and	⊥	⊥	⊥	⊥	⊥	⊥	⊥	
Kosovo	⊥	<PHYS	⊥	⊥	⊥	⊥	⊥	GPE
	US	forces	in	Somalia	Haiti	Bosnia	and	Kosovo

The graph shows the relationships between the entities: US (GPE), forces (EMP-ORG), in (ORG), Somalia (GPE), Haiti (GPE), Bosnia (GPE), and Kosovo (GPE). The relationships are labeled with PHYS (Physical) and EMP-ORG (Employment-Organization).

深層学習以前の事象構造抽出 EventMine [Miwa et al., 2012]

- 素性ベースの分類器 (SVM) を用いた逐次抽出



従来手法の素性の例

- 用語の素性
 - 文字 n -gram, 文字タイプ n -gram
 - 用語・用語ペア周辺の素性
 - 位置情報付き Bag-of-words, Bag-of- n -grams
 - 用語ペア間の構文木上での経路に関する素性
 - 経路の部分構造, 経路上での n -gramsなど
- ➔ 数百万から数十億の疎な素性空間
- 分類例一つに「発火」するのは数百から数万程度

深層学習による情報構造抽出

- 数年前までは素性を利用した機械学習が主流
 - 素性エンジニアリングが鍵



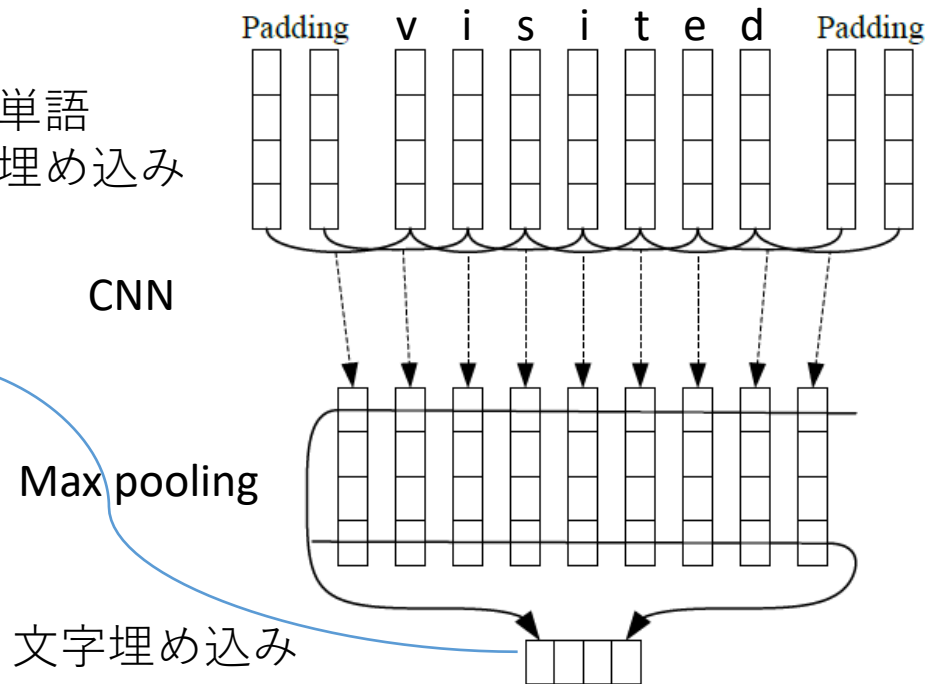
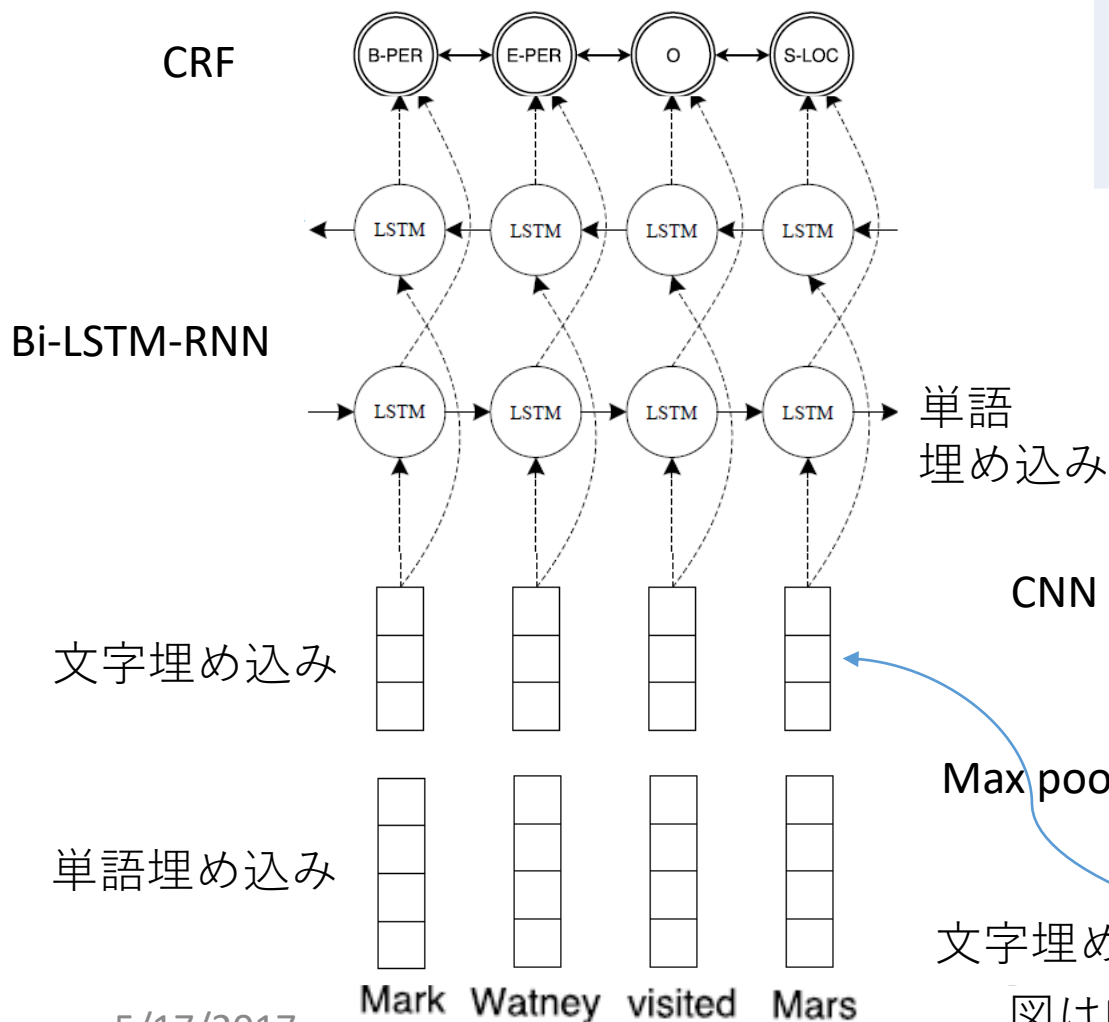
- 単語埋め込みの事前学習（表現学習） + 再帰型または畳み込みニューラルネットワークによる**半教師あり学習**
 - 従来の辞書・構文解析等を用いた従来の素性ベースの手法に比べて高い精度を達成
 - 従来の複雑な素性の計算・利用より安価な場合もある
 - モデルエンジニアリングが鍵？
- 以降，代表的なタスクの高精度なモデルを紹介

固有表現抽出

[Ma and Hovy, 2016]

	F1 [%]
Ando et al., 2005 素性ベースのCRF+辞書+ マルチタスク学習	89.31
Ma and Hovy, 2016 文字CNN単語埋め込み+ 事前学習単語埋め込み+ LSTM+CRF	91.21

CoNLL 2003での評価

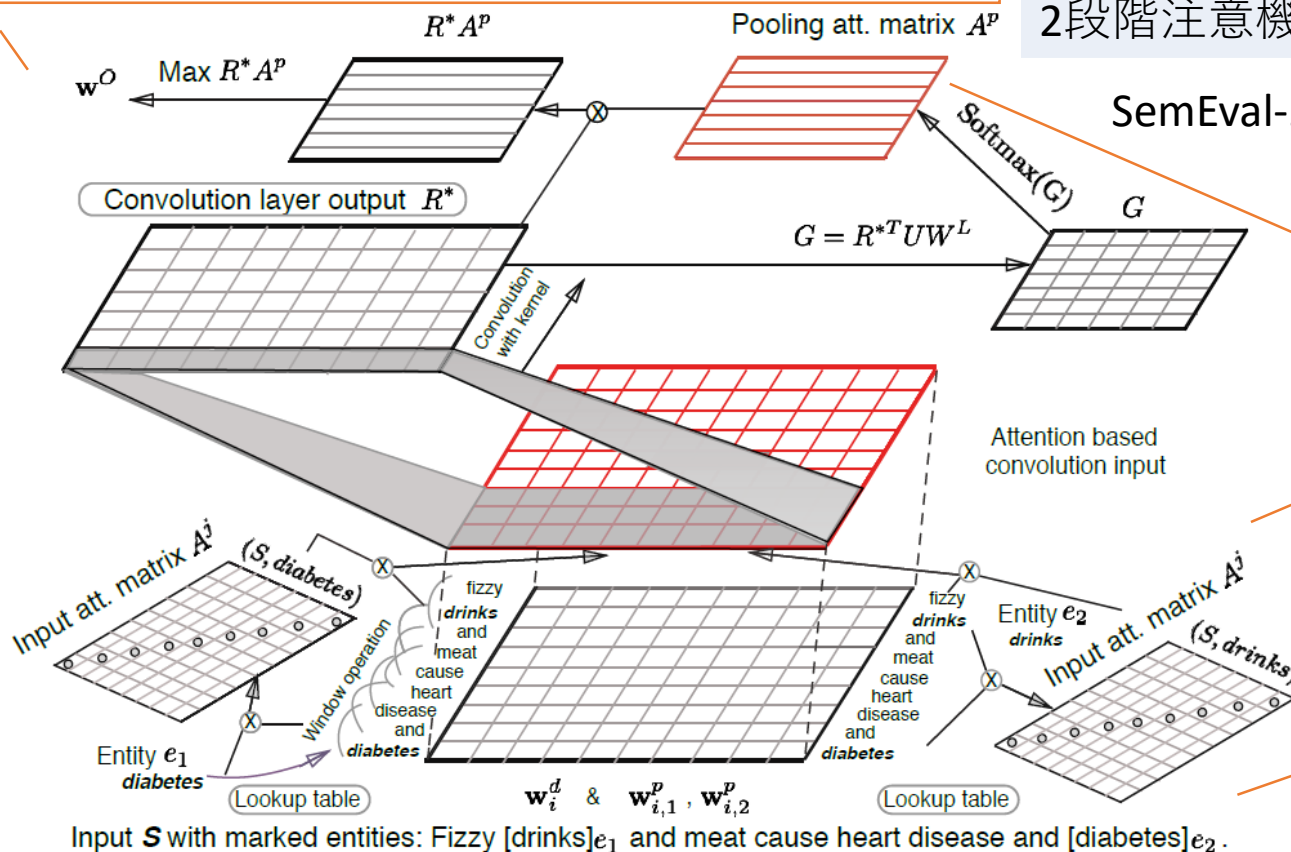


図は[Ma and Hovy, 2016]を改変

関係分類

[Wang et al., 2016]

出力された表現とラベル埋め込みの距離で分類



出力への
注意機構

入力への
注意機構

2つの表現が指定
された文を入力

	F1 [%]
Rink et al., 2010 SVM+辞書 + 構文解析	82.2
Wang et al, 2016 事前学習単語埋め込み+ 2段階注意機構付きCNN	88.0

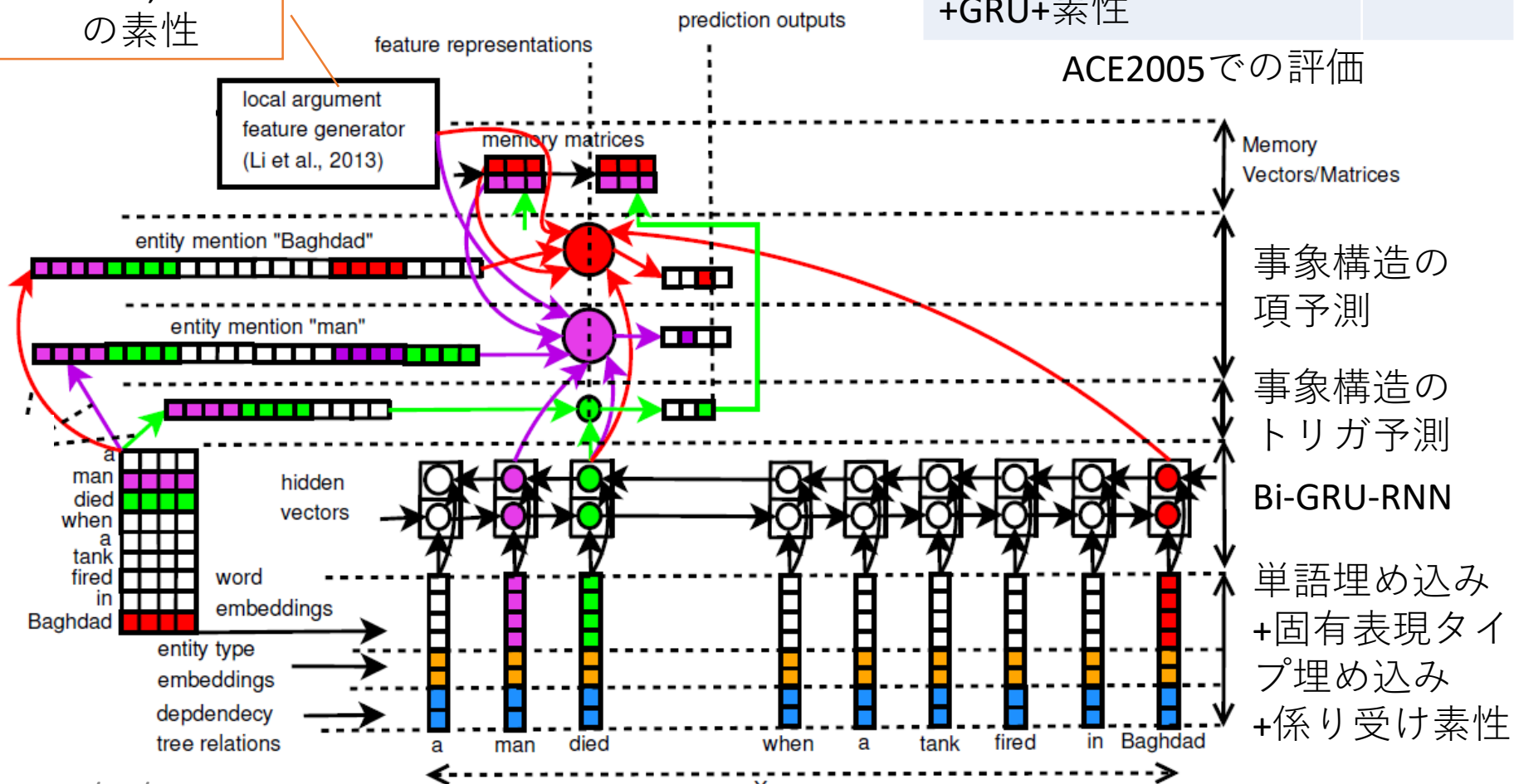
SemEval-2010 Task 8での評価

図は[Wang et al., 2016]を改変

事象構造抽出

[Nguyen et al., 2016]

Li et al., 2013
の素性



5/17/2017

図は[Nguyen et al., 2016]を改変 19

現在の深層学習による 情報抽出

- それぞれのタスクにおいて、深層学習により大きく精度向上
 - 固有表現・関係分類のモデルのみによる精度向上は昨年がピーク？
 - 基本モデルは決まってきた印象
 - 深層学習のパラダイムが変わらない限り、モジュールの入れ替え程度？
 - 事象構造についてはまだ少ない
 - ➔ タスクが複雑なため？
 - ⇔ 実際にアノテーションするところのくらの表現が欲しくなる

研究の紹介

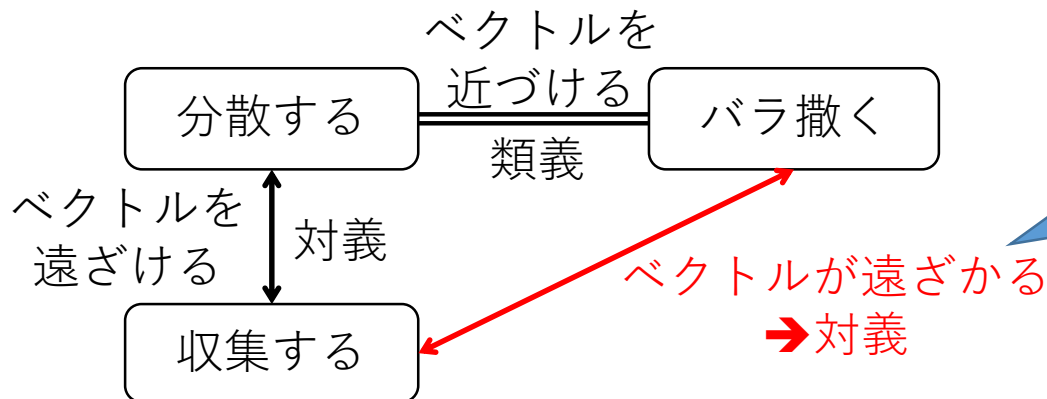
情報構造抽出へのアプローチ

- 外部知識を利用した表現学習
 - 文書内単語共起 (**word2vec**) 以外の表現学習
 - ➔ 獲得した表現を情報構造抽出モデルへの利用へ
- 情報構造抽出モデルの作成
 - 情報構造をより捉えたモデル
 - ➔ 情報構造全体を同時に発見・利用するモデルへ

表現学習

辞書情報の埋め込み: 単語類義・対義関係の学習 [Ono et al., 2015]

- 単語埋め込みベクトル間の距離が単語ペア間の類義・対義を表すようにベクトルを更新
- 辞書上の類義語ペアのベクトル間距離が近く、対義語ペアのベクトル間距離が遠くなるよう埋め込みベクトルを更新・学習



辞書上の関係から
(ベクトルを介して)
関係を推測

ベースモデル: Skip-Gram with Negative Sampling (SGNS)

[Mikolov et al., 2013]

- 目的関数

$$\ell_{\text{SGNS}} = \sum_{i=1}^N \sum_{i-C \leq j \leq i+C, j \neq i} \left\{ \log \sigma(\text{sim}(w_j, w_i)) + \sum_{\tilde{w} \sim P_0}^k \log \sigma(-\text{sim}(w_j, \tilde{w})) \right\}$$

文脈内の単語を対象単語と近づける

文脈外の単語を対象単語から遠ざける

$\text{sim}(w, c) = \mathbf{v}_w \cdot \mathbf{v}_c + b_w$ - 類似度

N - コーパス中の単語数

k - サンプルする負例の数

P_0 - (頻度補正した) ユニグラム分布

C - 文脈窓のサイズ

σ - ロジスティック関数

SGNSの目的関数の変換

- 目的関数

$$\ell_{\text{SGNS}} = \sum_{i=1}^N \sum_{i-C \leq j \leq i+C, j \neq i} \left\{ \log \sigma(\text{sim}(w_j, w_i)) + \sum_{\tilde{w} \sim P_0}^k \log \sigma(-\text{sim}(w_j, \tilde{w})) \right\}$$

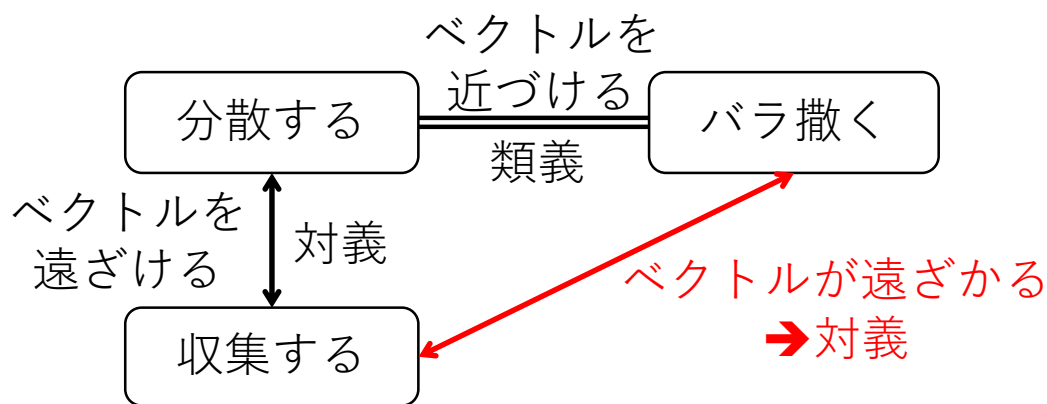
$$= \sum_{w \in V} \sum_{c \in V} \left\{ \#(w, c) \log \sigma(\text{sim}(w, c)) + k \#(w) P_0(c) \log \sigma(-\text{sim}(w, c)) \right\}$$

単語ペアに対する目的関数

$\#(w, c)$: w と c の共起頻度

$\#(w)$: w の頻度

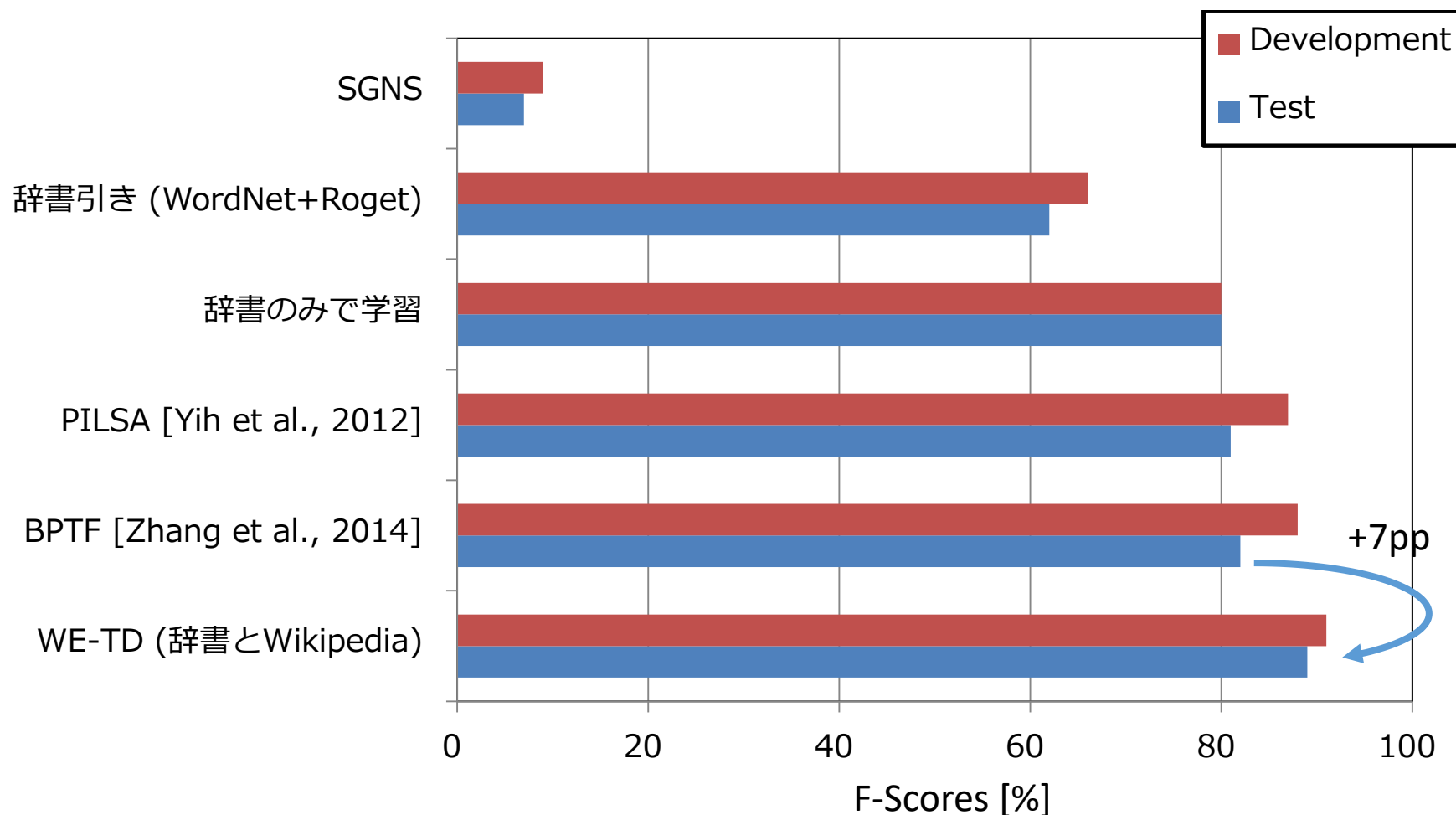
単語類義・対義関係の提案モデル



S_w : w の類義語
 A_w : w の対義語
 α, β : ハイパーパラメタ
 $A_{w,c}, B_{w,c}$ は係数をまとめたもの → 事前計算可能
 (小さい値は無視して計算を効率化)

$$\begin{aligned}
 \ell_{\text{WE-TD}} &= \sum_{w \in V} \sum_{s \in S_w} \log \sigma(\text{sim}(w, s)) + \alpha \sum_{w \in V} \sum_{a \in A_w} \log \sigma(-\text{sim}(w, a)) \\
 &+ \beta \ell_{\text{SGNS}} \quad \text{前ページのSGNSの目的関数} \\
 &= \sum_{w \in V} \sum_{c \in V} \{A_{w,c} \log \sigma(\text{sim}(w, c)) + B_{w,c} \log \sigma(-\text{sim}(w, c))\}
 \end{aligned}$$

評価 (GREの対義語問題950題)



類義語・対義語の例

入力: natural

デモは<http://www.tti-coin.jp/demo/antonym/>で公開

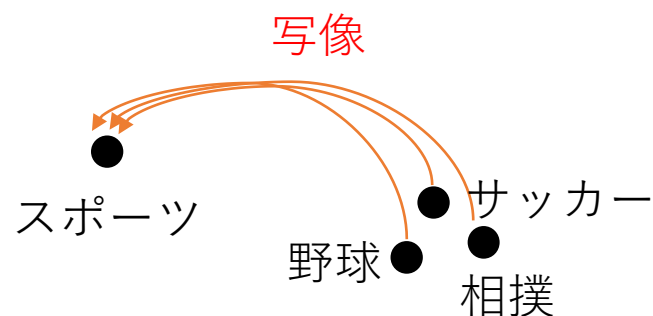
類義語		対義語	
normal	0.999999	unnatural	2.10E-07
simple	0.999998	artificial	1.42E-06
unmannered	0.999998	stilted	2.16E-06
physical	0.999998	forced	3.51E-06
regular	0.999998	unreal	4.82E-06
plain	0.999997	strained	8.55E-06
common	0.999997	false	9.01E-06
open	0.999997	supernatural	9.88E-06
genuine	0.999996	spectral	1.16E-05
unaffected	0.999993	elfin	1.35E-05

※数字は類似度0~1 (大きいほど類義, 小さいほど対義)

上位語の発見

[Yamane et al., 2016]

- 単語埋め込みを利用して,
下位語表現を上位語表現に対応付け
- 埋め込みベクトル間の写像行列を用意し, 下位語ベクトルの写像が上位語ベクトルに近づくように写像行列を学習



上位語発見モデル

Φ_k - 下位語から上位語への
 クラスタ k の写像行列
 P - 学習用上位語・下位語ペア
 (x, y) - x :下位語, y :上位語
 y'_i - 類似度 sim_k の大きい疑似負例
 σ - ロジスティック関数

目的関数

$$\sum_{\substack{(y,x) \in P, \\ k = \arg\max_k \text{sim}_k(y,x)}} \left\{ \log \sigma(\text{sim}_k(y, x)) + \sum_{y'_i (\neq y)} \log \sigma(-\text{sim}_k(y'_i, x)) \right\}$$

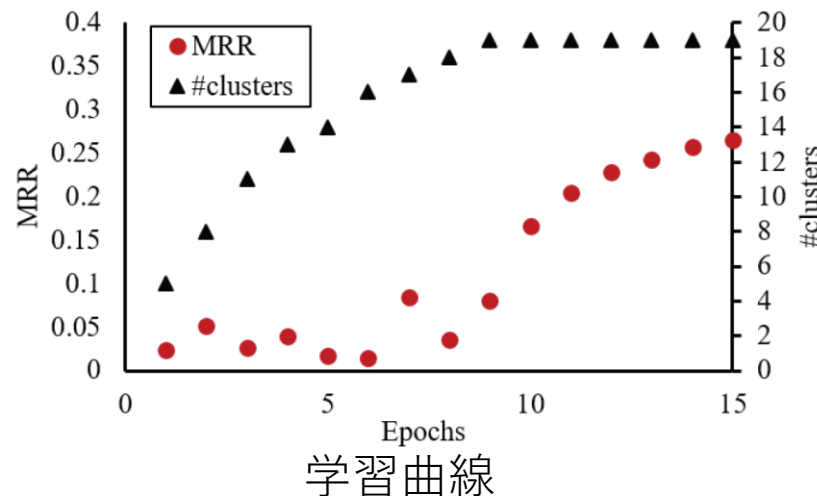
$$\text{sim}_k(y, x) = \Phi_k \mathbf{v}_x \cdot \mathbf{v}_y + b_k$$

• 特徴

- 写像行列 Φ_k を学習時に追加
 - 学習時に正例の $\text{sim}_k(y, x)$ が小さい場合に新しい Φ を追加
 - 上位下位ペアのクラスタリングを実現
- 類似度 sim_k の大きいものから疑似負例 y'_i を利用
 - 精度向上（識別）に有効な疑似負例を選択

上位語発見の評価

- 初期ベクトル
 - Yahoo知恵袋データ上で学習 (100万語彙)
- 日本語上位語下位語データで学習・評価
 - 訓練: 14,814ペア, 開発: 1,752ペア, テスト: 4,598ペア



クラスタ数19, 上位語を
平均3位以内 ($MRR=0.368$)
で発見

Fu et al., 2015の再現
($MRR=0.215$) より高性能

MRR: Mean reciprocal rank
平均逆順位

論文の属性情報の対応付けによる 検索システム [Yoneda et al., 2017]

- 論文の属性情報それぞれをベクトルに対応付けし、相互に対応付け
 - 属性情報 = 論文, 参照論文, 著者, 発表年,
タイトル・概要の用語
- ベクトル間の距離を用いて検索
 - 「この著者に"近い"論文は？」
 - 「この年に"近い"用語は？」

論文の属性情報の対応付け

e_i^a - カテゴリ a の i 番目のエンティティ
 $\mathbf{w}, \mathbf{v}$ - 対象と文脈埋め込み
 b - バイアス
 D - データ中のエンティティペア集合
 E_k - カテゴリ k のエンティティ集合

目的関数

$$- \sum_{(e_i^a, e_j^b) \in D} \log \frac{\exp(\mathbf{w}_{e_i^a} \cdot \mathbf{v}_{e_j^b} + b_{e_j^b})}{\sum_{e_k^b \in E_k} \exp(\mathbf{w}_{e_i^a} \cdot \mathbf{v}_{e_k^b} + b_{e_k^b})}$$

- タイトル・概要の用語については全体のベクトルを平均したものをエンティティとして利用
 - 用語間の関係は考慮しない
 - 高速化のため
- カテゴリには「論文・用語・著者・参照論文・年」を利用

論文の属性情報の対応付けによる 検索システム [Yoneda et al., 2017]

ACL anthology reference corpusのデータを利用

ACL Anthology Search System (Bib2vec)

検索クエリ

search query: Ryan McDonald [author]

論文

P10-1136: 2013
Zhichun Hsu; Maluoshan
In Discussion Forums
Proceedings of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies

P08-1059: 2014
Andrea Gesmundo; Keith Hall
Projecting the Knowledge Graph to Syntactic Parsing
Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, volume 2: Short Papers

P09-1097: 2009
Mehdi Manshadi; Xiao Li
Semantic Tagging of Web Search Queries
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

N10-1120: 2013
Mohammad Khan; Markus Dickinson; Sandra Kuebler
Does Size Matter? Text and Grammar Revision for Parsing Social Media Data
Proceedings of the Workshop on Language Analysis in Social Media

E09-1029: 2009
Katja Filippova; Mihai Surdeanu; Massimiliano Ciaramita; Hugo Zaragoza
Company-Oriented Extractive Summarization of Financial News
Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009)

N10-1060: 2010
Jennifer Foster
"cba to check the spelling": Investigating Parser Performance on Discussion Forum Posts
Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics

P06-1063: 2006
John Judge; Aoife Cahill; Josef van Genabith

www.aclweb.org/anthology/W/W07/W07-2216

用語

approximate inference
structured prediction
dependency parsing
transition-based
experimental evidence
dynamic programming
non-projective
spectral clustering
contrastive
latent variables
entity mentions
simplest
figure
hypothesized
machine

著者

Sina Zarrieß
Shravan Vasishth
Zdenek Zabokrtsky
David Talbot
Yoad Goldberg
Keith Hall
Mladen Vouk
Vincent Van Asch
Hao Zhang
Ryan McDonald
Kuzman Ganchev
Emily Pitler
Slav Petrov
Yonatan Belinkov
Sandra Kuebler
Kepa Bengio
Firoshi Ichikawa
Georgi Georgiev
Joakim Nivre
Oscar Täckström
Giorgio Satta
Nobuyuki Shimizu
Michael Collins

参照論文

P12-2047: 2012
Kuzman Ganchev; Keith Hall; Ryan McDonald; Slav Petrov
Using Search-Logs to Improve Query Tagging
Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)

W07-2216: 2007
Ryan McDonald; Giorgio Satta
On the Complexity of Non-Projective Data-Driven Dependency Parsing
Proceedings of the Tenth International Conference on Parsing Technologies

J14-2001: 2014
Joakim Nivre; Yoad Goldberg; Ryan McDonald
Squibs: Constrained Arc-Eager Dependency Parsing

E06-1038: 2006
Ryan McDonald
Discriminative Sentence Compression with Soft Syntactic Evidence

E09-1059: 2009
Kevin Lerman; Ryan McDonald
Contrastive Summarization: An Experiment with Consumer Reviews
Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Companion Volume: Short Papers

W06-1615: 2006
John Blitzer; Ryan McDonald; Fernando Pereira
Domain Adaptation with Structural Correspondence Learning
Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing

年

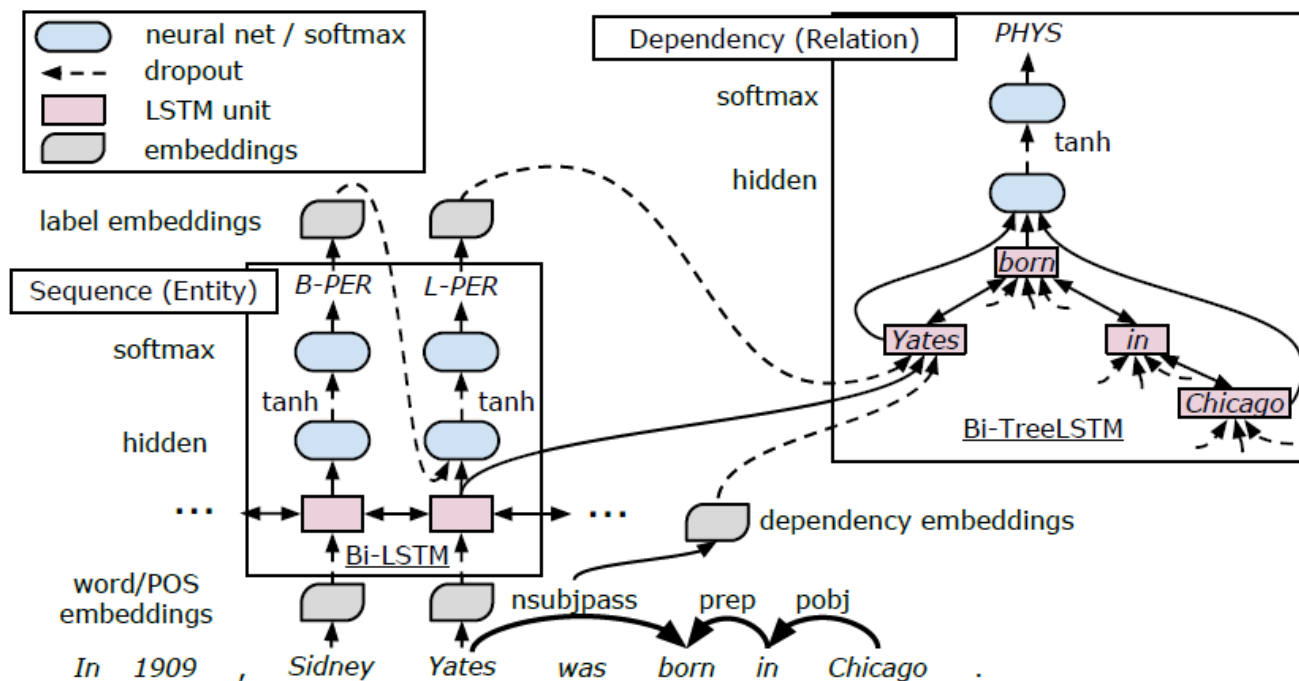
2000 2001 2002 2003 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015

情報構造抽出モデル

固有表現・関係の同時抽出

LSTM-ER [Miwa and Bansal, 2016]

- 固有表現用RNNと関係抽出用RNNを積み重ね
 - 構造学習ではなく **greedy** に予測するモデル

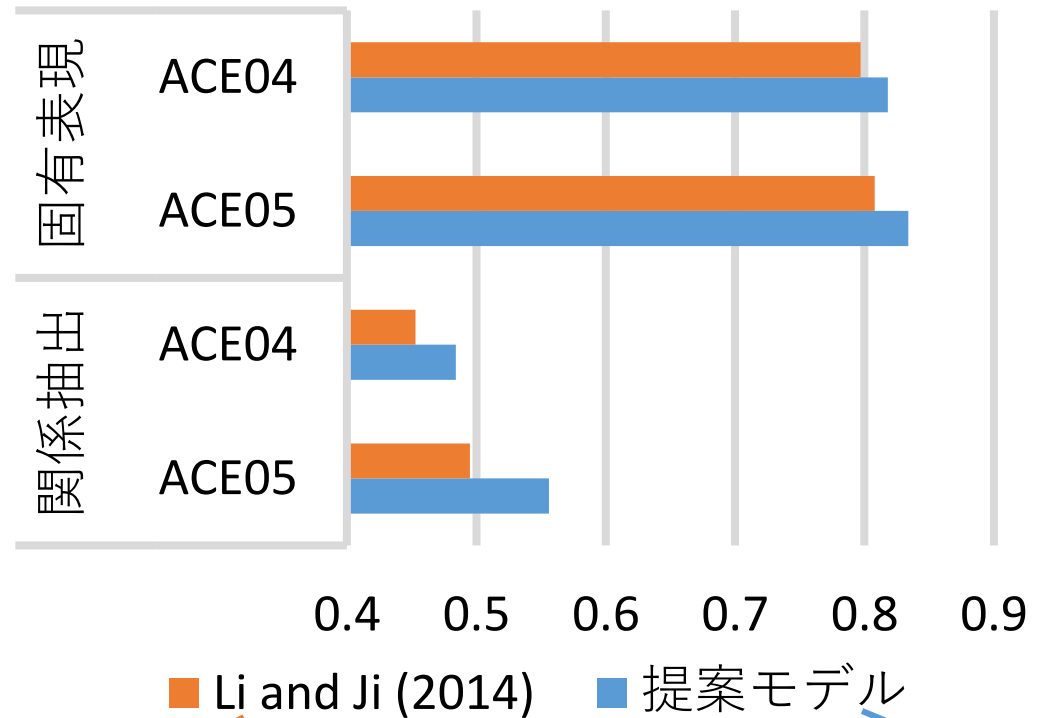


固有表現・関係の同時抽出

- 固有表現と関係のモデルを繋いで学習・予測
 - 予測ラベルに対応する埋め込みを次の予測の入力に利用
 - 関係は予測した固有表現を基に動的に作成
- モデル上の工夫
 - 固有表現の事前学習
 - 先に固有表現のみ学習 → 関係候補を学習初期から作成可能
 - 固有表現の学習の効率化 (scheduled sampling)
 - ある確率 (学習初期ほど大きい) で固有表現の予測を正解に置き換え → 序盤でも正解の周辺で学習
 - ドロップアウト+L2正則化+平均化
 - biasに正則化をかけない (出来ないライブラリが多い?)
 - モデルは平均化して利用 (かなり安定)

固有表現・関係の同時抽出

- ACE2004, 2005のタスクで評価
- 固有表現で2-3 pp, 関係抽出で3-5 ppの精度向上



<https://github.com/tticoin/LSTM-ER> で公開

構造化パーセプトロン

事前学習単語埋め込み+LSTM

まとめ

- 深層学習・表現学習による半教師あり情報構造抽出について紹介
 - 個別のタスクについての基本モデルは出来てきた
- 現状
 - マルチタスク学習・end-to-endモデルが少しずつ出ている
 - 弱教師あり・教師なし学習との連携はまだ
 - （深層学習以前から）分野が個別に発展しているが、「深くて狭い」構造と「浅くて広い」構造をどうつなぐかは大きな課題

今後の課題: 情報構造抽出の汎化

- 汎用なグラフ構造**表現**モデル
 - **個別の**外部知識・リソースは表現できるが,
全体を繋ぐにはどうすればよいか?
 - 例. 他の文, 文書中の図表, 固有表現データベース情報,
文書のメタ情報・関係 (citation network) など
- 汎用なグラフ構造**抽出**モデル
 - 文内については, 文から重複したグラフを取り出す
モデルを作れば十分
 - ➔ マルチタスク学習・転移学習などにも利用可能
 - 文を超えた場合は?

参考文献

[Miwa et al., 2012] Makoto Miwa, Paul Thompson and Sophia Ananiadou. Boosting automatic event extraction from the literature using domain adaptation and coreference resolution. Bioinformatics, Vol. 28. 2012.

[Miwa et al., 2013] Makoto Miwa, Tomoko Ohta, Rafal Rak, Andrew Rowley, Douglas B. Kell, Sampo Pyysalo and Sophia Ananiadou. A method for integrating and ranking the evidence for biochemical pathways by mining reactions from text. Bioinformatics, Vol. 29. 2013.

[Mikolov et al., 2013] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado and Jeff Dean. Distributed representations of words and phrases and their compositionality. NIPS. 2013.

[Miwa et al., 2014] Miwa, M, Sasaki, Y. 2014. Modeling Joint Entity and Relation Extraction with Table Representation. EMNLP. 2014.

[Ono et al., 2015] Masataka Ono, Makoto Miwa and Yutaka Sasaki. Word Embedding-based Antonym Detection using Thesauri and Distributional Information. NAACL-HLT. 2015.

[Nhung et al., 2016] Thien Huu Nguyen, Kyunghyun Cho and Ralph Grishman. Joint Event Extraction via Recurrent Neural Networks. NAACL. 2016.

[Ma and Hovy, 2016] Xuezhe Ma and Eduard Hovy. End-to-end Sequence Labeling via Bi-directional LSTM-CNNs-CRF. ACL. 2016.

[Wang et al., 2016] Linlin Wang, Zhu Cao, Gerard de Melo and Zhiyuan Liu. Relation Classification via Multi-Level Attention CNNs. ACL. 2016. [Miwa and Bansal, 2016] Makoto Miwa and Mohit Bansal. End-to-End Relation Extraction using LSTMs on Sequences and Tree Structures. ACL. 2016.

[Yamane et al., 2016] Josuke Yamane, Tomoya Takatani, Hitoshi Yamada, Makoto Miwa and Yutaka Sasaki. Distributional Hypernym Generation by Jointly Learning Clusters and Projections. COLING. 2016.

[Yoneda et al., 2017] Takuma Yoneda, Koki Mori, Makoto Miwa and Yutaka Sasaki. Bib2vec: Embedding-based Search System for Bibliographic Information. EACL Demo. 2017.