

深層強化学習を用いた意味依存構造解析は 自発的に平易優先戦略を学習する

栗田 修平

Anders Søgaard

京都大学 大学院情報学研究科 University of Copenhagen

kurita@nlp.ist.i.kyoto-u.ac.jp, soegaard@di.ku.dk

1 はじめに

SemEval 共有タスク 2014 [1] および 2015 [2] で提案された意味依存構造解析 (*semantic dependency parsing*, SDP) は、単語同士の意味的な関係を有向非巡回グラフで表現する意味解析で、従来の依存構造解析より柔軟に文の意味的な構造を表現できる。意味依存構造解析のグラフは、意味上の述語 (*semantic predicates*) となる単語から意味上の項 (*semantic arguments*) となる単語への有向エッジの集合である。その大きな特徴として、単語が複数の主辞 (*head*) 単語を持ちうる事が挙げられる。例えば、“The man went back and spoke to the desk clerk.” という英文について、単語 “man” は2つの動詞 “went” と “spoke” の意味上の主語となりうる。意味依存構造解析では、この関係は2つの述語から項である単語 “man” へ向かう2本の有向エッジとして表現される。つまり単語 “man” の主辞は複数あり、この点が文法的な構文木とは大きく異なる。意味依存構造解析は言語学上の違いから DM、PAS、PSD の3形式のデータセットが用意されている。PAS は Enju [3] から作られており、依存構造解析に比較的近いとされる [4]。図1にそれぞれの形式での意味依存構造グラフを示している。

本研究では、意味依存構造解析の解決のために反復的意味述語選択 (*iterative predicate selection*, IPS) による意味依存構造解析を提案する。IPS アルゴリズムは、従来手法であるグラフに基づく構造解析と遷移に基づく構造解析の両方の性質を併せ持ち、意味依存構造解析グラフ内部の複数の意味的な主辞を解決することができる。本研究では、この IPS アルゴリズムに基づくモデルを作成し、DM、PAS、PSD の3形式のマルチタスク学習 [4] を行うと同時に、方策勾配法 [5] による深層強化学習手法により学習した。IPS モデルに強化学習を用いることで、モデルは意味依存解析を行う順序を学習中に決定できる。提案手法では SemEval2015

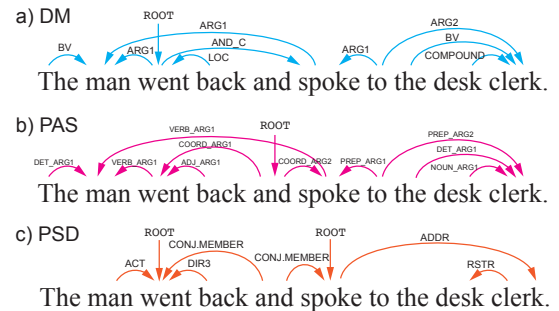


図1: DM、PAS、PSD の各形式における意味依存構造解析のグラフ

タスク 18 のドメイン内データセットで世界最高性能を達成した。また、この手法により深層強化学習させたモデルは解析において自発的に平易優先 (*easy-first*) 戦略を取ることを確認した。

本論文の貢献は以下のようにまとめられる。(i) 意味依存構造解析に対する新しい解析手法である IPS アルゴリズムを提案した、(ii) IPS アルゴリズムに対し方策勾配法による強化学習を提案した、(iii) 3つの意味依存構造形式に対する新しい世界最高性能を達成した、(iv) 強化学習により学習されたモデルは自発的に平易優先戦略を取ることを確認した。

2 関連研究

依存文法解析手法は、遷移に基づく解析とグラフに基づく解析の2つに大別される。2つの手法はそれぞれ異なる特徴を持つ。遷移に基づく解析では、文を先頭から処理していくため、最初の誤りが後の解析まで影響する誤差伝播の問題を抱えており、他方、グラフに基づく解析では、単語間の係り受けスコア付き完全グラフをまず作成し、最終的な構文木やグラフを得るためにはスコアの学習とは関係のない最大全域木や AD^3 のようなアルゴリズムを使用する必要がある。本論文で提案する IPS アルゴリズムは、Zhang et al.

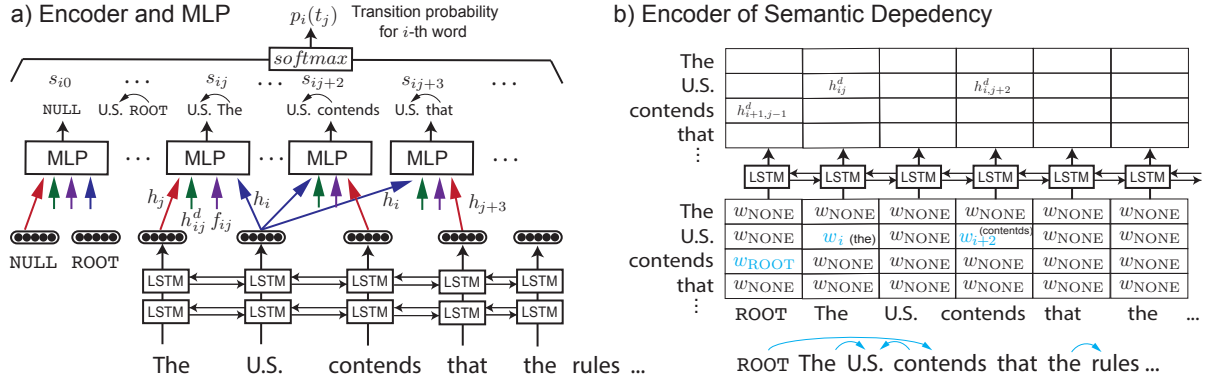


図 2: ニューラルネットワークモデル。(a) 単語の隠れ表現 h_i と主辞候補の隠れ表現 h_j のエンコーダと遷移確率を計算するための多層パーセプトロン。(b) 意味依存関係を表現するテンソルとそのエンコーダ。(b) では列に主辞候補、行に単語が並べられている。この各要素は、主辞候補の分散表現に対応するベクトルである。

(2017) 等による主辞選択 (*head-selection*) 構文解析 [6] を改良し、反復的な手法により意味依存解析グラフを作成するアルゴリズムである。本アルゴリズムでは遷移に基づく解析を行いながら、グラフに基づく解析のように単語間の大域的なスコア付けを行う。提案手法では、さらに深層強化学習を利用することで、誤差伝播問題に対処すると共に、モデル自身が文の解析順序を決定することができる。関連研究として Goldberg と Elhadad (2010) による平易優先構文解析 [7] がある。

3 提案手法

3.1 反復的意味述語選択

本論文では、依存構造解析における主辞選択モデル [6] に基づいて、意味依存構造解析モデルを提案する。意味依存構造解析においては、通常の依存構造解析と異なり、ある単語は複数の意味上の述語に対する意味上の項となりうる。つまりある単語に対する主辞の数は 0 を含む自然数個となる。この関係を処理するため、提案アルゴリズムである反復的意味述語選択では、各時刻に文中の全ての単語に対して遷移操作を行う。文中の i 番目の単語 w_i に対し、モデルは遷移操作 $t_i^\tau \in T_i^\tau$ を行う。 τ は遷移を行う時刻 (ただし $\tau \in \mathbb{N}$) である。単語 w_i に対する全遷移操作 T_i は

$$\{\text{NULL}, \text{ARC}_{i,\text{ROOT}}, \text{ARC}_{i,1}, \dots, \text{ARC}_{i,n}\}$$

と定義され $\text{ARC}_{i,j}$ は j 番目の単語から i 番目の単語へ有向エッジを作る操作、すなわち、意味的な述語 w_j から意味的な項 w_i への関係付けに対応する。NULL は有向エッジを作成しないという操作に対応する。実際に時刻 τ において単語 w_i に対して可能な遷移操作 T_i^τ は、 T_i の部分集合であり、それらの遷移操作は、どの

時刻 τ でも可能な NULL 操作を除いて、(i) 各単語は自分自身への有向エッジを作れない、(ii) 作成されたグラフは作成済みの部分グラフ \mathbf{y}^τ に含まれない、という条件を満たすものである。

モデルは意味依存構造解析のグラフを以下の順序で作成する。

- 1 全ての単語 w_* に対し、 T_*^τ から操作を選択する。
- 2 選択された操作に基づき、意味依存構造グラフを更新する。
- 3 全ての単語に対し NULL 操作が選択された場合、解析は終了する。そうでない場合、1 に戻る。

このアルゴリズムでは、最終的な解析グラフが得られるまでに複数の経路が存在する。例えば、図 1 の DM のグラフ作成にて、単語 “man” の意味的な主辞 “The”、“went” および “spoke” からのエッジのうち、どれを先に作成するかは任意である。ただし、意味依存構造解析においては、いくつかのエッジは比較的簡単に解決可能と考えられる。例えば “The” と “man” を結ぶエッジは比較的わかりやすい。他方、長い距離を結ぶエッジは、比較的に解決が難しいと考えられるが、同時に文の構造に関する重要な情報を含んでいると考えられ、どのようなエッジを先に作成させるかはモデルの判断に任せるべきであると言える。

3.2 ニューラルネットワークモデル

図 2 はニューラルネットワークモデル全体を表している。モデルは、文のエンコーダと構築中の意味依存構造解析のグラフのエンコーダ、および遷移操作を選択する多層パーセプトロンからなる。

文のエンコーダ 3 層双方向 LSTM を使用し、文内の単語の分散表現を得る。これは遷移により変化しない。**意味依存グラフのエンコーダ** 係り受けを持つ単語のベクトルを主辞の単語ごとに並べた 3 階のテンソルを作成し、各時刻ごとに双方向 LSTM を使用して各単語の主辞単語の表現を得る。遷移ごとに変化する。**意味上の述語選択モデル** 主辞選択モデルとして 2 層パーセプトロンと主辞候補全体への softmax を用いる。**ラベル付モデル** 多層双方向 LSTM のエンコーダーと 2 層パーセプトロンからなる係り受けエッジのラベル予測モデルである。上のモデルとは独立に学習する。

教師あり学習においては、複数の正解エッジが存在する場合、エポックごとにランダムに選択された順番で学習される。

3.3 強化学習

方策勾配法 強化学習は、ある動的な環境に対して、将来的に環境から得られる報酬を最大化させるために、各時刻に何らかの動作を行うことを繰り返して学習するモデルである。本提案手法においては、環境はモデルが作成している意味依存グラフを含み、各時刻の動作は遷移操作 t_i^τ に対応する。報酬は、モデルが作成している意味依存グラフおよび正解の構文グラフ \mathbf{y}^g から与えられる。本研究においては、IPS アルゴリズムの学習に Williams による方策勾配法 [5] の変種を使用する。強化学習の目的関数は、時刻 τ での i 番目の単語に対する報酬 r_i^τ の期待値として

$$J(\theta) = E_\pi [r_i^\tau] \quad (1)$$

である。この目的関数の微分の計算では、ある方策 π 下での期待値 E_π を計算する際に、ネットワークの出力 p_i からサンプルされた単一の遷移経路に対する平均で近似する:

$$\nabla J(\theta) \approx \sum_{t_i^\tau \in t} [r_i^\tau \nabla \log p_i(t_i^\tau | \mathbf{y}^\tau)] \quad (2)$$

なお、これは方策勾配定理の表現の一つである。

時刻 τ の i 番目の単語に対して、モデルは遷移確率 π に従い、可能な遷移 T_i のなかからひとつの遷移 t_i^τ をサンプルする。サンプルされた t_i^τ に基づき、モデルは意味依存グラフを $\mathbf{y}^{\tau+1}$ に更新し、報酬を計算する r_i^τ 。すべての単語に対し、NULL 遷移が最も確からしい遷移となった場合、もしくは、予め定められた最大遷移数 (典型的には 10) を超過した場合には、学習は終了する。その後、各時刻にサンプルされた遷移と報酬からパラメータを更新する。

報酬	遷移
$r_i^\tau = 1$	(1) モデルが i 番目の単語から主辞への正しいエッジを作成した。 (2) モデルが i 番目の単語から主辞への全ての正しいエッジのみ作成後、NULL を選択した際、その最初の 1 回目。
$r_i^\tau = -1$	(3) モデルが i 番目の単語に対する間違っ主辞へのエッジを作成した。
$r_i^\tau = 0$	(4) その他のすべての遷移操作

表 1: IPS モデルに対する強化学習の報酬。モデルへの報酬は単語単位で与えられる。

なお、通常の交差エントロピー誤差による学習との重要な違いとして、以下の点が挙げられる。(1) 強化学習は遷移に対してサンプリングを行う。これにより、教師ありモデルが決して辿らない遷移経路を探索できる。(2) 教師あり学習は、一つ一つの遷移に対してモデルを最適化するが、強化学習はモデルが全ての遷移を終えたあとに最適化を行う。(3) 強化学習の報酬は負値を取りうる。特に (1) と (2) の性質により、特定の遷移経路に対する報酬付けと最適化が可能となる。なお、本研究ではあらかじめ教師あり学習で事前学習を行った上で強化学習を用いる。

報酬 本研究では、各時刻に報酬を受け取る中間報酬を採用する。本モデルにて、 i 番目の単語に対し時刻 τ に与えられる報酬 r_i^τ を表 1 に示した。全体として、モデルは正しいエッジを作成した際に +1 の報酬、間違っているエッジを作成した際には -1 の報酬を受け取るが、NULL 遷移を選択した場合には、その単語に対する遷移操作が正しく終了した直後を除いて、正負いずれの報酬も受け取らない。

4 実験

モデルの実験には SemEval2015 タスク 18 データセット [2] を使用した。訓練データセットは WSJ コーパスから取られた 33,964 文からなり、開発データセット、ドメイン内テストセット、Brown コーパスから取られたドメイン外テストセットをそれぞれ 1,692 文、1,410 文、1,849 文ずつ含む。すべての文は DM、PAS、PSD の各形式でのアノテーションが付けられている。すべての評価は先行研究と同様にミクロ平均されたラベル付き F 値を使用する [4, 8]。

本研究での提案手法は、教師あり学習のみで学習された *IPS+ML*、および、強化学習 (RL) を用いて学習された *IPS+ML+RL* である。いずれの手法でも、文のエンコーダの学習に DM、PAS、PSD のマルチタスク

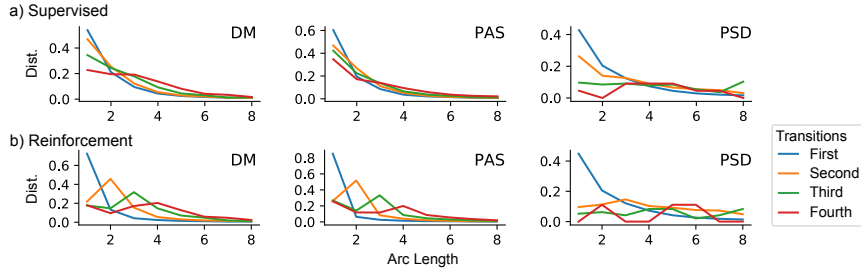


図 3: 各遷移における新規に作成された係り受けエッジの長さの分布。(a) 教師あり学習 (*IPS+ML*)。 (b) 強化学習 (*IPS+ML+RL*)。4本のグラフは、1および2、3、4回目の遷移にそれぞれ対応する。

ドメイン内テスト	DM	PAS	PSD	Avg.
Peng+ 17 Freda3	90.4	92.7	78.5	88.0
Peng+ 18	91.6	-	78.9	-
IPS +ML	91.2	92.4	78.8	88.3
IPS +ML +RL	91.9[‡]	92.8[‡]	79.3[‡]	88.8[‡]
ドメイン外テスト	DM	PAS	PSD	Avg.
Peng+ 17 Freda3	85.3	89.0	76.4	84.4
Peng+ 18	86.7	-	77.1	-
IPS +ML	86.0	88.2	77.2	84.6
IPS +ML +RL	87.2[‡]	88.8[‡]	77.7[‡]	85.3[‡]

表 2: SemEval 2015 データセットにおける実験結果。(上) ドメイン内テストと(下) ドメイン外テスト。Avg. は3つの形式のスコアの重み付き平均。強化学習 (RL) モデルの[‡]は強化学習を使わないモデルに対し、対応のある t 検定を行い $p < 10^{-3}$ で統計的に優位であったことを示す。

(ML) 学習を行っている。比較対象として、グラフに基づく解析である Peng *et al.* (2017) および Peng *et al.* (2018) を用いた。Peng *et al.* (2018) は構文情報を用いており、構文解析に比較的近いとされる PAS は対象としていない。提案手法では構文情報は用いていない。

表 2 に実験結果を示す。強化学習を用いた手法は、強化学習を用いない手法に対し、全ての形式で統計的に有意に上回った。また IPS アルゴリズムを用いない既存手法に対して、構文解析に近いとされる PAS 以外の形式でドメイン内テストで 0.3%、ドメイン外テストで 0.5% 以上も上回る性能を達成した。さらに驚くべきところに、強化学習にて学習されたモデルは、教師付き学習により学習されたモデルとは異なる解析法を行うことがわかった。図 3 では、強化学習によって学習されたモデルが、遷移毎に短い係り受けエッジから処理をしていくことを示している。これは、教師あり学習には見られず、学習時に暗黙的にも与えていないことから、モデルが自発的に学習したと考えられる。

実際には文法的な順序にも影響されると考えられ、発表では具体的な意味解析例を含めて詳しく議論する。

5 おわりに

本論文では、反復的意味述語選択 (IPS) アルゴリズムを提案し、意味依存構造解析においてその有効性を実証するとともに、深層強化学習を用いて学習することで、モデルが自発的に平易優先戦略をとって意味解析を行うようになることを実験において示した。将来研究においては深層強化学習を用いたこの手法を応用し、モデルが本当に必要な構文・意味解析のみを自ら選択し入力文に適用させる手法で、end-to-end モデルと構文・意味解析モデルを統合することを目指したい。

謝辞

本研究は JST ACT-I の支援を受けたものである。

参考文献

- [1] Stephan Oepen, Marco Kuhlmann, Yusuke Miyao, Daniel Zeman, Dan Flickinger, Jan Hajic, Angelina Ivanova, and Yi Zhang. Semeval 2014 task 8: Broad-coverage semantic dependency parsing. In *SemEval 2014*, pp. 63–72, Dublin, Ireland, August 2014.
- [2] Stephan Oepen, Marco Kuhlmann, Yusuke Miyao, Daniel Zeman, Silvie Cinkova, Dan Flickinger, Jan Hajic, and Zdenka Uresova. Semeval 2015 task 18: Broad-coverage semantic dependency parsing. pp. 915–926, Denver, Colorado, June 2015.
- [3] Yusuke Miyao, Takashi Ninomiya, and Jun’ichi. Tsujii. Corpus-oriented grammar development for acquiring a head-driven phrase structure grammar from the penn treebank. In *Proceedings of IJCNLP-04*, 2004.
- [4] Hao Peng, Sam Thomson, and Noah A. Smith. Deep multitask learning for semantic dependency parsing. In *ACL*, pp. 2037–2048, Vancouver, Canada, July 2017.
- [5] Ronald J Williams. *Simple statistical gradient-following algorithms for connectionist reinforcement learning*. Springer, 1992.
- [6] Xingxing Zhang, Jianpeng Cheng, and Mirella Lapata. Dependency parsing as head selection. In *EACL*, pp. 665–676, Valencia, Spain, April 2017.
- [7] Yoav Goldberg and Michael Elhadad. An efficient algorithm for easy-first non-directional dependency parsing. In *Human Language Technologies: NAACL*, pp. 742–750, Los Angeles, California, June 2010.
- [8] Hao Peng, Sam Thomson, and Noah A. Smith. Backpropagating through structured argmax using a spigot. In *ACL*, pp. 1863–1873, 2018.