# 複数の言語単位に対するスパン表現を用いた論述構造解析

栗林 樹生 [1] 大内 啓樹 [2,1] 井之上 直也 [1,2] Paul Reisert [2] 三好 利昇 [2,3] 鈴木 潤 [1,2] 乾 健太郎 [1,2]
[1] 東北大学 [2] 理化学研究所 AIP センター
[3] 日立製作所研究開発グループ基礎研究センタ
{kuribayashi, naoya-i, jun.suzuki, inui}@ecei.tohoku.ac.jp
{hiroki.ouchi, paul.reisert, toshinori.miyoshi}@riken.jp,
toshinori.miyoshi.pd@hitachi.com

## 1 はじめに

論述構造解析は論述文から論述構造を同定するタスクである．図 1に論述構造の例を示す．論述構造解析は以下の 4 つのサブタスクからなる[*1].

1. **論述要素の同定**: 論述文から論述要素の位置を同定する．

2. **論述要素分類**: 各論述要素を “Premise”，“Claim”，“Major-Claim” に分類する．

3. **論述関係認識**: 論述要素間の論述関係を予測する．

4. **論述関係分類**: 論述関係を ‘Support” か “Attack” に分類する．

本研究では既存研究 [7, 10, 12] と同様に，論述要素の同定は完了している設定で他 3 タスクを解く．

これまで論述構造解析のための様々な手法が提案されてきた．Potash ら [12] はニューラルベースのモデルを提案し，最高精度を達成した．また，文献 [10, 11, 14] では，論述文の言語的な性質を考慮した様々な素性が提案されてきた．例えば，論述文には論述要素 (Argumentative component (AC))，接続表現 (Conjunctive expression (CE))，論述的談話単位 (Argumentative discourse unit (ADU)) といった内部構造が存在し (2章), Stab らや Peldzsus ら [**Stab2017**，**Peldzsus2015**] の分析では，“In my point of view” などの機能的な役割を果たす接続表現の情報が重要であるとされてきた．そこで本研究では, ニューラルベースのモデルでもこのような言語的な特徴を導入することで，さらに性能が向上するという仮説に基づき，その方法論の提案と実験による検証を行う．

本研究では，以下のような性質を持つニューラルベースのモデルを提案する．

(a) 複数のレベルの言語単位 (論述要素，接続表現，論述的談話単位) に対して個別のスパン表現を割り当てる．

(b) 複数のレベルの言語単位の情報を用いて，論述全体の流れを異なる視点から捉える．

例えば CE の系列からは *in my opinion* → *admittedly* → *however* のような大まかな論述の流れを捉えることができる．本研究の大きなねらいは，役割の異なる言語単位を意識して複数の視点で論述文全体の流れを捉えるところにある．

実験結果から，複数の言語単位に対するスパン表現を活用することで優位な性能向上が観察され，取り組んだ全てのサブタスクで最高性能を達成した (5章)．また，木構造の深さ機能表現の有無という観点から性能の変化を分析し，提案モデルと既存のモデルで異なる傾向が見られた (6章).

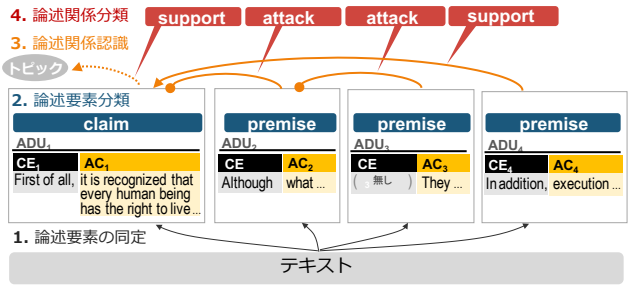


**図1:** 論述構造の例. 段落内には論述的談話単位 (ADU) が存在し，それぞれの ADU は更に論述要素 (AC) と接続表現 (CE) に分割ができる．論述構造解析では ADU 間の関係と，ADU の種類，関係の種類を予測する．

## 2 論述文における言語単位

論述文中の各段落には，以下の 3 種類の言語単位が存在する．

1. **論述要素 (AC)**: 著者の主張や根拠の内容に関連する箇所．論述文中の AC を読むことで著者の論述を構成する内容を把握することができる．

2. **接続表現 (CE)**: *by contrast*，*because* や，*I believe that* といった論述の流れを明示化する箇所．論述文中の CE を読むことで，論述の構造的な側面を捉えることができる．

3. **論述的談話単位 (ADU)**: AC と直前の CE からなる単位.

例えば図 1では，段落は 4 つの ADU に分割されており，更にそれぞれの ADU が CE と AC に分割可能である[*2]．本研究では，1章で説明した各タスクは ADU 単位の問題として解く[*3].

## 3 モデル

本セクションでは，提案モデルのアイデアを説明する．まず，トークンレベルの LSTM を用いて ADU スパン表現を獲得する ADU モデル (Figure 2 左 (a)) を説明する．次に ADU モデルを拡張し，CE，AC といったより詳細な言語単位同士で論述全体の情報を捉える AC-CE-ADU モデル (Figure 2 (b)) を説明する．

[*1] より詳細な定義は Stab ら [14]，Peldzsus ら [10] を参考にされたい.

[*2] CE のアノテーションがついていないデータセットでは，簡単なルールや Penn Discourse Treebank Annotation Manual[13] などの辞書を用いて CE のセグメンテーションをした.

[*3] Persuasive Essay Corpus 上では AC を基本単位としてタスクが定義されているが, AC と ADU は 1 対 1 に対応させることが可能であるため, ADU 単位の問題とみなす.

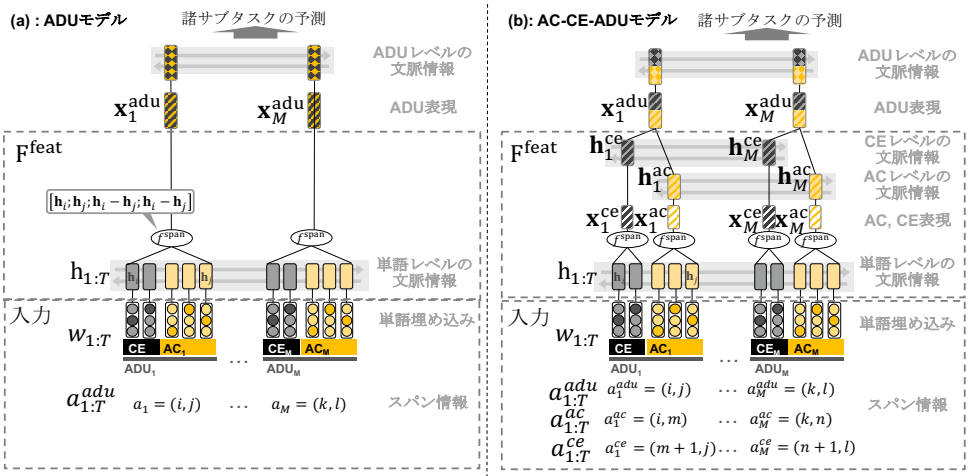


**図2:** ADU スパン表現の作成方法の概要である．(a) 単一スパン表現抽出関数は，ADU スパン表現を獲得するために ADU スパンの端の表現を用いる．すなわち，AC や CE といった言語単位を考慮していない．(b) 複数スパン表現抽出関数は AC，CE スパン表現から ADU スパン表現を獲得する．

### 3.1 モデルアーキテクチャの概要

始めに，$T$ 個のトークン $w_{1:T} = \{w_1, \cdots, w_T\}$ と $M$ 個の ADU スパン $a_{1:M} = \{a_1, \cdots, a_M\}$ を入力とし，ADU 表現の系列 $\mathbf{x}^{\text{adu}}_{1:M}$ を獲得する．

$$\mathbf{x}^{\text{adu}}_{1:M} = \mathrm{F}^{\text{feat}}(w_{1:T}, a_{1:M}) \ . \tag{1}$$

$a_m \in a_{1:M}$ は各 ADU の範囲を表しており，$a_m = (i, j)$ という 2 つのインデックスで表現される $(1 \leq i \leq j \leq T)$．各インデックスは段落内の単語列のインデックスに対応する．$\mathrm{F}^{\text{feat}}$ は，単語系列 $w_{1:T}$ と ADU スパン $a_{1:M}$ を入力とし，ADU 表現の系列 $\mathbf{x}^{\text{adu}}_{1:M}$ を返す．$\mathrm{F}^{\text{feat}}$ をどう設計するかが本研究の大きな焦点である．

次に双方向 LSTM を用いて ADU スパンの粒度で文脈情報を取り込む．

$$\mathbf{h}^{\text{adu}}_{1:M} = \text{BiLSTM}(\mathbf{x}^{\text{adu}}_{1:M}) \tag{2}$$

その後，$\mathbf{h}^{\text{adu}}_{1:M}$ を入力として，Softmax 関数を用いて AC/link type を予測する．論述関係認識では，各 ADU に対してソフトマックス関数を用いて論述関係の掛かり先を予測する．

### 3.2 ADU スパン表現抽出関数 $\mathrm{F}^{\text{feat}}$

式 1における $\mathrm{F}^{\text{feat}}$ として，2 つのアプローチを比較する．

#### 3.2.1 ADU モデル

図 2左が単一スパン表現による ADU 特徴抽出の概要である．始めに $f^{\text{emb}}$ が各単語に関して単語ベクトルを参照し，双方向 LSTM でそれらの単語ベクトルに文脈情報を取り込む．次に，$f^{\text{adu}}$ を用いて各 $a_m$ に対する ADU スパン表現を計算する．

$$\begin{aligned}
\mathbf{w}_{1:T} &= f^{\text{emb}}(w_{1:T}) \ , \\
\mathbf{h}_{1:T} &= \text{BiLSTM}(\mathbf{w}_{1:T}) \ , \\
\mathbf{x}^{\text{adu}}_m &= [\mathbf{h}_i \, ; \mathbf{h}_j \, ; \mathbf{h}_i - \mathbf{h}_j \, ; \mathbf{h}_j - \mathbf{h}_i] \ .
\end{aligned} \tag{3}$$

#### 3.2.2 AC-CE-ADU モデル

図 2右が複数スパン表現による ADU 特徴抽出の概要である．単一スパン表現による ADU 特徴抽出関数とは異なり，AC や CE といった異なる言語単位の情報を用いる．始めに AC と CE に対するスパン表現を得る．$T$ 個のトークン $w_{1:T}$ と $M$ 個の AC スパン $b_{1:M}$ から，式 3と同様に AC スパン表現 $\mathbf{x}^{\text{ac}}_{1:M}$ を獲得する．CE に対しても同様にスパン表現 $\mathbf{x}^{\text{ce}}_{1:M}$ を獲得する．その後，双方向 LSTM を用いて AC，CE 同士で文脈情報を取り入れる．

$$\mathbf{h}^{\text{ac}}_{1:M} = \text{BiLSTM}(\mathbf{x}^{\text{ac}}_{1:M}) \ .$$

$$\mathbf{h}^{\text{ce}}_{1:M} = \text{BiLSTM}(\mathbf{x}^{\text{ce}}_{1:M}) \ .$$

こうして得られた AC のスパン表現と CE のスパン表現を結合し，ADU のスパン表現を獲得する．

$$\mathbf{x}^{\text{adu}}_m = f^{\text{multi}}(\mathbf{h}_{1:T}, a_m) = [\mathbf{h}^{\text{ac}}_m \, ; \, \mathbf{h}^{\text{ce}}_m] \ ,$$

### 3.3 学習

本論文では，論述要素分類，論述関係認識，論述関係分類の 3 タスクをジョイントで解き，複数タスクを同時に解くことによる精度向上を狙う．3 タスクの交差エントロピーを線形補間した損失 $l$ を最小化する．

$$\ell = \alpha \, \ell^{\text{link}} + \beta \, \ell^{\text{ac-type}} + \gamma \, \ell^{\text{link-type}} \ ,$$

実験では，$\alpha = 0.5$，$\beta = 0.25$，$\gamma = 0.25$ とした．

## 4 実験設定

Persuasive Essay Corpus (PEC) [14] と Arg-Microtext Corpus (MTC) [9] を用いた．両コーパスに関して，著者と同じ方法で訓練データと評価データを分け，訓練データの 10% を開発データとした．PEC では 3 つの異なる初期値でモデルを学習し，それらのモデルのスコアの平均を用いた．MTC では 5 分割交差検証を 10 回繰り返した結果の平均を用いた．

既存研究に従い，各タスクの評価指標として F1 スコアを，モデルの総合的な評価指標として 3 タスクにおける MacroF1 スコアの平均を用いた．各タスクの MacroF1 スコアをブートストラップ検定 [3] によって検定した．また，単語表現として，異なる性質を持つ GloVE と ELMo を用いて実験した．

### 4.1 比較手法: Bag-of-Words 表現を用いたモデル

Potash ら [12] と同様の方法で，ADU スパン表現として BoW ベースの素性を用いた．この関数を用いたモデルを BoW モデルと呼ぶ．この手法と他の手法を比較することで，単語レベルの LSTM を用いたスパン表現が本タスクにおいて有効であることを示す．

## 5 結果

表 1から，複数スパンを考慮したモデルが最も良い性能を示したことが分かる．また，Link identification (LI) タスクと

**表1:** PEC データセット上での 3 タスクにおける結果．F1 スコアを評価指標として用いている．† は統計的に有意差があることを ($p < 0.05$) 示す．数値の右上のマークは BoW モデルと比較した際の有意差を，右下のマークは単一スパンモデルと比較した際の有意差を示す．

| | | Overall | 論述関係認識 | | | 論述関係分類 | | | 論述要素分類 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Avg. | Macro | リンクあり | リンクなし | Macro | Support | Attack | Macro | MajorClaim | Claim | Premise |
| ELMo | AC-CE-ADU モデル | **80.9** | **80.4**$^{\dagger}_{\dagger}$ | **67.3** | **93.6** | **75.8**$^{\dagger}$ | **96.4** | **55.3** | **86.5**$^{\dagger}_{\dagger}$ | **92.6** | **74.7** | **92.2** |
| | ADU モデル | 78.5 | 78.4$^{\dagger}$ | 63.8 | 92.9 | 73.7$^{\dagger}$ | 96.1 | 51.2 | 83.5 | 90.2 | 69.5 | 90.8 |
| | BoW モデル | 74.0 | 73.7 | 56.1 | 91.4 | 66.7 | 95.5 | 37.9 | 81.5 | 87.4 | 67.3 | 89.5 |
| GloVe | AC-CE-ADU モデル | **79.7** | **78.3** | **63.8** | **92.9** | **77.2**$^{\dagger}$ | **97.0** | **57.7** | **83.6** | **88.8** | **71.0** | **91.1** |
| | ADU モデル | 78.7 | 78.1$^{\dagger}$ | 63.3 | 92.8 | 75.5 | 96.3 | 54.7 | 82.6 | 86.7 | 70.1 | 91.0 |
| | BoW モデル | 76.1 | 74.2 | 56.7 | 91.5 | 71.3 | 96.0 | 46.6 | 82.8 | 90.0 | 68.2 | 90.3 |
| Pointer Net (Potash et al., 2017) | | - | 76.7 | 60.8 | 92.5 | - | - | - | 84.9 | 89.4 | 73.2 | 92.1 |
| Pointer Net (our reimplement.) | | - | 76.3 | 60.4 | 92.1 | - | - | - | 84.7 | 88.6 | 73.3 | 92.3 |
| ILP Joint (Stab and Gurevych 2017) | | 75.2 | 75.1 | 58.5 | 91.8 | 68.0 | 94.7 | 41.3 | 82.6 | 89.1 | 68.2 | 90.3 |

**表2:** アブレーションテストの結果．各数値は F1 スコアを示す．† は統計的に有意差があることを ($p < 0.05$) 示す．

| | 全体 | 関係認識 | 関係分類 | 要素分類 |
|---|---|---|---|---|
| | Avg. | Macro | Macro | Macro |
| AC-CE-ADU モデル (ELMo) | **80.9** | **80.4** | 75.8 | **86.5** |
| – ADU レベルの文脈情報 | 80.5 | **80.4** | 75.9 | 85.1$^{\dagger}$ |
| – AC レベルの文脈情報 | 79.3 | 79.3$^{\dagger}$ | 74.1 | 84.9$^{\dagger}$ |
| – CE レベルの文脈情報 | 80.0 | 78.7$^{\dagger}$ | **76.7** | 84.7$^{\dagger}$ |
| 論述関係認識のみ | - | 78.3$^{\dagger}$ | - | - |
| 論述関係分類のみ | - | - | 75.2 | - |
| 論述要素分類のみ | - | - | - | 85.3 |

**表3:** MTC データセット上での 3 タスクにおける結果．F1 スコアを評価指標として用いている．† は統計的に有意差があることを ($p < 0.05$) 示す．数値の右上のマークは BoW モデルと比較した際の有意差を，右下のマークは単一スパンモデルと比較した際の有意差を示す．

| | Overall | 関係認識 | 関係分類 | 要素分類 |
|---|---|---|---|---|
| | Avg. | Macro | Macro | Macro |
| AC-CE-ADU モデル (ELMo) | **78.0** | **72.8** | **80.8**$^{\dagger}_{\dagger}$ | 80.3 |
| ADU モデル (ELMo) | 76.4 | 71.9 | 76.9$^{\dagger}$ | 80.3 |
| BoW モデル (ELMo) | 70.6 | 71.6 | 57.4 | **82.7** |

AC type classification (ATC) タスクにおいて，ELMo を用いるとモデル間の性能の向上がより顕著になることが確認された．3 つのサブタスク全てにおいて提案手法が最高性能を達成した．

表 2にアブレーションテストの結果を示す．表 2の中央ブロックから，LI タスクと ATC タスクにおいて，それぞれのレベルの言語単位のスパン表現に関して段落全体の文脈の情報を取り入れることの有効性が示された．また，表 2の下のブロックから，複数のタスクを同時に予測することでそれぞれのタスクの性能が向上することも分かる．この傾向は Potash ら [12] や，Stab ら [14] の研究と一致する．

また，マイクロテキストコーパス [9] 上でも複数の言語単位を考慮することの有効性に関して同様の傾向が見られた (表 3)．

## 6 分析

特に予測が難しい論述関係予測タスクにおいて，PEC データセット上でエラー分析をした．接続表現や論述要素といった複数の言語単位に個別のスパン表現を割り当てた効果を分析するため，特に接続表現に着目して分析を行った．また，論述構造における木構造の深さという観点からも分析を行い，我々のアプローチによる予測の性質を調べた．各表の数値は，各 AC に対してかかり先を予測した結果の正答率とする．

### 6.1 接続表現の有無から見た分析

*because*，*I believe that* といった接続表現が直前に存在する AC とそうでない AC を比較し，リンクの掛かり先予測精度の分析をした．接続表現が無いほうが予測が難しいことが分かる．スパンを用いたモデルは，接続表現が存在しない AC に対しても頑健に予測ができている．また，接続表現が存在する AC に対しても大きな性能の向上が見られる．各接続表現にスパン表現を割り当てたり，接続表現同士で文脈の情報を共有することにより，接続表現の情報をより効果的に扱えていることが分かる．

**表4:** 接続表現の有無に注目して AC をカテゴリ分けし，それぞれのカテゴリの AC に関して論述関係予測の正解率を示す．

| モデル | 接続表現あり | 接続表現なし |
|---|---|---|
| AC-CE-ADU モデル (ELMo) | **77.7** | **71.6** |
| ADU モデル (ELMo) | 73.0 | 66.9 |
| BoW (ELMo) | 68.0 | 63.1 |
| BoW (GloVe) | 69.6 | 63.1 |
| Joint Pointer Net (Previous SOTA) | 71.5 | 68.2 |


**図3:** 文脈情報を考慮した CE スパン表現を 2 次元平面上に可視化したもの．

### 6.2 接続表現スパンの定性的分析

接続表現のスパン表現がどのような性質をもっているかを調べる．図 3は，開発データの段落内の接続表現に対して，複数スパンモデルが割り当てたベクトルを可視化したものである．次元削減には PCA を用いた．青い点は “Premise” に分類される AC の直前の CE を，黄色い点は “Claim” の直前の CE を，赤い点は “MajorClaim” の直前に出現する CE を指す．3 種類の CE が概ねクラスタリングされている様子が観察される．

論述文中で似た役割を果たす CE スパンは，表層が異なるもの ( *to cite an example* と *for instance* など) に対しても近いベクトルが付与されることが観察できた．また，全く同じフレーズに対しても文脈に応じて異なる表現を割り当てていることが観測された．例えば，“Premise” を述べるために用いられた *first of all* は図 3の左上の位置に，“Claim” を述べるために用いられた *first of all* は図 3の右下に位置している．各言語単位に対して文脈を考慮したスパン表現を割り当てる行為は，ELMo が文脈を考慮して各単語に表現を割り当てる行為とも類似している．

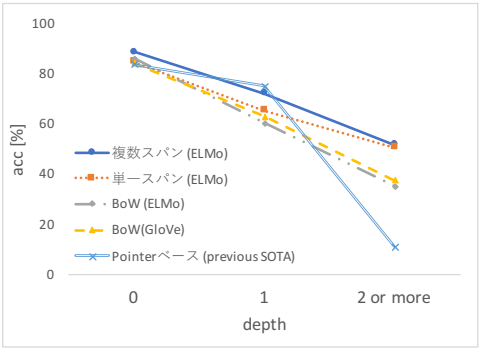


**図4:** AC の深さ別の論述関係予測における掛かり先予測正解率を示す．X 軸は AC の深さを Y 軸は正解率である．

**表5:** 段落構成別の論述関係予測の掛かり先予測正解率を示す.(a)，(b)，(c) はそれぞれ段落の構成の種類に対応する．

| Model | (a) | (b) | (c) |
|---|---|---|---|
| 複数スパン (ELMo) | 93.1 | **76.4** | **60.4** |
| 単一スパン (ELMo) | 92.3 | 74.4 | 47.8 |
| BoW (ELMo) | 91.5 | 62.9 | 54.3 |
| BoW (GloVe) | 88.7 | 67.6 | 50.7 |
| Joint Pointer Net (Previous SOTA) | **94.3** | 67.5 | 58.4 |

### 6.3 木構造の深さから見た分析

図 4から，深い階層に位置する AC から出るリンクは予測が難しいことが分かる．このことは，Stab ら [14] も報告している．既存研究のモデル [12] と比較して，複数スパン表現を用いたモデルは深い AC から出るリンクに対しても比較的頑健に予測できていることが分かる．

深い位置の論述関係を当てるには，例えば “of course”→“but”→ “therefore” といったマクロな論述の流れを捉える必要がある．AC-CE-ADU モデルでは，役割の異なる言語単位を意識して複数の視点で論述文全体の流れを考慮することができると考えられ，マクロな論述の構造を捉えることができていると示唆される．

### 6.4 文章の構成から見た分析

PEC コーパスは，複数の段落を持つ作文からなるコーパスであり，論述構造解析のタスクでは段落ごとに予測を行っている．コーパス中の段落は以下の 3 つのタイプに分類することが可能である．

(a) 作文中の導入，結論の段落．(20%)
(b) 導入，結論以外の段落で，主張から始まる段落 (トップダウン型)．(53%)
(c) 導入，結論以外の段落で，主張から始まらない段落 (ボトムアップ型)．(27%)

段落のタイプごとに，論述関係予測精度を分析した．傾向として (c) タイプの段落で論述関係の予測が難しいことが分かる．しかしながら，(c) タイプの段落に対しても複数スパンモデルは最も良い結果を出していることが分かる．

## 7 関連研究

論述構造予測において，離散的な素性を用いた予測モデル [6, 10, 14] や，ニューラルベースの手法を適用したモデル [1, 2, 12] など多くのアプローチが提案されてきた．この中でも特に本研究と関連が強い研究として，Potash ら [12] が挙げられる．ポインターネットワークの枠組みを論述構造解析に適用した．デコーダー側で予測結果の系列的な依存性を考慮している点，談話単位の表現として Bag-of-Words ベースの表現を用い単一のスパンのみを考慮している点で，我々の手法とは異なる．

談話構造解析の文脈では，Li ら [5] は談話単位の表現を得る際に，階層的に文脈の情報を取り入れている．しかし，複数の言語単位の情報を用いていることと，論述文のドメインに焦点を当てている点で我々のモデルは異なる性質を持つ．

スパン表現を用いたモデルは他の NLP タスクでも注目を浴びている [4, 8, 15, 16]．Wang ら [16] は LSTM を用いたスパン表現獲得の方法を提案しており，本研究は彼らのモデルの拡張とみなすことができる．

## 8 おわりに

複数の言語単位に対するスパン表現を用いた論述構造解析手法を提案した．我々のモデルは論述構造解析の 3 タスクで最先端の性能を記録した．エラー分析から，論述関係予測においてさらなる精度向上を実現するための課題を提示した．木構造中の深い位置の論述関係や，主張が論述文の頭に登場しないような構成の段落においてロバストに構造を予測することが重要と考えられる．また，AS タスクも含めた End-to-End のセッティングにおいても，高い予測精度を実現できるようにしたい．

## 謝辞

本研究の一部は JST 未来社会創造事業（JPMJMI17C7）および JST CREST（JPMJCR1513）の支援を受けて行った．

## 参考文献

[1] Oana Cocarascu and Francesca Toni. “Identifying attack and support argumentative relations using deep learning”. In: *ACL* (2017), pp. 1385–1390.
[2] Steffen Eger, Johannes Daxenberger, and Iryna Gurevych. “Neural End-to-End Learning for Computational Argumentation Mining”. In: *ACL*. Vancouver, Canada: Association for Computational Linguistics, 2017, pp. 11–22.
[3] Philipp Koehn. “Statistical Significance Tests for Machine Translation Evaluation”. In: *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*. 2004.
[4] Kenton Lee et al. “End-to-end Neural Coreference Resolution”. In: *Proceedings of EMNLP*. 2017, pp. 188–197.
[5] Qi Li, Tianshi Li, and Baobao Chang. “Discourse Parsing with Attention-based Hierarchical Neural Networks”. In: *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. Austin, Texas: Association for Computational Linguistics, 2016, pp. 362–371.
[6] Huy V Nguyen and Diane J Litman. “Context-aware Argumentative Relation Mining”. In: *ACL* (2016), pp. 1127–1137.
[7] Vlad Niculae, Joonsuk Park, and Claire Cardie. “Argument Mining with Structured SVMs and RNNs”. In: (2017), pp. 985–995.
[8] Hiroki Ouchi, Hiroyuki Shindo, and Yuji Matsumoto. “A Span Selection Model for Semantic Role Labeling”. In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Brussels, Belgium: Association for Computational Linguistics, 2018, pp. 1630–1642.
[9] Andreas Peldszus and Manfred Stede. “An Annotated Corpus of Argumentative Microtexts”. In: *Studies in Logic and Argumentation* (2016).
[10] Andreas Peldszus and Manfred Stede. “Joint prediction in MST-style discourse parsing for argumentation mining”. In: (2015), pp. 938–948.
[11] Isaac Persing and Vincent Ng. “End-to-End Argumentation Mining in Student Essays”. In: (2016), pp. 1384–1394.
[12] Peter Potash, Alexey Romanov, and Anna Rumshisky. “Here’s My Point: Joint Pointer Architecture for Argument Mining”. In: *Proceedings of EMNLP*. 2017, pp. 1375–1384.
[13] Rashmi Prasad et al. “The penn discourse treebank 2.0 annotation manual”. In: (2007).
[14] Christian Stab and Iryna Gurevych. “Parsing argumentation structures in persuasive essays”. In: *Proceedings of Computational Linguistics*. Vol. 43. 3. MIT Press, 2017, pp. 619–659.
[15] Mitchell Stern, Jacob Andreas, and Dan Klein. “A Minimal Span-Based Neural Constituency Parser”. In: *Proceedings of ACL*. 2017, pp. 818–827.
[16] Wenhui Wang and Baobao Chang. “Graph-based dependency parsing with bidirectional lstm”. In: *Proceedings of ACL*. 2016, pp. 2306–2315.