

回答の根拠を解釈可能な機械読解

西田 光甫 西田 京介 齊藤 いつみ 浅野 久子 富田 準二
日本電信電話株式会社 NTT メディアインテリジェンス研究所

kosuke.nishida.ap@hco.ntt.co.jp

1 はじめに

テキストを参照し読み解きながら質問に答える機械読解が注目を集めており, BERT [2] を始めとした深層学習モデルによって取り組まれている. しかし, 深層学習モデルは Black Box であり, モデルの予測の妥当性や予測の理由がわからないという課題が存在する.

本研究は解釈可能な機械読解モデルを構成することを目指す. 本研究では, モデルの解釈可能性を, テキスト分類における先行研究 [4, 1] 同様に予測の根拠を用いたパイプライン型のモデルによって定義する.

定義 1 (解釈可能な機械読解モデル). 機械読解モデルが解釈可能であるとは, モデルが下記の入出力を持つ 2つのモジュールから構成されることを言う.

入力 (根拠モジュール): 参照テキスト P , 質問文 Q

出力 (根拠モジュール): 根拠 R

入力 (回答モジュール): 根拠 R , 質問文 Q

出力 (回答モジュール): 回答 A

ただし, 根拠 R は参照テキスト P の部分文字列の集合である. 本研究では文の集合とする.

図 1 に解釈可能な機械読解モデルの概念図を示す. 解釈可能な機械読解モデルの特徴は, 回答モジュールが参照テキスト P のうち根拠 R のみを利用することである. 例えば根拠モジュールの隠れ状態は, 根拠 R 以外のテキストに由来する情報を含むため, 利用しない. 解釈可能な機械読解モデルのメリットに, (1) 回答を予測した理由を根拠 R のみに限定できる, (2) 回答モジュールの入力が短くなるため回答精度の向上が期待できる, の 2点がある.

我々は解釈可能な機械読解モデルと End-to-End での学習アルゴリズムを提案する. 提案モデルは根拠の教師データがある場合に教師あり学習が可能であるが, 教師あり学習では回答モジュールの損失を根拠モジュールに伝播することができない. そこで, gumbel-softmax trick [3, 5] を用いた End-to-End の教師なし追加学習を行うことで, 回答モジュールが回答しやすい根拠を出力するように根拠モジュールを学習できる. 評価実験により, 解釈可能な機械読解モデルの回答精度が根拠モジュールを用いない機械読解モデルの性能を上回ることで, End-to-End の追加学習をすることでさらに性能が向上することを示す.

2 提案手法

提案手法の概要を示す. それぞれのモジュールは BERT[2] などの事前学習済み言語モデルとモジュール

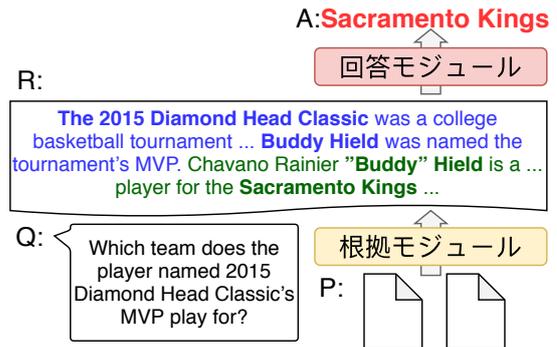


図 1 解釈可能な機械読解.

特有のネットワークで構成される. 図 2 に概略図を示す. なお, 提案手法は参照テキストから 1 区間を抽出することで回答する抽出型機械読解に限定して記述するが, 他の機械読解タスクにも適用できる.

2.1 根拠モジュール

入力: 根拠モジュールは参照テキスト P と質問 Q を文字列として受け取る. 参照テキスト P が複数のテキストから成る場合は連結した 1つのテキストとみなす.

言語理解層: 参照テキスト P と質問 Q をトークン系列 $[[CLS_Q]; \text{質問}; [SEP_Q]; [CLS_P]; \text{文 } 1; [SEP_P]; \dots; [CLS_P]; \text{文 } n; [SEP_P]]$ として BERT に入力する. $[CLS_Q]$, $[SEP_Q]$, $[CLS_P]$, $[SEP_P]$ は特殊トークンであり, n は参照テキスト P 中の文数である. BERT の出力のうち, $[CLS_Q]$ に相当する位置のベクトルを質問ベクトル $\mathbf{q} \in \mathbb{R}^d$, i 番目の $[CLS_P]$ に相当する位置のベクトルを文 i のベクトル $\mathbf{s}_i \in \mathbb{R}^d$ とする. d は BERT の出力の次元である. 言語理解層の出力は質問ベクトル \mathbf{q} と, 参照テキスト P 中の全ての文ベクトルの集合 $\{\mathbf{s}_i\}$ である.

根拠抽出層: 根拠抽出層は質問ベクトル \mathbf{q} , 文ベクトル集合 $\{\mathbf{s}_i\}$ を入力として受け付ける. 出力は, 根拠 R として用いる文の id の集合 \hat{R} である. 根拠抽出層は, 先行研究 [7] 同様に GRU を用いて再帰的に文を抽出する. 根拠抽出層が時刻 t に行う処理は, 文 \hat{r}^t を抽出しその文ベクトル $\mathbf{s}_{\hat{r}^t}$ を用いて質問ベクトル \mathbf{q}^t を更新する処理である. この処理を抽出が終了するまで繰り返し $\hat{R} = \{\hat{r}^t\}$ を出力する. $\mathbf{q}^0 = \mathbf{q}$ である.

時刻 t には, 初めに質問ベクトル \mathbf{q}^t を用いて, 文 i のスコア

$$p_i^t = \mathbf{q}^{t\top} W_p \mathbf{s}_i \in \mathbb{R}$$

を得る. $W_p \in \mathbb{R}^{d \times d}$ は訓練可能なパラメータである.

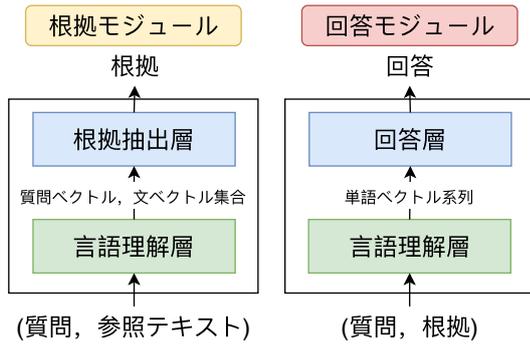


図2 モデル概略.

根拠抽出層の出力である、時刻 t に抽出する文の id \hat{r}^t は、文スコアに応じて決定する。推論時は、これまでに抽出されていない文の中で最もスコアの高い文

$$\hat{r}^t = \operatorname{argmax}_{i \in S \setminus \hat{R}^{t-1}} p_i^t$$

である。ただし、 S は文 id 全体の集合、 \hat{R}^{t-1} は時刻 $t-1$ までに抽出された文 id の集合である。

そして、抽出した文のベクトルを用いて、質問ベクトルを更新する

$$\mathbf{q}^{t+1} = \operatorname{GRU}(s_{\hat{r}^t}; \mathbf{q}^t) \in \mathbb{R}^d.$$

質問ベクトル \mathbf{q}^t は質問に答えるために不足している情報を表現する。初期状態 \mathbf{q}^0 は質問に答えるために必要な情報の全てであり、抽出した文 \hat{r}^t にある情報が GRU によって \mathbf{q}^t から除かれることが期待される。

抽出の終了を判定するため、文ベクトルの集合 $\{s_i\}$ にダミーの文 (EOE 文と呼ぶ) のベクトル $s_{EOE} \in \mathbb{R}^d$ を追加する。推論時は EOE 文を他の文と同様に扱い、EOE 文が選択された際に抽出を終了する。EOE 文ベクトル s_{EOE} は学習可能なパラメータであり全データで共通とする。

2.2 回答モジュール

言語理解層： 予測された根拠 \hat{R} と質問 Q をトークン系列 $['[CLS]'; \text{質問}; '[SEP]'; \text{文 } r^1; \dots; \text{文 } r^T; '[SEP]']$ として BERT に入力する。 T は抽出した文数、 $['[CLS]'$, $['[SEP]'$ は特殊トークンである。出力は各トークンのベクトルの系列 $H \in \mathbb{R}^{l \times d}$ である。 l は系列長である。

回答層： 言語理解層の出力を線形変換層によって回答のスコアに変換する

$$[a_{s,i}; a_{e,i}]^\top = W_a h_i + b_a \in \mathbb{R}^2.$$

$a_{s,i}, a_{e,i}$ は第 i トークンが回答の始端と終端になるスコアを表す。 $W_a \in \mathbb{R}^{2 \times d}, b_a \in \mathbb{R}^2$ は学習可能なパラメータである。

2.3 教師あり学習

損失関数： 教師あり学習では、回答モジュールと根拠モジュールを別々に学習する。回答モジュールの損失 L_A は線形変換層出力の Cross-Entropy 損失の和

$$L_A = -\log \operatorname{softmax}(\mathbf{a}_s)_i - \log \operatorname{softmax}(\mathbf{a}_e)_j$$

である。ただし、 i は真の回答範囲の始端、 j は終端である。根拠抽出モジュールの損失 L_R は各時刻の根拠抽出に関する負の対数尤度の平均を用いる

$$L_R = -\frac{1}{T} \sum_{t=1}^T \log \left(\Pr(\hat{r}^t; \hat{R}^{t-1}) \right).$$

$\Pr(i; \hat{R}^{t-1})$ は時刻 t に文 i が出力される確率である

$$\Pr(i; \hat{R}^{t-1}) = \operatorname{softmax}_i(p_i^t).$$

抽出文の決定： 時刻 t に出力する文は Teacher-Forcing を用いて決定する。即ち、時刻 t に出力する文 \hat{r}^t は真の根拠 R 中から選択する

$$\hat{r}^t = \operatorname{argmax}_{i \in R \setminus \hat{R}^{t-1}} p_i^t.$$

根拠の教師データには順序が存在しないため、上式の argmax 操作によって、まだ選ばれていない根拠の中で最もスコアが高い文を時刻 t の正解データとみなす。EOE 文は真の根拠 R 全てが抽出された後に選択する。

2.4 End-to-End 追加学習

2.3 節で説明した教師あり学習では回答モジュールと根拠モジュールが別々に学習されるため、相互の依存関係を学習できない。そこで、教師あり学習後に回答の損失を用いて両モジュールを End-to-End に学習する手法を提案する。根拠の抽出操作 argmax は微分不可なので、gumbel-softmax trick によって勾配の伝播を可能にする。なお、追加学習では根拠の教師データを用いない教師なし学習を行う。教師あり学習中の根拠 \hat{R} は必ず真の根拠だが、追加学習中は真の根拠とは限らない。

抽出文の決定： g_i ($i = 1, \dots, n$) を一様独立な gumbel 分布からサンプリングした確率変数とする¹。gumbel-softmax trick では、時刻 t に抽出する文 \hat{r}^t を

$$\hat{r}^t = \operatorname{argmax}_i \left(g_i + \log \Pr(i; \hat{R}^{t-1}) \right)$$

に従って決定する。one-hot ベクトル $1_{\hat{r}^t}$ の連続緩和は τ を温度パラメータとして

$$y_i = \operatorname{softmax}_i \left(\frac{g_i + \log \Pr(i; \hat{R}^{t-1})}{\tau} \right)$$

であり、逆伝播時は straight-through gumbel-softmax estimator $\nabla 1_{\hat{r}^t} \approx \nabla \mathbf{y}$ を近似値として用いる。

損失関数： End-to-End の学習では、入力が P, Q , 潜在変数が R , 出力が A のモデルの学習を行う。そのため、損失関数として、本来は確率分布

$$\Pr(A|P, Q) = \sum_{\hat{R}} \Pr(A|\hat{R}, Q) \Pr(\hat{R}|P, Q)$$

に対応する損失 $-\log \Pr(A|P, Q)$ を用いることが望ましい。そこで、 $-\log \Pr(A|P, Q)$ の近似として回答モジュールの損失 L_A を用いる。また、学習を補助する

¹ $u_i \sim \operatorname{Uniform}(0, 1)$ として $g_i = -\log(-\log(u_i))$ である。

目的で2種類の正則化項 $L_{\text{NoAns}}, L_{\text{Entropy}}$ を用意し、和 $L = L_A + \lambda_1 L_{\text{NoAns}} + \lambda_2 L_{\text{Entropy}}$ を損失関数とする。 λ_1, λ_2 はハイパーパラメータである。

回答モジュールの損失 $L_A = -\log \Pr(A|\hat{R}, Q)$ は、 $-\log \Pr(A|P, Q)$ の上限に相当することが Jensen の不等式により示される。即ち、

$$\begin{aligned} & -\log \Pr(A|P, Q) \\ &= -\log \sum_{\hat{R}} \Pr(A|\hat{R}, Q) \Pr(\hat{R}|P, Q) \\ &\leq -\sum_{\hat{R}} \Pr(\hat{R}|P, Q) \log \Pr(A|\hat{R}, Q) \\ &= -\mathbb{E}_{\hat{R} \sim \Pr(R|P, Q)} \left[\log \Pr(A|\hat{R}, Q) \right] \\ &\approx -\log \Pr(A|\hat{R}, Q) = L_A \end{aligned}$$

である。最後の近似は gumbel-softmax trick による。根拠の抽出に失敗した場合、回答 A が抽出した根拠 \hat{R} に含まれないことがある。このとき、回答位置に関する Cross-Entropy 損失によって回答の損失 L_A を計算することができない。そこで、回答が抽出した根拠に含まれない場合は損失計算用の根拠

$$\hat{r}_{\text{Ans}} = \operatorname{argmax}_{i \in S_A} p_i^0$$

を \hat{R} に追加して回答モジュールに渡す。ただし、 S_A は回答を含む文の集合である。

損失計算用の根拠を追加することで根拠に回答が含まれることが保証されてしまうため、根拠モジュールが回答を含む文を抽出する動機付けができなくなる。そこで、回答がないことへの罰則

$$\begin{aligned} L_{\text{NoAns}} &= \min_{1 \leq t \leq T} \max(0, p_{\hat{r}^t}^t - p_{\hat{r}_{\text{Ans}}^t}^t), \\ \hat{r}_{\text{Ans}}^t &= \operatorname{argmax}_{i \in S_A} p_i^t \end{aligned}$$

を損失関数に加える。 \hat{r}_{Ans}^t は回答を含む文のうち時刻 t に最も選ばれやすかった文である。1回でも回答を含む文が他の文より選ばれやすかった場合、その時刻で $p_{\hat{r}^t}^t \leq p_{\hat{r}_{\text{Ans}}^t}^t$ となるため罰則 $L_{\text{NoAns}} = 0$ となり、それ以外の場合は $L_{\text{NoAns}} > 0$ である。

エントロピー正則化は強化学習でよく用いられる。

$$L_{\text{Entropy}} = \frac{1}{T} \sum_t \log \Pr(\hat{r}^t; \hat{R}^{t-1})$$

は1時刻の根拠抽出に関する負のエントロピーに相当する。エントロピーを大きくすることで根拠抽出の探索範囲を広げ、学習を安定させる効果がある。

3 評価実験

3.1 データセット

本研究では、根拠の教師データを持つデータセットである HotpotQA [8] を用いて評価した。HotpotQA では、質問文 Q は Wikipedia 中の2段落に跨る内容を問うように作成される。参照テキスト P は上記2

	回答	根拠
ベースライン	55.0/69.5	—
教師あり学習	56.5/70.5	61.3/87.2
追加学習	57.6/72.3	29.6/80.3

表1 回答と根拠の EM/F1 での評価結果。

段落を繋げたテキストとした²。出力は回答 A と根拠 R である。回答 A は回答ラベル {yes, no, span} のいずれかと回答区間である。回答区間は、回答ラベルが 'span' のときのみ存在する。よって、回答層では回答区間の抽出に加えて回答ラベルの分類も行った。質問が2段落に跨る内容を問う質問に限定されているため、根拠 R は2文以上である。議論のため、回答を含む根拠を回答文、回答を含まないが回答するために必要な根拠を補助文と呼ぶ。例を図1に示す。回答文は緑で、補助文は青で示す。

3.2 実験設定

本研究では、BERT_{Base} を用いた3つの手法を比較した。ベースラインモデルは根拠モジュールなしのモデルであり、回答モジュールに直接参照テキスト P と質問文 Q を入力する。提案手法は、教師あり学習と End-to-End の追加学習を評価した。

実験は NVIDIA Quadro P6000 1枚で行った。教師あり学習ではバッチサイズを60、エポック数を3、最適化手法を Adam、学習率を $5e-5$ とした。追加学習では、エポック数を1、学習率を $5e-6$ 、 τ を0.001、 λ_1 を1、 λ_2 を0.001とした。

3.3 評価と議論

解釈可能な機械読解モデルと End-to-End 追加学習の有効性検証のため、以下のリサーチクエスチョンについて議論する。

提案手法は回答精度の向上に貢献するか。表1に回答の評価結果を示す。回答精度に関して、提案手法はベースラインモデルの性能を上回った。ベースラインモデルの性能を上回ったことで、回答モジュールの前に根拠モジュールを用いる解釈可能な機械読解モデルが、回答モジュール単体に比べて精度よく回答できることを確認した。これは、根拠モジュールが回答に必要なテキストを取り除く効果があり、回答モジュールでの推論が容易になるためであると考えられる。

また、End-to-End の追加学習によって回答精度がさらに向上することを確認した。これは、End-to-End の追加学習によって回答モジュールが正答しやすい根拠を根拠モジュールが学習したためであると考えられる。次に各モジュールの挙動の変化を議論する。

End-to-End 追加学習で根拠モジュールの挙動がどう変化したか。表1では、End-to-End の追加学習によって根拠の教師データに対する精度が下がっている。よって、人間のアノテーションの再現という観点では、教師あり学習が適していると言える。

²HotpotQA の本来のタスク設定では参照テキストに10個の段落を用意する。本研究では入力長の制限から2段落のみを用いた。

	Precision	Recall	F1	文数
教師あり学習	91.5	85.2	87.2	2.19
追加学習	71.6	96.1	80.3	3.44

表2 根拠の抽出結果.

	回答文	補助文
教師あり学習	86.5	84.9
追加学習	96.2	96.8

表3 真の根拠文の種類ごとの抽出率 (Recall).

表2で根拠の出力の変化について調べた。EOE文によって抽出の終了を動的に判定する提案手法は、教師あり学習時にはPrecisionを重視して抽出する。一方で追加学習後はRecall重視に傾向が変化している。抽出する文数の増加は平均で1.25文である。

Recall重視に変わった理由を調査するため、3.1節で定義した根拠文の種類(回答文・補助文)別でRecallを評価した。表3に評価結果を示す。教師あり学習では、回答文が補助文よりも抽出されやすかった。これは、質問文が回答文の表現を流用することが多いためだと考えられる。

End-to-Endの追加学習でRecallが増大した理由は、回答の損失 L_A と根拠に回答がないことの罰則 L_{NoAns} に求めることができる。表3では、追加学習の後で回答文よりも補助文が選ばれやすくなっている。根拠に回答がないことの罰則 L_{NoAns} は回答文が選ばれる確率を高める効果を持つ。回答の損失 L_A は、回答モジュールが回答しやすい根拠を抽出するように学習を進める効果を持つ。補助文の抽出率が上がった本結果は、回答モジュールが回答するにあたって根拠が過剰であることよりも不足することの方が悪影響を持つことを、根拠モジュールが回答の損失 L_A から学習していることを示唆している。

End-to-End追加学習で回答モジュールの挙動がどう変化したか。 回答モジュール単体としての性能を評価するため、根拠の予測結果で開発データを分類して評価を行った。両モデルで根拠の抽出結果 \hat{R} が真の根拠 R に対して完全一致($R = \hat{R}$)・過剰($R \subsetneq \hat{R}$)・不足($R \supsetneq \hat{R}$)となるサンプルに限定したデータ群で評価した。回答ラベルが’span’かつ回答文を抽出していないサンプルでは回答が不可能であるため、分析に用いなかった。表4に評価結果を示す。

End-to-End追加学習では、全データ群で性能が向上している。これは、根拠モジュールがRecall重視になったこととは別に、回答モジュール単体で見た場合もEnd-to-End追加学習によって性能が向上したことを示す。追加学習では回答モジュールの入力が予測した根拠になるので、推論時に近い状況でロバストに学習ができ、精度よく回答ができると考えられる。

一方で、根拠不足のデータ群では、回答するために必要な情報を根拠 \hat{R} が持たない。よって、根拠不足のデータ群では本来は回答は不可能である。根拠不足である場合も精度が上がっている本結果は、解釈可能性の観点からは望ましくない。回答不能の認識は今後の課題であると言える。

	全て	完全一致	過剰	不足
教師あり学習	57.2/72.3	61.5/75.4	57.3/72.0	52.7/60.3
追加学習	58.6/74.4	63.0/77.0	58.3/74.0	61.5/67.6

表4 根拠の抽出結果別の回答結果 (EM/F1).

4 おわりに

本研究は機械読解を解釈可能にすることに取り組んだ。本研究の貢献を以下に示す。

本研究の独自性 解釈可能な機械読解モデルを定義し、End-to-Endでの学習手法を提案した初めての研究である。解釈可能な機械読解モデルの関連研究には[6]がある。彼らは機械読解の計算コストを削減するために、回答の前に文選択をするモデルを提案した。彼らの手法では事前学習済み言語モデルを用いない文選択モジュールを疑似ラベルで回答モジュールと独立に学習する。彼らはモデルの解釈可能性には言及していない。

本研究の重要性 解釈可能な機械読解モデルとEnd-to-Endでの追加学習の性能を評価した。回答モジュールのみを使う通常の機械読解モデルに比べて、解釈可能な機械読解モデルは前段で根拠の抽出を行うことで精度が向上することを確認した。さらにEnd-to-Endの学習を追加することで精度が改善した。End-to-Endの追加学習では、回答モジュール・根拠モジュールそれぞれで回答精度が改善する学習が進んでいることが確認できた。

機械読解モデルを解釈可能にすることで、ユーザーに納得感を与えることができる、ソースを明示化でき事実検証が可能になる、のように従来の機械読解が抱える社会的課題を解決できる。また、本研究で行ったEnd-to-Endの追加学習をスクラッチからの教師なし学習に拡張することで、根拠の教師データを持たないデータセットでも根拠を抽出することが可能となる。

参考文献

- [1] J. Bastings, W. Aziz, and I. Titov. Interpretable neural predictions with differentiable binary variables. In *ACL*, pp. 2963–2977, 2019.
- [2] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. *NAACL-HLT*, pp. 4171–4186, 2019.
- [3] E. Jang, S. Gu, and B. Poole. Categorical reparameterization with gumbel-softmax. In *ICLR*, 2017.
- [4] T. Lei, R. Barzilay, and T. Jaakkola. Rationalizing neural predictions. In *EMNLP*, pp. 107–117, 2016.
- [5] C. J. Maddison, A. Mnih, and Y. W. Teh. The concrete distribution: A continuous relaxation of discrete random variables. In *ICLR*, 2017.
- [6] S. Min, V. Zhong, R. Socher, and C. Xiong. Efficient and robust question answering from minimal context over documents. In *ACL*, pp. 1725–1735, 2018.
- [7] K. Nishida, K. Nishida, M. Nagata, A. Otsuka, I. Saito, H. Asano, and J. Tomita. Answering while summarizing: Multi-task learning for multi-hop QA with evidence extraction. In *ACL*, pp. 2335–2345, 2019.
- [8] Z. Yang, P. Qi, S. Zhang, Y. Bengio, W. W. Cohen, R. Salakhutdinov, and C. D. Manning. HotpotQA: A dataset for diverse, explainable multi-hop question answering. In *EMNLP*, pp. 2369–2380, 2018.