

強化学習を用いた自動要約における基礎研究

慶留間 諒大†

當間 愛晃‡

† 琉球大学工学部情報工学科

1 はじめに

1.1 自動要約の歴史

自動要約とは与えられた単一、または複数の文書から、その内容を簡潔に表した短い文書を自動的に生成する研究課題である。自動要約の歴史は1950年代に行われたLuhnの研究から始まる[1]。Luhnの手法では3つの手順で要約を作成する。はじめに、文書中から単語の頻度に基づき重要語を抽出する。次に、重要語を含む比率によって文に重要度を与える。最後に指定された要約率(要約の長さ÷原文の長さ)に達するまで重要度の高い順に文を抽出するといったものである。その後も様々な研究が行われているが、重要文を抽出して要約を作成するという方法は現在でも代表的な要約手法である。

しかし、現在では重要文を抽出する以外の要約手法の研究も行われており、その中の一つに文圧縮というものがある。文圧縮とは文を対象にして、単語、またはフレーズを抽出単位とする要約手法である。文圧縮は文中の冗長な表現を削除できるため、文の抽出と文圧縮の両方を逐次的に利用して要約を行う研究が多くなっている[5]。

1.2 研究背景

自動要約の代表的な手法として、MMR(Maximal Marginal Relevance)という貪欲的な手法が有る。この手法は局所的に良い要約を作成できても全体としてよい要約を作れる保証はなく[3]、大域的に良い要約を作成しなければならない。大域的に最適化された要約を作成するために、ナップザック問題として定式化した上で整数計画問題を最適化したり、最大被覆問題として定式化して最適な要約を作成するというアプローチが行われている[2]。しかしこれらの整数計画問題や最大被覆問題と言った最適化問題はNP-Hardであり、文圧縮も含めると膨大な計算時間が必要であるという問題が有る。MMRや最適化問題の他に、自動要約を強化学習問題として定式化して要約を得る研究が近年行われ始めた[3]。本研究でも自動要約を強化学習問題としてとり扱う。

1.3 研究目的

自動要約を強化学習として定式化し、実際に要約を行う研究は既に行われている。しかし、先行研究の実験で使用された学習手法はTD()のみであり、他の学

習手法による実験は行われていない。TD()以外との学習法による比較を行うことでより要約に適した学習手法を探していくことが本研究の目的である。

2 強化学習

強化学習とは試行錯誤を通して環境に適応する学習の枠組みである[4]。教師有り学習のように正しい行動を明確に示す教師信号は存在せず、報酬という名のスカラ値を基に行動した結果の善し悪しを判断していく。

学習主体となるエージェントは、(1)環境から与えられた状態 s_t を観測し、(2)それに対してなんらかの行動 a_t を取る。行動の結果、(3)環境は状態 s_{t+1} へ遷移し、遷移に応じた報酬 r_t がエージェントに与えられる。この3ステップを繰り返し行うことで、各状態においてどのような行動を取れば良いのかを学習していく。

強化学習の利点は、何をさせたいかを報酬という形で与えておくだけでエージェントがどのようにすべきかを自動的に学習していく点である。

3 先行研究による自動要約定式化

梁ら[3]は、自動要約を強化学習問題として定式化した。梁らの研究では文抽出と文圧縮の両方を考慮した問題を強化学習として定式化している。以下に文抽出のみを考慮して定式化されたものについて紹介する。文抽出のみを考慮した方法について紹介する理由は、文抽出と文圧縮の両方を考慮した手法が文抽出のみを考慮した手法を拡張したものであるためである。状態 s は現在の要約に含まれる文集合を保持しており、任意の要約 S の状態を表現する。状態集合全体 S は、可能な任意の要約全体である。

$$s_t = \begin{pmatrix} S_t \\ A_t \\ f_t \end{pmatrix} \quad (1)$$

ここで、 S_t は時刻 t における要約であり、 A_t は、状態 s_t に至るまでに取った行動の集合である。また、 f_t は終了状態を表すビットである。終了状態集合は $S_G = \{s \mid s \in S, f=1\}$ であり、初期状態はどの文も追加されていない空の状態である。

次に行動について説明する。行動 a は、途中過程の要約に対して何らかの手を加える操作である。文抽出の問題であると考えられる場合、行動集合 A は以下のように考えることができる。

$$A = \{insert_1, insert_2, \dots, insert_N, finish\} \quad (2)$$

‡A survey and basic study on automatic summarization using reinforcement learning.

†Ryouta KERUMA ‡Naruaki TOMA

†Department of Information Engineering, Faculty of Engineering, University of the Ryukyus

行動集合中の insert は要約に文を加える操作である。それぞれ番号に対応した位置の文を要約に加える。また、指定された要約長を超えていなければ、finish とこれまでに要約に加えられていない文を加える操作が可能であり、要約長を超えていれば強制的に終了する (finish を行う) ことになる。

なんらかの行動を取ることで、それぞれ以下のような決定的な状態遷移をする。

$$\begin{pmatrix} S_t \\ A_t \\ 0 \end{pmatrix} \xrightarrow{a_t \text{ insert}_i} \begin{pmatrix} S_{t+1} \\ S_t \cup \{x_i\} \\ A_t \cup \{\text{insert}_i\} \\ 0 \end{pmatrix} \quad (3)$$

$$\begin{pmatrix} S_t \\ A_t \\ 0 \end{pmatrix} \xrightarrow{\text{finish}} \begin{pmatrix} S_t \\ A_t \cup \{\text{finish}\} \\ 1 \end{pmatrix} \quad (4)$$

最後に、報酬の与え方について説明する。環境は、エージェントが終了状態に達したとき、つまり、行動 finish を行った時に以下のように報酬 r_t を与える。

$$r_t = \begin{cases} \text{score}(S_t) & (K-C \leq L(S_t) \leq K) \\ -R_{\text{penalty}} & (\text{otherwise}) \end{cases} \quad (5)$$

K は要約長の上限、 C は要約長の許容範囲を定める定数である。終了状態に達したときに要約長が許容範囲に収まっていれば、あらかじめ定義されたスコア関数によって要約の内容に応じたスコアが与えられ、要約が短すぎる、または長過ぎる場合には負の報酬が与えられる。

目的スコア関数は次のように定義される。

$$\text{score}(S) = \alpha \text{Sim}(S, D) - (1 - \alpha) \max_{x_i, x_j} \min_{S, i, j} \text{Sim}(x_i, x_j) \quad (6)$$

D とは要約の対象となる文書集合である。Sim はコサイン類似度であり、各ベクトルは D における tf-idf 表現ベクトルの上位 d 要素 (d は $d > 0$ であるパラメータ) までのみを考慮する。 α は $0 \leq \alpha \leq 1$ を満たすトレードオフ定数である。第一項は要約の妥当性、第二項は冗長性を表している。

4 先行研究の問題に対する検討項目

先行研究では TD() 学習での実験が行われているが、その他の学習方法での実験結果は示されていない。強化学習の学習手法は環境同定型と経験強化型の二種類に分類され、TD() は環境同定型に含まれる。環境同定型の学習法を用いた場合と、経験強化型の手法を用いた場合の違いを比較するためにも経験強化型の手法で実験を行う必要があると考えられる。そこで、我々は経験強化型の学習法である Profit Sharing を用いて自動要約を行う。

Profit Sharing は学習速度が速いため、TD() より高速に要約を作成することが期待できる。反面、TD() のように最適性が保証されている訳ではないので、作成された要約の質が落ちる恐れもある。

5 現時点での課題

強化学習を用いた自動要約には現状多くの課題が残っている。まずは前節で挙げた TD() とは異なるタイプの強化学習法での実験が挙げられる。先行研究においては、文圧縮パラメータが連続値を取る可能性がある場合が考えられ、連続値行動を取れる枠組みの適用を考える必要があることと、評価関数の表現方法の妥当性の検討が課題であると指摘されている [3]。

6 今後の予定

現在は強化学習を用いた自動要約に関するサーベイと自動要約のプログラミングを行っている段階である。自動要約のプログラムが完成次第 Profit Sharing のプログラムを自動要約に落とし込む予定である。追実験を行うので、実験方法や評価方法は先行研究 [3] を踏襲する予定である。

参考文献

- [1] 言語処理学会編:“デジタル言語処理学事典”, 共立出版 (2010)
- [2] 高村大也, 奥村学: “最大被覆問題とその変種による文書要約モデル”, 人工知能学会論文誌, Vol. 23, No.6, pp.505-513 (2008)
- [3] 梁成基, 阿辺川 武: “強化学習によるテキスト自動要約手法の提案”, 言語処理学会 言語処理学会第 18 回年次大会発表論文集, pp.1067-1070 (2012)
- [4] 木村 元, 宮崎 和光, 小林 重信: “強化学習システムの設計指針”, 計測と制御, Vol.38, No.10, pp.618-623 (1999)
- [5] 富田紘平, 高村大也, 奥村学: “重要文抽出と文圧縮を組み合わせた新たな抽出的手法”, 情報処理学会研究報告, 自然言語処理研究会報告, 2009(2), pp.13-20 (2009)
- [6] 宮崎和光, 木村元, 小林重信: “Profit Sharing に基づく強化学習の理論と応用”, 人工知能学会誌, Vol.14, No.5, pp.800-807 (1999)