

文書要約：数理的手法の発展

人工知能学会合同研究会2016

高村大也 東工大
<http://www.lr.pi.titech.ac.jp/~takamura/>
 takamura@pi.titech.ac.jp
 Nov. 10, 2016

文書要約とは

与えられた文書に記述された情報を簡潔にまとめた短い文書を、自動的に生成すること

反政府勢力と治安部隊の衝突が続くA国の国会は、Y大統領の罷免を決定した。EUとの関係強化を訴える野党側は国会を掌握し、Y大統領の罷免を決定した。これに対しY大統領側は、議会の決定は無効であると主張している。しかし、Y大統領の所在は掴めず、また与党を離脱する議員が相次ぎ、大津A国で大統領が解任されて政権崩壊状態である。つまり、事実上の政権崩壊状態に陥った

通常、150文字以内などの長さ制限が与えられる

A国で大統領が解任されて政権崩壊状態である。

2

このトークで必要となる知識

最適化問題とは？

目的関数

決定変数

max. $f(x_1, x_2, \dots, x_m)$

s.t. $g_1(x_1, x_2, \dots, x_m) \leq 0$
 $g_2(x_1, x_2, \dots, x_m) \leq 0$
 $\vdots$
 $g_p(x_1, x_2, \dots, x_m) \leq 0$

制約

例

max. $x_1^2 + x_2^2$

s.t. $x_1 + x_2 \leq 1$,
 $x_1 \geq 0$,
 $x_2 \geq 0$.

3

Integer Linear Programming (ILP) problems

整数計画問題とは？

なぜか日本語では「線形(linear)」を入れないでよぶことが多い

目的関数

決定変数

max. $\sum b_i x_i$

s.t. $\sum c_i x_i \leq C$
 $\forall i, x_i \in \{0, 1\}$

制約

整数制約

- 最適化問題の一種
 - 決定変数に整数制約
 - 目的関数と制約式は線形
- 整数計画問題は解法が盛んに研究されている

つまり、整数計画問題で表せると嬉しい

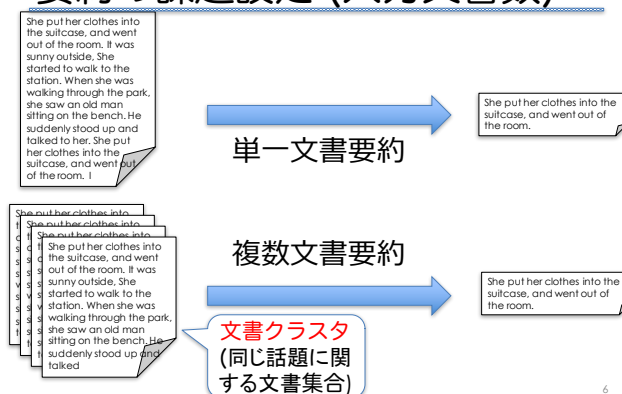
4

Table of Contents

- 文書要約の基礎
 - 要約の種類
 - 単一文書要約から複数文書要約へ
- 複数文書要約手法の発展
 - 手続きの記述から最適化問題へ
- より柔軟な要約生成へ向けて
 - 文圧縮、非抽出的要約
- ニューラルネットワークによる要約生成

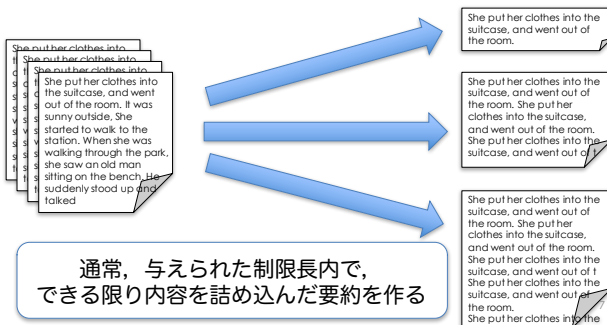
5

要約の課題設定 (入力文書数)



6

要約の課題設定 (要約長)

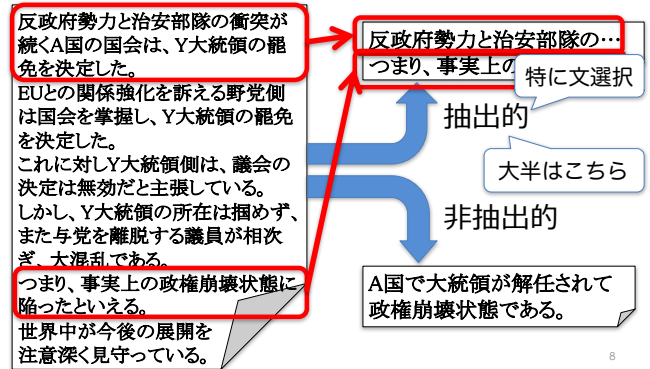


抽出的要約手法

- 元文書(クラスタ)の一部を抽出して、適切に並べることで、要約を作る方法
- 要約生成は、どれを選ぶか、という問題に帰着
- 多くは文抽出を利用
⇒ 文レベルの文法性は保証される

※ 今日は、抽出的要約の話から始めて、後で非抽出的要約の話をしてします

抽出的及び非抽出的要約手法

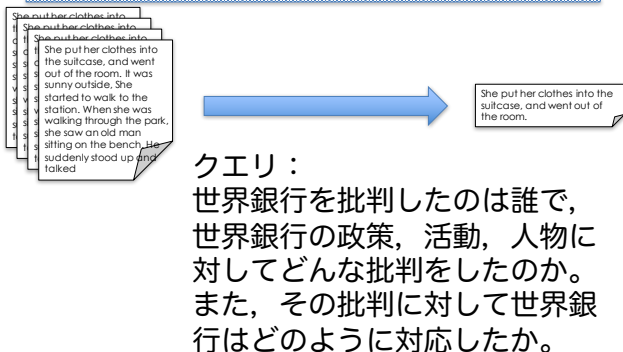


特殊な要約課題

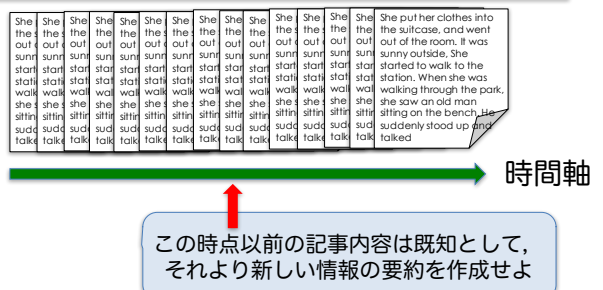
- クエリ指向型要約
- Update summarization
- Guided summarization

これに対して、通常の文書要約課題は **generic summarization** とよばれる

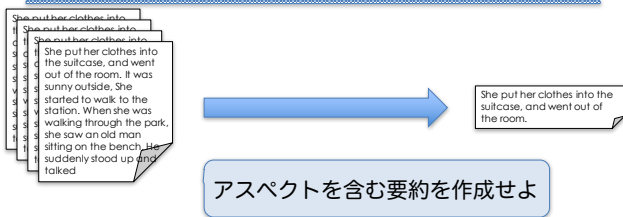
クエリ指向型要約



Update Summarization



Guided Summarization



アスペクト(「自然災害」の場合) :
災害内容, 発生場所, 発生時刻, 発理由,
影響を受ける人, 死傷者数, . . .

13

Table of Contents

- 文書要約の基礎
 - 要約の種類
 - 単一文書要約から複数文書要約へ
- 複数文書要約手法の発展
 - 手続きの記述から最適化問題へ
- より柔軟な要約生成へ向けて
 - 文圧縮, 非抽出的要約
- ニューラルネットワークによる要約生成

14

単一文書要約のためのベースライン手法

リード法

反政府勢力と治安部隊の衝突が続くA国の国会は、Y大統領の罷免を決定した。
EUとの関係強化を訴える野党側は国会を掌握し、同大統領の罷免を決定した。
これに対しY大統領側は、議会の決定は無効だと主張している。しかし、Y大統領の所在は掴めず、また与党を離脱する議員が相次ぎ、大混乱である。つまり、事実上の政権崩壊状態に陥ったといえる。世界中が今後の展開を注意深く見守っている。

文の出現位置

新聞記事では、文書の最初に簡単なまとめ(リード文)がしばしば書かれる

文書の最初の数文を選択する手法をリード法という

15

その他の手がかり

反政府勢力と治安部隊の衝突が続くA国の国会は、Y大統領の罷免を決定した。
EUとの関係強化を訴える野党側は国会を掌握し、同大統領の罷免を決定した。
これに対しY大統領側は、議会の決定は無効だと主張している。しかし、Y大統領の所在は掴めず、また与党を離脱する議員が相次ぎ、大混乱である。つまり、事実上の政権崩壊状態に陥ったといえる。世界中が今後の展開を注意深く見守っている。

文の出現位置

単語の重要度
(元文書での単語頻度)

手がかり語句
(e.g., まとめて)

タイトルとの類似度

$$(\text{文のスコア}) = \sum_i (\text{手がかり}i\text{の重み}) \times (\text{手がかり}i\text{の数値})$$

16

文の二値分類による単一文書要約

ナイーブベイズ分類器

[Kupiec+ 1995,
Aone+ 1999]

決定木

[Lin+ 1999]

対数線形モデル

[Osborne+ 2002]

SVM

[Hirao+ 2002]

(人手)

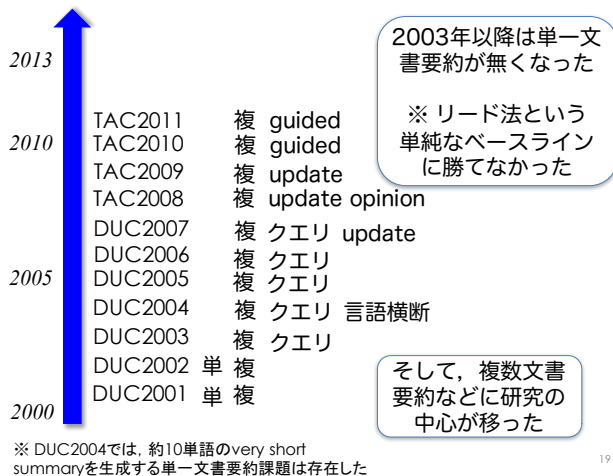
[Edmundson 1969]

17

DUC/TAC

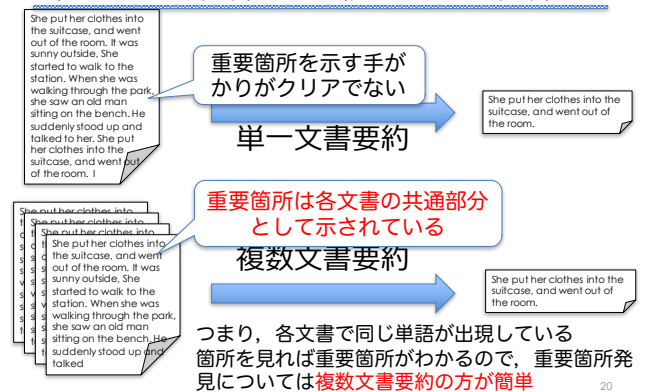
- DUC (Document Understanding Conference) :
要約を共通課題とした評価型国際会議。
2001~2007年
<http://duc.nist.gov/>
- TAC (Text Analysis Conference) :
DUCの後継となる評価型国際会議。
要約以外に質問応答, 含意関係認識などの研究課題も扱う。2008年~
<http://www.nist.gov/tac/>

18



19

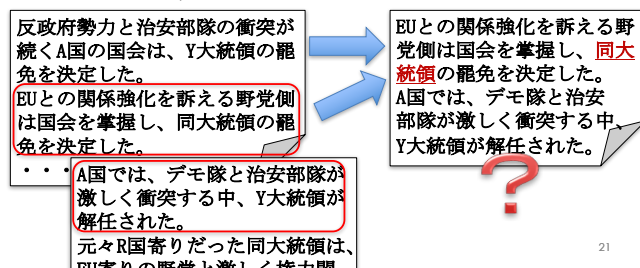
単一文書要約 vs 複数文書要約



20

複数文書要約の難しさ (選択文の整列)

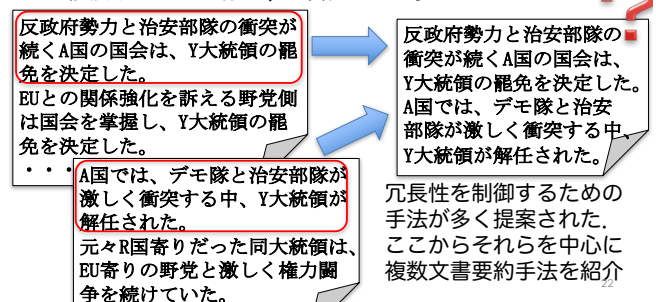
- 選択した文の整列が必要：単一文書要約では、元文書中での出現順に並べるのが自然だが、複数文書要約では使えない



21

複数文書要約の難しさ (冗長性の制御)

- 冗長性の制御が必要：複数文書の場合、類似文が多い



22

Table of Contents

- 文書要約の基礎
 - 要約の種類
 - 単一文書要約から複数文書要約へ
- 複数文書要約手法の発展
 - 手続きの記述から最適化問題へ
- より柔軟な要約生成へ向けて
 - 文圧縮、非抽出的要約
- ニューラルネットワークによる要約生成

23

MMR (Maximal Marginal Relevance)

スコアの高い方から1文ずつ選択。
ただし、一つ選択する度に、スコアを更新

(文のスコア)=(元のスコア)-(現時点での要約との類似度)

似た文が選ばれにくくなる

反政府勢力と治安部隊の衝突…	1.3	---	---	---
EUとの関係強化を訴える野党…	1.0	0.3	0.3	0.3
これに対しY大統領側は、…	0.3	0.3	0.3	0.3
しかし、Y大統領の所在は掴…	0.5	0.5	0.4	0.4
つまり、事実上の政権崩壊状…	0.7	0.7	---	---
世界中が今後の展開を注意深…	0.2	0.2	0.2	0.2
A国では、デモ隊と治安部隊…	1.2	0.5	0.4	0.3
元々R国寄りだった同大統領…	0.6	0.6	0.6	---

24

手続きの記述から最適化問題へ

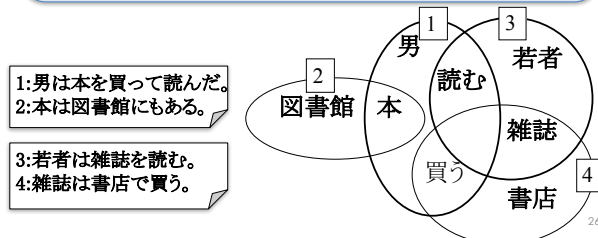
- 初期の複数文書要約手法は、処理の手続きだけが記述されたものであった
- しかし、手続きだけでは、どういう出力に辿り着こうとしているかが不明である
- 要約の良さを目的関数として表現し、要約生成を最適化問題(特に整数計画問題)で表現する試みが2007年位から出てきた
- 多くの変数の値(文を選ぶか否か)を、全体として良くなるように決定する

25

最適化問題による定式化の例

(重み無し)最大被覆モデル

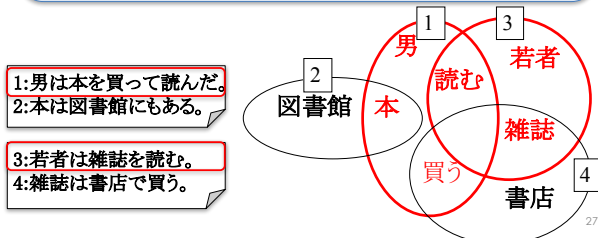
単語はなんらかの概念を近似的に表わす。
文書クラスに記述されている概念をできる限り被覆するために、制限長内でなるべく多くの種類の単語を含むような文集合を選択



最適化問題による定式化の例

(重み無し)最大被覆モデル

単語はなんらかの概念を近似的に表わす。
文書クラスに記述されている概念をできる限り被覆するために、制限長内でなるべく多くの種類の単語を含むような文集合を選択



27

表記

今は文選択アプローチなので、単語 j を含む文を選択することで単語 j を要約に含めることができる

- 決定変数
 - x_i : 文 i が要約に含まれれば1,それ以外は0
 - z_j : 単語 j が要約に含まれれば1,それ以外は0
- パラメータ(定数)
 - K : 最大要約長
 - c_i : 文 i の長さ
 - a_{ij} : 文 i が単語 j を含んでいれば1,それ以外は0

28

(重み無し)最大被覆モデル

なるべく多くの種類の単語を含むような文集合を選択

x_i : 文 i が選択されたか
 z_j : 単語 j が被覆されたか
 K : 最大要約長
 c_i : 文 i の長さ
 a_{ij} : 文 i が単語 j を含むか

$$\begin{aligned} \max. \quad & \sum_j z_j \\ \text{s.t.} \quad & \sum_i c_i x_i \leq K, \\ & \forall j, \sum_i a_{ij} x_i \geq z_j, \\ & \forall i, x_i \in \{0,1\}, \\ & \forall j, z_j \in \{0,1\}. \end{aligned}$$

被覆された単語の種類数

要約長制約

変数 x_i と z_j の整合性:
もし単語 j を含む文が一つでも選択されていたら左辺は1以上。その場合、目的関数をmaxにするので、必ず $z_j=1$ に

29

(重み無し)最大被覆モデル

なるべく多くの種類の単語を含むような文集合を選択

x_i : 文 i が選択されたか
 z_j : 単語 j が被覆されたか
 K : 最大要約長
 c_i : 文 i の長さ
 a_{ij} : 文 i が単語 j を含むか

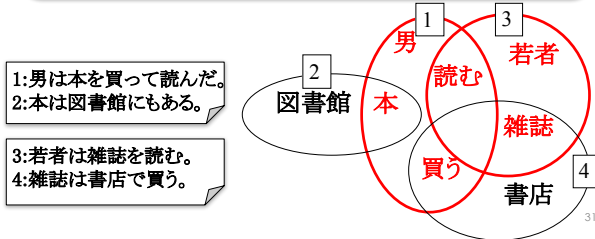
$$\begin{aligned} \max. \quad & \sum_j z_j \\ \text{s.t.} \quad & \sum_i c_i x_i \leq K, \\ & \forall j, \sum_i a_{ij} x_i \geq z_j, \\ & \forall i, x_i \in \{0,1\}, \\ & \forall j, z_j \in \{0,1\}. \end{aligned}$$

$x_1=1, x_2=0, x_3=1, x_4=0, \sum_j z_j=6.$

30

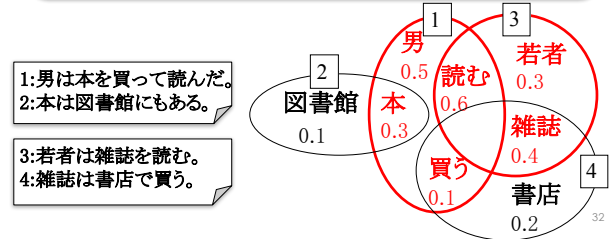
最大被覆モデル [Gillick+ 2008, Takamura+ 2009]

被覆すべき単語と、そうでない単語がある。
各単語(概念)に重要度が与えられているとして、
制限長内で、被覆した単語の重要度の和が最大
になるように文集合を選択



最大被覆モデル [Gillick+ 2008, Takamura+ 2009]

被覆すべき単語と、そうでない単語がある。
各単語(概念)に重要度が与えられているとして、
制限長内で、被覆した単語の重要度の和が最大
になるように文集合を選択



最大被覆モデル

被覆された単語の重要度の和が
最大になるように文集合を選択

$$\begin{aligned} \max. \quad & \sum_j b_j z_j \\ \text{s.t.} \quad & \sum_i c_i x_i \leq K, \\ & \forall j, \sum_i a_{ij} x_i \geq z_j, \\ & \forall i, x_i \in \{0,1\}, \\ & \forall j, z_j \in \{0,1\}. \end{aligned}$$

被覆された単語の重要度の和

x_i : 文*i*が選択されたか
 z_j : 単語*j*が被覆されたか
 K : 最大要約長
 c_i : 文*i*の長さ
 a_{ij} : 文*i*が単語*j*を含むか
 b_j : 単語*j*の重要度

最大被覆モデルと冗長性

- 最大被覆モデルでは、冗長性は直接的に表現されていない
- しかし、制限長内でできる限り多くの内容を被覆しようとするれば、自然に冗長性を減らすことになる

最大被覆モデルのパラメータ

- 単語*j*の重要度 b_j
 - 文書クラスタにおける単語*j*の頻度を用いる
 - 訓練データを用いて、各単語が正解要約に出現するか否かを予測する分類器を学習し、その出力スコアを用いる
- 詳細はYihら(2007)の論文を参照

※ Yihら(2007)の手法は、最大被覆モデルをスタックデコーディングで解いたものと、おおよそ位置づけられる。ただし、最後の一文だけ分割して加えているので、純粋な最大被覆モデルではない。

最大被覆モデルのデコーディング

デコーディング:
実際に最大値を与える変数値を見つけること

$$\begin{aligned} \max. \quad & \sum_j b_j z_j \\ \text{s.t.} \quad & \sum_i c_i x_i \leq K, \\ & \forall j, \sum_i a_{ij} x_i \geq z_j, \\ & \forall i, x_i \in \{0,1\}, \\ & \forall j, z_j \in \{0,1\}. \end{aligned}$$

- 分枝限定法 (CPLEXなどのツールを用いる) により厳密解を得る
- 貪欲法などの近似解法を用いて近似解を得る [Khuller+ 1999]

その他の要約モデル

- 各々は主題への高い関連性を持つが、互いに似ていない文から成る文集合を選択 [McDonald+ 2007]
- 施設配置問題 [Takamura+ 2009]
- 冗長性制約付きナップサック問題 [Nishikawa+ 2012]

37

複数文書要約手法における知見

何ができるようになったのか？

- 重要な箇所が少しわかるようになった
- 冗長性は最適化問題でモデル化できた
- パラメータ推定、デコーディングもできるようになった
- 非抽出的な方法はほとんどできていない
- 意味理解ができているわけではない

38

Table of Contents

- 文書要約の基礎
 - 要約の種類
 - 単一文書要約から複数文書要約へ
- 複数文書要約手法の発展
 - 手続きの記述から最適化問題へ
- より柔軟な要約生成へ向けて
 - 文圧縮, 非抽出的要約
- ニューラルネットワークによる要約生成

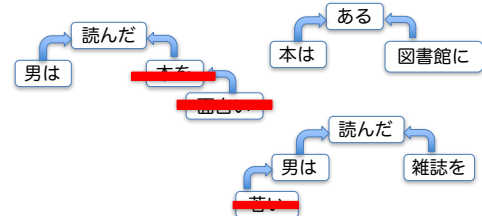
39

文圧縮を組み込んだ要約手法

単語選択による
抽出的な

[Martins+ 2008, Berg-Kirkpatrick+ 2011, Qian+ 2013, Almeida+ 2013, Morita+ 2013]

文法的な文を生成するために構文木を枝刈りすることが多い

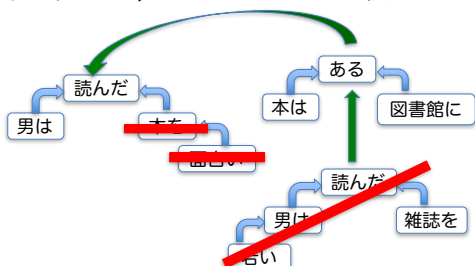


40

談話構造を利用した要約手法

[菊池+ 2014]

文間の依存関係と、単語間の依存関係を同時に利用し、nested treeの枝刈りをする



41

その他の話題

- 劣モジュラ関数の利用 [Lin+ 2010, 2011]
- 選択した文を適切に並べる方法 [Lapata 2005, Bollegala+ 2010, Nishikawa+ 2010, Christensen+ 2013]
- 文間の関係を利用したモデル [Takamura+ 2009 (CIKM)]
- ラグランジュ緩和によるデコーディング [西川+ 2012, Yasuda+ 2012, Nishino+ 2013, Qian+ 2013, Almeida+ 2013]
- 談話構造の利用 [Hirao+ 2013, 菊池+ 2014]
- 評判要約 [Lerman+ 2009, Nishikawa+ 2010]
- 調停要約 [中野+ 2011, 渋谷+ 2013]
- 学術論文の要約, 引用要約 [Nanba+ 1999, Teufel+ 2002, Qazvinian+ 2008, Qazvinian+ 2010]

42

現在の要約手法の問題点

$$b_j = \mathbf{w} \cdot \mathbf{v}_j$$

$$b_j \propto \exp(\mathbf{w} \cdot \mathbf{v}_j)$$

- 学習に大量の素性を投入できていない
- 構造学習などの大域的学習手法が十分に力を発揮できていない

訓練データが不足していることが一因

機械翻訳のための対訳コーパスならば、古典的なHansardsコーパスでも約130万文ペア。一方、例えばDUC2004は約4,000文の要約

とりあえず、訓練データを作ろう (2014年3月時点) 43

Table of Contents

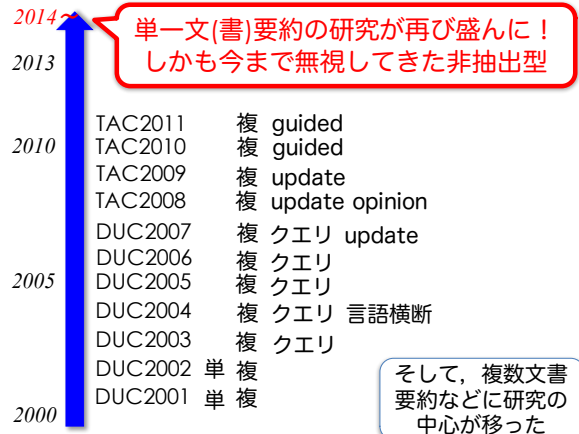
- 文書要約の基礎
 - 要約の種類
 - 単一文書要約から複数文書要約へ
- 複数文書要約手法の発展
 - 手続きの記述から最適化問題へ
- より柔軟な要約生成へ向けて
 - 文圧縮、非抽出的要約
- ニューラルネットワークによる要約生成

Neural Network の隆盛

(Deep Learning, Representation Learning)

- 感情極性分類 (ポジティブ/ネガティブ)
- 単語ベクトル表現 (word2vecなど)
- 機械翻訳
- 対話
- **要約**
 - ⇒ 単一文書要約の研究が盛んになってきた

45

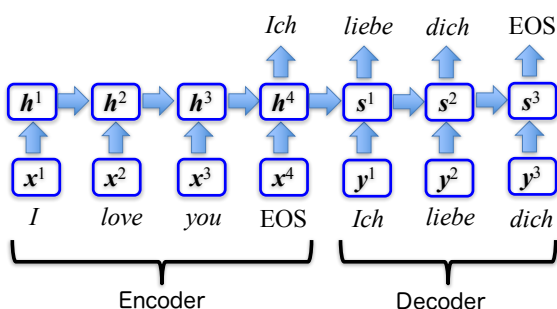


※ DUC2004では、約10単語のvery short summaryを生成する単一文書要約課題は存在した

46

[Sutskever+ 2014, Bahdanau+ 2015]

Encoder-decoderモデルによる機械翻訳



※ 隠れ層を多層にする、Encoderを双方向にする、注視機構を加える、隠れ層のノードにLSTMを用いる、などの工夫が加えられる 47

Encoder-decoderモデルによる文要約

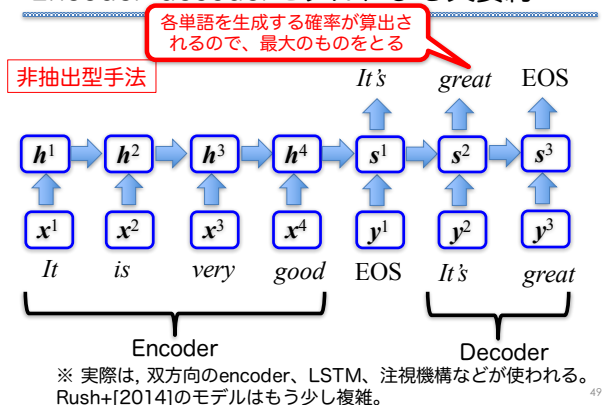
- データがあれば、同じモデルを要約に適用可能
 - ⇒ 新聞記事のヘッドラインと先頭文を用いて訓練データを作成 [Rush+ 2015]
 - (ノイズ除去の結果、約400万訓練事例)

I(I): a detained iranian-american academic accused of acting against national security has been released from a tehran prison after a hefty bail was posted , a to p judiciary official said tuesday .
G: iranian-american academic held in tehran released on bail

先頭文(上)とヘッドラインの例 (文献[Rush+ 2015]より)

48

Encoder-decoderモデルによる文要約



Encoder-decoderモデルによる出力例

I(1): brazilian defender pepe is out for the rest of the season with a knee injury , his porto coach jesualdo ferreira said saturday .
G: football : pepe out for season
A+: ferreira out for rest of season with knee injury
R: brazilian defender pepe out for rest of season with knee injury

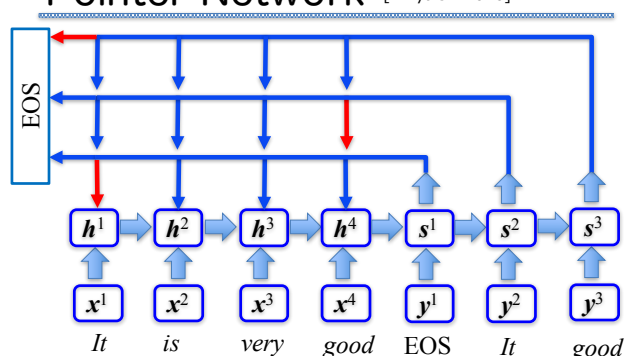
I(2): economic growth in toronto will suffer this year because of sars , a think tank said friday as health authorities insisted the illness was under control in canada 's largest city .
G: sars toll on toronto economy estimated at c\$ # billion
A+: think tank under control in canada 's largest city
R: think tank says economic growth in toronto will suffer this year

I(3): colin l. powell said nothing – a silence that spoke volumes to many in the white house on thursday morning .
G: in meeting with former officials bush defends iraq policy
A+: colin powell speaks volumes about silence in white house
R: powell speaks volumes on the white house

50

抽出型手法

Pointer Network [Vinyals+ 2016]

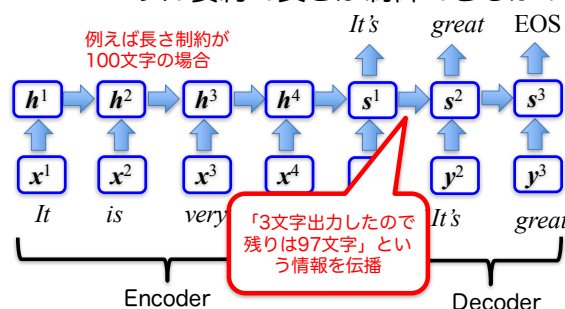


Gulcehreら[2016]やGuら[2016]のモデルは、各単語の生成に、pointingと通常のRNNのいずれを使うかをアダプティブに決定する非抽出型の手法⁵¹

[Kikuchi+ 2016]

Encoder-decoderモデルにおける出力長制御

- 要約には通常長さ制約がある。
ニューラル要約で長さが制御できるか？



52

出力長制御の例

source	five-time world champion michelle kwan withdrew from the ##### us figure skating championships on wednesday , but will petition us skating officials for the chance to compete at the ##### turin olympics .
reference	injury leaves kwan 's olympic hopes in limbo
<i>fixLen</i> (30)	kwan withdraws from us gp
(50)	kwan withdraws from us skating championships
(75)	kwan pulls out of us figure skating championships for turin olympics
<i>fixRng</i> (30)	kwan withdraws from us gp
(50)	kwan withdraws from figure skating championships
(75)	kwan pulls out of us figure skating championships for turin olympics bid
<i>LenEmb</i> (30)	kwan withdraws from us skating
(50)	kwan withdraws from us figure skating championships
(75)	world champion kwan withdraws from ##### olympic figure skating championships
<i>LenInit</i> (30)	kwan quits us figure skating
(50)	kwan withdraws from ##### us figure skating worlds
(75)	kwan withdraws from ##### us figure skating championships for ##### olympics

Table 2: Examples of the output of each method with various specified lengths.

53

ニューラル文要約からニューラル文書要約へ

- ニューラルネットによる文要約は発展しつつあるが、ニューラル文書要約はまだこれから
- ニューラル文要約モデル + distraction機構が単一文書要約に有効との報告[Chen+, 2016] (distraction機構：入力の一部だけを利用するのでなく、なるべくいろいろな箇所を考慮して出力する仕組み)
- ニューラルネットワークの複数文書要約への有効な利用方法はまだ知られていない

54

おわりに

- 重要箇所同定は、複数文書要約の方が簡単である
- 最適化問題として要約モデルを表そう
- ニューラル文要約が流行っている
- ニューラル文書要約、特に複数文書要約はまだこれから

55

参考文献

書籍, サーベイ, チュートリアル資料など

- 奥村学, 難波英嗣「テキスト自動要約」オーム社, 2005.
- Ani Nenkova and Kathleen McKeown, "Automatic Summarization", *Foundations and Trends in Information Retrieval*, 5(2-3), 103-233, 2011.
- 西川仁, "自動要約技術の研究動向: これまでとこれから", *SIGNL*-213, 2013.

論文

- Aone+, "A scalable summarization system using robust NLP", *Workshop on intelligent scalable text summarization*, 1997.
- Almeida+, "Fast and robust compressive summarization with dual decomposition and multi-task learning", *ACL*, 2013.
- Bahdanau+, "Neural machine translation by jointly learning to align and translate", *ICLR*, 2015.
- Berg-Kirkpatrick+, "Jointly Learning to Extract and Compress", *ACL*, 2011.
- Bollegala+, "A bottom-up approach to sentence ordering for multi-document summarization", *Information processing and management*, vol. 46, 2010.
- Carbonell+, "The use of MMR, diversity-based rerunning for reordering documents and producing summaries", *SIGIR*, 1998.
- Chen+, "Distraction-based neural networks for modeling documents", *IJCAI*, 2016.
- Cheung+, "Towards robust abstractive multi-document summarization: a caseframe analysis of centrality and domain", *ACL*, 2013.

56

- Chopra+, "Abstractive sentence summarization with attentive recurrent neural networks", *NAACL*, 2016.
- Christensen+, "Towards coherent multi-document summarization", *NAACL*, 2013.
- Edmundson, "New methods in automatic extracting", *J. of the ACM*, 1969.
- Erkan+, "LexRank: graph-based centrality as salience in text summarization", *JAIR*, 2004.
- Gillick+, "The ICSI Summarization System at TAC 2008", *TAC*, 2008.
- Gulcehre+, "Pointing the unknown words", arXiv:1603.08148, 2016.
- Gu+, "Incorporating copying mechanism in sequence-to-sequence learning", *ACL*, 2016.
- Hirao+, "Extracting important sentences with support vector machines", *COLING*, 2002.
- Hirao+, "Single-document summarization as a tree-knapsack problem", *EMNLP*, 2013.
- Khuller+, "The budgeted maximum coverage problem", *Information processing letters*, vol. 70, no. 1, 1999.
- 菊池+, "修辞構造と係り受け構造を制約とした単一文書要約手法", 言語処理学会年次大会, 2014.
- Kikuchi+, "Controlling output length in neural encoder-decoders", *EMNLP*, 2016.

57

- Krause+, "A note on the budgeted maximization of submodular functions", Technical Report, CMU-CALD-05, 2005.
- Kupiec+, "A trainable document summarizer", *SIGIR*, 1995.
- Lapata, "Automatic evaluation of text coherence: models and representations", *IJCAI*, 2005.
- Lerman+, "Contrastive summarization: an experiment with consumer reviews", *NAACL*, 2009.
- Li+, "Enhancing Diversity, Coverage and Balance for Summarization through Structure Learning", *WWW*, 2009.
- Lin+, "Training a selection function for extraction", *CIKM*, 1999.
- Lin+, "Multi-document summarization via budgeted summarization of submodular functions", *NAACL*, 2010.
- Lin+, "A class of submodular functions for document summarization", *NAACL*, 2011.
- 牧野+, "アスペクトの被覆を実現するための最小値最大化問題に基づく文書要約モデル", 言語処理学会年次大会, 2012.
- Martins+, "Summarization with a joint model for sentence extraction and compression", *Workshop on ILP for NLP*, 2009.
- McDonald, "A study of global inference algorithms in multi-document summarization", *ECIR*, 2007.
- Mihalcea+, "TextRank: Bringing order into texts", *EMNLP*, 2004.

58

- Mihalcea+, "A language independent algorithm for single and multiple document summarization", *IJCNLP*, 2005.
- Morita+, "Subtree extractive summarization via submodular maximization", *ACL*, 2013.
- Nallapati+, "Abstractive text summarization using sequence-to-sequence RNNs and beyond", *CoNLL*, 2016.
- 中野+, "情報信憑性判断支援のための直接調停要約生成手法", 電気情報通信学会論文誌D, 2011.
- Nanba+, "Towards multi-paper summarization using reference information", *IJCAI*, 1999.
- Nenkova, "Automatic text summarization of newswire: lessons learned from the Document Understanding Conference", *AAAI*, 2005.
- Nishikawa+, "Opinion summarization with integer linear programming formulation for sentence extraction and ordering", *COLING*, 2010.
- Nishikawa+, "Text summarization model based on redundancy-constrained knapsack problem", *COLING*, 2012.
- 西川+, "ラグランジュ緩和による複数文書要約の高速求解", 言語処理学会年次大会, 2012.
- Nishino+, "Text summarization while maximizing multiple objectives with Lagrangian relaxation", *ECIR*, 2013.
- Osborne, "Using maximum entropy for sentence extraction", *Workshop on automatic summarization*, 2002.

59

- Qazvinian+, "Citation summarization through keyphrase extraction", *COLING*, 2010.
- Qazvinian+, "Scientific paper summarization using citation summary networks", *COLING*, 2008.
- Qian+, "Fast joint compression and summarization via graph cuts", *EMNLP*, 2013.
- 洪木+, "情報信憑性判断支援のための対話型調停要約生成手法", 自然言語処理, 2013.
- Sipos+, "Large-margin learning of submodular summarization models", *EACL*, 2012.
- Sutskever+, "Sequence to sequence learning with neural networks", *NIPS*, 2014.
- Takamura+, "Text summarization model based on maximum coverage problem and its variant", *EACL*, 2009.
- Takamura+, "Text summarization model based on budgeted median problem", *CIKM*, 2009.
- Takamura+, "Learning to generate summary as structured output", *CIKM*, 2010.
- Teufel, "Summarizing scientific articles: experiments with relevance and rhetorical status", *Computational Linguistics*, 2002.
- Vinyals+, "Pointer networks", *NIPS*, 2015.
- Yasuda+, "A query-focused summarization method that guarantees the inclusion of query words", *DEXA*, 2012.
- Yih+, "Multi-document summarization by maximizing informative content-words", *IJCAI*, 2007.