

August 7, 2019

バリューエンジニアリングにおける アイデア発想支援システム

1. はじめに
2. 分野
3. 手順
4. フローチャート

山元 悠貴

富山県立大学 情報基盤工学講座

1. はじめに
2. 分野
3. 手順
4. フローチャート
5. おわりに

1. はじめに

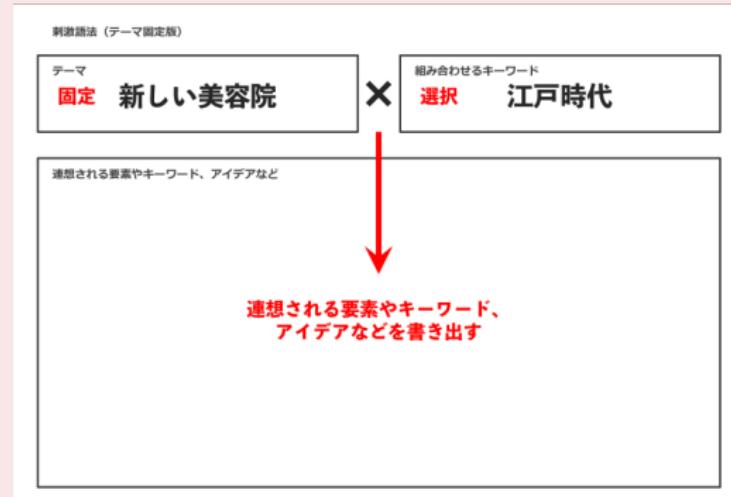
2/11

目的

1. はじめに
2. 分野
3. 手順
4. フローチャート

複数の文章からキーワードを抽出し、そのキーワードの中でアイデア発想につながる二つのキーワードの組み合わせを出力する。

イメージ



2. 分野

自然言語処理

自然言語処理とは私たちが使う言語をコンピュータコンピュータに処理させる技術のこと。

- はじめに
- 分野
- 手順
- フローチャート

- コーパス
- 辞書
- 形態素解析
- 構造解析



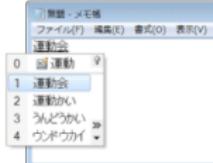
2. 分野

4/11

身近な例

- はじめに
- 分野
- 手順
- フローチャート

■ 文字変換



■ 機械翻訳



■ 検索エンジン



■ 対話システム



3. 手順

5/11

ステップ 1

クローリング、スクレイピングを行ってウェブ上にある文章（ウェブページ）を集める。あるキーワードからその内容に関するページを集めてくる。

クローリングとはウェブサイトから HTML や任意の情報を取得する技術・行為のことで、スクレイピングとは取得した HTML から任意の情報を抽出する技術・行為のこと。

1. はじめに
2. 分野
3. 手順
4. フローチャート

3. 手順

ステップ1の続き

- はじめに
- 分野
- 手順
- フローチャート

ステップ1で集めるページは、有用なページであるほうがいい。
そこで、リンクを受けている数が多いもの優先的に集めるように
する。

$$p_i = \sum_{j=i \text{へのリンクがある番号}} \frac{p_j}{\text{ページ } j \text{ が出しているリンク数}}$$

3. 手順

ステップ 2

ステップ 1 で集めた文章を形態素解析し、重要度が高いと想定される単語（キーワード）を見つける。

形態素解析とは文を単語に分解する手法のこと。
単語の重要度を調べるために tf-idf 法を使う。

$$tf(t, d) = \frac{\text{文書 } d \text{ における単語 } t \text{ の出現回数}}{\text{文書 } d \text{ における総単語数}}$$

$$idf(t) = 1 + \log \left(\frac{\text{全文書数}}{\text{単語 } t \text{ が出現する文書数}} \right)$$

$$tf-idf(t, d) = tf(t, d) \times idf(t)$$

これは簡単に説明すると、他の文章に出てきてない単語が多くなるほど重要度が高くなるということ。

3. 手順

8/11

ステップ 3

1. はじめに
2. 分野
3. 手順
4. フローチャート

ステップ 2 で集めたキーワードに意味を持たせるためキーワードをベクトル化する。

word2vec という手法でキーワードの意味をベクトル表現化することができる。ベクトル表現した二つのキーワードの選び方について検討中。

word2vec を使うと単語間の関係性が数値で表される。

3. 手順

9/11

ステップ4

ステップ3からキーワードの組み合わせからアイデアを発想する。
出てきたキーワードか発想したアイデアについて評価をする。
評価方法については検討する必要がある。

1. はじめに
2. 分野
3. 手順
4. フローチャート

4. フローチャート

10/11

フローチャート

① これまでの手順をフローチャートにすると

- はじめに
- 分野
- 手順
- フローチャート



5. おわりに

11/11

まとめ

1. はじめに
2. 分野
3. 手順
4. フローチャート

今後の課題

- ① 実際にシステムを構築する.
- ② キーワードの組み合わせからシステムがアイデアを提案する.
- ③ システムの評価方法についての考察をする.