

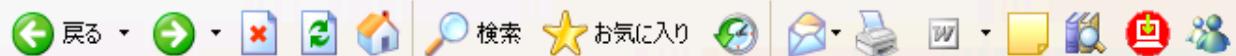
中間発表

06T4007A 伊藤輝将

研究テーマ

“グラフクラスタリングによる
単語用例クラスタリング”

- ・分類したい単語(キーワード)を入力すると、用例文を語義ごとに分類し、結果を各クラスタごとに表示するシステムの改良



検索クエリを入力して検索

keywords: 走る

clusters:

検索に活用形も含める

- 概念を考慮しない
- 概念を考慮する

- MCL
- Graclus

- 重複許可
- 重複禁止

検索 リセット

走る - Search Results - Microsoft Internet Explorer

ファイル(E) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H)

戻る 検索 お気に入り

アドレス(D) http://localhost/search.cgi

Google 検索

グループ数: 86 走る の検索結果 1 - 151 件 (92秒)

用例検索

グループ 0: 8件 ▾

1. 北へおよそ四〇キロ走ると湿地帯になり、呼蘭川に至る。
2. 廊下にでると、他の隊員たちの走る後ろ姿が見える。
3. 細く深い、溝のような道を青い自転車で走る少年がはっきり見えた。
4. リムジンバスは、目に美しい、広大な緑の田園地帯を走る。
5. 敵兵たちは次に、サフィラの横を走るエラゴンを目にとめる。
6. 「大丈夫だよ、ナギ」彼女は目を潤ませて意味不明なことを口走る。
7. 『黙れ、黙れ、黙れーッ』孫はチャンの後ろ姿から涙を浮かべた目を背け、頭の中に走る出来事を消そうとしていた。
8. パラディーノの丘の北側に広がる低地は、それまでは溝が網の目のように走る湿地帯だった。

△ グループ 1: 7件 ▾

9. 車で走るほどの広い町ではないわけだ。
10. 「これまでだ。もう限界だ」私は走る車の上でそう思った。
11. 細く深い、溝のような道を青い自転車で走る少年がはっきり見えた。
12. 『車で北海道まで行くんだから東北道を走るのには間違いはない。
13. 24「どこへいくの？」マディソン・ストリートをフリーウェイに向かって走る車の中で、彼女は訊く。
14. 7みんなの家で、一番さむい所って、どこかな？8朝、早くから、細い道を、走るもの、これ、なーに。
15. 登山口であるヤクスギランドまで、舗装された道を車で30分ほどくねくねと走る。

△ グループ 2: 5件 ▾

16. 深い森の中を、黒い影が走る。
17. 「豊くん、いる？」その声に、豊があわてて立ちあがり、戸口へ走る。
18. 「動くな！！」たちまち、海渡の声に緊張感が走る。
19. 湖水の中に魚の影がいくつも走るのが見えた。
20. 宿直室をぬけだし、声の方へむかったら、「クワン」ひと声のあとサッサーと走る影。

△ グループ 3: 5件 ▾

21. 麻子、心の中に電流が走る。
22. 逆トして、なにが恐ろしい自殺的な行為に走るか判じれないからだ。

http://localhost/search.cgi#anchor3

イントラネット

処理手順

1. キーワード入力
2. 用例文集合の抽出
3. 形態素解析しキーワード以外の名詞を抽出
4. 共起単語その概念からなるグラフの作成
5. クラスタリング
6. 単語クラスタから用例文の導出
7. 分類結果の表示

今後

- ・プログラムの理解
- ・グラフ構造作成部の改良
UBC-ASシステム(Agirre, Soroa, 2007)

UBC-ASシステム

- 単語共起によるグラフの作成
 - ノードは単語、エッジには相対頻度
- 高密度なノード(ハブ)を求める
 - HyperLexアルゴリズム
- 最小スパニング木を求める
- 各文に対してスコアベクトルを計算
- Markov Clustering(MCL)でクラスタリング