進捗報告

1 学位申請書草稿

github にアップロードしましたので，ご確認お願い致します.

2 色に関する単語の分散表現の演算について

基本色 ['red', 'green', 'blue', 'yellow', 'orange', 'purple', 'white', 'black', 'pink', 'skyblue'] と形容詞 ["dark", "light", "deep", "medium", "royal"] の組み合わせのうち， matplotlib にカラーコードとして登録されている色名は，[darkred，darkgreen，lightgreen，darkblue，lightblue，mediumblue，royalblue，lightyellow，darkorange，mediumpurple，lightpink，deeppink，lightskyblue，deepskyblue] の 14 色.

表 1 に，組み合わせた造語に対して，分散表現が最も近いカラーコードの中に，もととなった基本色名が含まれていないものを不一致として，各手法ごとの不一致数をまとめた，

特に "royal" が含まれるのは 'royalblue' のみ．そのためか，'royal' が含まれる造語は青系の単語に勘違いされやすい傾向がある（'royalblue' や， 'navy' , 'cornflowerblue' など）.

結論としては，subwords に分割している fastText が基本色とのずれが少なく，共起行列も参照する GloVe のずれが大きいという，あまり予想に反しない結果になりました.

ただ，各手法ごとに無難なものというか,「困ったときはこれを出しとけ」というような単語が見られたのは興味深い気もします．が，興味深いだけで情報工学的に理論的な考察はだいぶ難しい気がしています.

手法	不一致数	特色
Word2Vec	35	'yellow' と 'blue' を間違いやすい． 灰色系の単語が候補としてあげられることが多い.
GloVe	43	'lightpurple' や 'lightwhite' が 'lightgreen' と間違えられる．青緑系に引っ張られている印象.
fastText	28	'lightwhite' や 'lightblack' が 'lightsalmon' と間違えられる． subwords に分割している影響か.

3 DAMSM の同期的学習あり・SWD なしの実験

比較実験として引き続き（DAMSM の同期的学習あり・SWD なし）の実験を動かしています.