We employ diachronic word embedding models in the task of predicting the events of armed conflicts escalating or calming down in various geographical locations, spanning over 16 years (1994–2010). The task is similar to that of detecting semantic shifts [Kulkarni et al., 2015; Hamilton et al., 2016a], but focused more on subtle changes of perspective instead of full-scale meaning changes.

We monitor changes in the local semantic neighborhoods of country names, applying it to the downstream task of predicting changes in the state of conflict:

1. Nothing has changed in the country conflict state year-to-year (class ‘stable’);
2. Armed conflicts have escalated in the country year-to-year (class ‘war’);
3. Armed conflicts have calmed down in the country year-to-year (class ‘peace’).

Example: given two distributional models trained on news texts from 2002 and 2003, predict in what direction the conflicts state moved in these years in Senegal (the correct answer is ‘it escalated’)?

We employ diachronic word embedding models in the task of predicting the events of armed conflicts, in the time period from 1946 to the present (Gleditsch et al., 2002). The resulting test set mentions 52 unique armed conflicts, in the country-year pairs.

As an example, for Congo, the transition from 2001 to 2002 was accompanied by the ending of armed conflicts. Thus, for the ‘Congo’ test point, \( \delta = 0 - 1 = -1 \). Then, there were no changes (each new \( \delta = 0 \) until 2006, when armed conflicts resumed with the intensity of 1. Thus, for the ‘Congo’ 2006 data point, \( \delta = 1 - 0 = 1 \). Then \( \delta \) values were transformed to classes:

\[
\text{class} = \begin{cases} 
\text{war} & \text{if } \delta \geq 0.5 \\
\text{peace} & \text{if } -0.5 \leq \delta < 0.5 \\
\text{stable} & \text{otherwise}
\end{cases}
\]

The approaches proposed in the previous work for large-scale shifts observed over decades or even centuries are not very successful in this more fine-grained task. Hamilton et al. (2016b) report almost perfect accuracy for the Procrustes transformation when detecting the direction of semantic change. However, our time periods are much more granular and we attempt to detect subtle associative shifts (e.g., predator-prey link) rather than full-scale shifts of the meaning.

The proposed ‘anchor words’ method outperforms previous work approaches by large margins, but it still achieves a macro F1 measure of only 0.26 on the task of ternary classification (stable, escalating, calming down).

See also our forthcoming EMNLP’17 paper on tracing semantic relations in diachronic models.

Methods

To actually detect semantic shifts for the word \( w_q \), one can either:

1. align two models (\( M_{cur} \) and \( M_{prev} \)) using the orthogonal Procrustes transformation, and then measure cosine similarity between the \( w_q \) vectors in both models, as proposed in [Hamilton et al., 2016b]
2. alternatively, define a set of anchor words related to the semantic categories we are interested in, and then measure the ‘shift’ of \( w_q \) towards or away from these ‘anchors’ in \( M_{cur} \) compared against \( M_{prev} \).

The anchor words method can provide information about the exact direction of the shift. It can be quantified in 2 ways:

1. for each anchor, calculate its cosine similarity against \( w_q \) in \( M_{cur} \) and \( M_{prev} \) (Sim).
2. as above, but use the position of each anchor in the models’ vocabulary sorted by similarity to \( w_q \) (Rank).

These methods produce two vectors \( R_{M_{cur}} \) and \( R_{M_{prev}} \) corresponding to the models \( M_{cur} \) and \( M_{prev} \), with the size is equal to the number of the anchor words. Then we can either:

1. calculate the cosine distance between these ‘second-order vectors’ (SimDist or RankDist).
2. element-wise subtract \( R_{M_{prev}} \) from \( R_{M_{cur}} \) to get the idea of whether \( w_q \) drifted towards or away from the anchors (SimSub or RankSub).

These features are then fed to SVM classifier.