A Comparison of Vector-based Representations for Semantic Composition

William Blacoe and Mirella Lapata
Institute for Language, Cognition and Computation
School of Informatics, University of Edinburgh
10 Crichton Street, Edinburgh EH8 9AB
w.b.blacoe@sms.ed.ac.uk, mlap@inf.ed.ac.uk

| dim.  | c.m. | Adj-N | N-N | V-Obj |
|-------|------|-------|-----|-------|
| SDS (BNC) |
| 2000 | +    | 0.37  | 0.38 | 0.28  |
| 2000 | ⊙    | 0.48  | 0.50 | 0.35  |
| 100  | RAE   | 0.34  | 0.29 | 0.32  |
| DM (3-BWC) |
| vary | +    | 0.37  | 0.30 | 0.29  |
| vary | ⊙    | 0.21  | 0.37 | 0.33  |
| 100  | RAE   | 0.21  | 0.29 | 0.06  |
| NLM (BNC) |
| 50   | +    | 0.28  | 0.26 | 0.24  |
| 50   | ⊙    | 0.26  | 0.22 | 0.18  |
| 100  | RAE   | 0.20  | 0.18 | 0.14  |

Table 3: Correlation coefficients of model predictions with subject similarity ratings (Spearman’s $\rho$); columns show dimensionality: fixed or varying (see Section 2.1), composition method: + is additive vector composition, ⊙ is component-wise multiplicative vector composition, RAE is Socher et al. (2011)’s recursive auto-encoder.

|       | NLM (BNC) | DM (3-BWC) | SDS (BNC) |
|-------|-----------|------------|-----------|
| +     | 69.04     | 73.51      | 72.93     |
| (con, other) |          | (other)    | (other)   |
| ⊙     | 67.83     | 67.54      | 73.04     |
| (sub, other) |         | (other)    | (other)   |
| RAE   | 71.42     | 68.23      | 73.28     |
| (sub, other) |         | (other)    | (other)   |

Table 4: Paraphrase classification accuracy in %. Included features are in parentheses: “con” is sentence vector concatenation, “sub” is sentence vector subtraction, “other” stands for 4 other features (see Section 4).

|       | NLM (BNC) | DM (3-BWC) | SDS (BNC) |
|-------|-----------|------------|-----------|
| +     | 81.00     | 82.16      | 80.76     |
| (con, other) |        | (other)    | (other)   |
| ⊙     | 80.41     | 80.18      | 82.33     |
| (sub, other) |       | (other)    | (other)   |
| RAE   | 80.16     | 80.55      | 82.14     |
| (sub, other) |       | (other)    | (other)   |

Table 5: Paraphrase classification F1-score in %. The involved features are exactly the same as in Table 4.

ERRATUM: It has come to our attention that there was a flaw in our use of Socher et al. (2011)’s recursive auto-encoder (RAE). Whereas the original version of this paper used one and the same pair of compositional matrices $W^{(1)}, W^{(2)}$ for composing word vectors from all involved sources, in this corrected version we retrained the RAE’s compositional matrices for each word vector source. This somewhat improves accuracy in our paraphrase detection experiment for sentence vectors obtained from the RAE. The outcome of the phrase similarity experiment, however, remains largely unchanged.
References
Richard Socher, Eric H. Huang, Jeffrey Pennin, Andrew Y. Ng, and Christopher D. Manning. 2011. Dynamic Pooling and Unfolding Recursive Autoencoders for Paraphrase Detection. In Advances in Neural Information Processing Systems 24, pages 801–809. Granada, Spain.