Appendix A. Data Set Descriptions

UIUC: ACE, a newswire subset of the ACE coreference data set (Mitchell et al., 2005), was introduced in (Ratinov et al., 2011). The first nominal mention of each gold coreference chain is annotated by Amazon's Mechanical Turk workers. MSNBC was developed by (Cucerzan, 2007), which consists of 20 MSNBC news articles on different topics.

AIDA: Based on CoNLL 2003 Named Entity Recognition data, Hoffart *et al.* (2011) hand-annotated all these proper nouns with corresponding entities in YAGO2. Both the dev set (AIDA-dev) and the test set (AIDA-test) are included in the benchmark.

TAC-KBP: From the annual TAC-KBP competitions[20], the evaluation sets from 2009 to 2012 are included (as well as a training set from 2010, TAC10T). Each data set consists of a series of linking queries for named entities. A query provides the surface form of the mention and the source document id. The source documents mainly come from newswire and web documents.

Appendix B. VINCULUM Algorithm

Subroutines: A mention extractor E, a candidate generator D, an entity type predictor TP, a coreference system R and a coherence function ϕ.
Input: Document d.
Output: Entity Link Annotations $\{(m, l_m)\}$

Extract mentions $M = E(d)$.
Run coreference resolution $R(d)$ and obtain coreference clusters of mentions. Denote the cluster containing a mention m as $r(m)$ and the representative mention of a cluster r as $rep(r)$.
for $m \in M$ **do**
  if $m = rep(r(m))$ **then**
    Generate candidates $C_m = D(m)$ (Sec. 4.2) ;
    use TP to predict the entity types (Sec. 4.3);
    for $c \in C_m$ **do**
      Compute the probability of each candidate $p(c|m, s_m)$ based on the predicted types.
    end
  else
    use the representative mention $rep(r(m))$ for linking (Sec. 4.4).
  end
  Set $p_m = \arg\max_{c \in C_m} p(c|m, s_m)$;
end
Let $P_d = \cup_{m_i \in M} \{p_{m_i}\}$ (Sec. 4.5) ;
for $m \in M$ **do**
  for $c \in C_m$ **do**
    Compute $p_\phi(c|P_d)$ using the given coherence function ϕ and the final score $s(c|m, d) = p(c|m, s_m) + p_\phi(c|P_d)$;
  end
  Set the final link $l_m = \arg\max_{c \in C_m} s(c|m, d)$
end
return $\{(m, l_m) : m \in M\}$

[20] http://www.nist.gov/tac/2014/KBP/