1 Annotation Guideline

1.1 Entity Category

- **Task**: Applications, problems to solve, systems to construct.
  E.g. information extraction, machine reading system, image segmentation, etc.

- **Method**: Methods, models, systems to use, or tools, components of a system, frameworks.
  E.g. language model, CORENLP, POS parser, kernel method, etc.

- **Evaluation Metric**: Metrics, measures, or entities that can express quality of a system/method.
  E.g. F1, BLEU, Precision, Recall, ROC curve, mean reciprocal rank, mean-squared error, robustness, time complexity, etc.

- **Material**: Data, datasets, resources, Corpus, Knowledge base.
  E.g. image data, speech data, stereo images, bilingual dictionary, paraphrased questions, CoNLL, Panntreebank, WordNet, Wikipedia, etc.

- **Evaluation Metric**: Metric measure or term that can express quality of a system/method.
  E.g. F1, BLEU, Precision, Recall, ROC curve, mean reciprocal rank, mean-squared error, robustness, compile time, time complexity...

- **Generic**: General terms or pronouns that may refer to an entity but are not themselves informative, often used as connection words.
  E.g. model, approach, prior knowledge, them, it...

1.2 Relation Category

Relation link cannot go beyond sentence boundary. We define 4 asymmetric relation types (Used-for, Feature-of, Hyponym-of, Part-of), together with 2 symmetric relation types (Compare, Conjunction). B always points to A for asymmetric relations.

- **Used-for**: B is used for A, B models A, A is trained on B, B exploits A, A is based on B. E.g.
  
  The TISPER system has been designed to enable many text applications.
  Our method models user proficiency.
  Our algorithms exploits local soothness.

- **Feature-of**: B belongs to A, B is a feature of A, B is under A domain. E.g.
  
  prior knowledge of the model
genre-specific regularities of discourse structure
English text in science domain

- **Hyponym-of**: B is a hyponym of A, B is a type of A. E.g. 
  
  TUIT is a software library
  
  NLP applications such as machine translation and language generation

- **Part-of**: B is a part of A... E.g.
  
  The system includes two models: speech recognition and natural language understanding
  
  We incorporate NLU module to the system.

- **Compare**: Symmetric relation (use blue to denote entity). Opposite of conjunction, compare two models/methods, or listing two opposing entities. E.g.
  
  Unlike the quantitative prior, the qualitative prior is often ignored...
  
  We compare our system with previous sequential tagging systems...

- **Conjunction**: Symmetric relation (use blue to denote entity). Function as similar role or use/ incorporate with. E.g.
  
  obtained from human expert or knowledge base
  
  NLP applications such as machine translation and language generation

1.3 **Coreference**

Two Entities that points to the same concept.

- **Anaphora and Cataphora**:
  
  We introduce a machine reading system... The system...
  
  The prior knowledge include...Such knowledge can be applied to...

- **Coreferring noun phrase**:
  
  We develop a part-of-speech tagging system...The POS tagger...

1.4 **Notes**

1. Entity boundary annotation follows the ACL RD-TEC Annotation Guideline ¹, with the extention that spans can be embedded in longer spans, only if the shorter span is involved in a relation.

2. Do not include determinators (such as the, a), or adjective pronouns (such as this, its, these, such) to the span. If generic phrases are not involved in a relation, do not tag them.

3. Do not tag relation if one entity is:
   
   - Variable bound:
     
     We introduce a neural based approach.. Its benefit is...
   
   - The word which:
     
     We introduce a neural based approach, which is a...

---

4. Do not tag coreference if the entity is

- Generically-used Other-ScientificTerm:
  ...advantage gained from local smoothness which... We present algorithms exploiting local smoothness in more aggressive ways...

- Same scientific term but refer to different examples:
  We use a data structure, we also use another data structure...

5. Do not label negative relations:
X is not used in Y or X is hard to be applied in Y

2  Annotation and Knowledge Graph Examples

Here we take a screen shot of the BRAT interface for randomly selected five papers. Following the annotation examples, we also attach five knowledge graphs examples, which are automatically constructed from our system including statistical machine translation (Figure 6), convolutional neural network (Figure 7), conditional random field (Figure 8), named entity recognition (Figure 9) and image segmentation (Figure 10).
1. We propose a unified variational formulation for joint motion estimation and segmentation with explicit occlusion handling.

2. This is done by a multi-label representation of the flow field, where each label corresponds to a parametric representation of the motion.

3. We use a convex formulation of the multi-label Potts model with label costs and show that the asymmetric map-uniqueness criterion can be integrated into our formulation by means of convex constraints.

4. Explicit occlusion handling eliminates errors otherwise created by the regularization.

5. As occlusions can occur only at object boundaries, a large number of objects may be required.

6. By using a fast primal-dual algorithm we are able to handle several hundred motion segments.

7. Results are shown on several classical motion segmentation and optical flow examples.

---

Figure 2: Annotation example 2 from CVPR

---

The problem of blind separation of underdetermined instantaneous mixtures of independent signals is addressed through a method relying on a nonstationarity of the original signals.

2. The signals are assumed to be piecewise stationary with varying variances in different epochs.

3. In comparison with previous works, in this paper it is assumed that the signals are not i.i.d. in each epoch, but obey a first-order autoregressive model.

4. This model was shown to be more appropriate for blind separation of natural speech signals.

5. A separation method is proposed that is nearly statistically efficient (approaching the corresponding Cramér–Rao lower bound), if the separated signals obey the assumed model.

6. In the case of natural speech signals, the method is shown to have separation accuracy better than the state-of-the-art methods.

---

Figure 3: Annotation example 3 from ICASSP
Figure 4: Annotation example 4 from ICML

This paper solves a specialized regression problem to obtain sampling probabilities for records in databases.

The goal is to sample a small set of records over which evaluating aggregate queries can be done both efficiently and accurately.

We provide a principled and provable solution for this problem; it is parameterless and requires no data insights.

Unlike standard regression problems, the loss is inversely proportional to the regressed-to values.

Moreover, a cost zero solution always exists and can only be excluded by hard budget constraints.

A unique form of regularization is also needed.

We provide an efficient and simple regularized Empirical Risk Minimization (ERM) algorithm along with a theoretical generalization result.

Our extensive experimental results significantly improve over both uniform sampling and standard stratified sampling which are de-facto the industry standards.

Figure 5: Annotation example 4 from IJCAI

Visitors who browse the web from wireless PDAs, cell phones, and pagers are frequently stymied by web interfaces optimized for desktop PCs.

Simply replacing graphics with text and reformatting tables does not solve the problem, because deep link structures can still require minutes to traverse.

In this paper we develop an algorithm, MINPATH, that automatically improves wireless web navigation by suggesting useful shortcut links in real time.

MINPATH finds shortcuts by using a learned model of web visitor behavior to estimate the savings of shortcut links, and suggests only the few best links.

We explore a variety of predictive models, including Naive Bayes mixture models and mixtures of Markov models, and report empirical evidence that
Figure 6: An example of our automatically generated knowledge graph centered on *statistical machine translation*. This is the original figure of Figure 3. in the paper.
Figure 7: Automatically generated knowledge graph example centered on *convolutional neural network*
Figure 8: Automatically generated knowledge graph example centered on *conditional random field*
Figure 9: Automatically generated knowledge graph example centered on *named entity recognition*
Figure 10: Automatically generated knowledge graph example centered on *image segmentation*