1. Overview of 2011 System

- Three components
  1) Baseline: 2010 System
     Lucene is used to do document retrieval;
     Slots are primarily filled using patterns (hand-coded and created semi-
     automatically using bootstrapping)
  2) Distant Learning
     A set of MaxEnt classifiers trained with Freebase and the KBP corpus
  3) Passage Retrieval (QA)
     Slot-specific passage retrieval system combined with name-type-based
     answer selection (implemented for 2 slots due to time limitations; used
     as a fall-back strategy)

2. Distant Learning for Slot Filling

- Class Label Refinement
  How to refine?
  1) Represent a training instance by its dependency pattern, the shortest path
     connecting the two names in the dependency tree of the sentence
  2) Estimate precision of the pattern by
     \[ \text{prec}(p,c) = \frac{\text{count}(p,c)}{\sum \text{count}(p,c)} \]
     Precision of a pattern \( p \) for the class \( c \) is defined as
     the number of occurrences of \( p \) in the class \( c \) divided by
     the number of occurrences of \( p \) in any of the classes \( C \)
  3) Assign the instance the class that its pattern is most precise about

- Undersampling the Majority Classes
  Why? Class distribution is extremely unbalanced. Too many OTHER instances
  How?
  Effort one:
  multiple n-way instead of single n-way classification
  Train an n-way classifier for each pair of name types
  Effort two:
  Downsize OTHER by randomly selecting a subset of them
  On average (10 runs on 2011 evaluation data)
  single n-way: 180 fills for 8 slots
  multiple n-way: 240 fills for 15 slots

- Contribution of Coreference
  No use of co-reference during training
  Use co-referred names of the query when filling slots
  P/R/F of the distant filler itself
  With co-reference: 36.4/11.4/17.4
  Without co-reference: 28.8/10.0/14.3

1. Overview of 2011 System

- Results on 2011 Evaluation Data
  New York University 2011 System for
  Knowledge Base Population Slot Filling
  Ang Sun, Ralph Grishman, Wei Xu and Bonan Min
  Computer Science Department, New York University
  {asun, grishman, xuwei, min} @ cs.nyu.edu

- Class Label Refinement
  Why refine?
  Class labels are noisy!
  Many False Positives because name pairs are often connected by
  non-relational contexts
  "Seattle" while a token
  represents most of information of interest.
  "is able to
  represent
  "if a Freebase instance but
  not a Freebase instance but
  "Seattle" while a token
  represents most of information of interest.
  "is able to
  represent
  "if a Freebase instance but
  not a Freebase instance but

- Ablation Study
  Undersampling Ratio
  \begin{tabular}{|c|c|c|c|c|c|c|c|c|c|}
    \hline
    Undersampling Ratio & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
    \hline
    F-Measure & .02 & .04 & .06 & .08 & .10 & .12 & .14 & .16 & .18 \\
    \hline
  \end{tabular}

- Class Label Refinement
  Why refine?
  Class labels are noisy!
  Many False Positives because name pairs are often connected by
  non-relational contexts
  "Seattle" while a token
  represents most of information of interest.
  "is able to
  represent
  "if a Freebase instance but
  not a Freebase instance but

- Ablation Study
  Undersampling Ratio
  \begin{tabular}{|c|c|c|c|c|c|c|c|c|c|}
    \hline
    Undersampling Ratio & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
    \hline
    F-Measure & .02 & .04 & .06 & .08 & .10 & .12 & .14 & .16 & .18 \\
    \hline
  \end{tabular}

- Class Label Refinement
  Why refine?
  Class labels are noisy!
  Many False Positives because name pairs are often connected by
  non-relational contexts
  "Seattle" while a token
  represents most of information of interest.
  "is able to
  represent
  "if a Freebase instance but
  not a Freebase instance but

- Ablation Study
  Undersampling Ratio
  \begin{tabular}{|c|c|c|c|c|c|c|c|c|c|}
    \hline
    Undersampling Ratio & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
    \hline
    F-Measure & .02 & .04 & .06 & .08 & .10 & .12 & .14 & .16 & .18 \\
    \hline
  \end{tabular}

- Class Label Refinement
  Why refine?
  Class labels are noisy!
  Many False Positives because name pairs are often connected by
  non-relational contexts
  "Seattle" while a token
  represents most of information of interest.
  "is able to
  represent
  "if a Freebase instance but
  not a Freebase instance but

- Ablation Study
  Undersampling Ratio
  \begin{tabular}{|c|c|c|c|c|c|c|c|c|c|}
    \hline
    Undersampling Ratio & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
    \hline
    F-Measure & .02 & .04 & .06 & .08 & .10 & .12 & .14 & .16 & .18 \\
    \hline
  \end{tabular}

- Ablation Study
  Undersampling Ratio
  \begin{tabular}{|c|c|c|c|c|c|c|c|c|c|}
    \hline
    Undersampling Ratio & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
    \hline
    F-Measure & .02 & .04 & .06 & .08 & .10 & .12 & .14 & .16 & .18 \\
    \hline
  \end{tabular}

- Ablation Study
  Undersampling Ratio
  \begin{tabular}{|c|c|c|c|c|c|c|c|c|c|}
    \hline
    Undersampling Ratio & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
    \hline
    F-Measure & .02 & .04 & .06 & .08 & .10 & .12 & .14 & .16 & .18 \\
    \hline
  \end{tabular}