Overview
Task: knowledge graph based simple question answering (KBSQA)
Knowledge Graph: multi-entity multi-relation directed graph containing fact triples (subject, relation, object)
Simple Question: can be answered by a single fact from knowledge graph
Our Method: subgraph ranking + joint scoring model + well-order loss
Result: new state of the art on SimpleQuestions dataset

Motivation
Challenges
(1) massive size of knowledge graph (billions of facts)
(2) variability of questions in natural language
Two-Step Solution
(1) subgraph selection
(2) fact selection

Conventional Approaches
(1) sequence labeling with BiLSTM-CRF + subgraph selection with n-grams
(2) match-scoring model + ranking loss

Problems
(1) subgraph is not ranked by relevance
(2) need leverage dependency between mention–subjects and pattern–relations
(3) ranking loss is suboptimal

Features
(1) jointly consider both input pairs and their dependency
(2) dependency dynamically adjusted by ||I|| and ||J||
(3) subject mismatch induces larger loss
(4) penalize subject mismatch to prune incorrect relations

Proposed Methods
(1) A subgraph ranking method with combined literal and semantic score
score(s, m) = |m(s, m) + (1 − ||I||) log P(s, m)|
|length of longest common subsequence|
= T(|m||I|) |m|−|I|N
= T(|m|−|I|)W

(2) A low-complexity joint-scoring CNN model

(3) A well-order loss

Experiments
Dataset
SimpleQuestions: 108,442 questions
Train/Valid/Test: 75,910/10,845/21,687
Knowledge Graph
Freebase (FB2M): 2,150,604 entities, 6,701 relations, 14,180,937 facts

Results
Table 1. Subgraph Selection Results

<table>
<thead>
<tr>
<th>Method</th>
<th>Spearman</th>
<th>Kendall</th>
<th>MRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>BiLSTM-CRF</td>
<td>0.732</td>
<td>0.802</td>
<td>0.754</td>
</tr>
<tr>
<td>Our method</td>
<td>0.803</td>
<td>0.839</td>
<td>0.806</td>
</tr>
</tbody>
</table>

Table 2. Fact Selection Accuracy

<table>
<thead>
<tr>
<th>Approach</th>
<th>Object (%)</th>
<th>Subject (%)</th>
<th>Relation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BiLSTM-CRF</td>
<td>77.45</td>
<td>78.36</td>
<td>71.60</td>
</tr>
<tr>
<td>Our method</td>
<td>80.48</td>
<td>81.47</td>
<td>75.26</td>
</tr>
</tbody>
</table>

Table 3. Error Decomposition (%) (total 3,157 errors)

Conclusions
(1) our ranking method improves subgraph selection
(2) our joint-scoring model with well-order loss improves fact selection
(3) incorrect subject or relation can still lead to correct answer