TEAM 10

Milestone 1 Report 11-791 Project

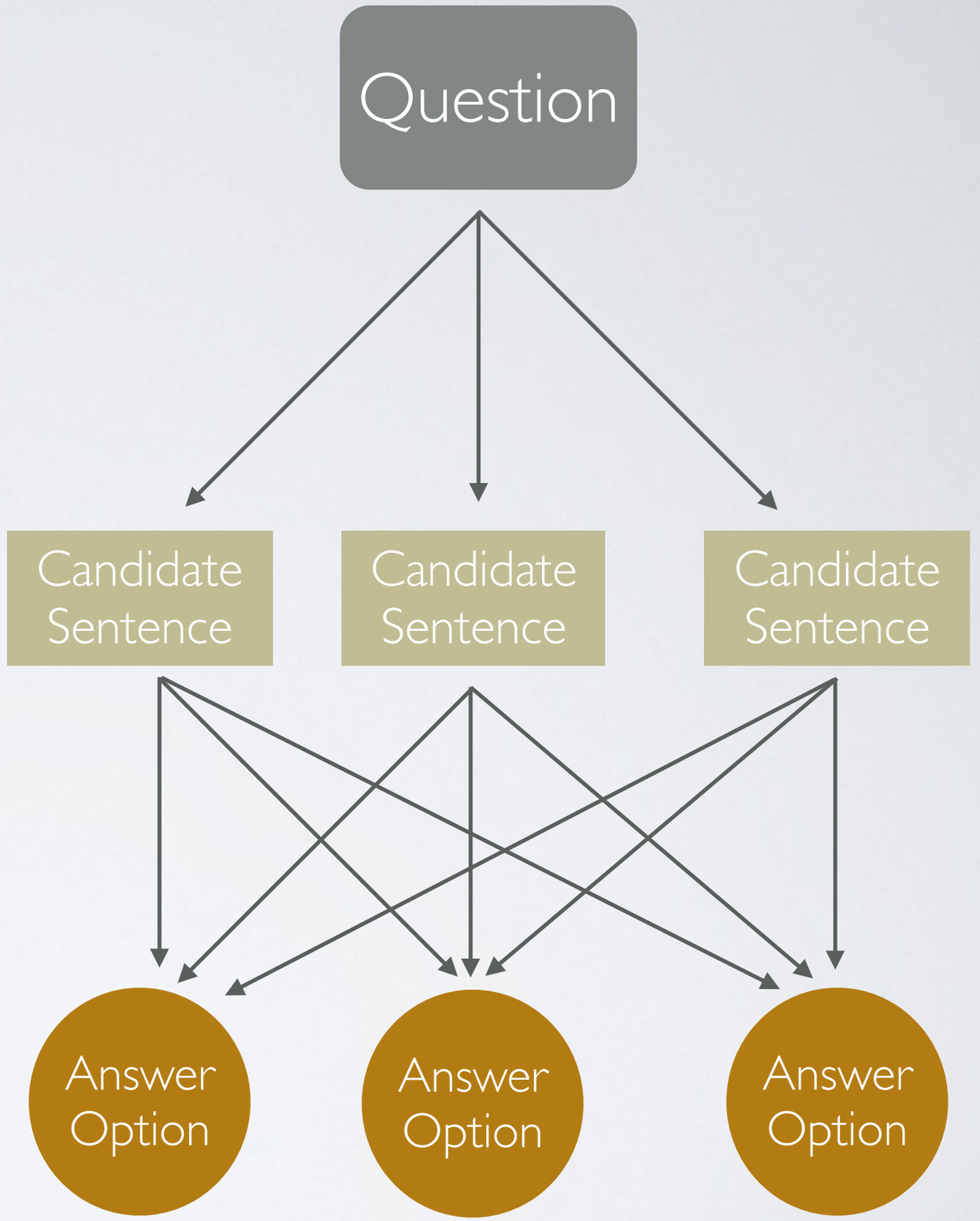
Nov. 13, 2013

Troy Hua Chenyan Xiong Vinay Vemuri Bo Ma

https://github.com/troyhua/hw5-team10

GENERAL FRAMEWORK

what we learn from the baseline



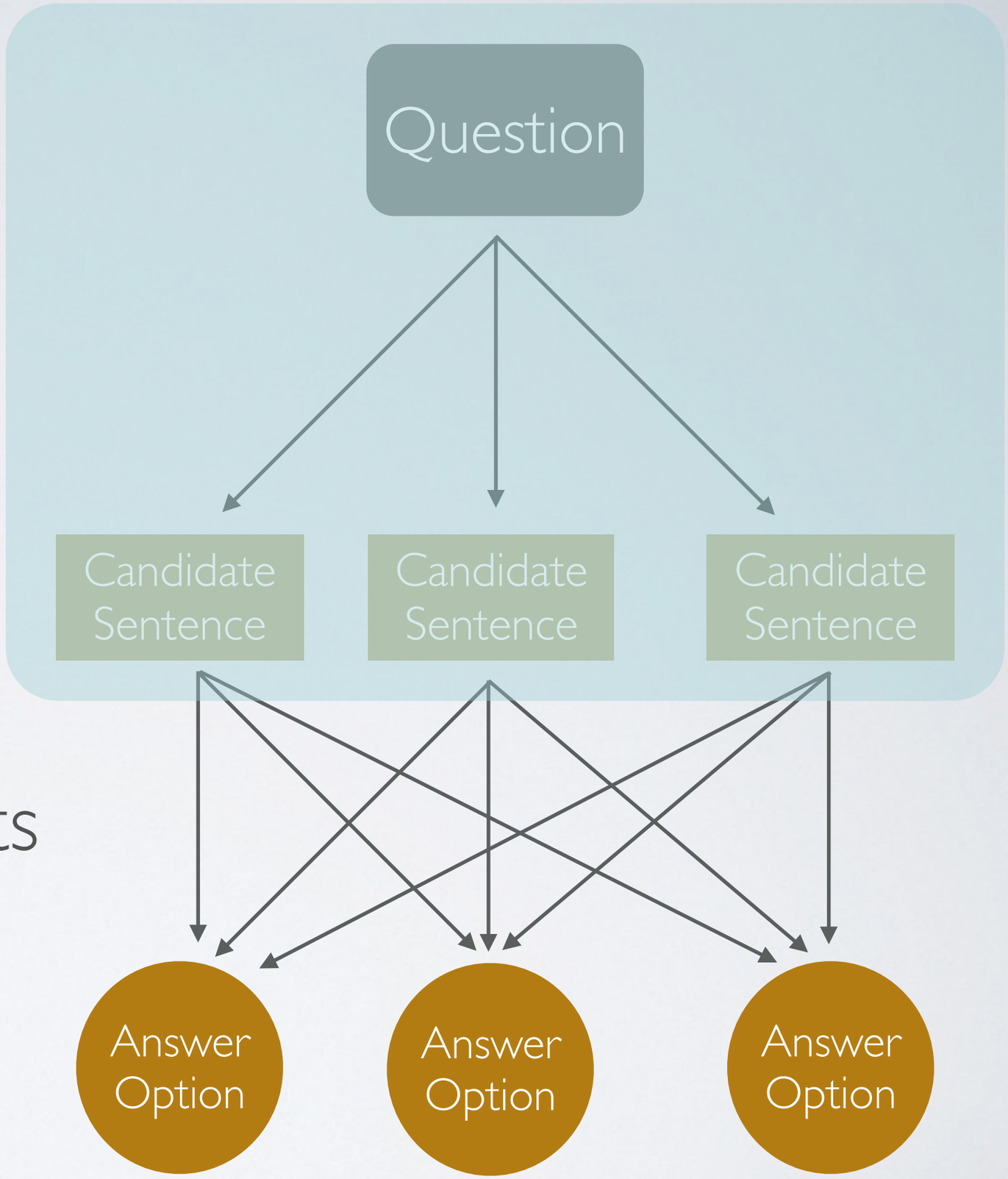
CANDIDATE SENTENCE SELECTION

Baseline Method:

Solr with NP, NER, etc.

Drawbacks:

* a black-box tool
* hard to tune weights for different features
* hard to use non-BoW features



Proposed Method:
Language Model Framework with supervised weights

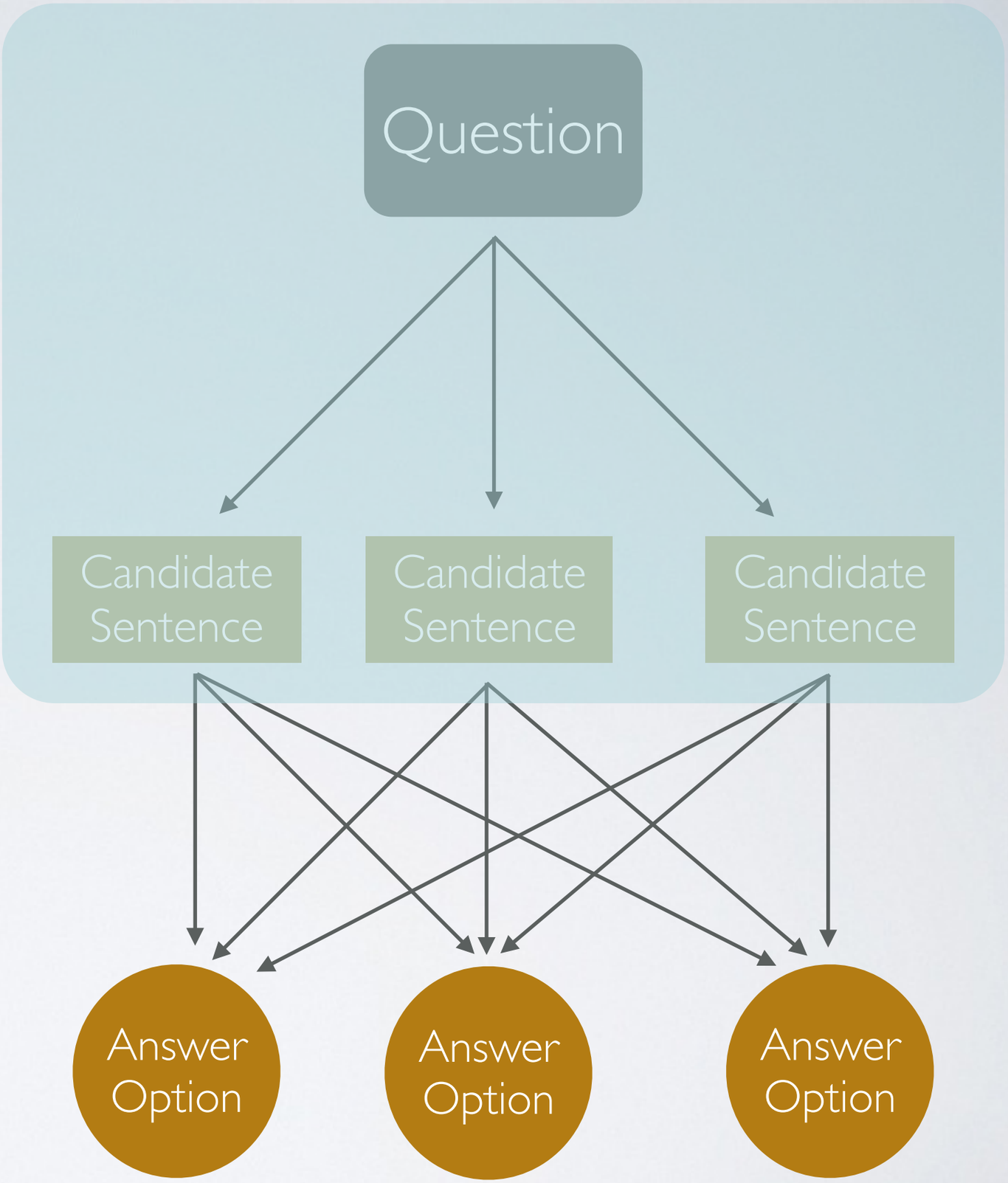
$$f(q, s_i) = p(q|s_i) = \sum_k \alpha_k p(v_k(q)|v_k(s_i))$$

Features:

Lexical features: unigram, bigram,
NER, NP, dependency parsing

Syntactic features: pos tags, parsing tree

Semantic features: adjacency in WordNet, freebase, other BIO background database



Proposed by Chenyan

ANSWER SCORE BASED ON CANDIDATE SENTENCES

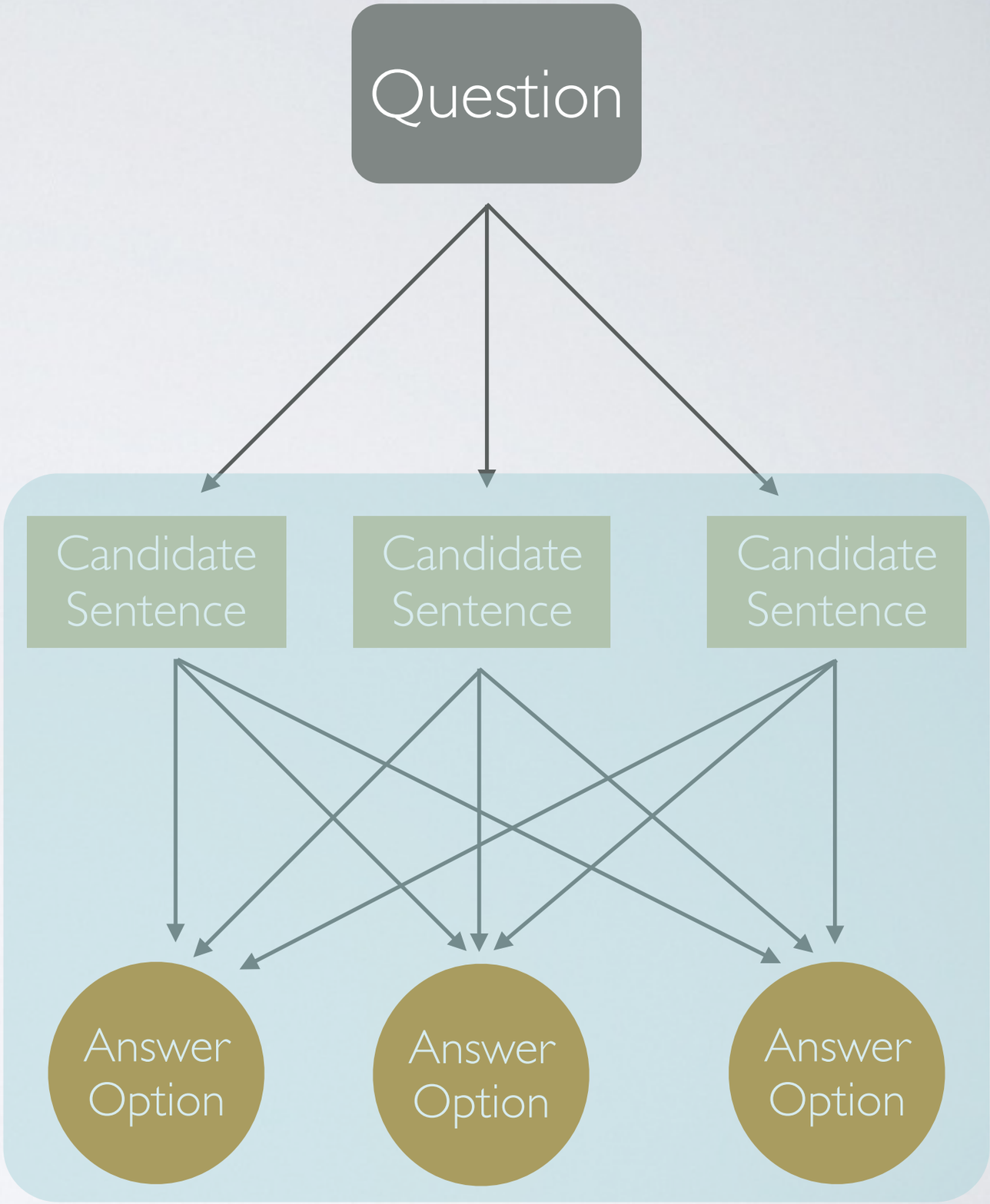
Baseline Method:

Joint voting with Similarity Score and PIM score in background documents

Drawbacks:

* Joint voting has a bias on answers with frequent terms

* Relevance with Question is ignored here



ANSWER SCORE BASED ON CANDIDATE SENTENCES

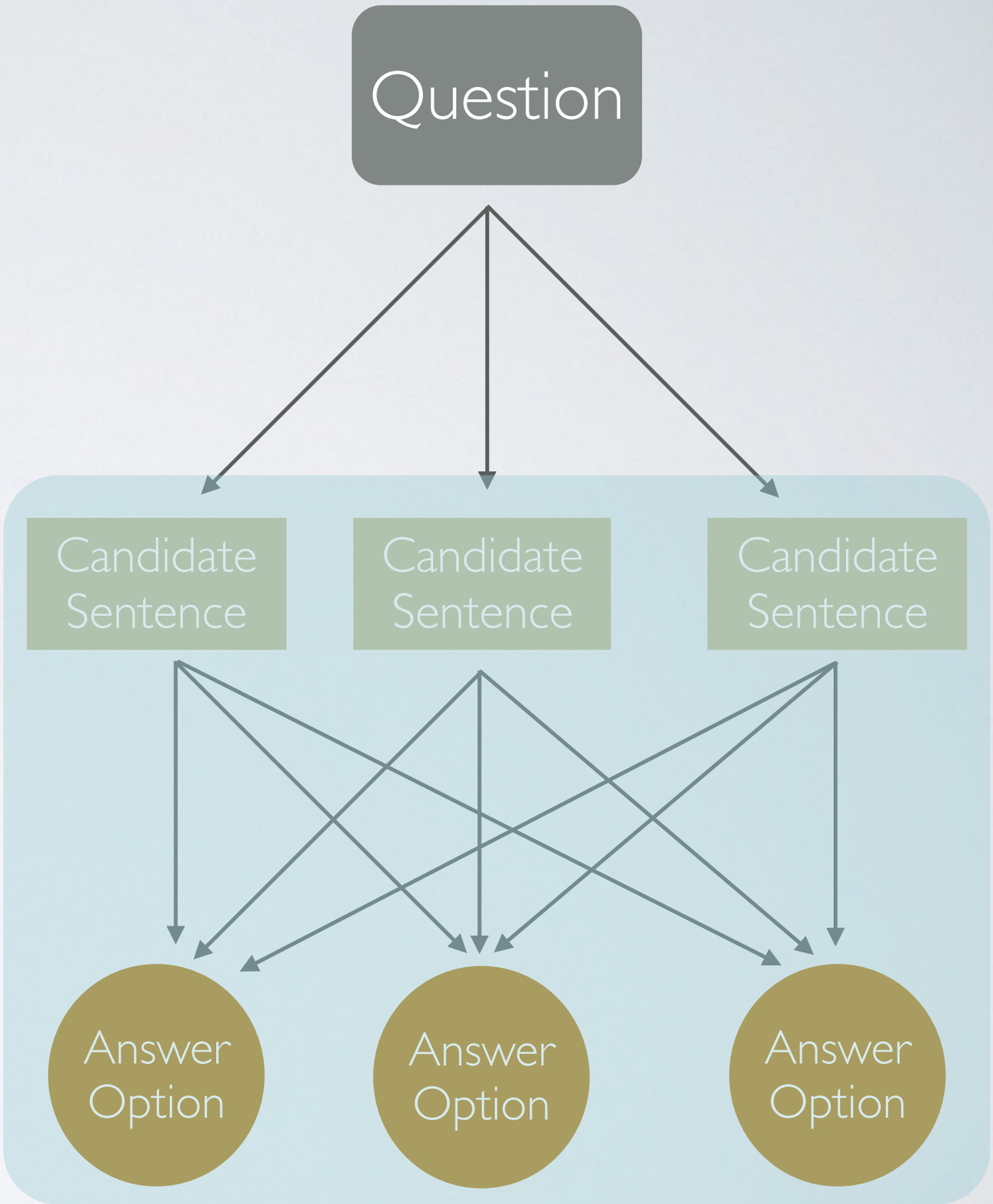
Proposed Method:

Lexical Features:

* Cosine Similarity
* Jaccard Similarity
* Dice Similarity

Semantic Features:

* Semantic Role of the "what/how/why" word compared with the answer word in the candidate sentence

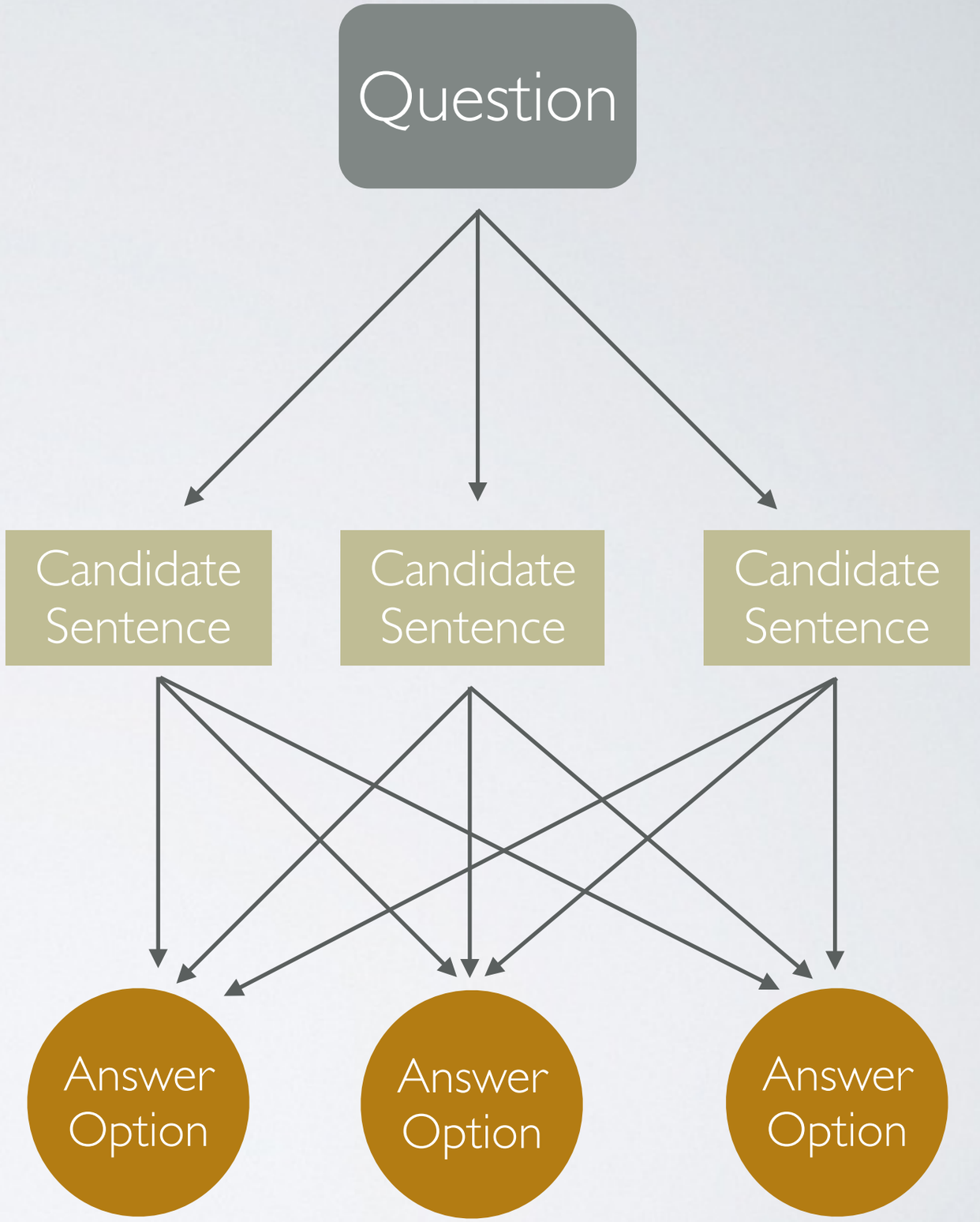


Proposed by Vinay and Bo

PROBLEMS WITH THE FRAMEWORK

* Good candidate sentences **cannot** be selected from only the question

* The relation between the question and answer is never measured directly.



Question

Answer Option

Answer Option

Answer Option

Candidate Sentence

Candidate Sentence

Candidate Sentence

Candidate Sentence

Candidate Sentence

Candidate Sentence

Candidate Sentence

Candidate Sentence

Candidate Sentence

max

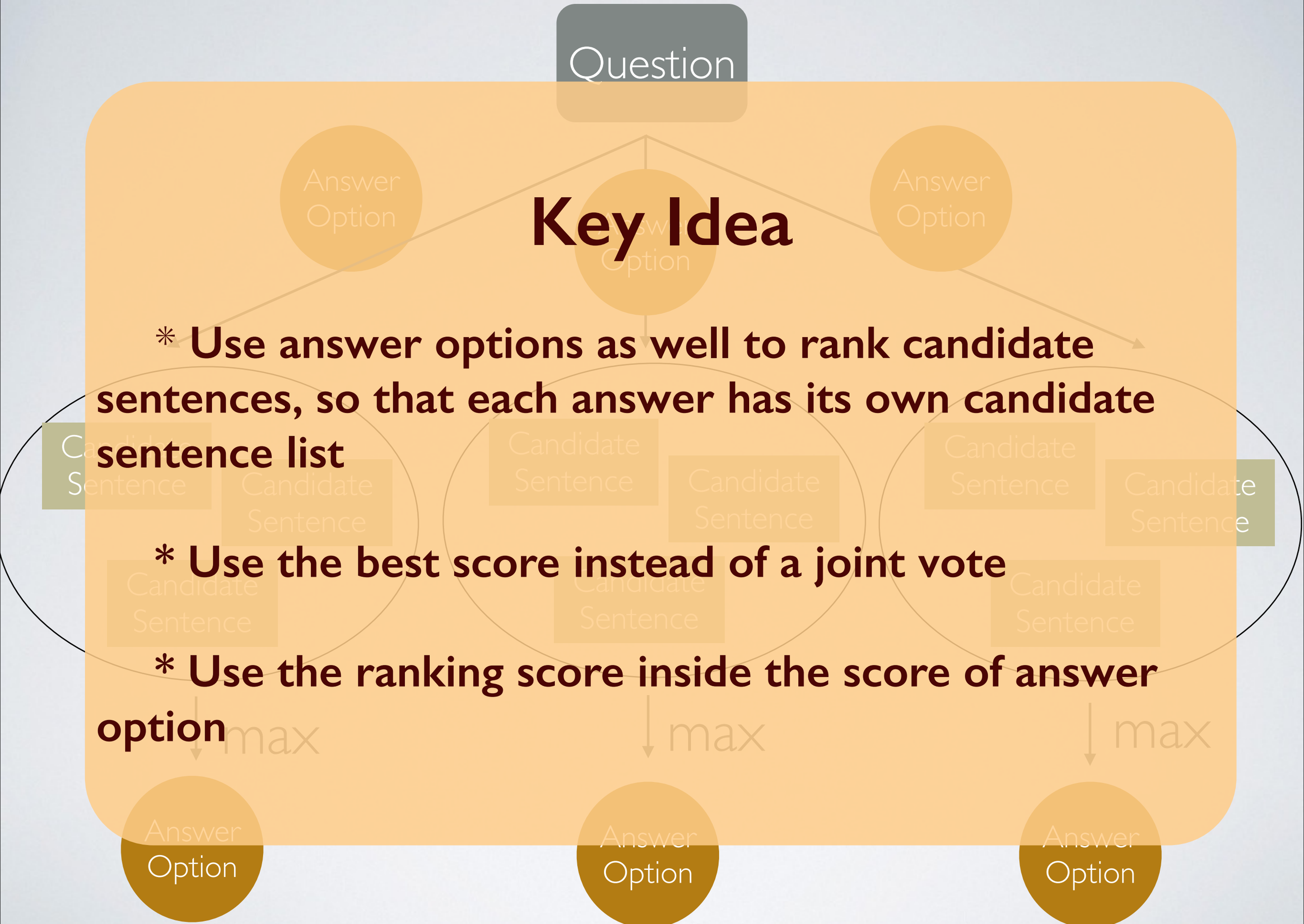
max

max

Answer Option

Answer Option

Answer Option



Key Idea

* Use answer options as well to rank candidate sentences, so that each answer has its own candidate sentence list

* Use the best score instead of a joint vote

* Use the ranking score inside the score of answer option

CURRENT STATUS

- Changed the type system to support our design on the framework
- Still using Solr query to find candidate sentence, but add answer to the query
- Performance on 2012 test set:
  - Baseline:
    - 20% (3/10, 1/10, 1/10, 3/10)
  - Our current methods:
    - 27.5% (3/10, 4/10, 2/10, 2/10)

TEAM 10

Milestone 1 Report 11-791 Project

Nov. 13, 2013

Troy Hua Chenyan Xiong Vinay Vemuri Bo Ma

https://github.com/troyhua/hw5-team10