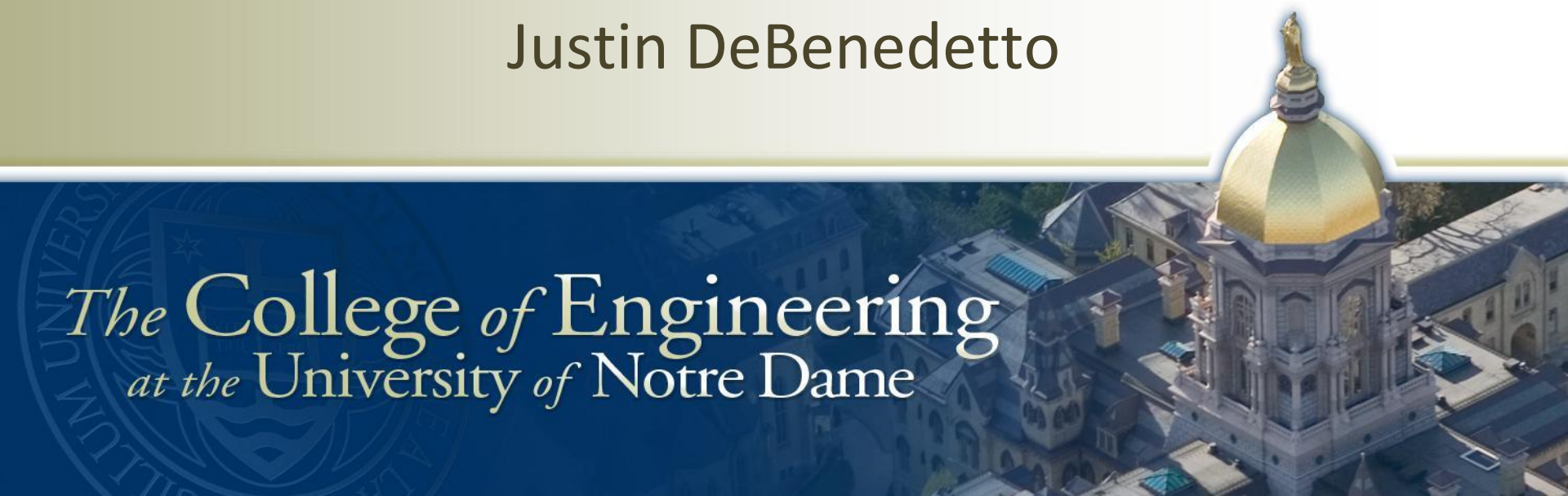
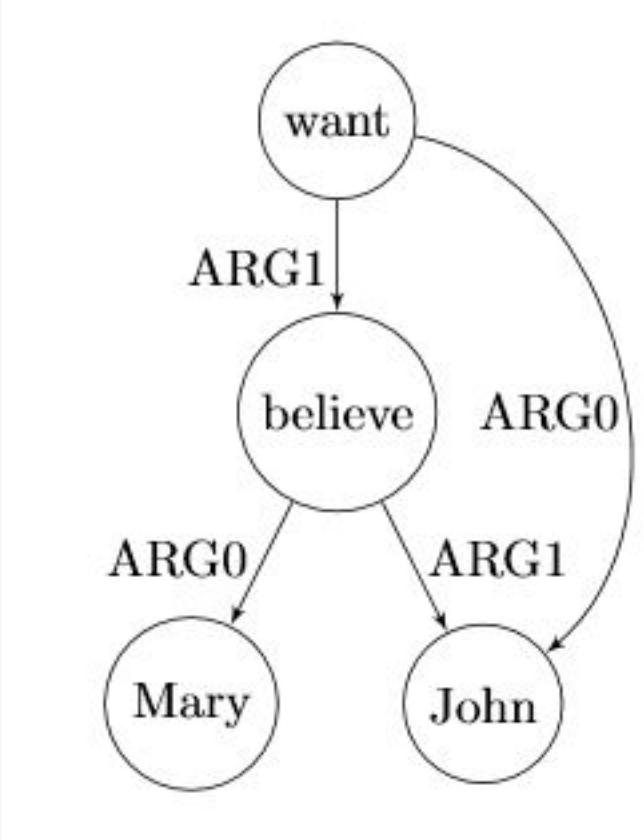
Graph Similarity Scoring
Applied to
Abstract Meaning Representation

Justin DeBenedetto



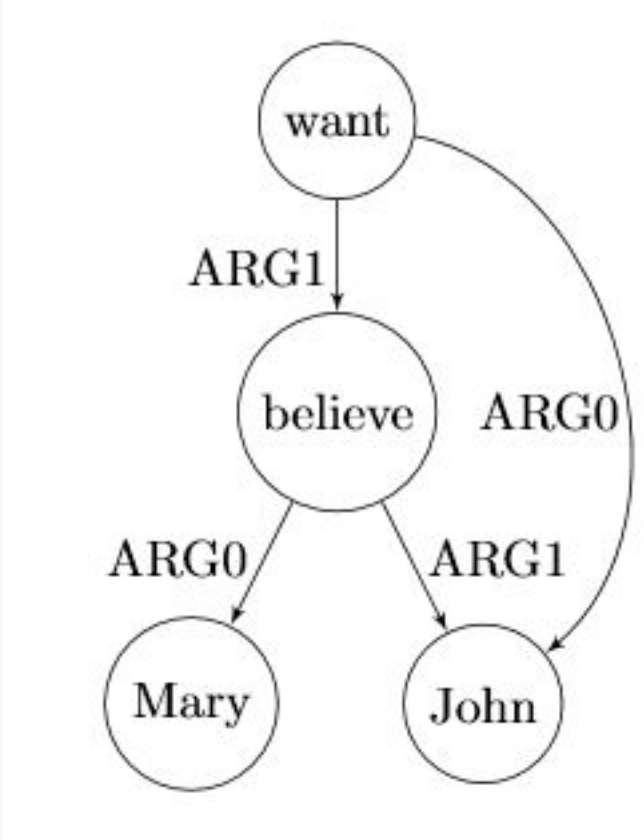
Abstract Meaning Representation (AMR)

- AMRs are a semantic formalism which models sentences



Abstract Meaning Representation (AMR)

- AMRs are a semantic formalism which models sentences
  - Nodes represent concepts
  - Edges represent relations between concepts
    - Semantic roles
    - ARG0 = Agent
    - ARG1 = Patient
    - Example AMR for sentence: “John wants Mary to believe him.”



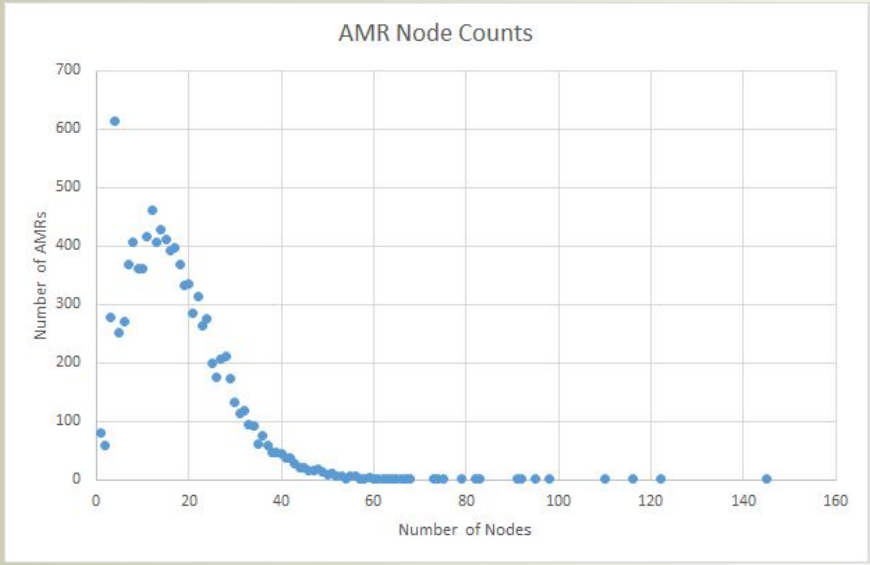
Properties of AMRS as Graphs

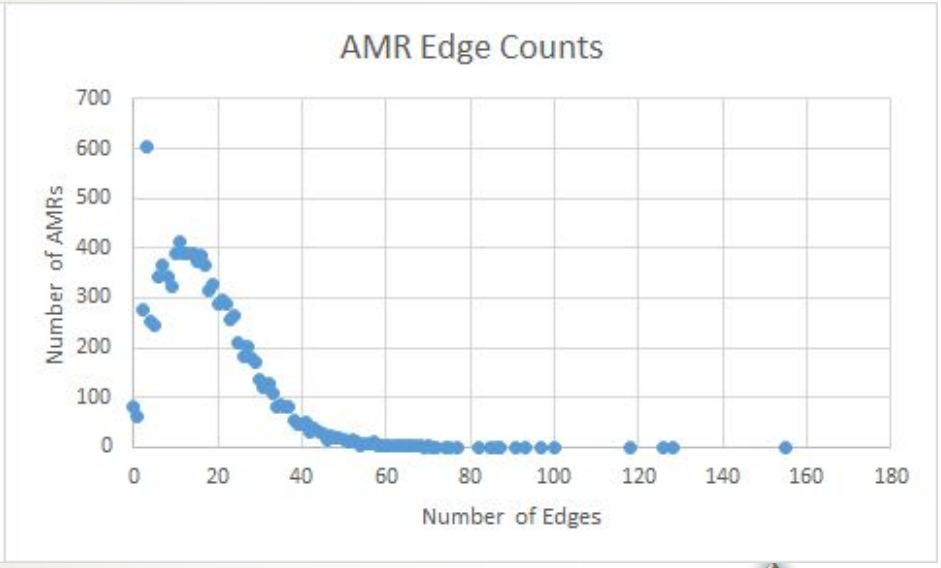
- Some properties of AMRs
  - Directed Acyclic Graphs (DAGs)
  - Single rooted (focus of sentence)
  - Each AMR represents a sentence

Dataset

- Set of 10,312 AMRs from various news sources
- Average number of nodes is: 17.1
- Average number of edges is: 17.1
- More than half are trees

Dataset


AMR Node Counts
Number of AMRs
Number of Nodes
0
100
200
300
400
500
600
700
0
20
40
60
80
100
120
140
160


AMR Edge Counts
Number of AMRs
Number of Edges
0
100
200
300
400
500
600
700
0
20
40
60
80
100
120
140
160
180

Application

- Given multiple candidate AMRs, find best one
- Use some AMRs for training
  - Need a way to score each choice
  - Want pairwise digraph similarity score

Kernel: Graph Similarity Scoring

- Want to assess similarity of a pair of graphs
- Several measures exist:
  - Degree distribution
  - Diameter
  - Clustering coefficient
- We have node and edge labels
  - Typical for AMR is SMATCH

SMATCH

- Semantic Match score
  - Find best matching of nodes
  - Score based on node and edge labels
  - F1 score
    - Node label
    - For each edge: edge type and end points

Pseudocode

```
For every node mapping:
    For each node pairing:
        If labels match: correct++
        Else: wrong++
        For each edge from nodes:
            If endpoint matches: correct++
            Else: wrong++
```

Complexity

- Most direct way (previous slide) has complexity ~O(N!|N+E|)
  - N = number of nodes in graph
  - E = number of edges
- In practice, we want to prioritize matching correct labels together
  - ~O((N-k)!|N+E|)
    - k = number of matched labels

SMATCH Evaluation

- SMATCH is used as an evaluation metric for AMR generation
- Only works when we have a “gold” AMR to evaluate against
- Can be made efficient

My Research

- Scoring without “gold” AMR
- Learn local weights to score likelihood of nodes and edges
- Combine local weights efficiently into a global score
- Use this to rerank
- Evaluate test AMRs scored this way using SMATCH score