Contents

• Background
• Outline of the Paper
• Sequence-to-sequence Model
• Abstract Meaning Representation
• Key Takeaways
What is AMR

AMR – Abstract Meaning Representation

A method to define “Who did what to whom?”

Forms:

- Conjunctions of logical triples
- Rooted, labeled, directed, graph
English sentence: “I do not understand”, said the little prince.

Logic format:
\[
\exists s, p, l, u, -:
\text{instance}(s, \text{say-01}) \land \text{instance}(p, \text{prince}) \land \text{instance}(l, \text{little}) \land \text{instance}(u, \text{understand}) \land \text{instance}(-, -) \land \text{arg0}(s, p) \land \text{arg1}(s, u) \land (u, p) \land \text{mod}(p, l) \land \text{polarity}(u, -)
\]

AMR format (PENMAN notation):

\[
(s / \text{say-01}
  :\text{arg0} (p / \text{prince}
    :\text{mod} (l / \text{little}))
  :\text{arg1} (u / \text{understand-01}
    :\text{arg0} p
    :\text{polarity} -))
\]

Graph format:

1. Variables (graph nodes) for entities, events, and states.
2. Each node in the graph represents a semantic concept.
3. Concepts can either be English words (prince), PropBank framesets (say-01), or special keywords
AMR – Example

Obama was elected and his voters celebrated.
Seq2seq Model

Sequence-to-sequence learning (Seq2Seq) is about training models to convert sequences from one domain to sequences in another domain, by constructing an encoder and decoder.
Method Outline - Tasks

With a pair of natural sentence s, and AMR a, train an AMR parser to predict AMR a for sentence s, and a AMR generator to predict sentence s with AMR a.
Method Outline – Seq2seq Model

• Stacked bidirectional-LSTM Encoder and Decoder

• Encode an input sequence and to decode from the hidden states produced by the encoder.

• Concatenate the forward and backward hidden states at every level of the stack instead of at the top of the stack.

• Introduce dropout in the first layer of the encoder
Method Outline - Pair Training

1. Input: Training set of sentences and associated AMR graphs

2. Output: AMR parser and AMR generator

3. Self-training: (1) parse samples from a large, unlabeled corpus, (2) create a new set of parameters by training on the previous iteration, and (3) tuning parameters. AMR Parser

4. Use Parser to label AMRs for corpus
Method Outline - Pair Training

1. Generated expensive AMR associated corpus
2. Increased the sample size for Seq2Seq Model
3. Reduced Sparsity
Method Outline – AMR Preparation

1. Graph Simplization
2. Dates Anonymization
3. Name Entity Clustering
Methods Outline – AMR Preparation

US officials held an expert group meeting in January 2002 in New York.

(a) hold
:ARG0 person :ARG0-of have-org-role :ARG1 country :name name :op1
United :op2 States :ARG2 official
:ARG1 meet :ARG0 person :ARG1-of expert :ARG2-of group
:time date-entity :year 2002 :month 1
:location city :name name :op1 New :op2 York

(b) country_0 officials held an expert group meeting in month_0 year_0 in city_1.
hold
:ARG0 person :ARG0-of have-org-role :ARG1 country_0 :ARG2 official
:ARG1 meet :ARG0 person :ARG1-of expert :ARG2-of group
:time date-entity year_0 month_0
:location city_1

(c) loc_0 officials held an expert group meeting in month_0 year_0 in loc_1.
hold
:ARG0 person :ARG0-of have-org-role :ARG1 loc_0 :ARG2 official
:ARG1 meet :ARG0 person :ARG1-of expert :ARG2-of group
:time date-entity year_0 month_0
:location loc_1

(d) loc_0 officials held an expert group meeting in month_0 year_0 in loc_1.
hold
:ARG0 ( person :ARG0-of ( have-org-role :ARG1 loc_0 :ARG2 official ) )
:ARG1 ( meet :ARG0 ( person :ARG1-of expert :ARG2-of group ) )
:time ( date-entity year_0 month_0 )
:location loc_1
Methods Outline – AMR Preparation

1. Reduced complexity

2. Addressed open domain vocabulary entries, such as named entities.
1. A novel approach of using Seq2seq model on AMR decoding and encoding though details to be discussed.

2. Reduced Sparsity by paired training

3. Open-domain capability for unlabeled dataset
Thank You