Deep Learning Methods for Reading Comprehension Question Answering

Nick Landy, Anand Natu, Megan Knight
Stanford University

Motivation

- Reading comprehension is a fundamental and multi-faceted problem in AI applied to NLP
- Design a deep learning model for context-intelligent question answering, and examine drivers of performance for specific QA tasks

Problem Definition

The model reads a passage of text…

…and answers questions about the contents

Passage (Input):
Sandra is in the ***kitchen****. Sandra* ***picked up*** *the* ***football*** *there. Sandra traveled to the* ***hallway****. Sandra* ***discarded*** *the* ***football*** *there.*

Question (Input):
Where is the football?

Answer (Output):
Hallway

Dataset and Context

Dataset: Facebook bAbI

Baseline Model: Attention-based model with pre-trained GloVe word embeddings and bi-directional GRU for context embedding

Evaluation Metrics: Exact Match (EM) accuracy, cross-entropy loss function, Adam optimizer

Methodology / Model Definition

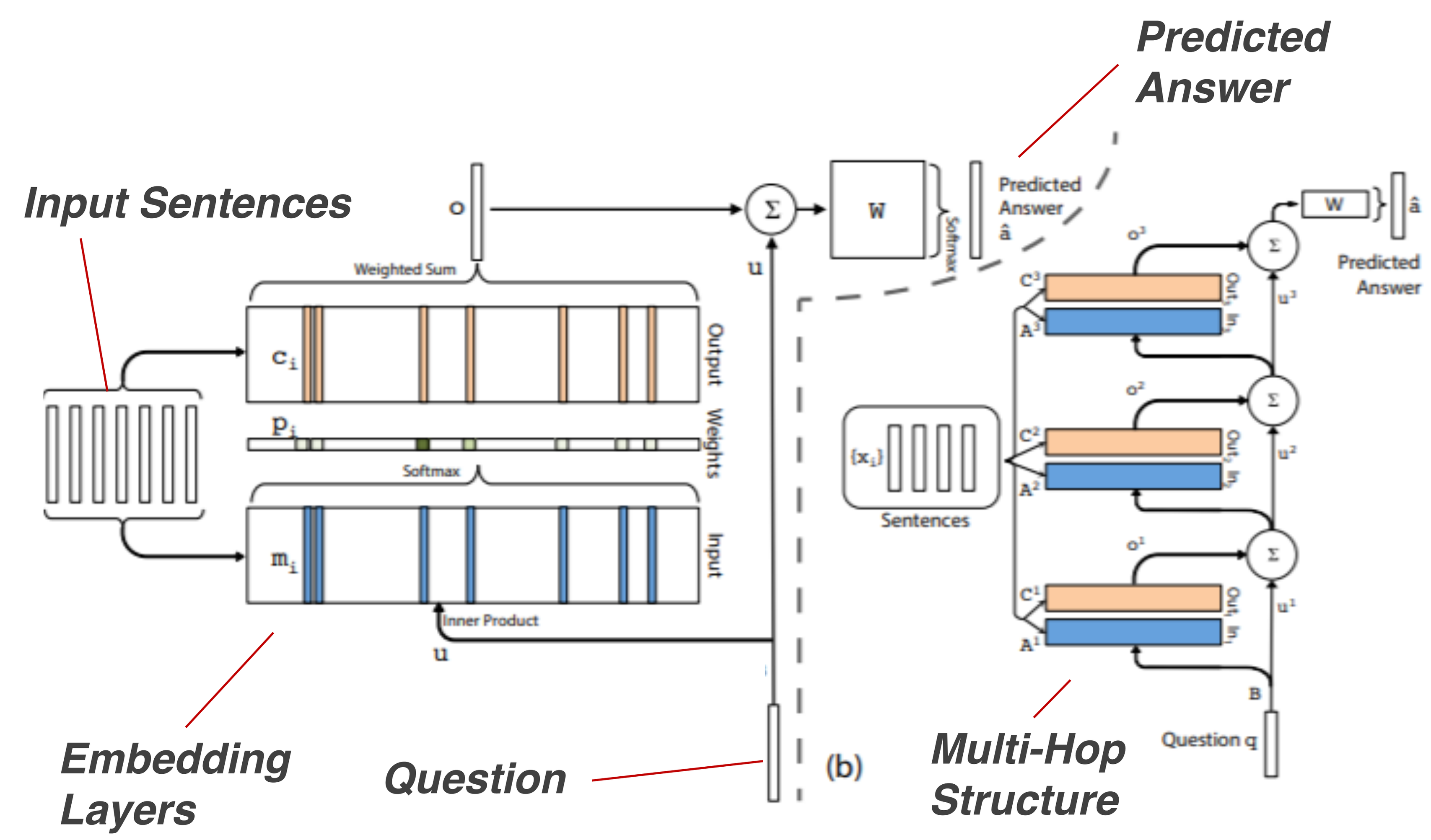
MemNN: A multi-hop memory network replacing the contextual RNN embedding of our baseline with learning from an external-memory representation

1. Probability of the encoded question u across inputs $\{m_1, m_2 \dots m_i\}$:
$$p_i = Softmax(u^T m_i)$$

2. Output memory representation of encoded inputs c_i
$$o = \sum_i p_i c_i$$

3. Final prediction and activation:
$$\hat{a} = Softmax(W(o + u))$$

4. K-stacking of "hops" in the memory network:
$$u^{k+1} = u^k + o^k$$



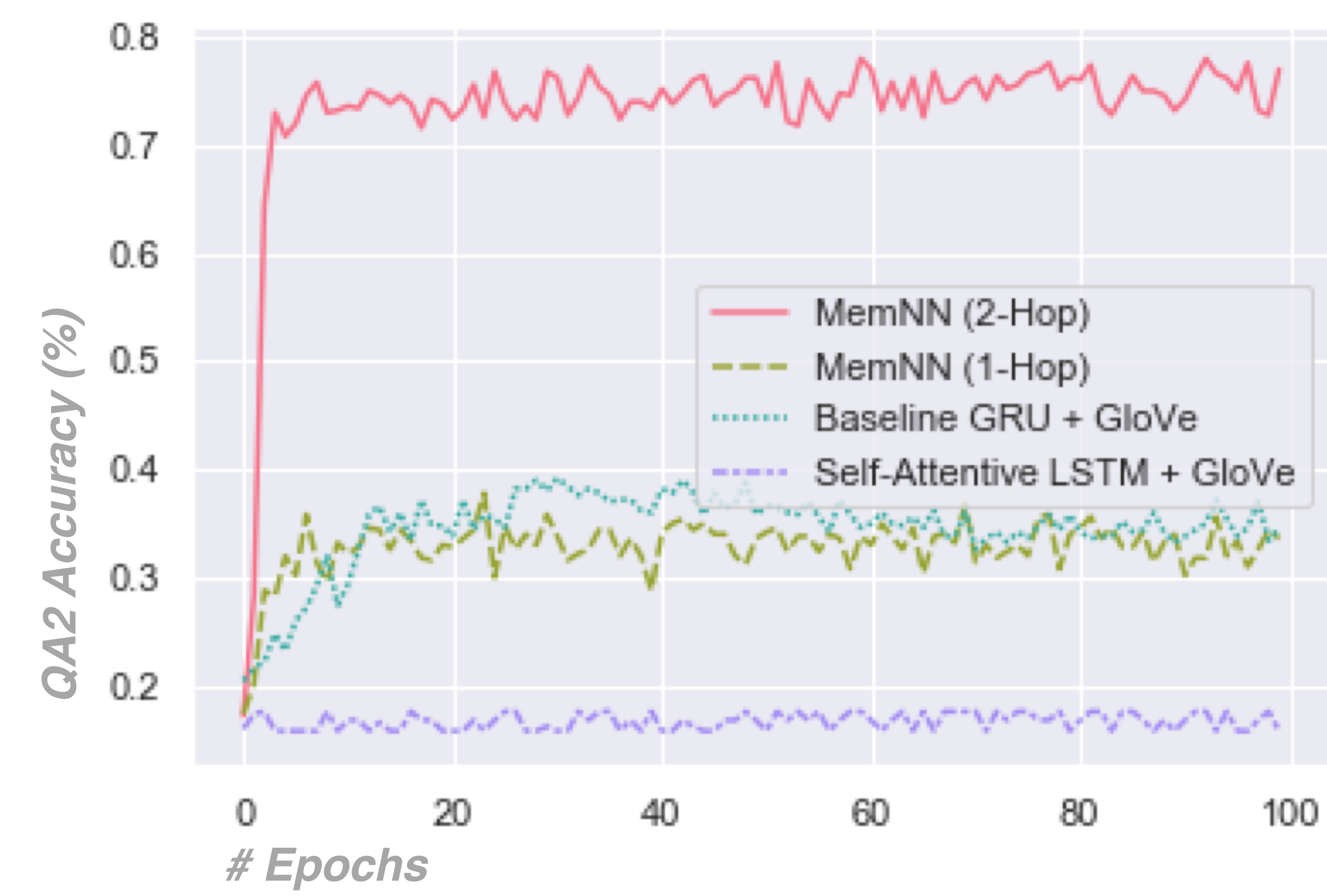
Self-Attention: Based on Microsoft R-Net – "dense" attention calculation for each representation $\{v_1 \dots v_N\}$

1. Dot multiplication
$$e_j^i = v^T tanh(W_1 v_j + W_2 v_i)$$

2. Activation
$$\alpha_i = Softmax(e^i)$$

3. Self-attention output
$$a^i = \sum_{j=1}^{N} \alpha_j^i v_j$$

Results and Analysis



QA2 accuracies of ~80% reached at convergence with MemNN-2Hop, but single-hop mostly mirrored baseline

RNN vs. external memory exhibit clear pros and cons depending on task (contextual vs. factual intelligence)

QA Challenge	Baseline	LSTM + Self-Attn	MemNN One-Hop	MemNN Two-Hop
QA1 Single Supporting Fact	42%	51%	99%	99%
QA2 Two Supporting Facts	33%	17%	36%	75%
QA4 Two Argument Relations	79%	70%	67%	68%
QA6 Yes/No Questions	74%	50%	50%	50%
QA7 Counting	76%	57%	80%	83%
QA8 Lists Sets	72%	77%	86%	89%
QA9 Simple Negation	71%	64%	90%	90%
QA10 Indefinite Knowledge	62%	44%	82%	82%