SEMANTIC REASONING USING DIFFERENTIABLE NEURAL COMPUTERS AND ENTNET

Jonathan Külz & Omar Dahroug

Team

Jonathan

M.Sc. Robotics

Omar

M.Sc. Robotics

Agenda

- Project Motivation & Objectives

- Introduction to DNC and EntNet

- Datasets

- Current Training Results

- Next Steps & Open Questions

Project Motivation and Objectives

- Goal:

  *"**Understand** and **implement** the Differentiable Neural Computer (DNC) and the Entity Network (EntNet) on the bAbI dataset in the context of uncertainty aware predictions."*

- Objectives:

  1. *Implement existing solutions: DNC, EntNet and LSTM on bAbI dataset*
  2. *Add datasets: Penn Treebank and Wikitext-103*
  3. *Elaborate on built models*
     - Improve models on additional dataset
     - Improve uncertainty awareness (e.g. MCDropout, temperature scaling)

Introduction to DNC

- Goal: **combine** ability to detect long-time or large-scale structure with microscopic variability.

- Processing done by **controller** (LSTM).

- Interaction with **external memory**, using read and/or write operations.

- **Temporal links** keep track of the order things were written in.

- **Writing** locations are based on stored content and a usage weighting of memory address.

- **Reading** is are based on content similarity and temporal relations between memory entries.
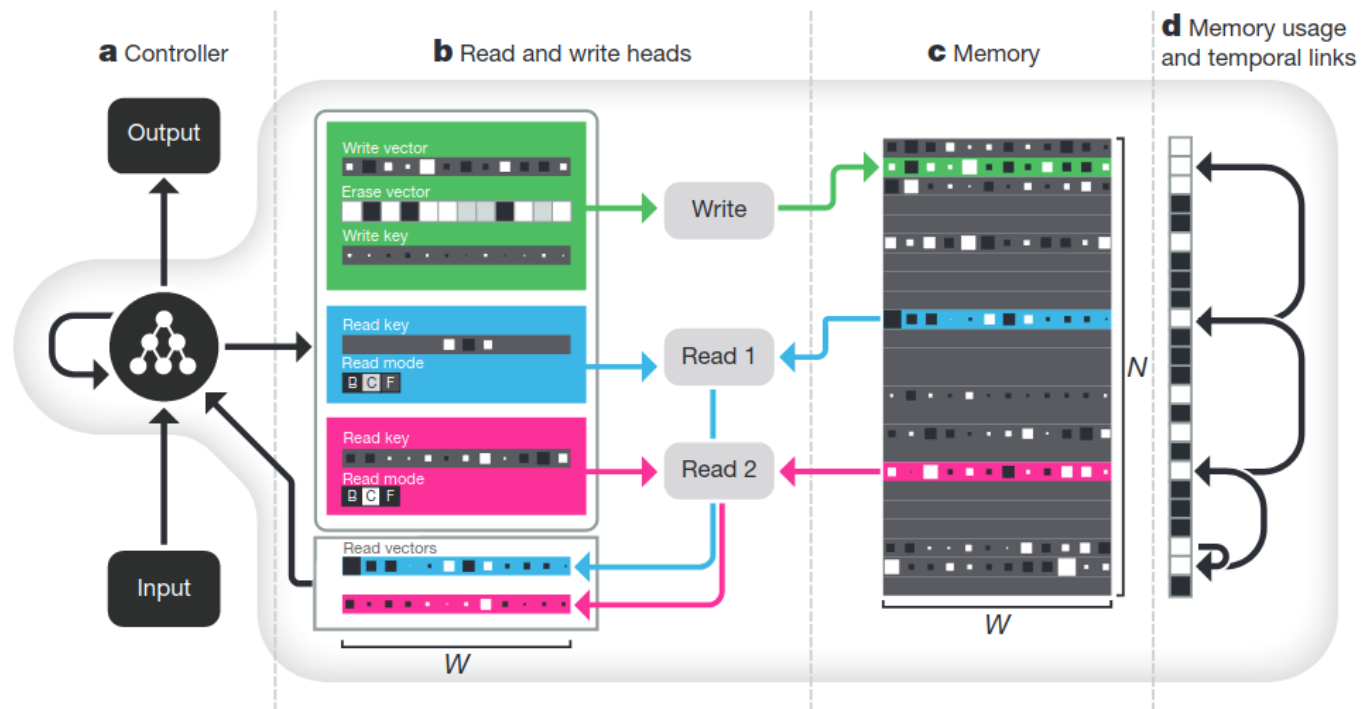
Source: Google DeepMind, 2016

Introduction to EntNet

- Goal: **Track** and **capture** the story dynamics over long timescales.

- Each **memory cell** is associated with a concept or entity in the world.

- EntNet performs location and content-based read and write operations on a fixed size **external memory.**
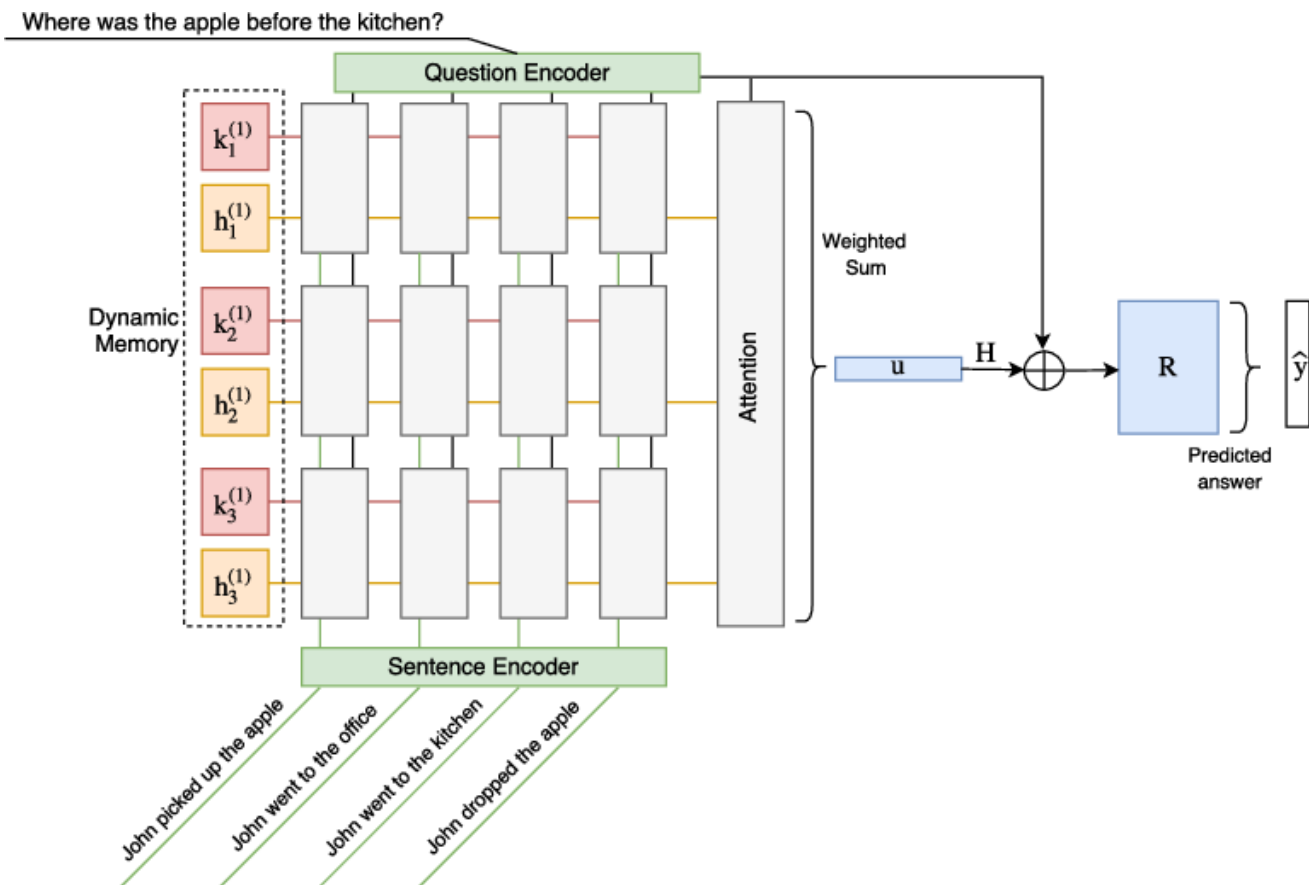
Source: Facebook AI Research, 2017
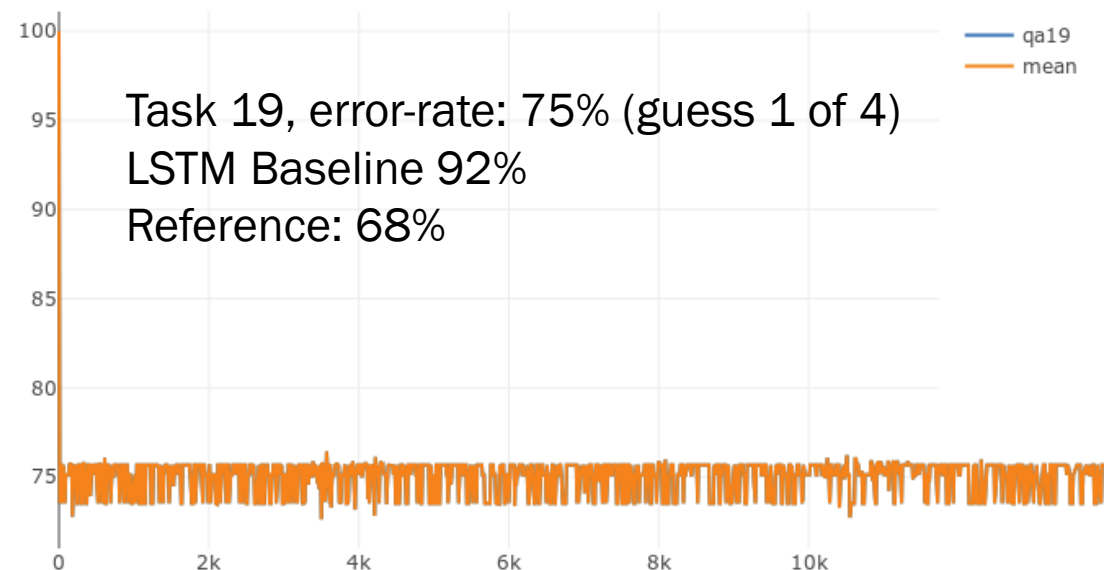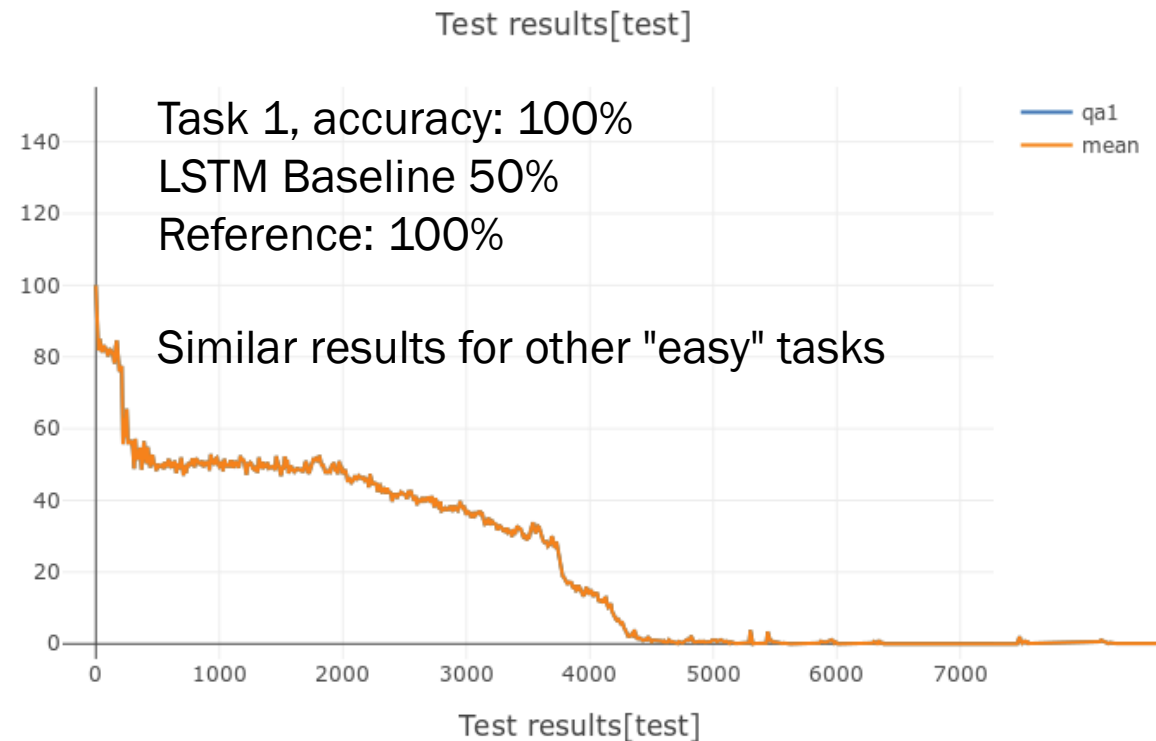
Datasets

bAbi

- Contains 20 different tasks for semantic reasoning

- Example task: (basic coreference)

  - Mary went to the kitchen.

  - Then she went to the bedroom.

  - Where is Mary?   **bedroom**

Penn Treebank & WikiText-103

- Two large NLP datasets known for quite some years, but so far not evaluated for a DNC

- PT: Syntactic annotations

- Wiki: Semantic reasoning, find the right token

# Current Results DNC

Works as expected

Test results[test]

Task 1, accuracy: 100%
LSTM Baseline 50%
Reference: 100%

Similar results for other "easy" tasks

Test results[test]

Task 19, error-rate: 75% (guess 1 of 4)
LSTM Baseline 92%
Reference: 68%

Current Results EntNet

- **Cyclic learning rate:** [1e-4; 1e-2]

- 100 **Epoch**

- **Bach size** 32

- **Task 1:** Low accuracy achieved: max of 32% at Epoch 60

- **Task 11:** Both training and validation loss decreases at Epoch 100

- **Task 19:** constant loss at low epoch

val_task1

val_task19
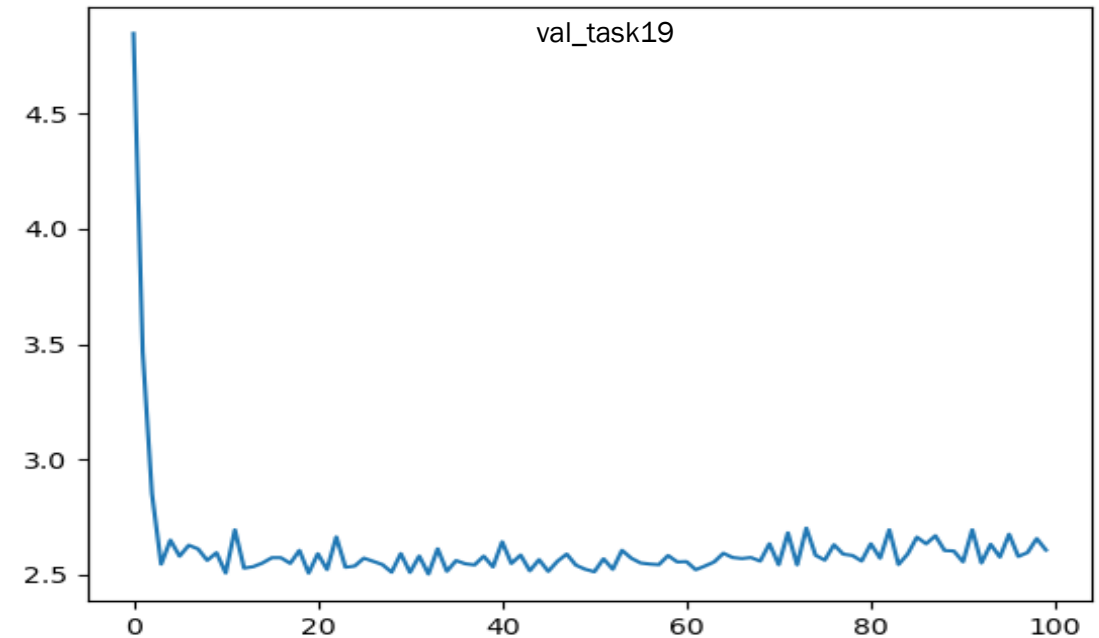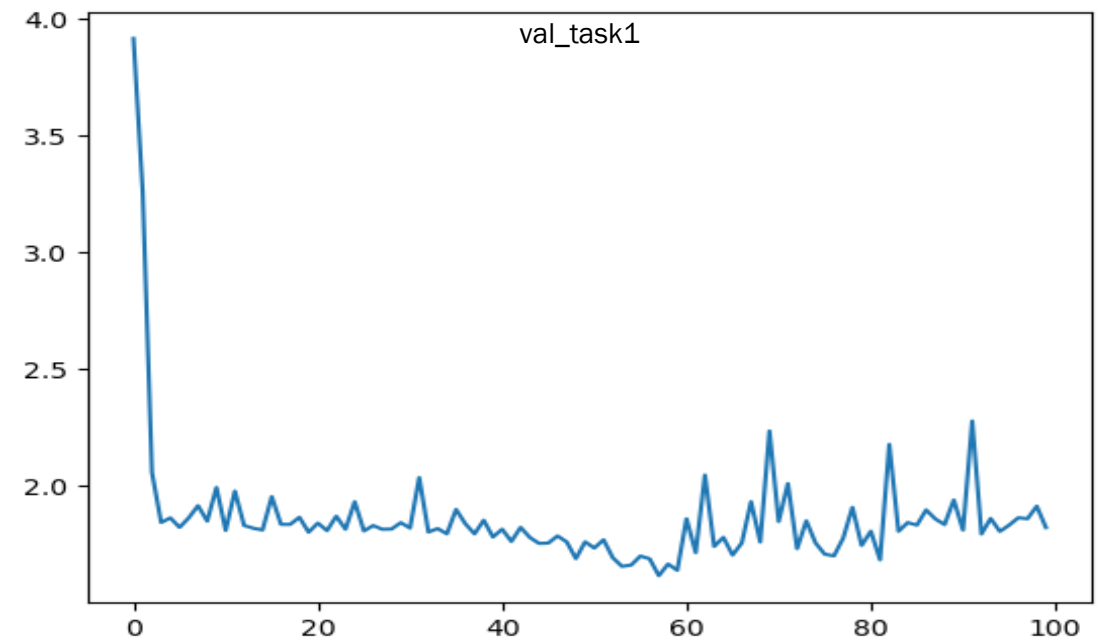
Next Steps & Open questions

■ Debug our implementations of EntNet and LSTM

■ Prepare pipeline for additional datasets (partly done)

■ Train and evaluate models on Penn Treebank and Wikitext-103

■ Either:

    – *Adapt the networks using own improvements for those tasks*

    – *Integrate Uncertainty Awareness*

■ Does anyone here have a successful remote debugging pipeline?