Machine Reading and Reasoning with Neural Program Interpreters

Sebastian Riedel

@riedelcastro
Collaborators

Johannes Welbl (UCL)

Pontus Stenetorp (UCL)

Matko Bosnjak (UCL)

Tim Rocktäschel (now at Oxford)

Jason Naradowsky (Johns Hopkins University)
“Should we separate meaning from language?” [Maybe not?]

Chris Manning
@AKBC 2013
End-to-End Reading and Comprehension

(Hermann et. al 2015, Seo et al., 2016, Rajpurkar et al., 2016, Weissenborn 2016…)

[Language] → ![Diagram] → [Information Need]
In January 1880, two of Tesla’s uncles put together enough money to help him leave Gospić for Prague where he was to study. Unfortunately, he arrived too late to enroll at Charles-Ferdinand University; he never studied Greek, a required subject; and he was illiterate in Czech, another required subject. Tesla did, however, attend lectures at the university, although, as an auditor, he did not receive grades for the courses.

What city did Tesla move to in 1880?

Prague

Why was he unable to enroll at the university?

arrived too late to enroll
How to read and reason?
Which medical specialty deals with pituitary ACTH hypersecretion?

Endocrinology

**Pituitary ACTH hypersecretion** is a form of hyperpituitarism characterized by an abnormally high level of ACTH produced by the **anterior pituitary**.

A major organ of the **endocrine system**, the **anterior pituitary** is the glandular, anterior lobe that...

The endocrine system is ... ... The field of study dealing with the **endocrine system** and its disorders is **endocrinology**, a branch of internal medicine.
How many pictures were in each of the albums?

Isabel uploaded 2 pictures from her phone and 4 from her camera to facebook. She sorted the pics into 3 different albums with the same amount of pics in each album.
Can we learn this end-to-end?
Part 1: A Read & Reason Dataset
What is the nationality of Jamie Burnett?

- Scotland (correct)
- China (incorrect)
- …

Jamie Burnett (born 16 September 1975) is a professional snooker player from Hamilton, South Lanarkshire…He began the 2014/2015 season with a quarter-final showing at the Yixing Open…

…Hamilton is a town in South Lanarkshire, in the central Lowlands of Scotland…

The Yixing Open was a professional minor-ranking snooker tournament that took place at the Yixing Sports Centre in Yixing, China.
Dataset Construction Method

Unlabelled Text

Knowledge Base

Dataset Construction Method

Multihop Dataset

[Conditionally Accepted to TACL]
What is the nationality of Jamie Burnett?

KB Triple: Jamie Burnett, citizenship, Scotland

Entities: Jamie Burnett, Scotland, Yixing Open, China

Documents:
- Jamie Burnett
- Hamilton (described in)
- Scotland (mentioned)
- Yixing Open (described in)
- China (mentioned)

Instance: Scottland

Correct Answer: Scotland
Dataset Construction Method

Unlabelled Text → Dataset Construction Method → Multihop Dataset

Knowledge Base
Dataset Construction Method

WikiHop

Unlabelled Text

Knowledge Base

Dataset Construction Method
Dataset Construction Method

Unlabelled Text

Knowledge Base

PubMed

DrugBank

Dataset Construction Method

MedHop
Baseline Results

Accuracy [%]

WikiHop
- Random: 11.5
- Max-Mention: 10.6
- TF-IDF: 25.6
- BiDAF: 42.9

MedHop
- Random: 13.9
- Max-Mention: 9.5
- TF-IDF: 9
- BiDAF: 47.8
Which medical specialty deals with pituitary ACTH hypersecretion?

Endocrinology

Pituitary ACTH hypersecretion ... is a form of hyperpituitarism characterized by an abnormally high level of ACTH produced by the anterior pituitary ...

A major organ of the endocrine system, the anterior pituitary is the glandular, anterior lobe that ...

The endocrine system is ... ... The field of study dealing with the endocrine system and its disorders is endocrinology, a branch of internal medicine.
Reduction to Traditional Machine Comprehension

Which medical specialty deals with pituitary ACTH hypersecretion?

Endocrinology

Pituitary ACTH hypersecretion ... is a form of hyperpituitarism characterized by an abnormally high level of ACTH produced by the anterior pituitary ...

A major organ of the endocrine system, the anterior pituitary is the glandular, anterior lobe that ...

The endocrine system is ... ... The field of study dealing with the endocrine system and its disorders is endocrinology, a branch of internal medicine.
A major organ of the endocrine system, the anterior pituitary is the glandular, anterior lobe that ...

The endocrine system is ... ... The field of study dealing with the endocrine system and its disorders is **endocrinology**, a branch of internal medicine.
Removing Relevant Documents, Keep Answer Documents

Accuracy [%]

WikiHop

MedHop

BIDAF

BIDAF doc-rem

54.5
44.6
33.7
30.4

10%
Part 2: Learning to Read and Calculate
Isabel uploaded 2 pictures from her phone and 4 from her camera to Facebook. She sorted the pics into 3 different albums with the same amount of pics in each album.
Differentiable Program Interpreters

How many pictures were in each of the albums?

Isabel uploaded 2 pictures from her phone and 4 from her camera to Facebook. She sorted the pics into 3 different albums with the same amount of pics in each album.

Bosnjak et al. ICML 2017
How many pictures were in each of the albums?

Isabel uploaded 2 pictures from her phone and 4 from her camera to Facebook. She sorted the pics into 3 different albums with the same amount of pics in each album.
Isabel uploaded 2 pictures from her phone and 4 from her camera to Facebook. She sorted the pics into 3 different albums with the same amount of pics in each album.

How many pictures were in each of the albums?
How many pictures were in each of the albums?

Isabel uploaded 2 pictures from her phone and 4 from her camera to Facebook. She sorted the pics into 3 different albums with the same amount of pics in each album.
How many pictures were in each of the albums?

Isabel uploaded 2 pictures from her phone and 4 from her camera to Facebook. She sorted the pics into 3 different albums with the same amount of pics in each album.
How many pictures were in each of the albums?

Isabel uploaded 2 pictures from her phone and 4 from her camera to Facebook. She sorted the pics into 3 different albums with the same amount of pics in each album.
How many pictures were in each of the albums?

Isabel uploaded 2 pictures from her phone and 4 from her camera to Facebook. She sorted the pics into 3 different albums with the same amount of pics in each album.
Isabel uploaded 2 pictures from her phone and 4 from her camera to facebook. She sorted the pics into 3 different albums with the same amount of pics in each album.

How many pictures were in each of the albums?
Training

How many pictures were in each of the albums?

Isabel uploaded 2 pictures from her phone and 4 from her camera to Facebook. She sorted the pics into 3 different albums with the same amount of pics in each album.
Results on Benchmark (Accuracy)

Roy & Roth (2015) | Ours | Seq2Seq
---|---|---
50 | 55.5 | 95
62.5 | 96 | 95
75 | 87.5 | 96
87.5 | 75 |
100 | 0 |

Seq2Seq solves simpler problem

How many pictures were in each of the albums?

(2 + 4) / 3

Bosnjak et al. ICML 2017
Limitations

- Continuous relaxations difficult to train
  - Hard to learn long programs
  - Hard to learn with recursive function calls
- Need better gradients in presence of discrete variables
Part 3: Learning to Aggregate
Which medical specialty deals with pituitary ACTH hypersecretion?

Pituitary ACTH hypersecretion ... is a form of hyperpituitarism characterized by an abnormally high level of ACTH produced by the anterior pituitary ...

A major organ of the endocrine system, the anterior pituitary is the glandular, anterior lobe that ...

The endocrine system is ... ... The field of study dealing with the endocrine system and its disorders is endocrinology, a branch of internal medicine.
Pituitary ACTH hypersecretion ... is a form of hyperpituitarism characterized by an abnormally high level of ACTH produced by the anterior pituitary ...

A major organ of the endocrine system, the anterior pituitary is the glandular, anterior lobe that ...

The endocrine system is ... ... The field of study dealing with the endocrine system and its disorders is endocrinology, a branch of internal medicine.

Which medical specialty deals with pituitary ACTH hypersecretion?
Neural Theorem Provers

Which medical specialty deals with pituitary ACTH hypersecretion?

Pituitary ACTH hypersecretion ... is a form of hyperpituitarism characterized by an abnormally high level of ACTH produced by the anterior pituitary ...

A major organ of the endocrine system, the anterior pituitary is the glandular, anterior lobe that ...

The endocrine system is ... ... The field of study dealing with the endocrine system and its disorders is endocrinology, a branch of internal medicine.

Which medical specialty deals with pituitary ACTH hypersecretion?

What is ACTH hypersecretion created by?

Rocktäschel and Riedel NIPS 2017
Pituitary ACTH hypersecretion is a form of hyperpituitarism characterized by an abnormally high level of ACTH produced by the anterior pituitary.

A major organ of the endocrine system, the anterior pituitary is the glandular, anterior lobe that...

The endocrine system is... The field of study dealing with the endocrine system and its disorders is endocrinology, a branch of internal medicine.
Pituitary ACTH hypersecretion is a form of hyperpituitarism characterized by an abnormally high level of ACTH produced by the anterior pituitary gland.

A major organ of the endocrine system, the anterior pituitary is the glandular, anterior lobe that... The endocrine system is... The field of study dealing with the endocrine system and its disorders is endocrinology, a branch of internal medicine.

The field of study dealing with the endocrine system and its disorders is endocrinology, a branch of internal medicine.
Pituitary ACTH hypersecretion is a form of hyperpituitarism characterized by an abnormally high level of ACTH produced by the anterior pituitary.

A major organ of the endocrine system, the anterior pituitary is the glandular, anterior lobe that...

The endocrine system is ... ... The field of study dealing with the endocrine system and its disorders is endocrinology, a branch of internal medicine.
Rocktäschel and Riedel

Question: Which specialty deals with pituitary ACTH hypersecretion?

Question: What is the pituitary a part of?

Agent

Program Interpreter

Question

Reader

X deals with Y if
Y produced by Z
Z is a part of U
X deals with U

dealsWith(hypersecretion, X)

dealsWith(endocrine system, endocrinology)

covers(endocrine system, X)

createdBy(hypersecretion, anterior pituitary)

partOf(endocrine system, anterior pituitary)
Catch: Currently only Works on Relational Data

Amounts to Differentiable Version of the Backward Chaining algorithm used in Prolog

Rocktäschel and Riedel NIPS 2017
supports soft unification

dealsWith(hypersecretion, X)

createdBy(hypersecretion, anterior pituitary)

partOf(endocrine system, anterior pituitary)

dealsWith(endocrine system, endocrinology)

岩塚慶和 and Riedel NIPS 2017
Results on Benchmark (Rank of Correct Answer)

Comparable or Better than Baselines, and **interpretable**
Interpretability: Learnt Rules

- if $X$ is located in $Y$ and $Y$ is located in $Z$ then $X$ is located in $Z$
- if $X$ expels diplomats of $Y$ then $X$ shows negative behaviour towards $Y$
- if $X$ interacts with $Y$ and $Y$ interacts with $Z$ then $X$ interacts with $Z$
Limitations

- Scalability
  - Currently only works for KBs with < 10k facts
  - Small proof depth
- Still requires relational representation
Conclusion

- Great Progress in End-to-End Reading Comprehension

- **Reasoning** (aggregation, calculation etc.) end-to-end is still very challenging

- **Our Approaches**
  - create datasets
  - cast reasoning as **program learning** and **execution**
  - are **end-to-end differentiable** (can be trained on downstream loss)
  - are inspired and tied to **traditional symbolic formalisms** (Forth, Prolog/Datalog)
  - are learnt models are **interpretable**
  - allow injection of **prior knowledge**