

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)



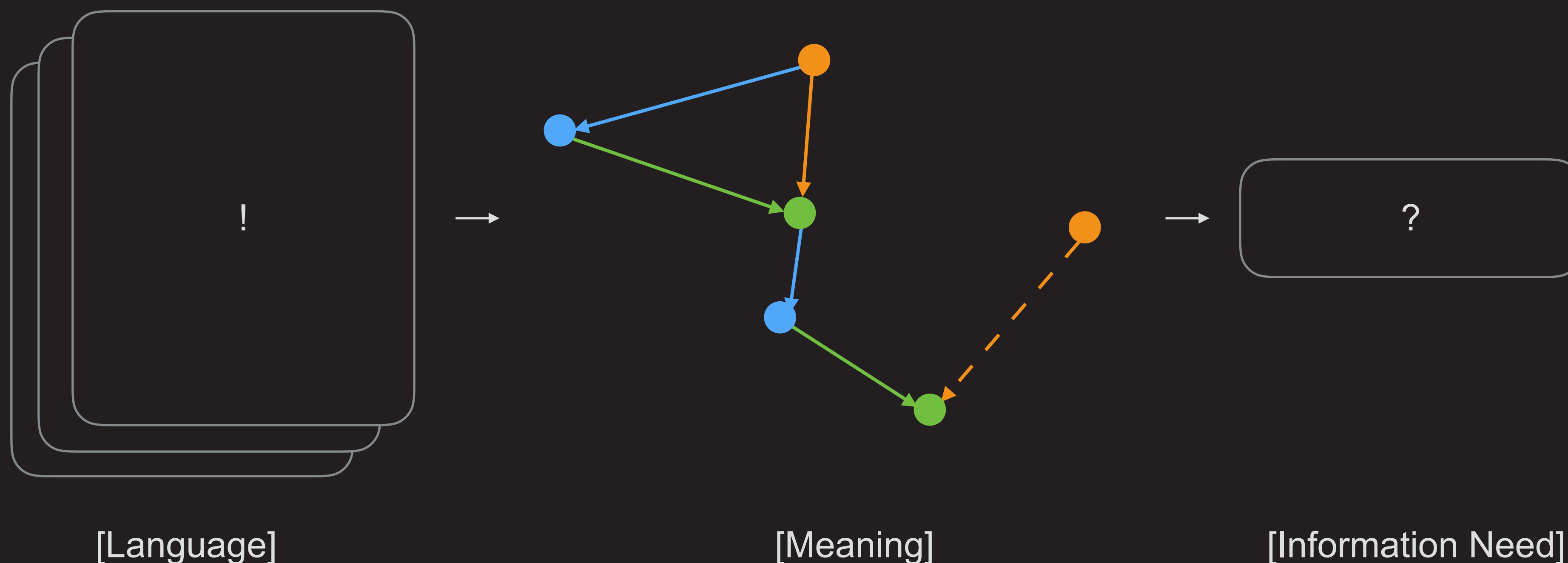
Jason Naradowsky

(Johns Hopkins University)



“Should we separate meaning from language?” **[Maybe not?]**

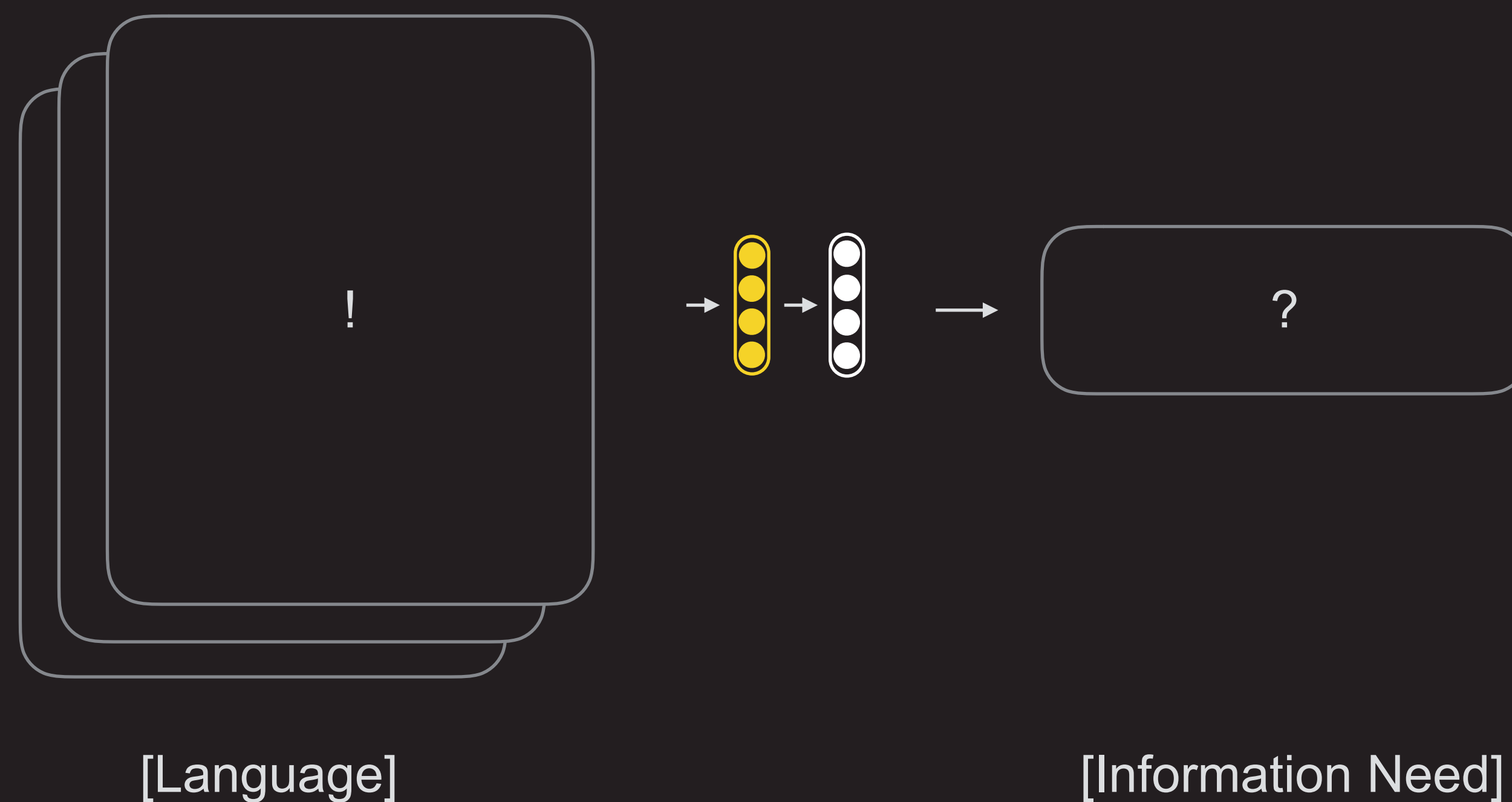
Chris Manning
@AKBC 2013





End-to-End Reading and Comprehension

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

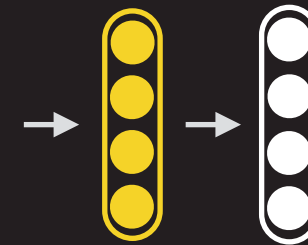


Machine Reading



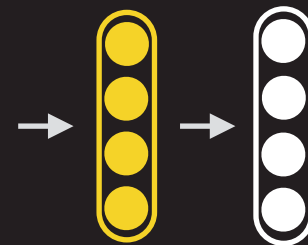
What city did Tesla move to in 1880?

Prague



Why was he unable to enroll at the university?

arrived too late to enroll



Nicola Tesla

...

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.

How to read and reason?

Machine Reading and Reasoning



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.

Machine Reading and Reasoning



How many pictures were in each of the albums?

2



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**?



Johannes Welbl



Pontus Stenetorp

Part 1: A Read & Reason Dataset



QAngaroo

Reading Comprehension with Multiple Hops

 Paper

 Dataset Download

Dataset Distribution under [CC BY-SA 3.0](#)

A Single Instance

What is the nationality of Jamie Burnett?

► Candidates:

► 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



Dataset Construction Method

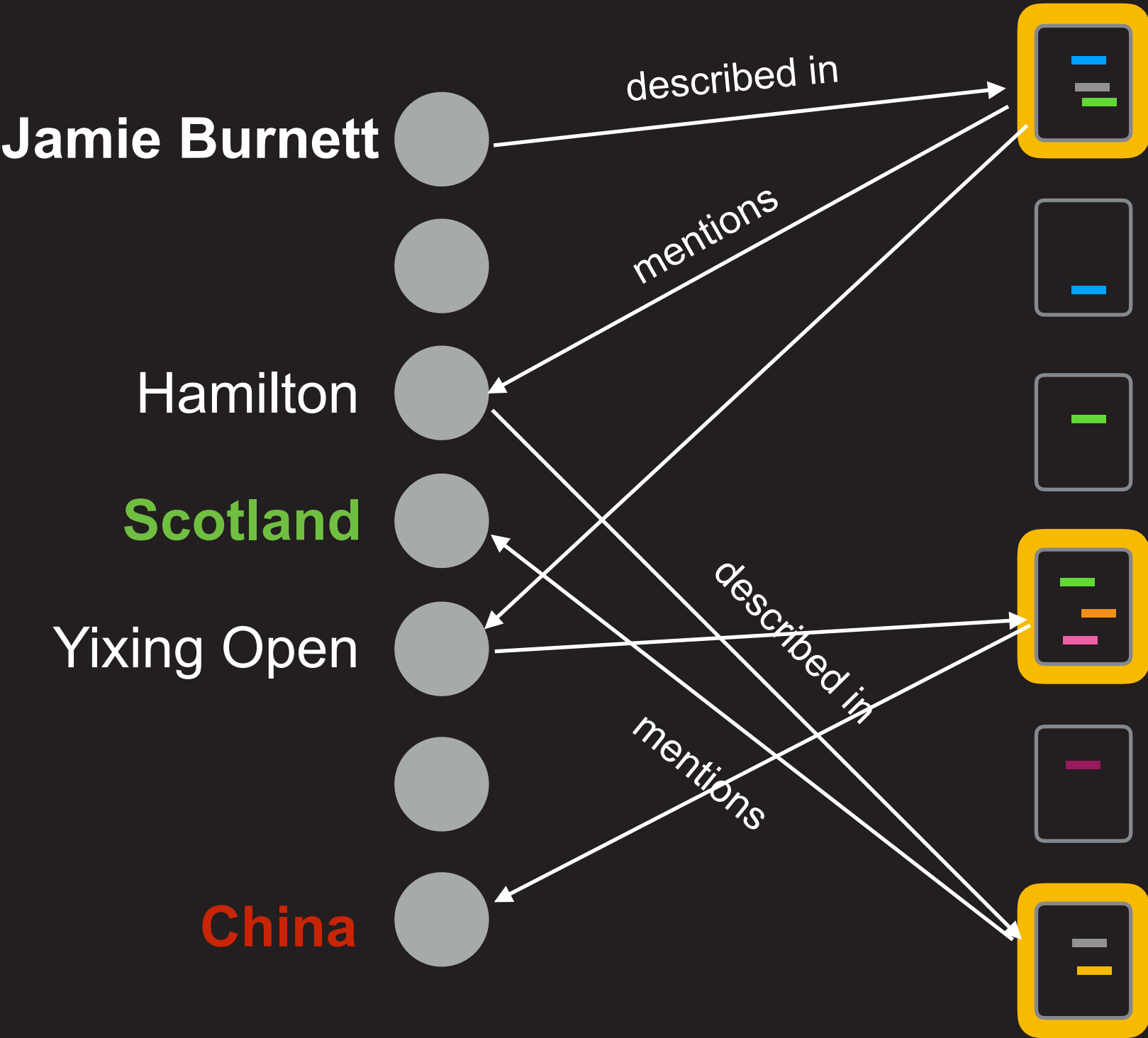
KB Triple

Entities

Documents

Instance

Jamie Burnett, citizenship, Scotland



What is the nationality of Jamie Burnett?

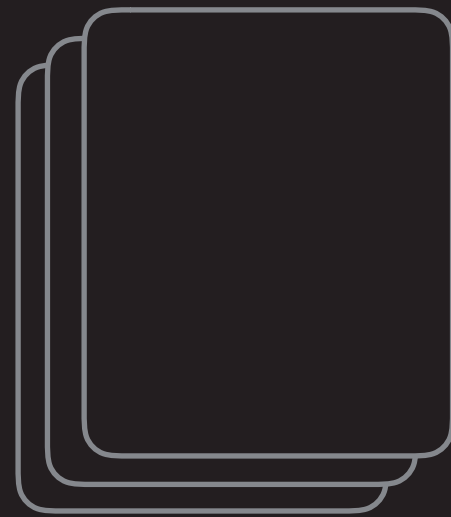
Three empty boxes for inputting the answer.

Scotland
China

Dataset Construction Method



Dataset Construction Method

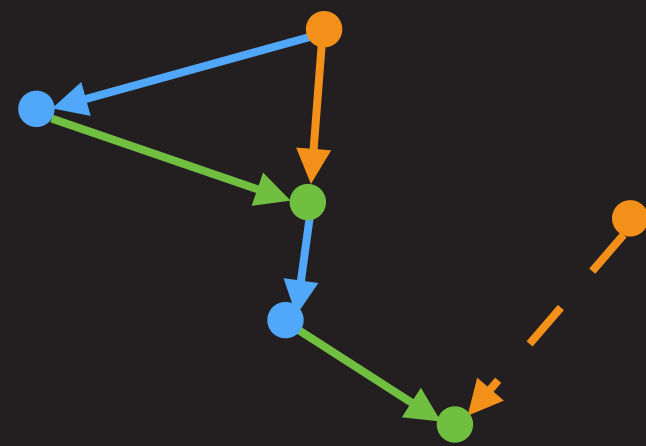


Unlabelled Text

Dataset Construction Method

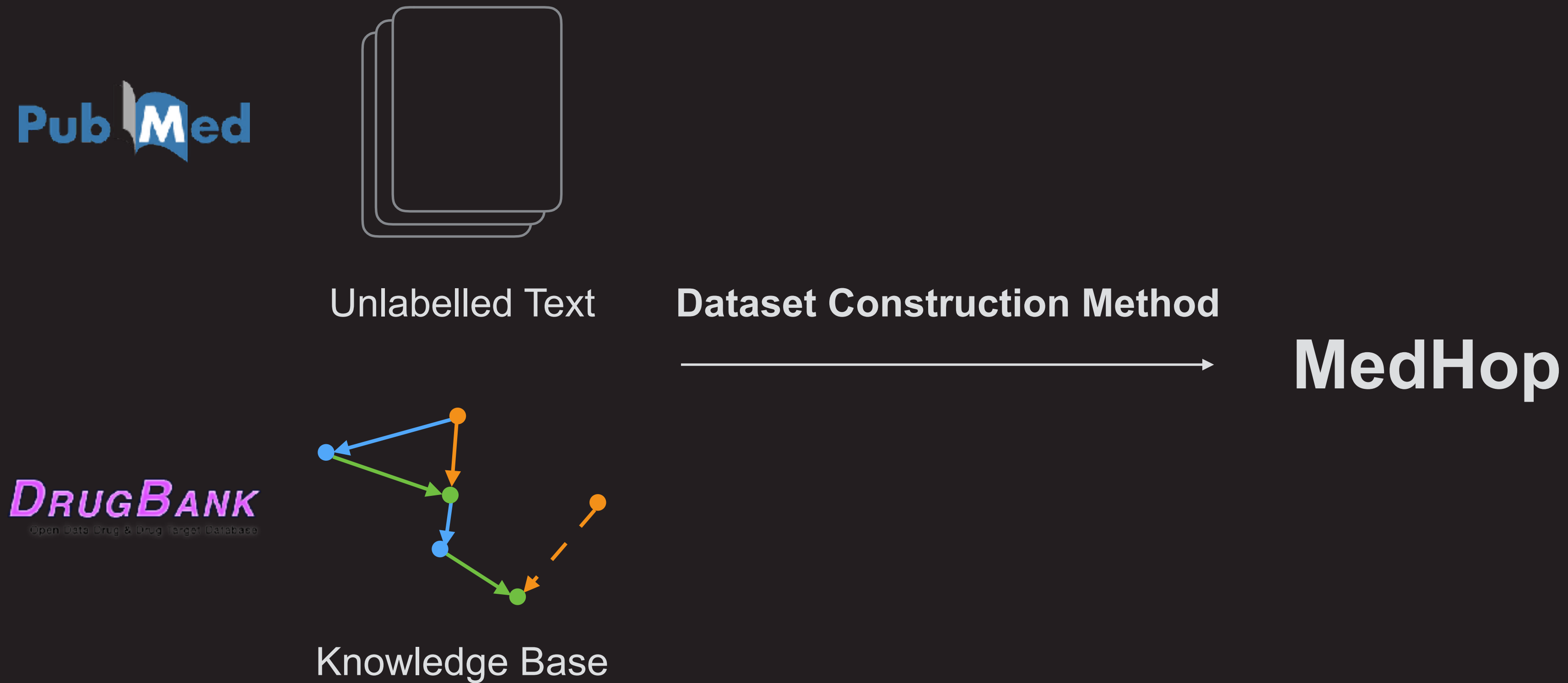


WikiHop

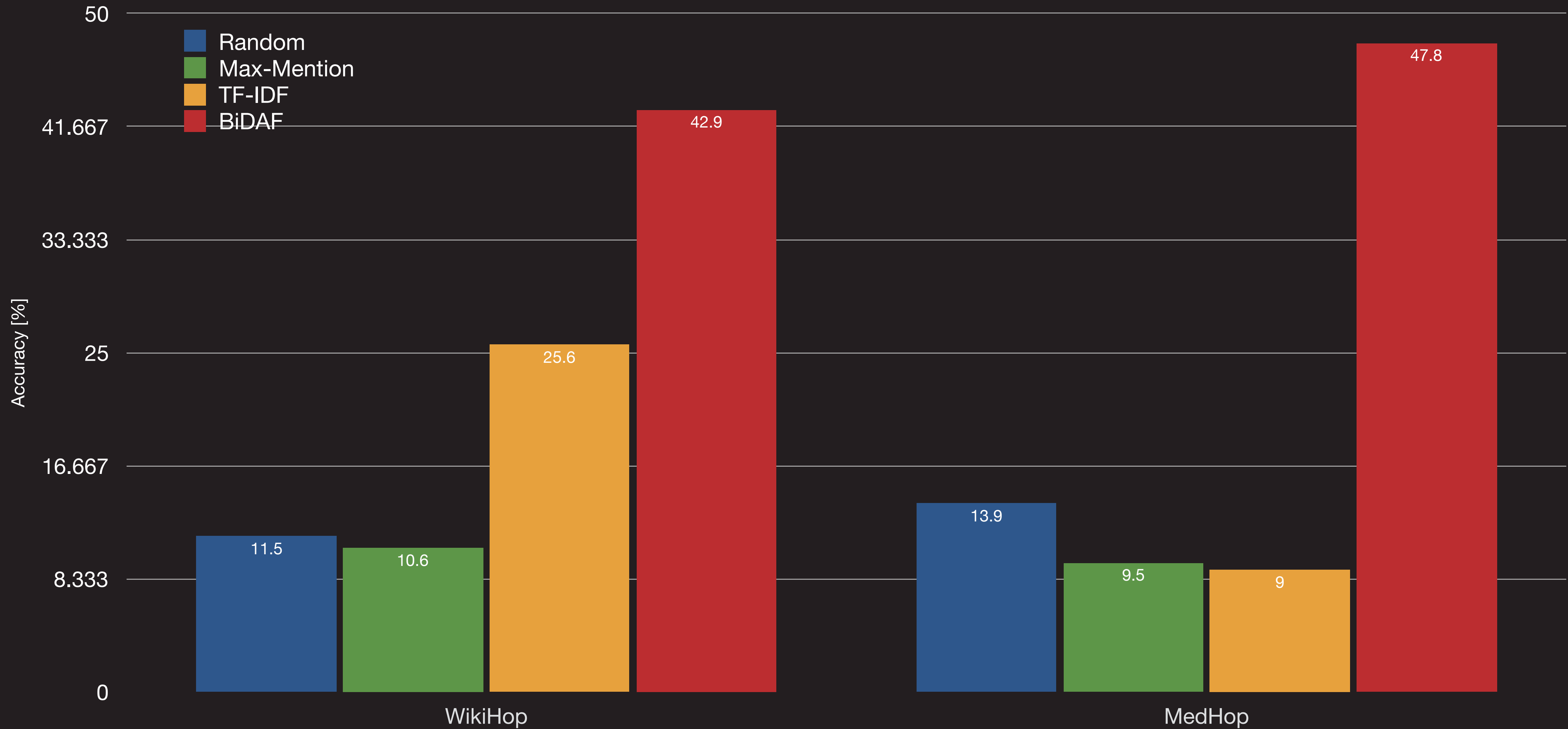


Knowledge Base

Dataset Construction Method



Baseline Results

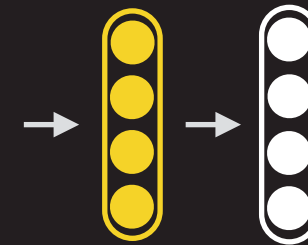


Reduction to Traditional Machine Comprehension



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 ...

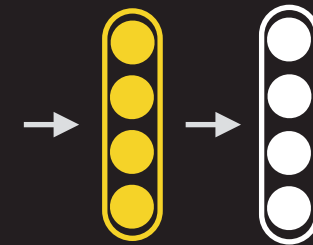
The endocrine system is ... The field of study dealing with the **endocrine system** and its disorders is **endocrinology**, a branch of internal medicine.

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 ...

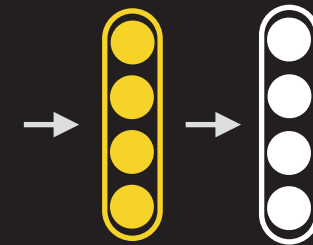
The endocrine system is ... The field of study dealing with the **endocrine system** and its disorders is **endocrinology**, a branch of internal medicine.

Does BiDaf Aggregate?



Which medical specialty deals with pituitary ACTH hypersecretion?

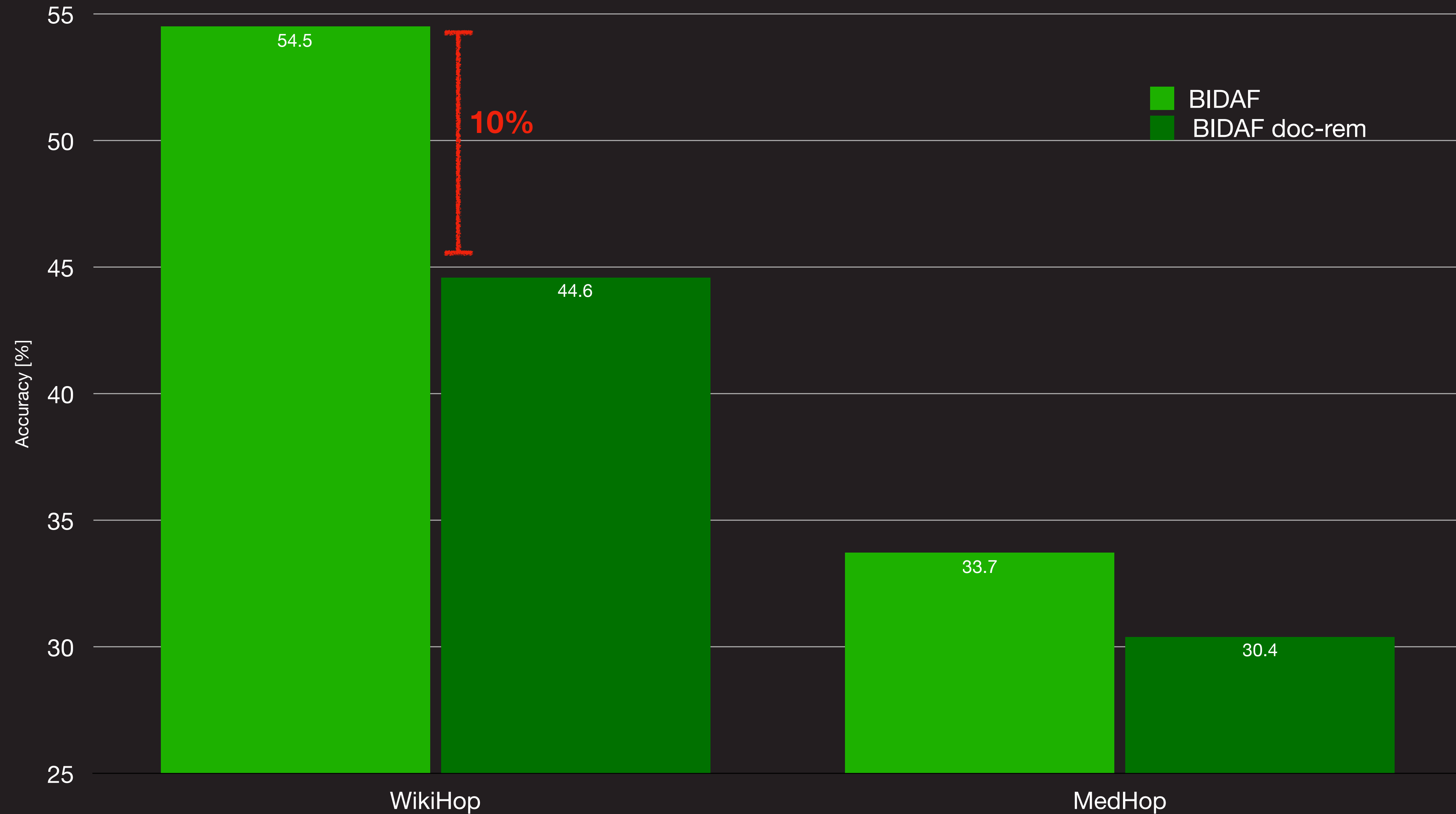
Endocrinology



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





Matko Bosnjak



Tim Rocktäschel



Jason Naradowsky

Part 2: Learning to Read and Calculate

Machine Reading and Reasoning: Math



How many pictures were in each of the albums?

2



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.

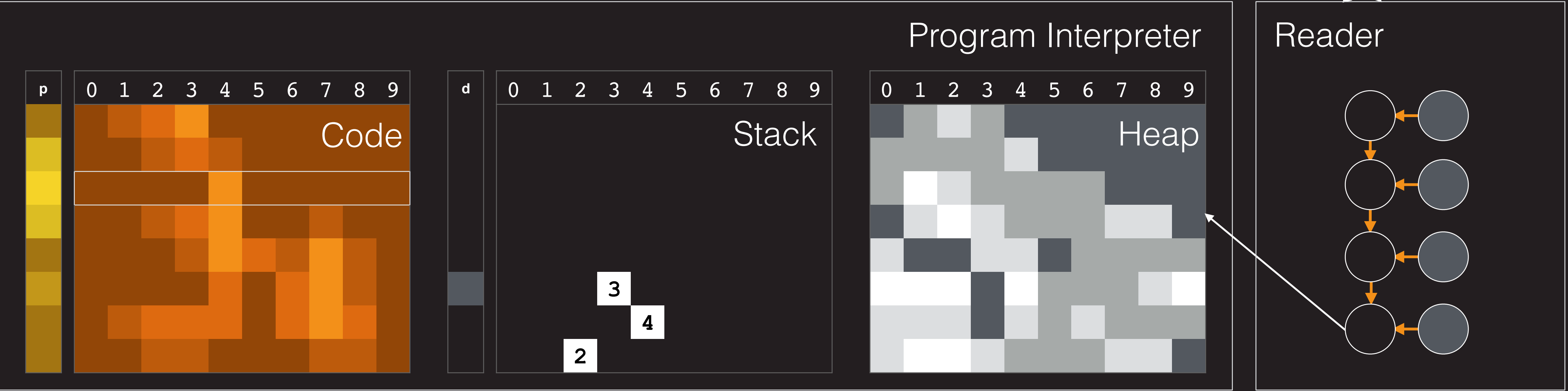
Differentiable Program Interpreters



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.

Model

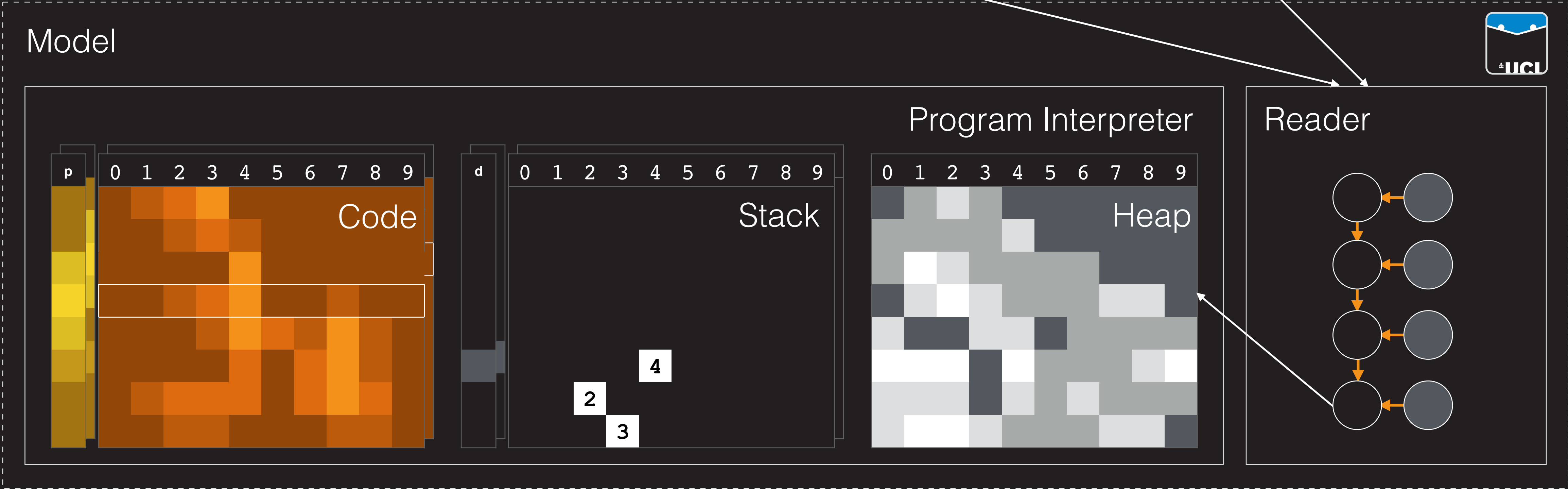


Differentiable Program Interpreters



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.



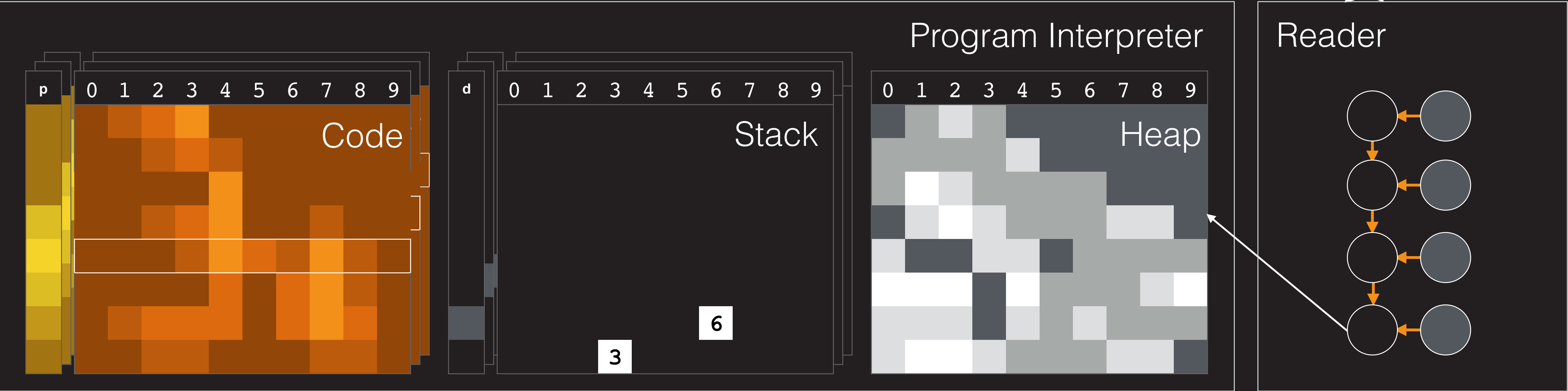
Differentiable Program Interpreters



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.

Model



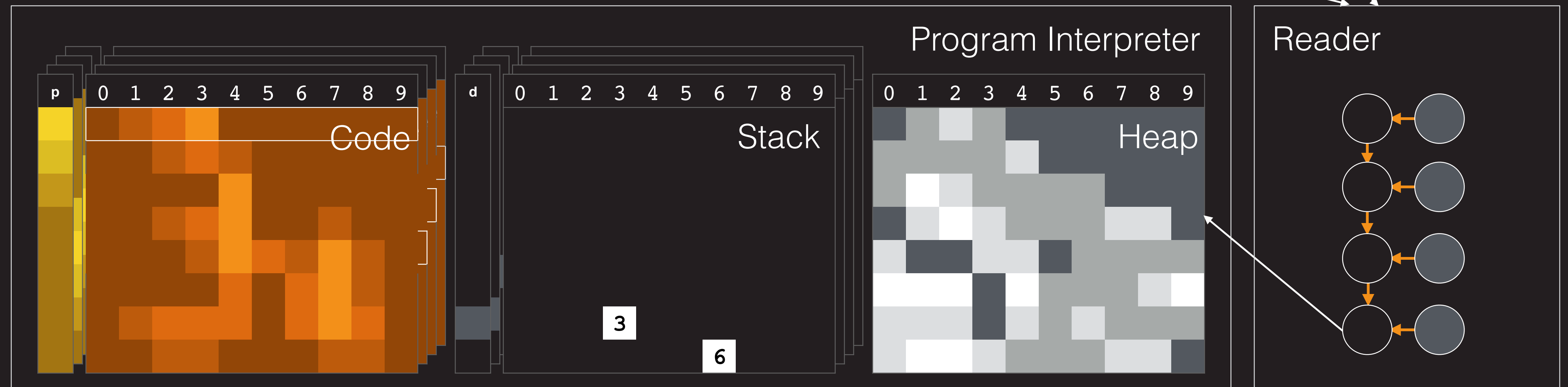
Differentiable Program Interpreters



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.

Model

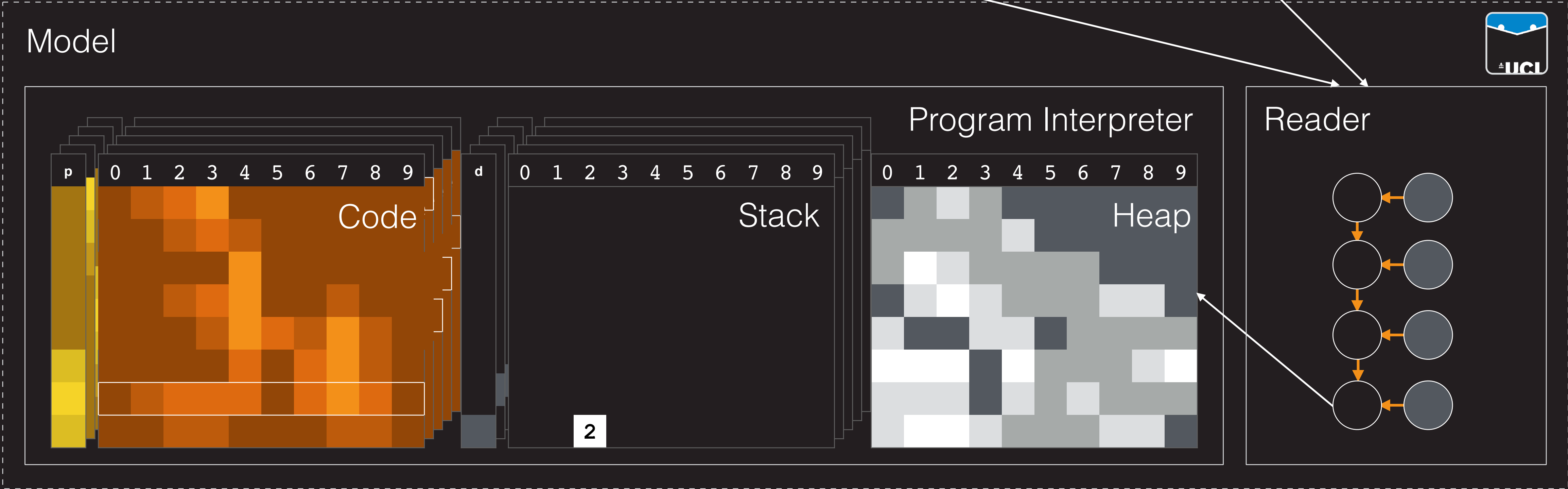


Differentiable Program Interpreters



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.



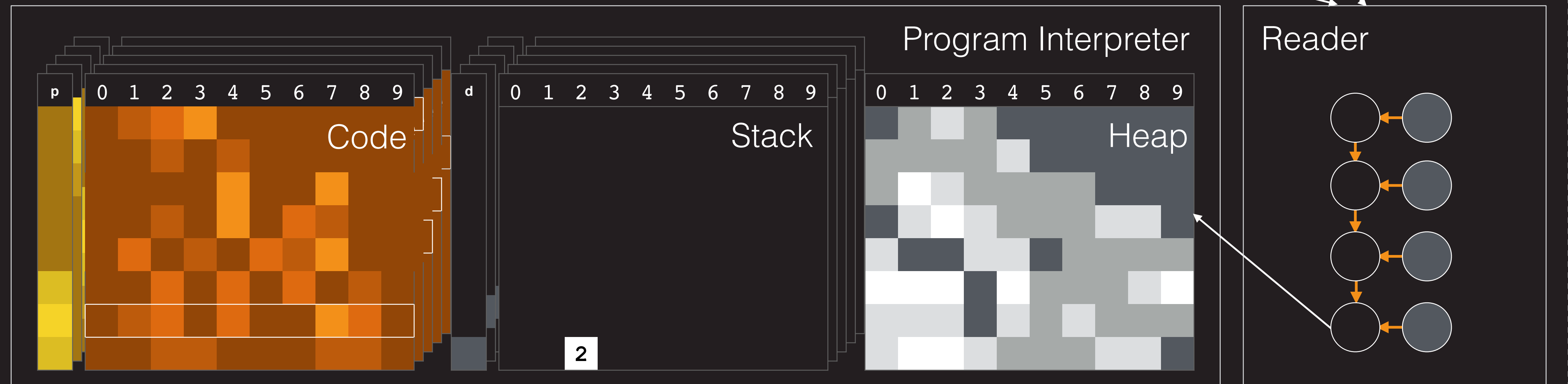
Training



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.

Model



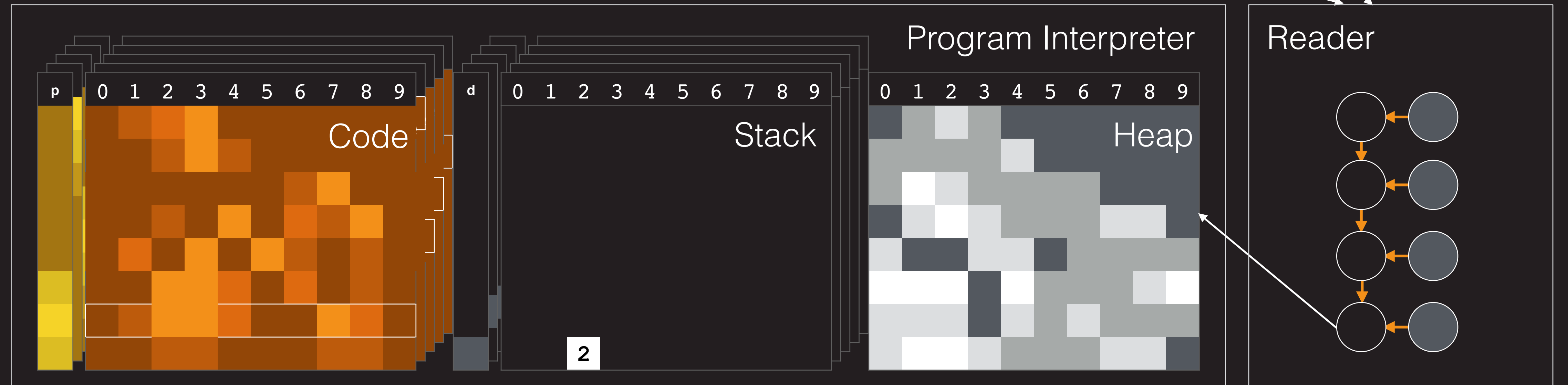
Training



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.

Model



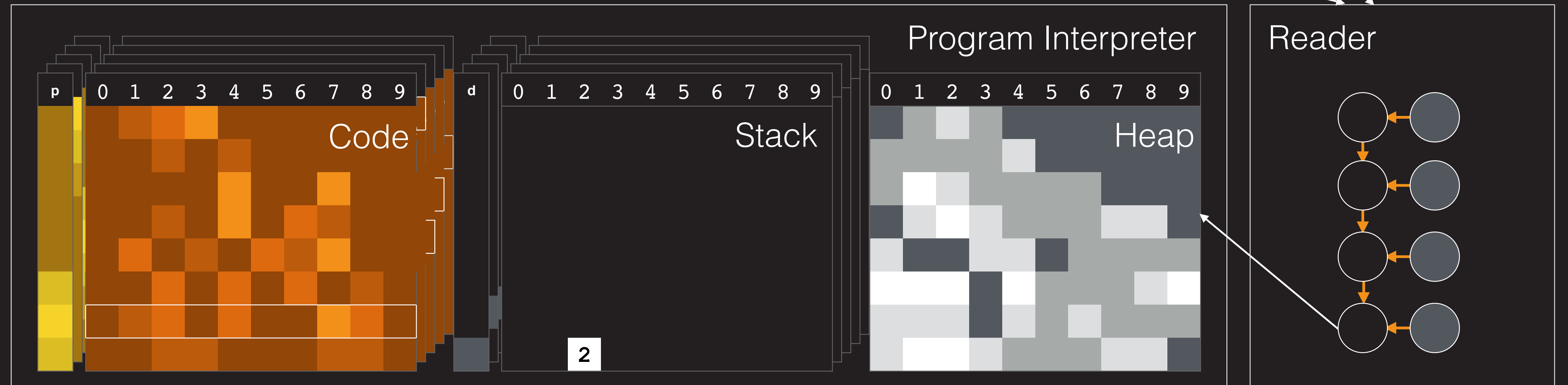
Training



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.

Model



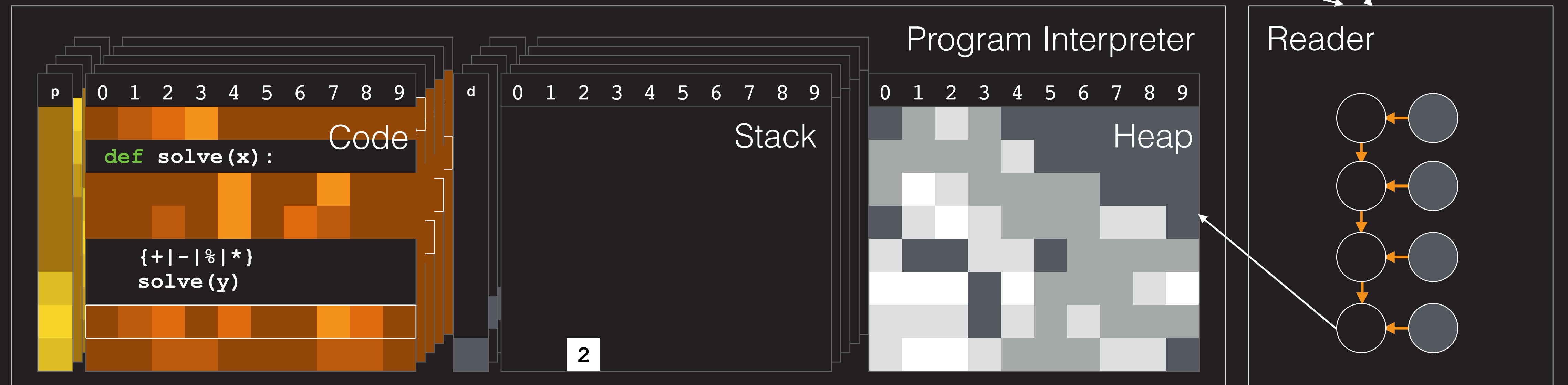
Training



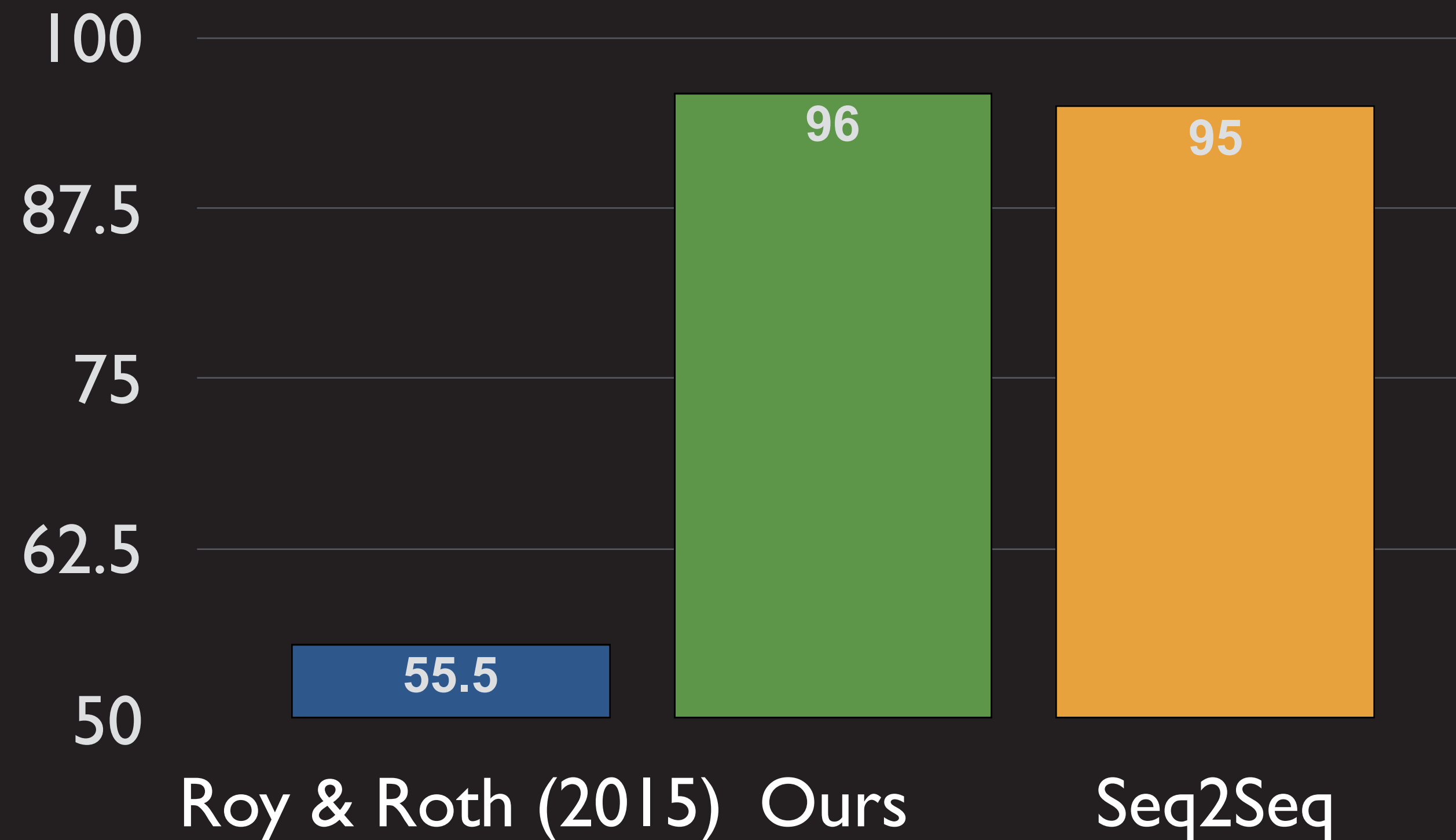
How many pictures were in each of the albums?

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.

Model



Results on Benchmark (Accuracy)



Seq2Seq solves simpler problem



How many pictures were in each of the albums?

$(2 + 4) / 3$

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



Tim Rocktäschel

Part 3: Learning to Aggregate

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.

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.

Agent



Program Interpreter

Code

Question

Which specialty deals with pituitary ACTH hypersecretion?

Reader

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** ...

Agent



Program Interpreter



Question

What is ACTH hypersecretion created by?

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.

Reader

Neural Theorem Provers



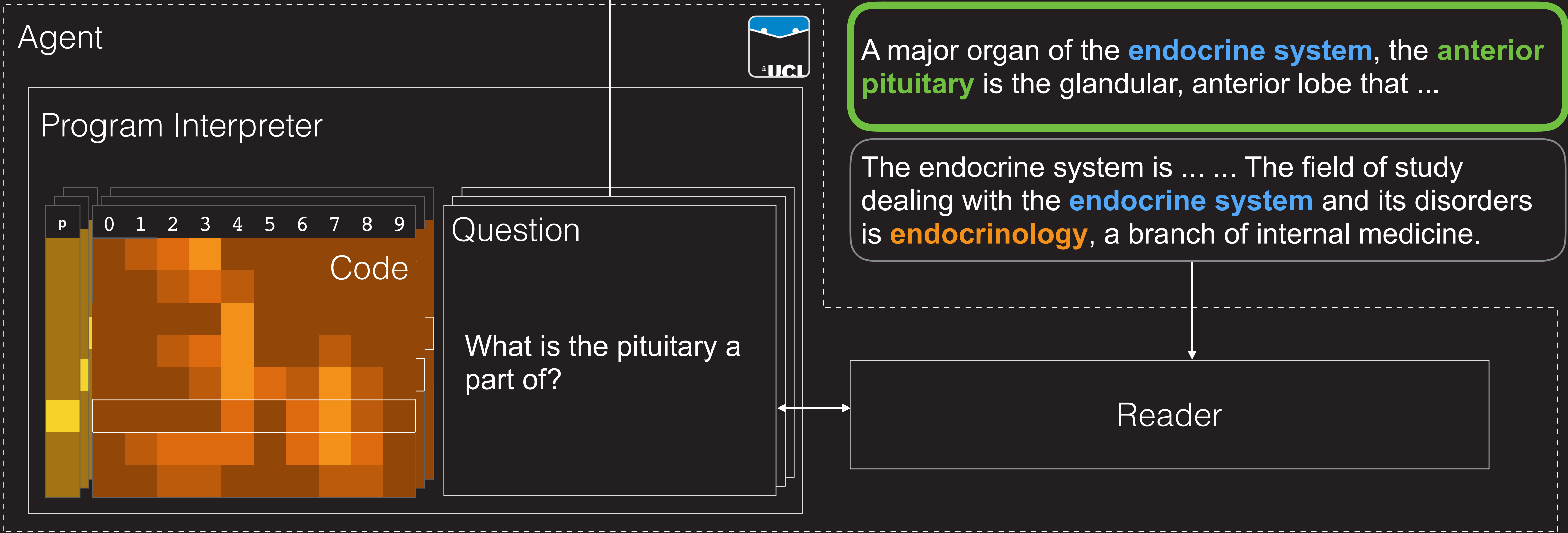
Which medical specialty deals with pituitary ACTH hypersecretion?



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.



Neural Theorem Provers



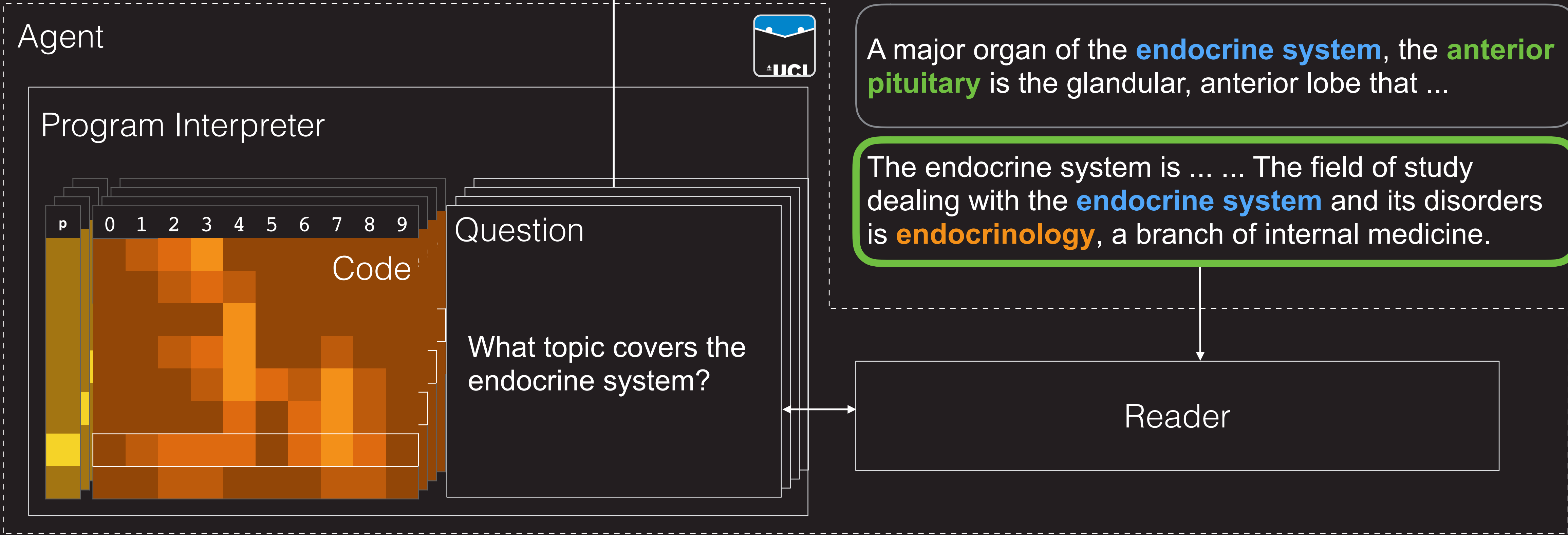
Which medical specialty deals with pituitary ACTH hypersecretion?



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.



Vectors Correspond to Interpretable Rules



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.

Agent



Program Interpreter



Question

What topic covers the endocrine system?

Reader

Catch: Currently only Works on Relational Data



dealsWith(**hypersecretion**, X)



createdBy(**hypersecretion, anterior pituitary**)

```
partOf(endocrine system, anterior pituitary)
```

```
dealsWith(endocrine system, endocrinology)
```

Agent



Program Interpreter

Question

```
covers(endocrine
system, X)
```

Reader

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

Catch: Currently only Works on Relational Data



dealsWith(**hypersecretion**, X)

createdBy(**hypersecretion**, **anterior pituitary**)

partOf(**endocrine system**, **anterior pituitary**)

dealsWith(**endocrine system**, **endocrinology**)

Agent



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

Supports Soft Unification



dealsWith(hypersecretion, X)



createdBy(hypersecretion, anterior pituitary)

partOf(endocrine system, anterior pituitary)

dealsWith(endocrine system, endocrinology)

Reader

Agent

Program Interpreter

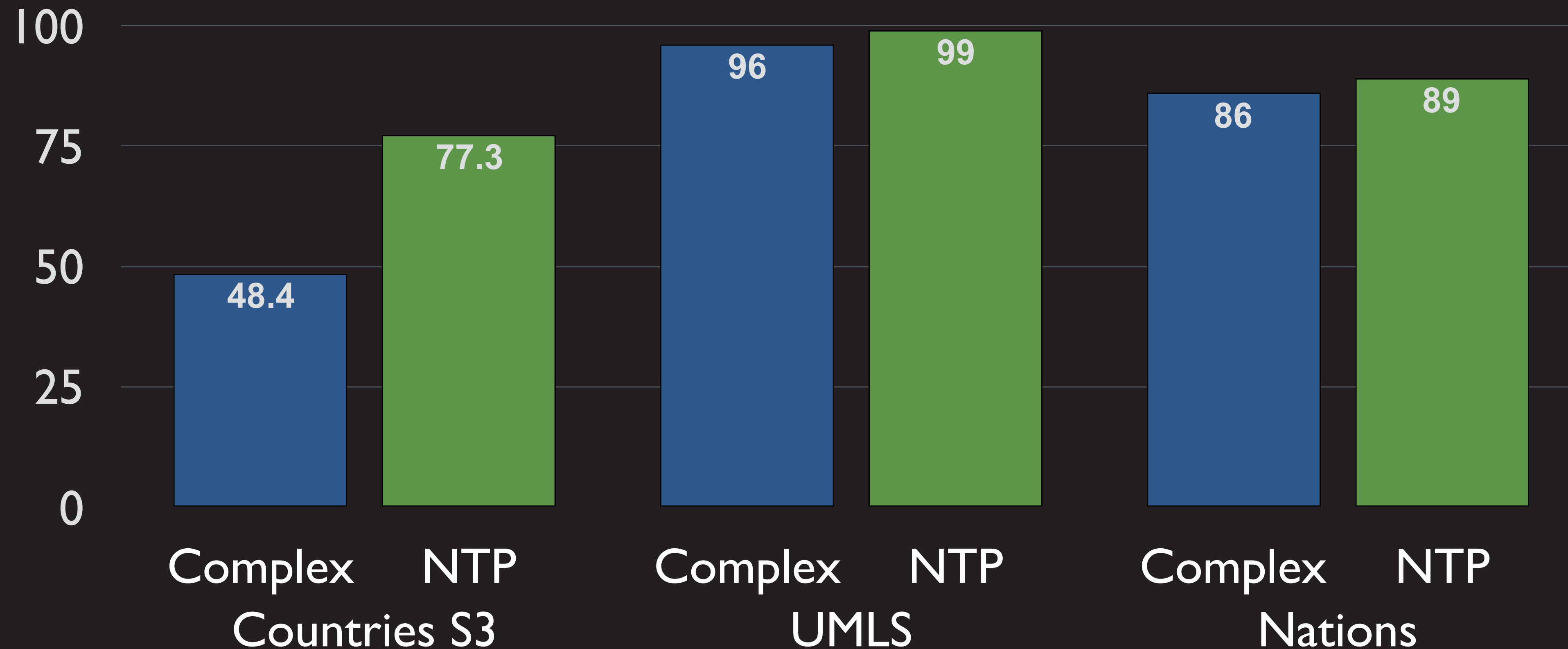
Question

covers(endocrine system, X)

Code

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

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**