Leveraging Structures for Machine Reasoning

Karan Samel

ksamel@gatech.edu

2/8/21

Motivation

How can we combine deep learning and logic reasoning to leverage existing model biases?

<u>Symbolic Execution</u>	<u>Neural + Symbolic</u>	<u>Neural Execution</u>
Interpretable		Comp. Efficient
Extensible	?	Data Hungry
Not Comp. Efficient		Not Interpretable

Research Directions

More specifically we will look at:

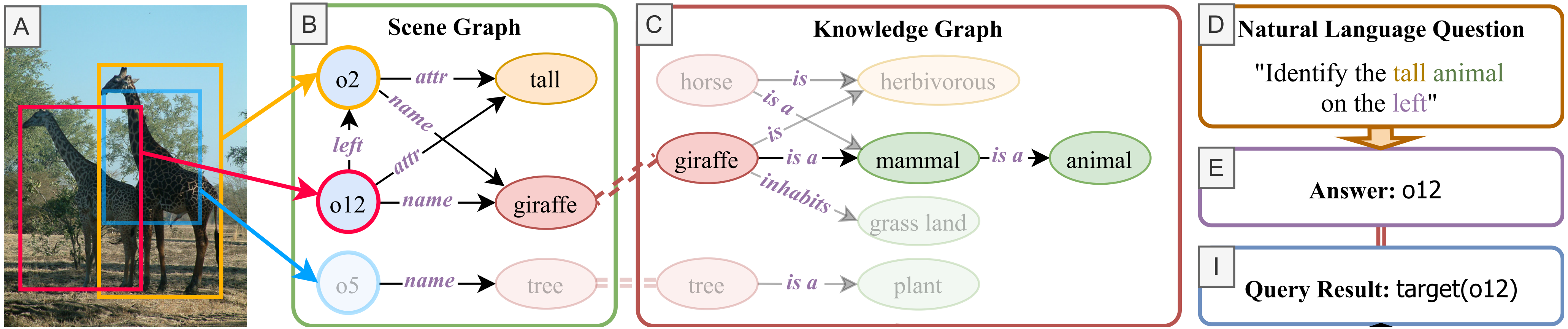
- Image Question Answering
    - How do we run real world inference and training given symbolic representations?
    - Given natural inputs, can we quantify and leverage the uncertainty in their underlying symbolic representations?
- Rule learning and inference over videos

Real world symbolic execution and training

(Mostly symbolic)

Real World Symbolic Execution - Example

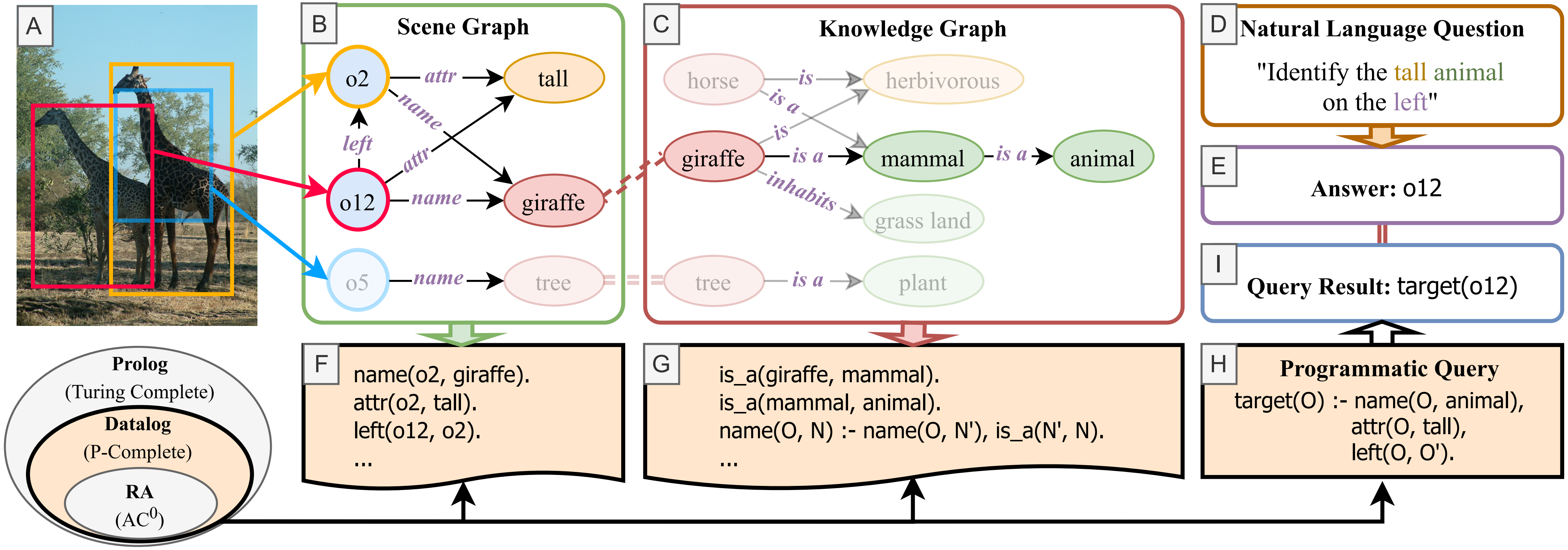
Visual question answering with common sense.



Huang, et al. Scallop: End-to-End Differentiable Reasoning at Scale. Under review at ICML'21

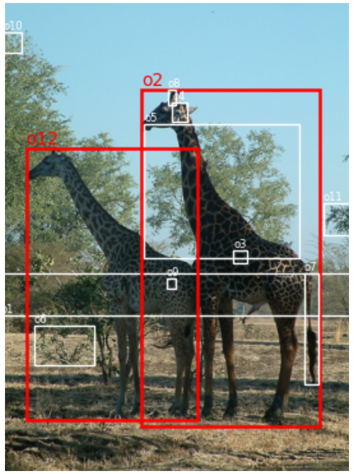
Real World Symbolic Execution - Example

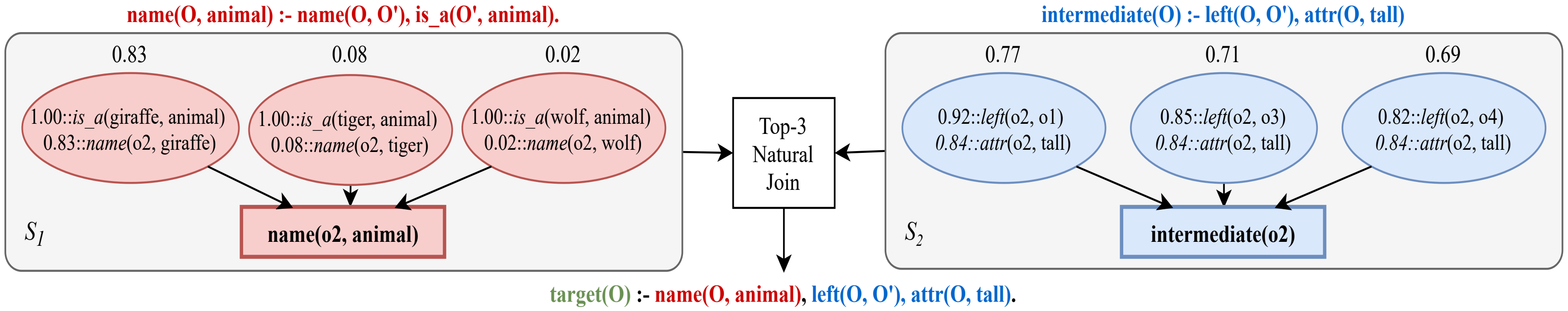
Logical programming execution



Real World Symbolic Execution - Issues

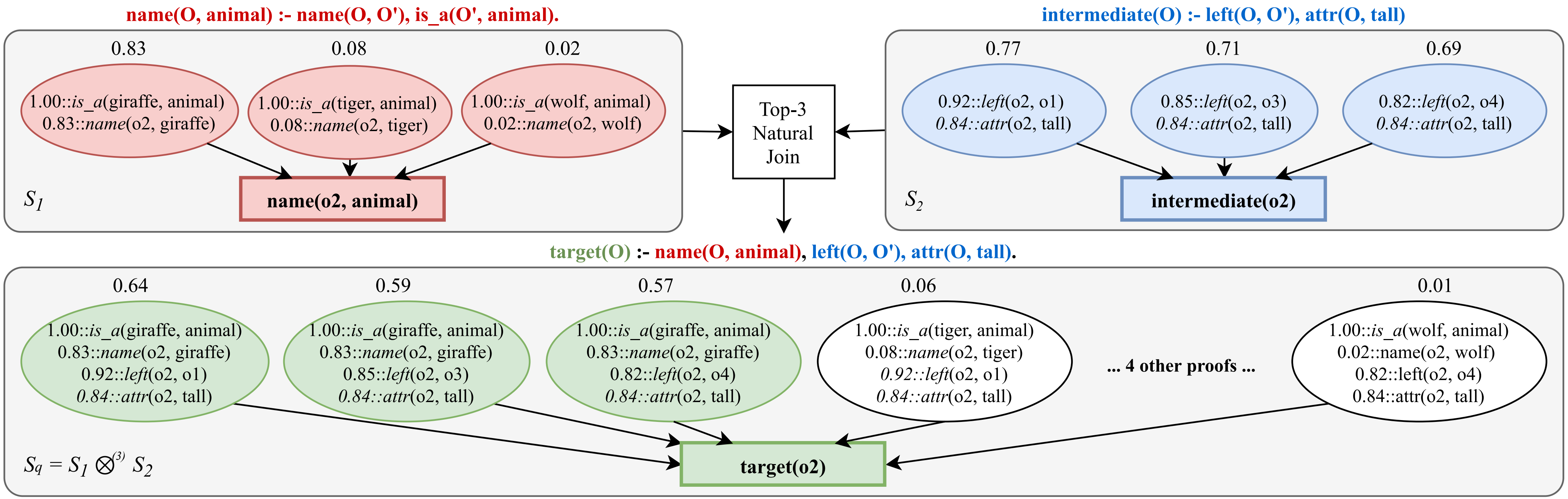
Probabilities for selecting an object conditioned on the query requires analyzing all possible combinations that satisfy the query.





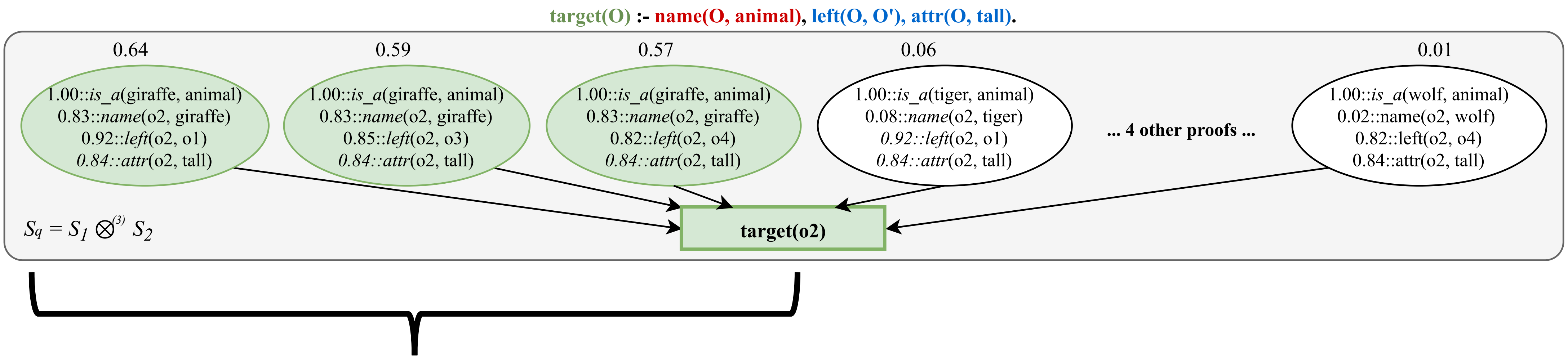
Real World Symbolic Execution - Issues

Probabilities for selecting an object conditioned on the query requires analyzing all possible combinations that satisfy the query.



Real World Symbolic Execution - Issues

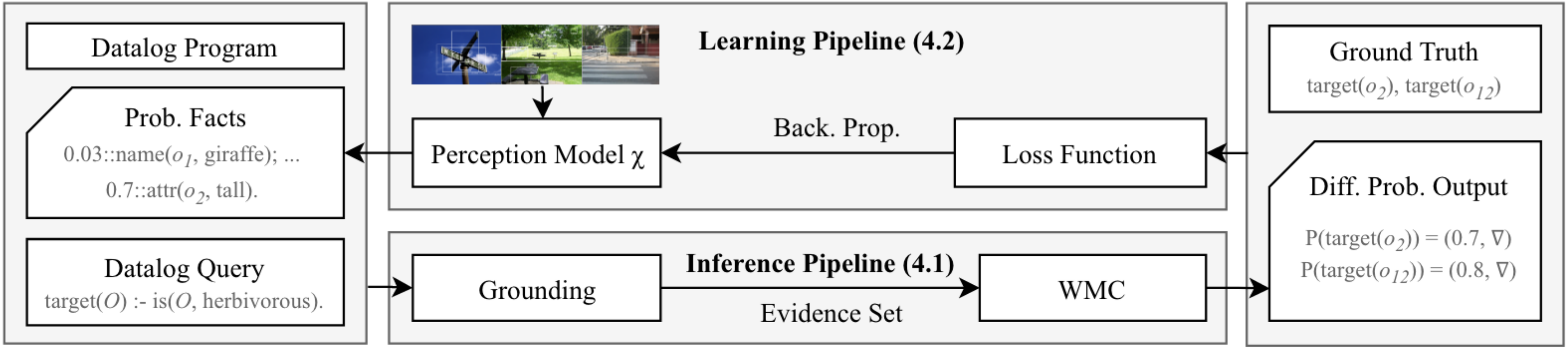
As the number of objects, knowledge, facts, and query length grows, this computation is exponential!



Intuition: take the top-k combinations at each step.

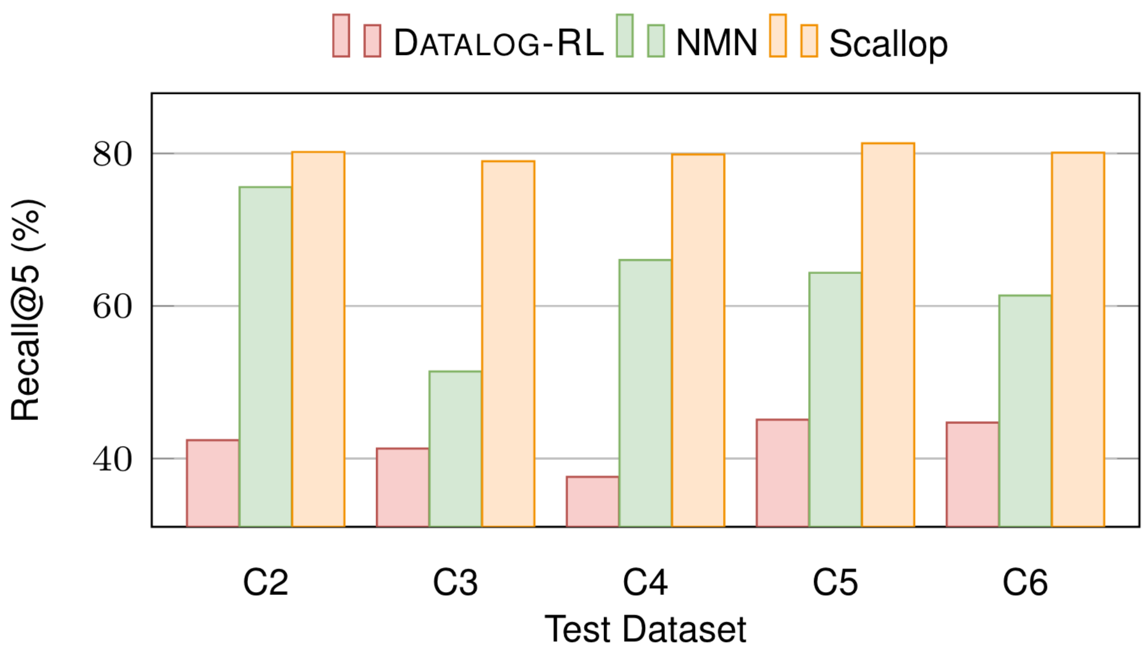
Real World Symbolic Execution - Scallop

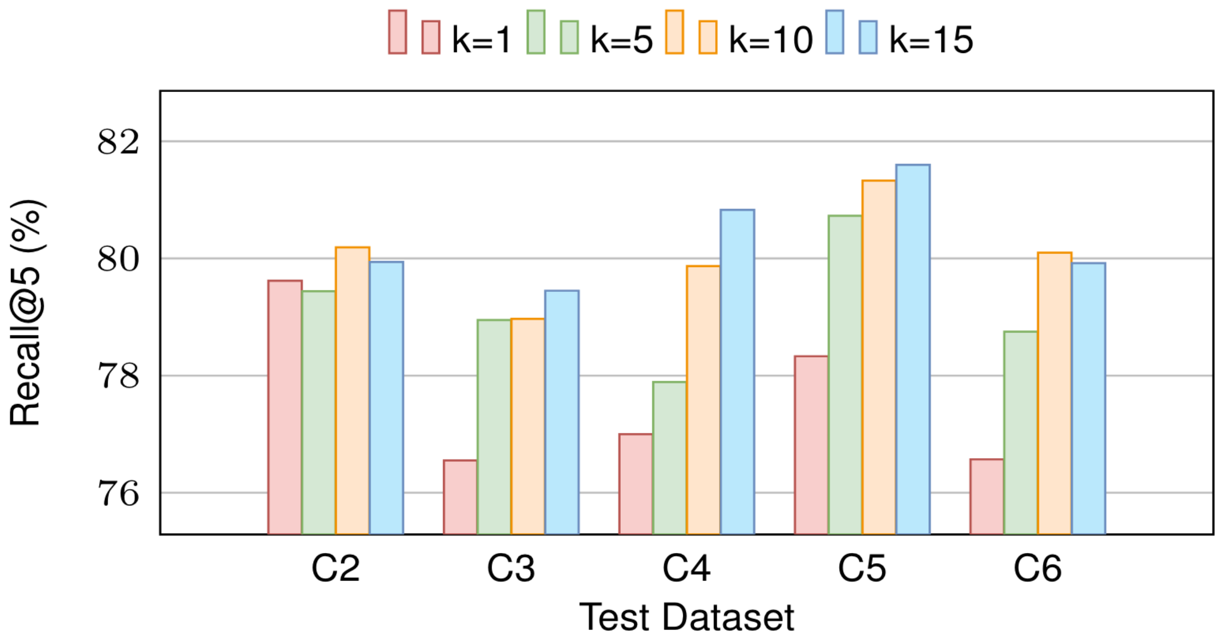
Pass through gradients in addition to the combination probabilities to enable end-to-end training.



Scallop - Results

On a Dataset leveraging GQA and ConceptNet for the knowledge base.





Scallop – Future Work

- How to integrate natural queries?

- How does it scale to noisy knowledge bases?

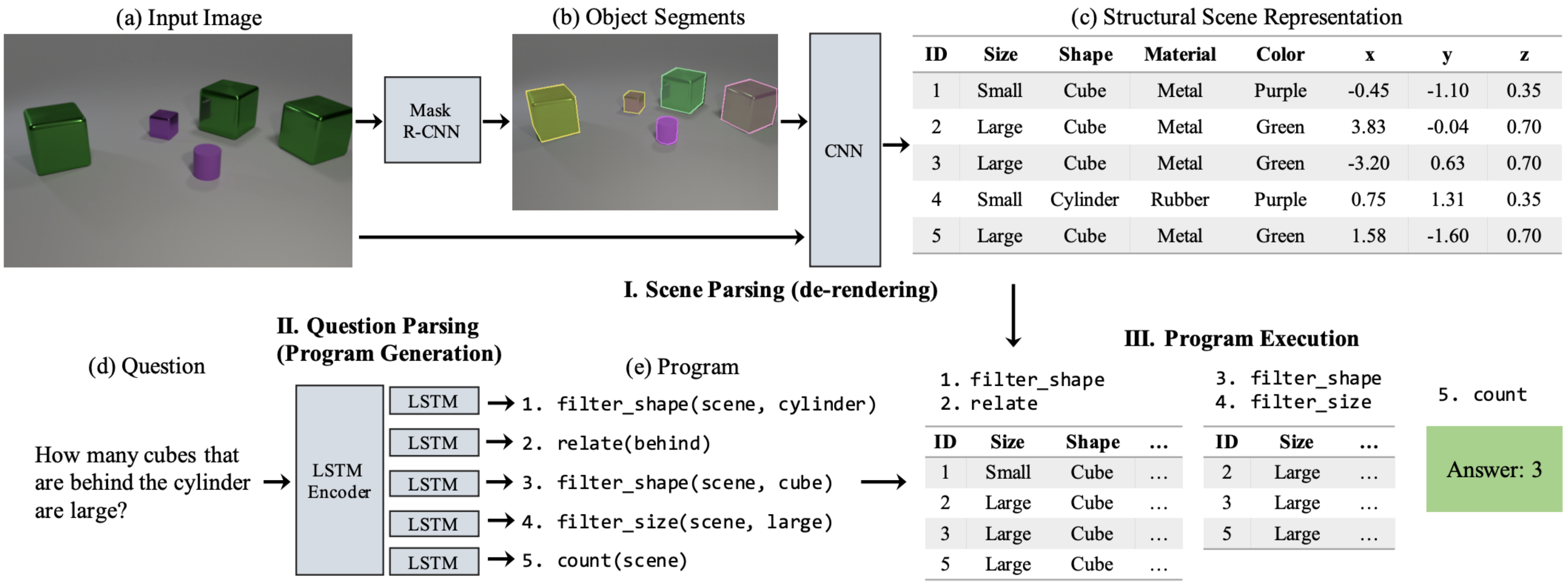
- Can we better estimate the gradient?

Recovering probabilistic symbolic representations from natural data

(A bit more neural)

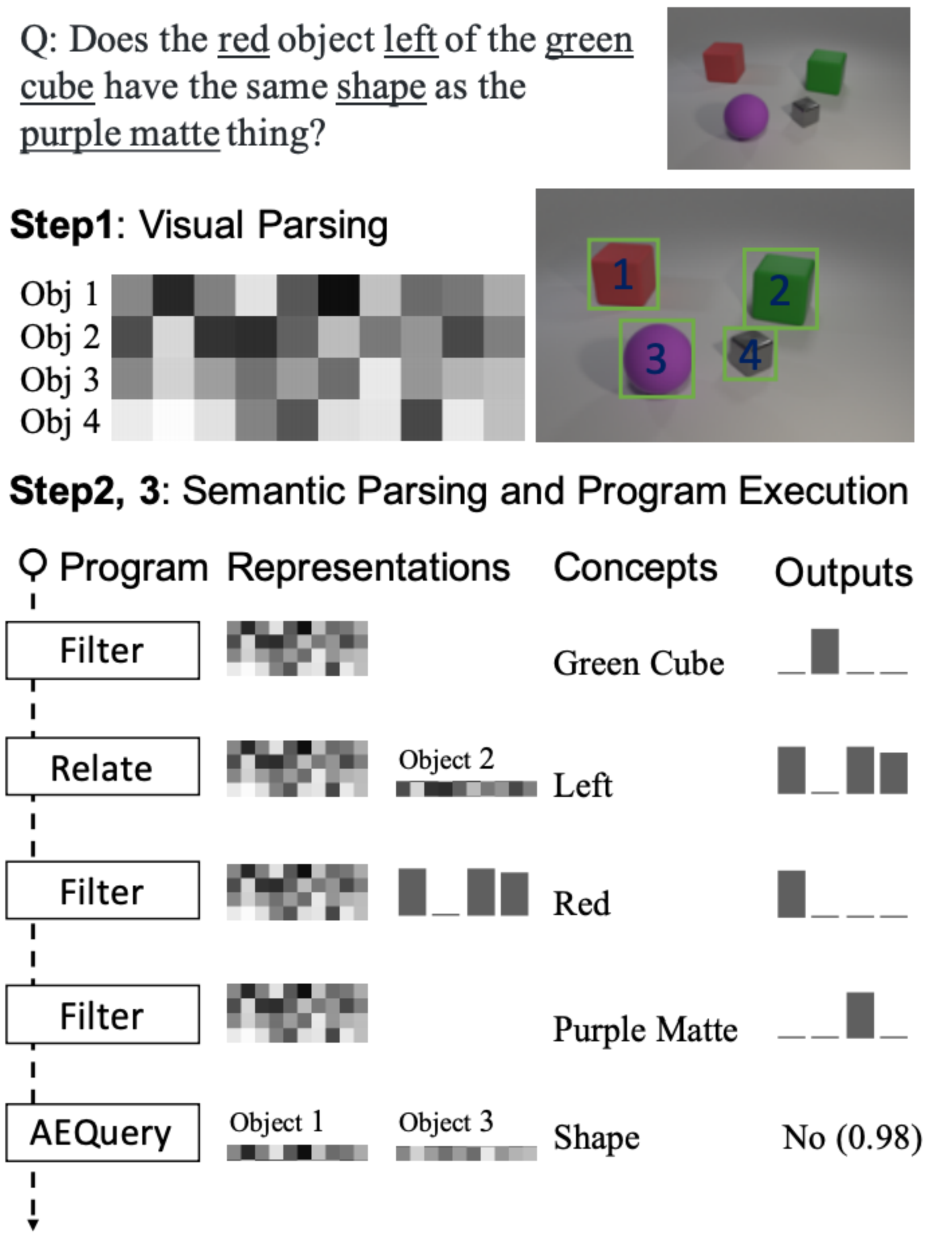
Symbolic Execution - NS-VQA

Discretize the query as well as the objects.



Yi. et al. Neural-Symbolic VQA: Disentangling Reasoning from Vision and Language Understanding. NeurIPS'18

Symbolic Query – NS-CL

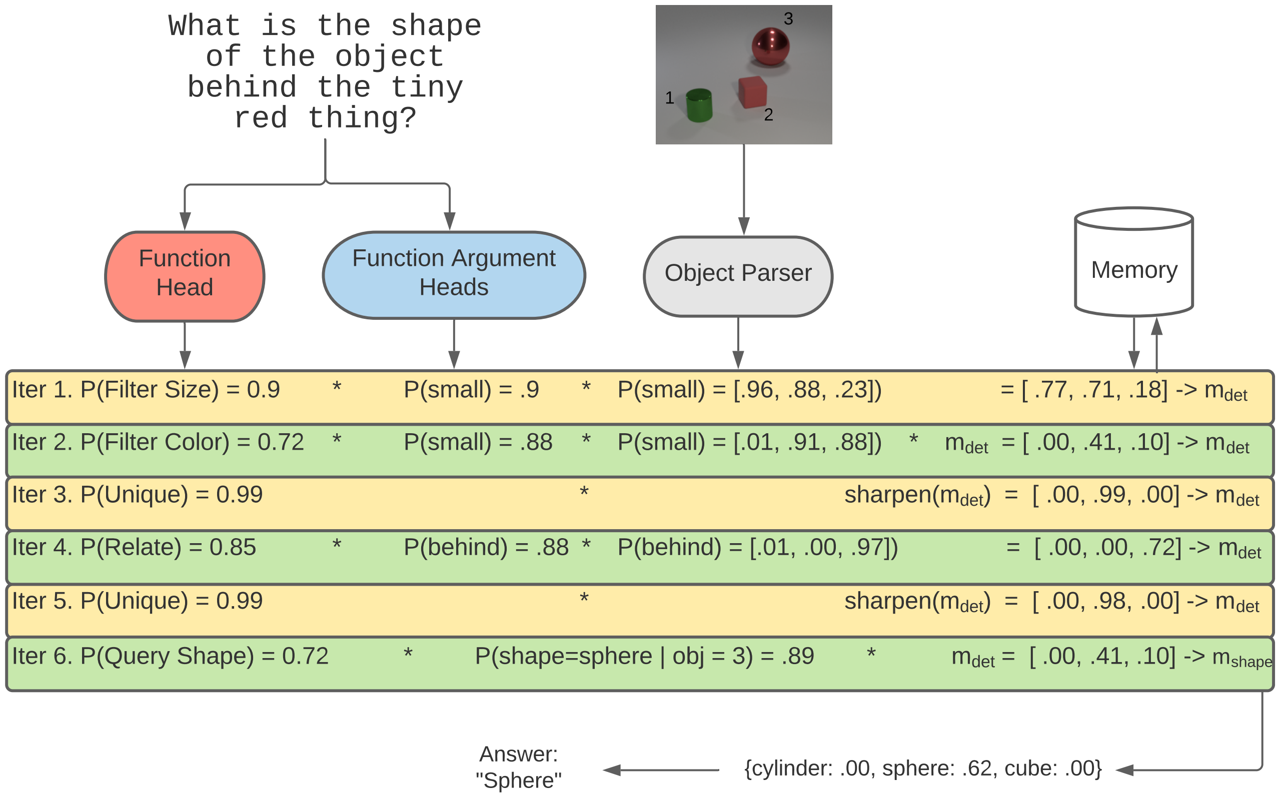


Discretize the query but keep the uncertainty in the vision.

Mao. et al. The Neuro-Symbolic Concept Learner: Interpreting Scenes, Words, and Sentences From Natural Supervision. ICLR'19

Prob. Query and Vision

What if we carry the uncertainty for both the vision and the logical representations?

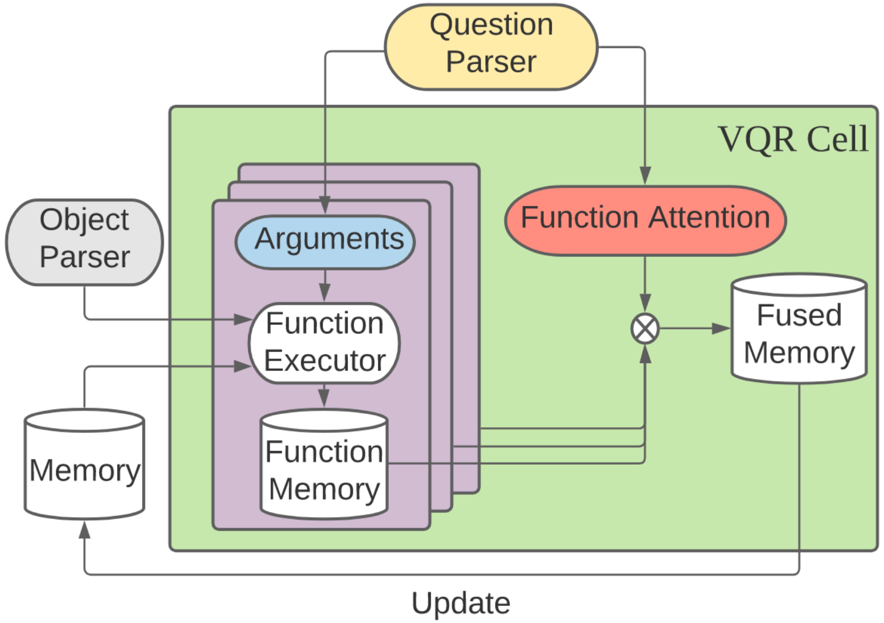


Samel, et al. How to Design Sample and Computationally Efficient VQA Models. In Progress

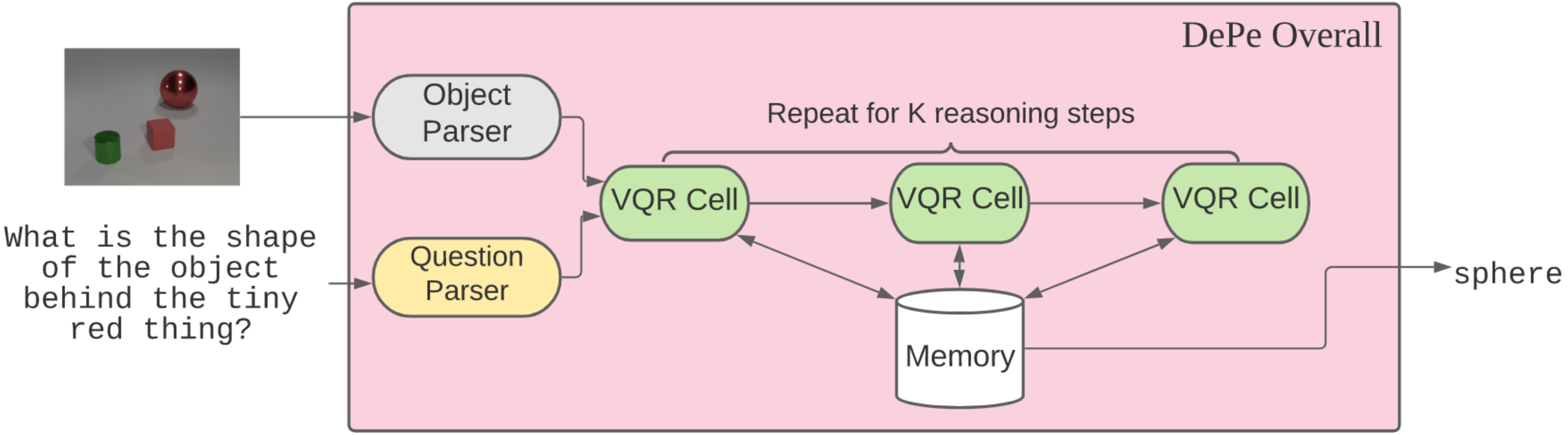
Why is this hard?

Exponential space of possible queries and solutions as the length grows.

Intuition: Compute the expected output conditioned on the current query function and the vision.



Differentiable End-to-End Program Executor (DePe)



The key is to design the memory can handle:

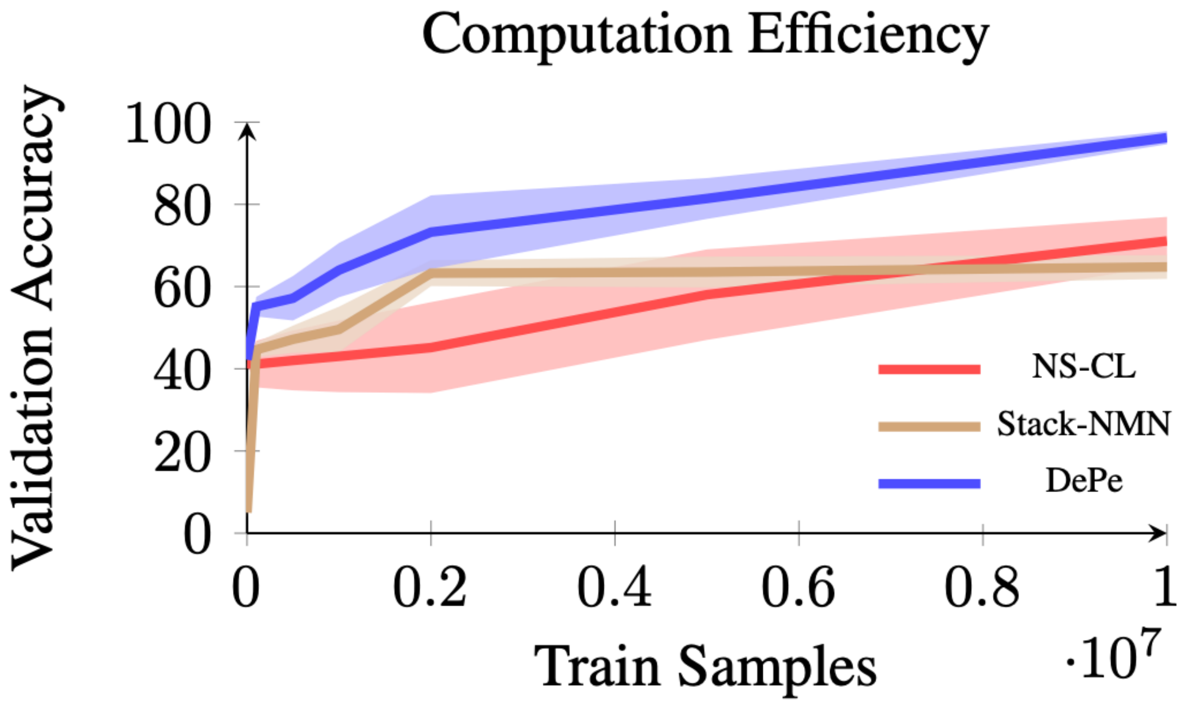
- **Symbolic representations**
- **Sub-queries**
- **Gradient computation throughout the entire execution**

DePe - Results

Results on CLEVR. Results on GQA in progress.

Sample Efficiency

QA Data %	Stack-NMN	NS-CL	DePe
1%	43.6	**68.1**	65.4
5%	66.6	86.7	**87.7**
10%	80.6	**98.8**	98.0



DePe – Future Work

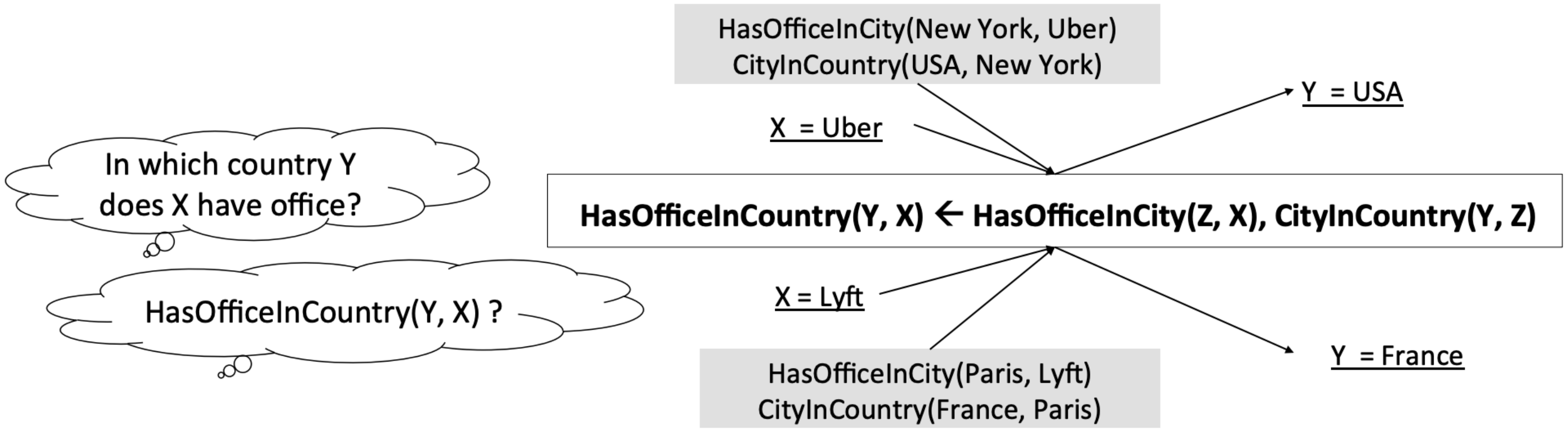
- Does the DSL cover natural questions?
- Can we use entire image based features for non-object centric reasoning?
- Instead of image embeddings, or symbolic attributes can we store attribute embeddings in memory?

Recovering logic rules from timeseries data

(Even more neural)

Background– Inductive Logic Programming

Given data on entities and their relations, can we derive new composite relations (rules)?



Given X can we walk along a knowledge graph to arrive at Y, and vice versa.

Yang, et al. Differentiable Learning of Logical Rules for Knowledge Base Reasoning. NeurIPS'17

Background– Inductive Logic Programming

Given sparse representations of entities X, Y and relations R, determine which relations are required to arrive from X to Y

$$\mathbf{s} = \sum_l \left(\alpha_l \left(\Pi_{\mathbf{k} \in \beta_l} \mathbf{M}_{\mathrm{R_k}} \mathbf{v_x} \right) \right), \ \mathrm{score}(\mathrm{y} \mid \mathrm{x}) = \mathbf{v}_{\mathbf{y}}^T \mathbf{s}$$

Here we assume that that the entities, as well as their relations are provided. Can we assume this in timeseries data?

Yang, et al. Differentiable Learning of Logical Rules for Knowledge Base Reasoning. NeurIPS'17

Problem Formulation – Time Series

- **Investigating timeseries inference where observed events induced by a compact subset of other events:**

$$\text{long jump} \coloneqq \text{before}(\text{run}, \text{jump})$$

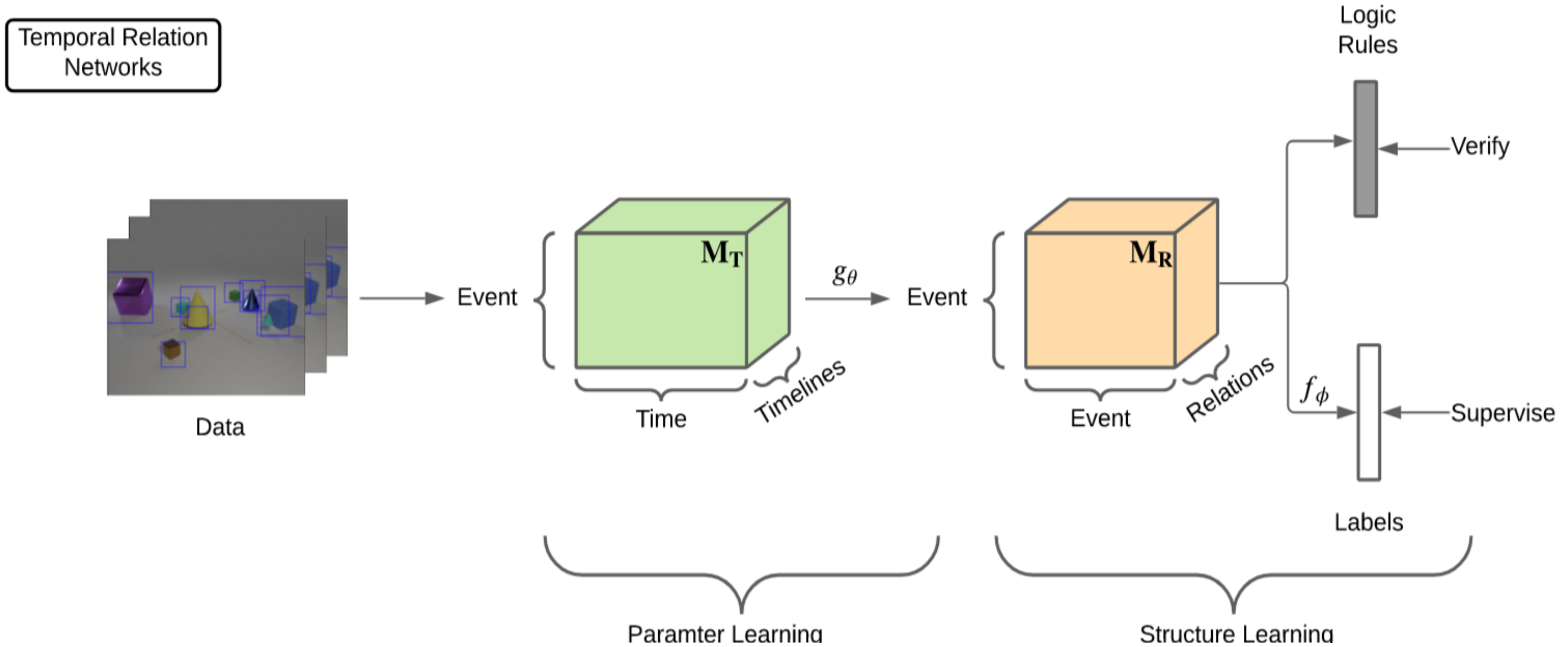
- **We may have labeled atomic event data, but rarely explicit relational data.**

- **Can we learn both the rule structure <u>and</u> the relation parameters for logic time series?**

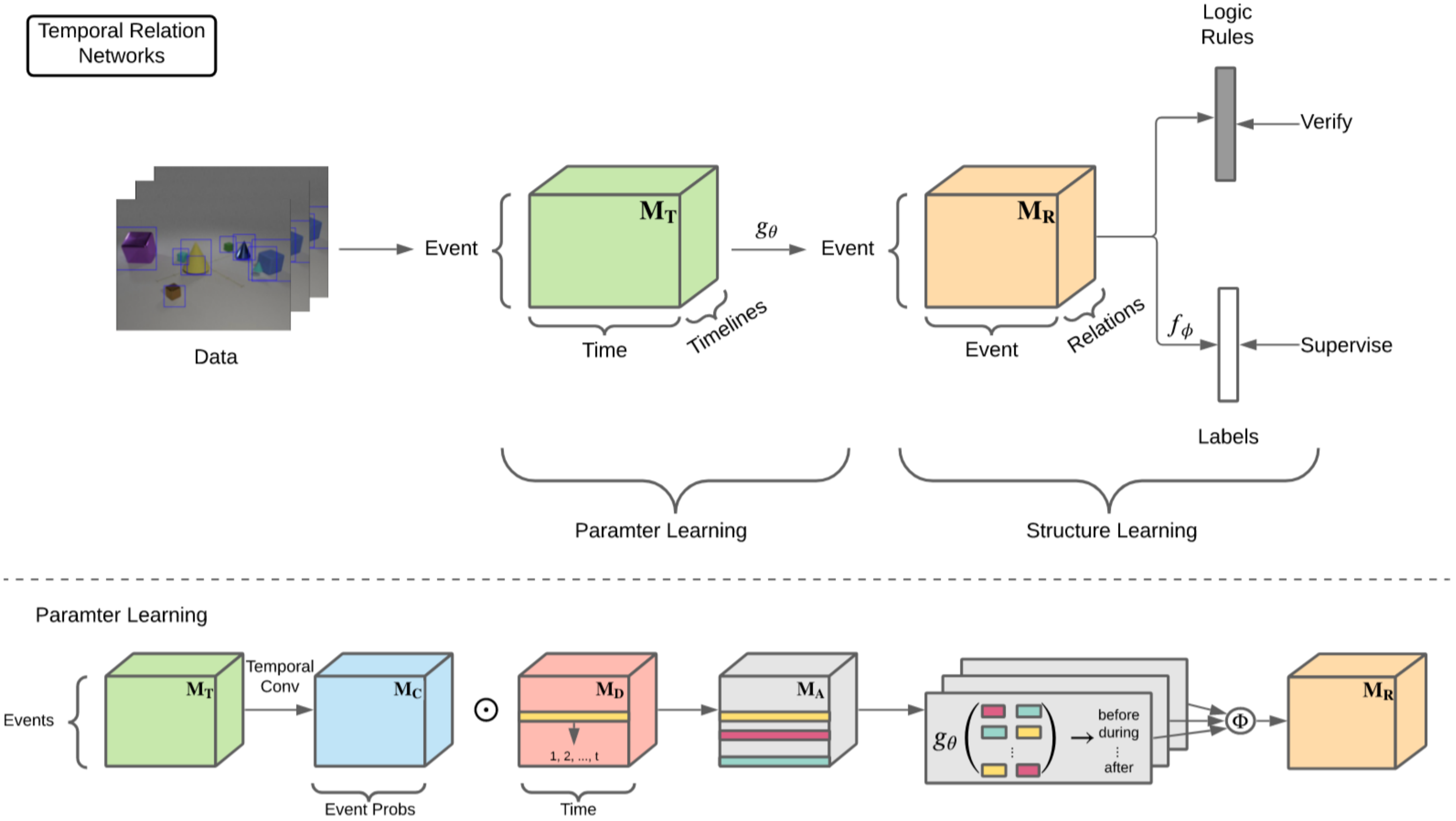
High Level Framework

Can we implicitly leverage the temporal relations between events?:

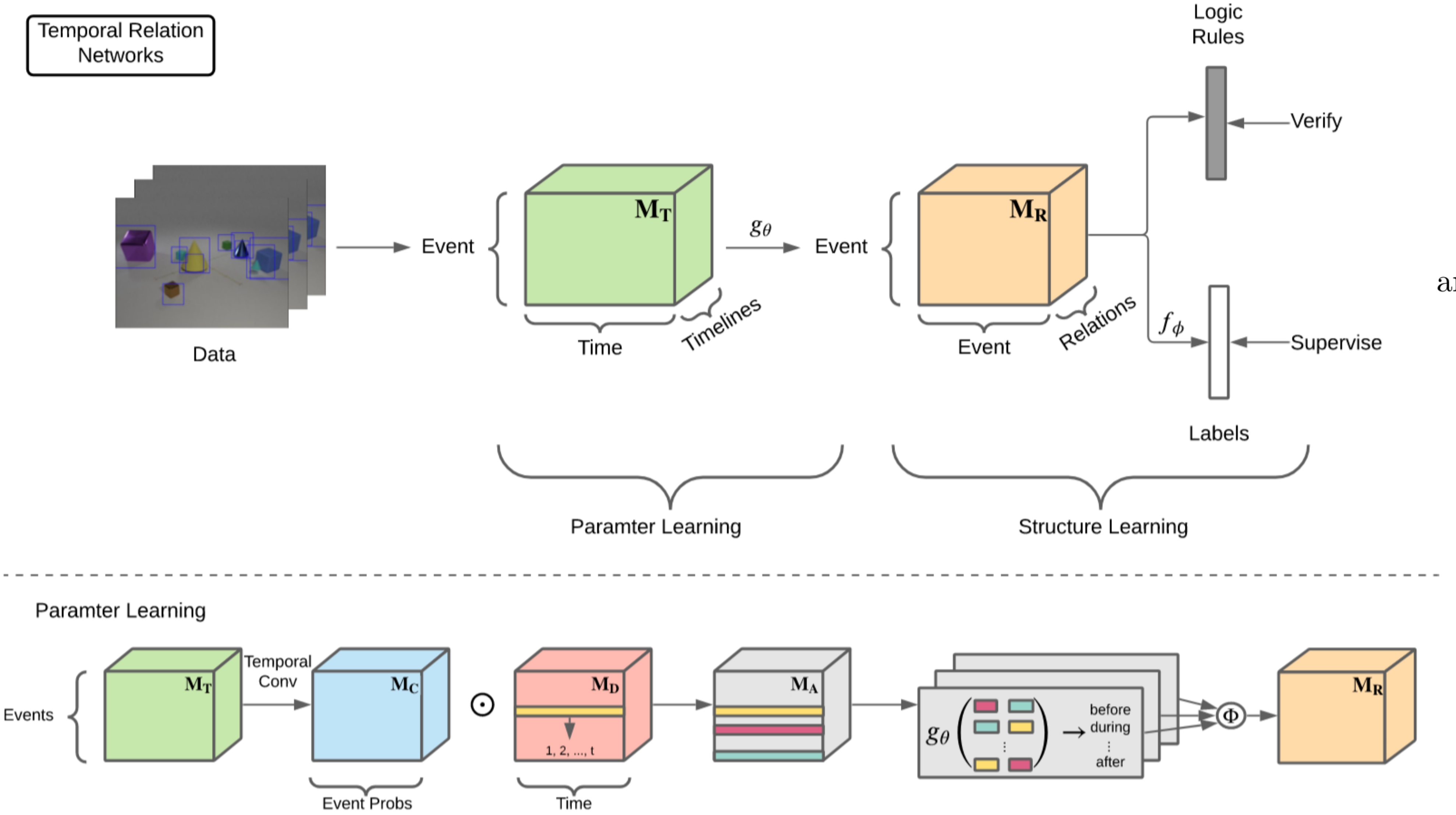
- **Optimize relations and rule structure end-to-end.**
- **Extract the underlying temporal logic rules for verification or discovery.**



Temporal Relation Networks - Learning



Temporal Relation Networks - Learning



$$\arg\min_{\theta,\phi} - \sum_{(\hat{\mathbf{y}},\mathbf{y})} \sum_{i=1}^{|\mathbf{y}|} \mathbf{y}_i \cdot \log(\hat{\mathbf{y}}_i)$$

$$\text{s.t. } \min \left\| \mathbf{M_R} - \bar{\mathbf{M}}_\mathbf{R} \right\|_F$$

$$\min \|\mathbf{M}_\alpha\|_1$$

Temporal Relation Networks - Extraction

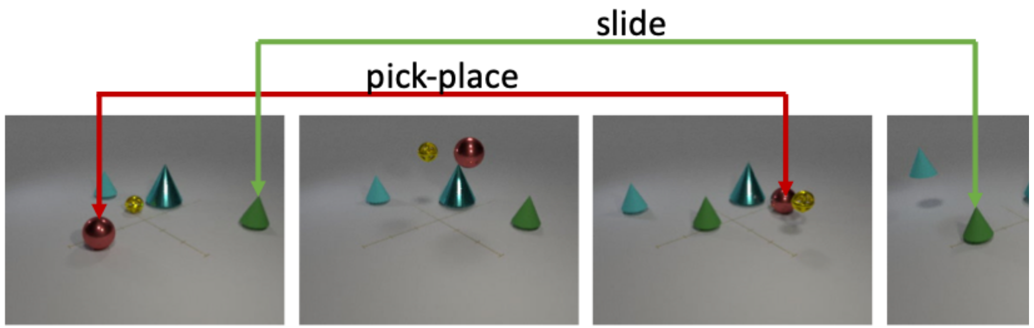
- Given the weights of the structure model, we take the highest weighted fact triplets to generate the rules of length k.

$$f(y_f) := \bigwedge_{i=1}^{k} r_t(x_u, x_v) \ ; \ P(f(y_f)) = \prod_{i=1}^{k} \mathbf{a}_i^{y_f}$$

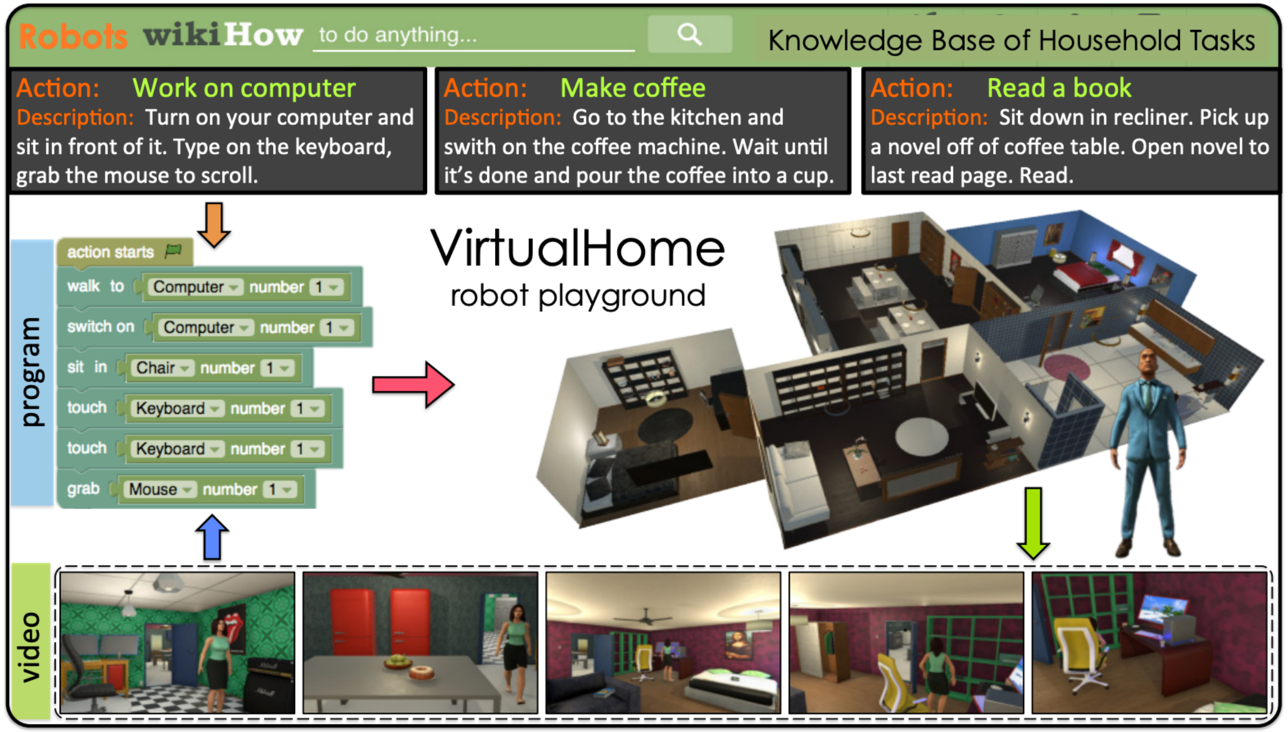
- Variable length rules we can explore tree based methods.

Current Work – Synthetic Verification

On synthetic data we can objectively evaluate our temporal rule extraction.



Event 5 :- during(pick-place, slide)

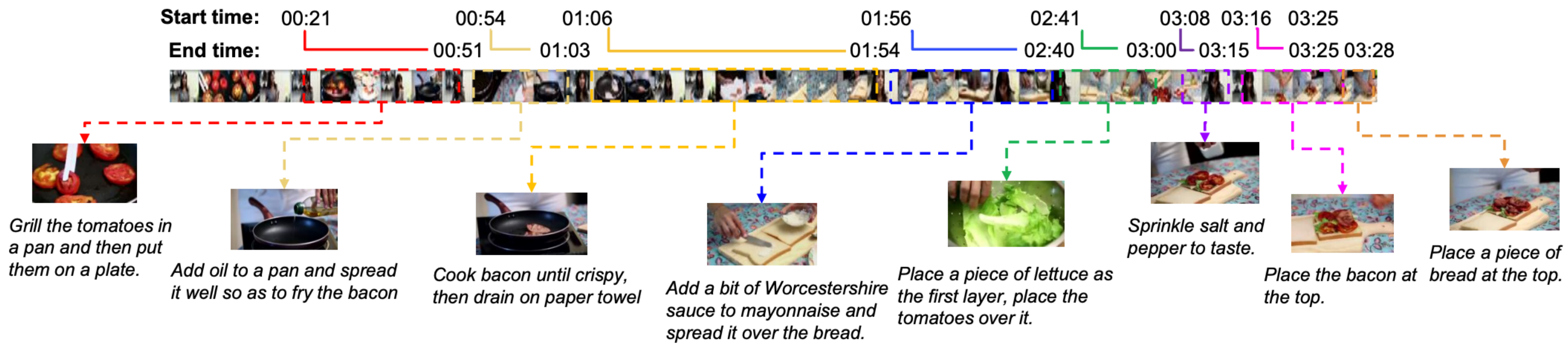


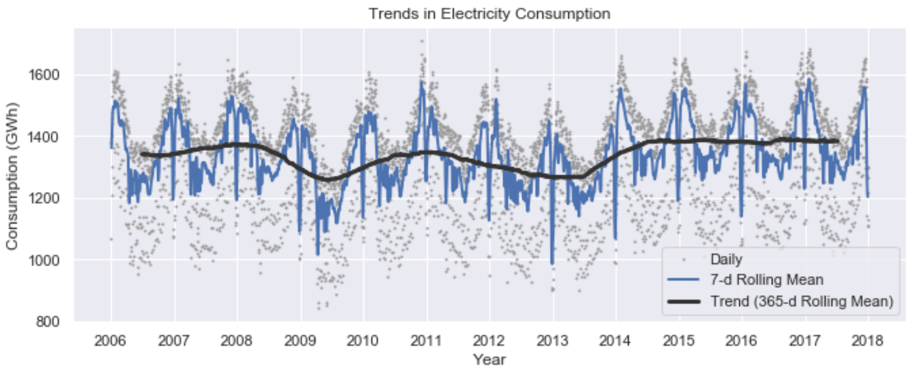
Puig et al. Virutalhome: Simulating household activities via programs. CVPR'18

Rohit Girdhar and Deva Ramanan. CATER: A diagnostic dataset for Compositional Actions and TEmporal Reasoning. ICLR'20

Future Work – Proposed Datasets

On real world timeseries data we can <u>subjectively</u> evaluate our rule proposals.





Zhou et al. Towards Automatic Learning of Procedures from Web Instructional Videos. AAAI'18

Future Directions

Multi-Modal Knowledge Construction

- We leverage a small percentage of structured data for knowledge based methods
- Most recent works focus on construction based on single modalities, ie text.
- How do we leverage jointly leverage unstructured text, images, or videos to construct this knowledge?

Multi-Modal Knowledge Construction

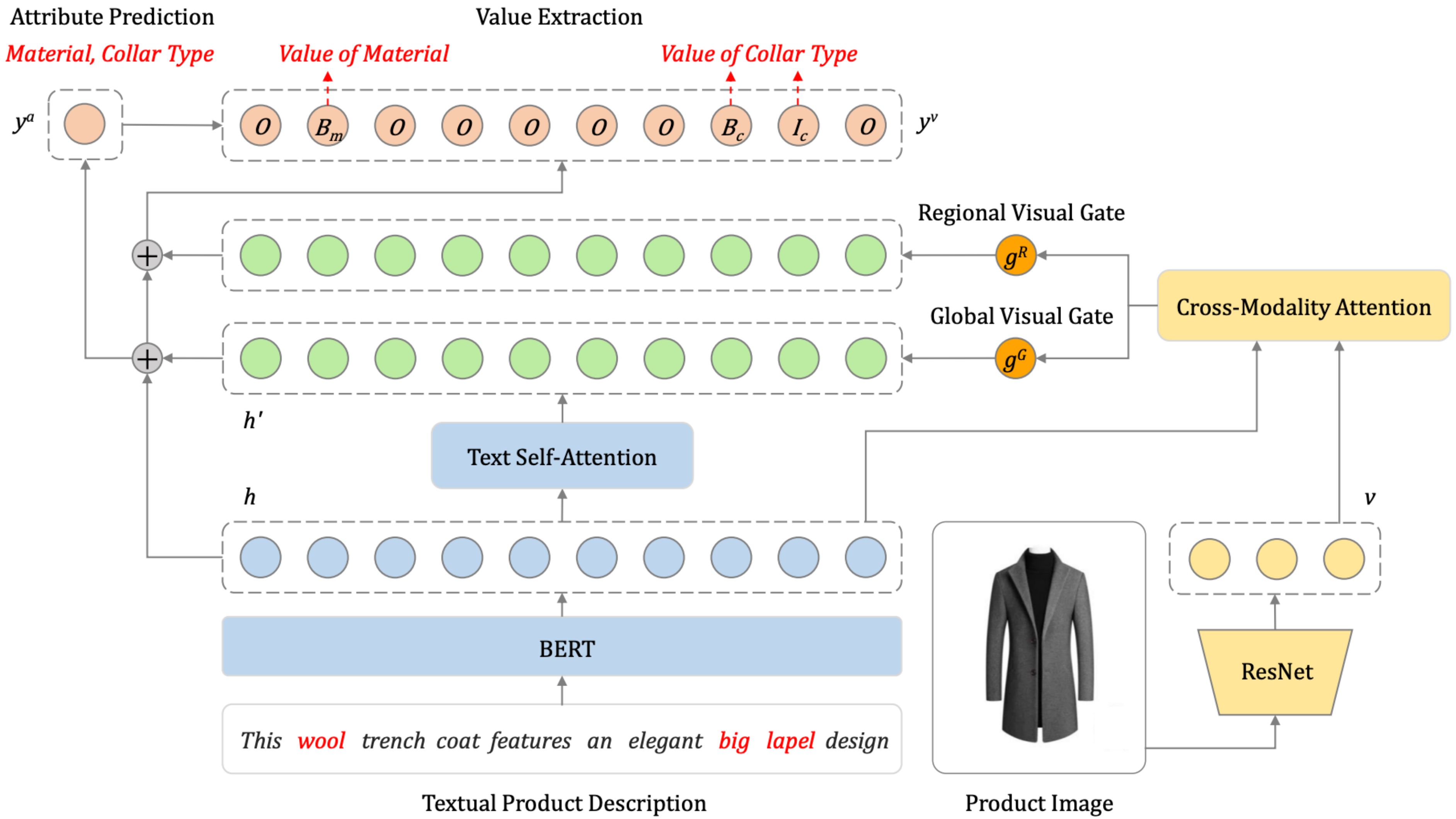
Fill in knowledge gaps in specific domains.

"
This golden lapel shirt can be paired with black shoes
……
"

Attribute	*Value*
Collar Type	*lapel*
Color	*golden*

Zhu et al. Multimodal Joint Attribute Prediction and Value Extraction for E-commerce Product. EMNLP'20

Multi-Modal Knowledge Construction



Zhu et al. Multimodal Joint Attribute Prediction and Value Extraction for E-commerce Product. EMNLP'20

Constructed Knowledge for Real World Tasks

- We have seen methods leveraging a DSL to reason over scene graphs and knowledge graphs.
- How effective are these functions given natural queries?
- Are there effective ways to determine a compact subset of these functions?
- Are there more effective methods besides just pure discrete reasoning vs neural reasoning (GNN)?

Research Summary

Covered a range of works in the neuro-symbolic spectrum:

- Image Question Answering
  - Real world symbolic execution and training
  - Handling uncertainty in symbolic learning
- Rule learning and inference over videos
- Future Directions
  - Multi-modal knowledge construction
  - Real world reasoning over inferred knowledge

Leveraging Structures for Machine Reasoning

Karan Samel

ksamel@gatech.edu

2/8/21