

Relationships from Entity Stream

Martin Andrews
martin@RedDragon.ai

Sam Witteveen
sam@RedDragon.ai

Summary

Task :

- ▲ Q&A about relationships between objects in an image (Sort-of-CLEVR)

Inspiration :

- ▲ Relation Network paper

Idea :

- ▲ Focus attention on entities in image
- ▲ Stream relevant entities
- ▲ Learn relationships from stream

Results :

- ▲ Similar accuracy to previous work

Qualitative Advantages :

- ▲ Only entities are reasoned about
- ▲ Greater interpretability
- ▲ Fewer parameters
- ▲ End-to-end trainable grounding

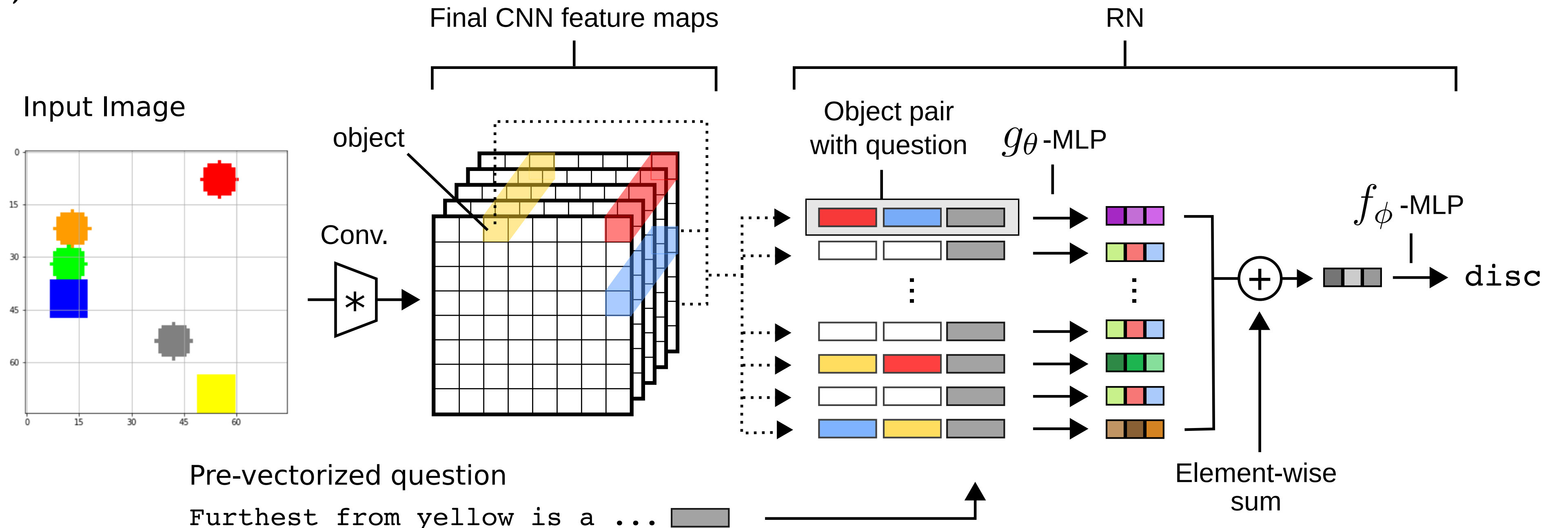
Source Code with Data Generator :

Available via GitHub :

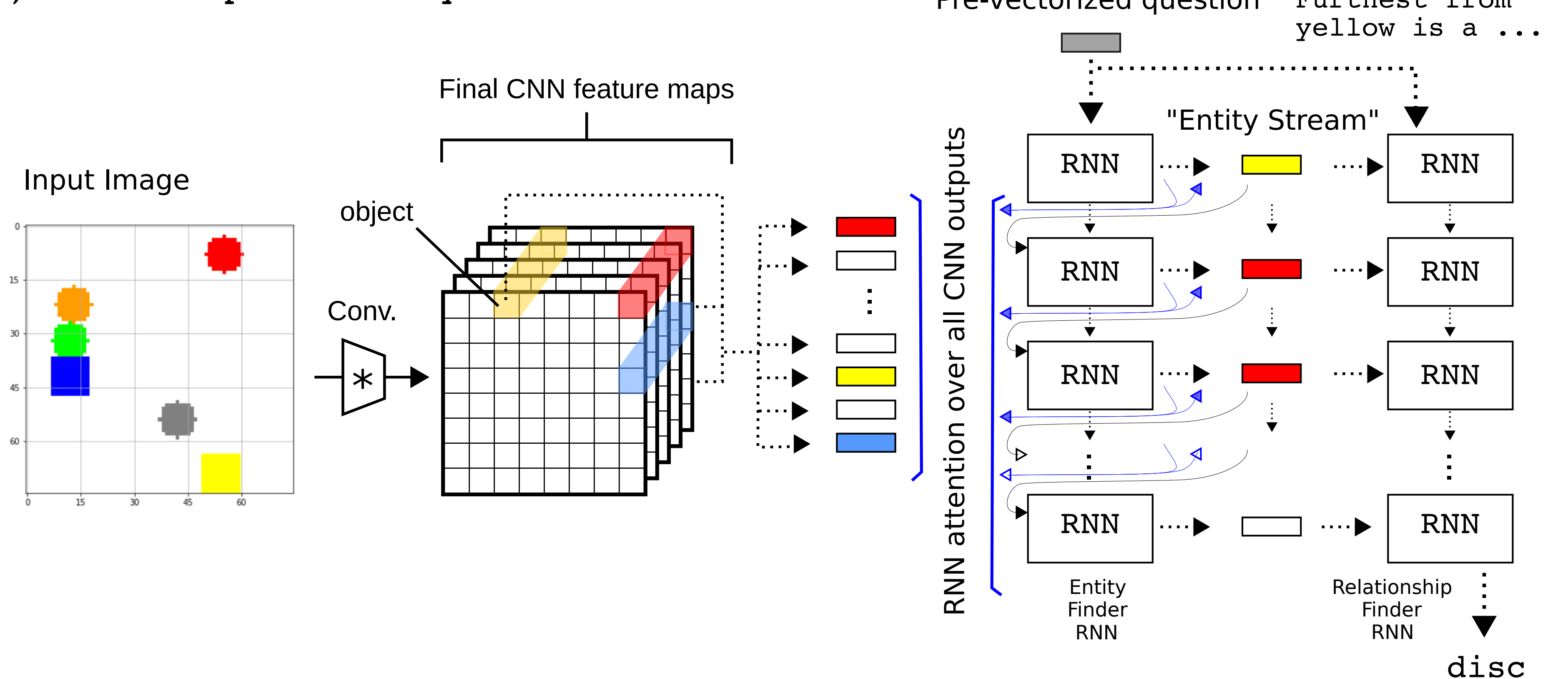
- ▲ <https://github.com/mdda/relationships-from-entity-stream>

Model Comparison

(a) Relation Network



(b) Relationships from Entity Stream



Model Detail

Input CNNs :

- ▲ 4 layers, 3x3 conv 24 channels + BN

Question & Answer :

- ▲ Pre-encoded as 11d & 10d binary
- ▲ Q:[r, g, b, o, k, y, q1, q2, s1, s2, s3]
- ▲ A:[yes, no, rect, disc, 1, 2, 3, 4, 5, 6]

Entity Finder RNN :

- ▲ 2-layer GRU, 32/64 hidden dims
- ▲ Output is query 'q' for attending
- ▲ Found 'entities' input to next timestep

Attention-is-all-you-need :

- ▲ Concatenate location to CNN output
- ▲ Split into 'k' and 'v' pieces
- ▲ Soft attention : weighted 'v' as entity
- ▲ Gumbel trick to learn hard attention

Relationship Finder RNN :

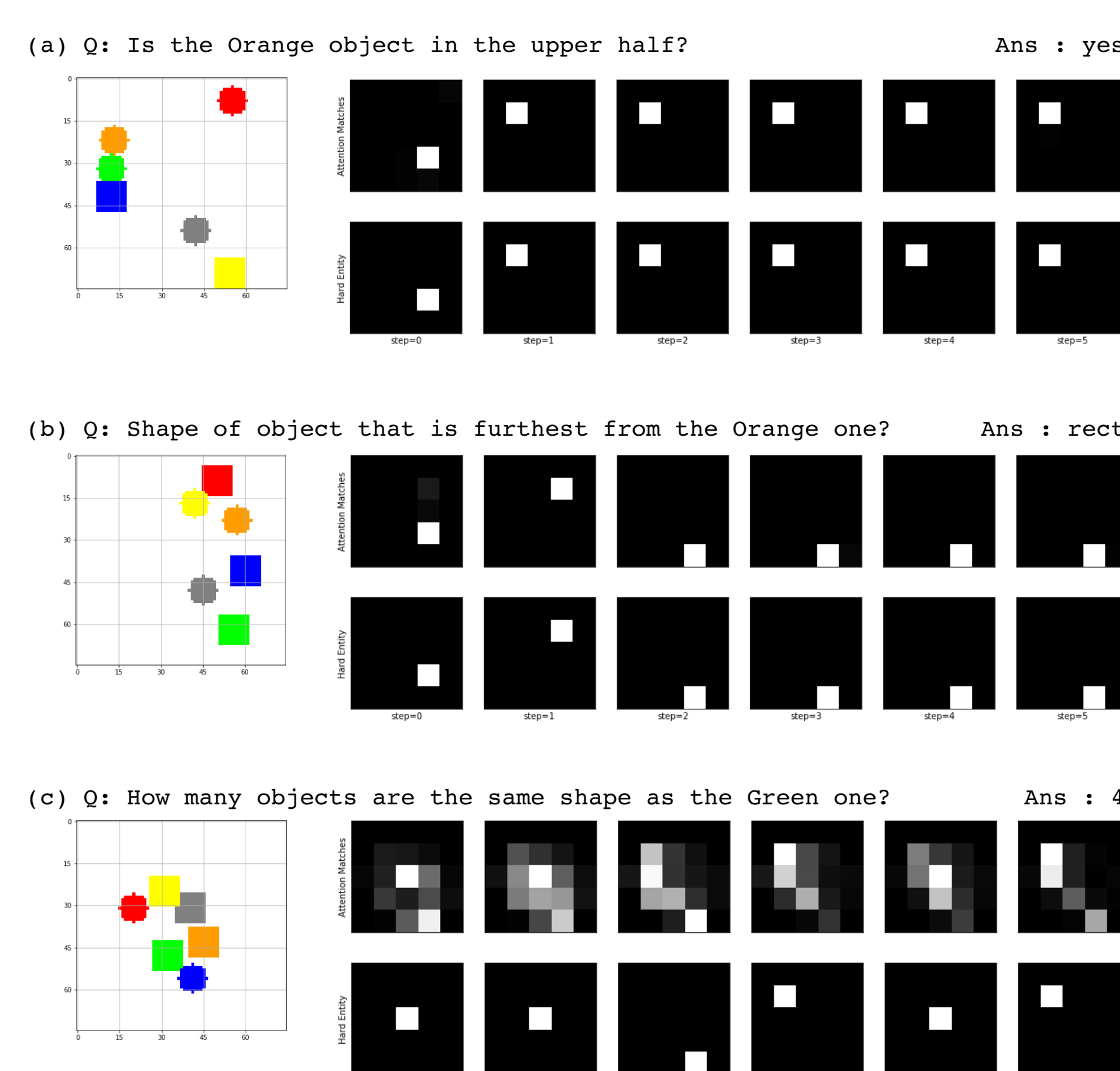
- ▲ Inputs are entities ('v' from attention)
- ▲ 2-layer GRU, 32/64 hidden dims
- ▲ Answer from last hidden state

Results

Quantitative

Model	NonRel fraction correct	BiRel fraction correct	hidden_dim	Size in bytes
RN	99%	93%	-	1,463,513
CNN	98%	63%	-	970,874
Our RFES-S	99%	95%	32	166,380
Our RFES-H	99%	93%	64	408,364

Qualitative



Discussion

Motivation :

- ▲ Avoid comparing all locations pairwise
- ▲ Want sequential reasoning
- ▲ Potential to mix & match streams

SoftMax Attention :

- ▲ Works quickly, more robust
- ▲ Capable of 'cheating' by snapshotting entity scene (e.g. to count)

Hard Attention :

- ▲ Some tuning required (larger GRU)
- ▲ Gumbel has self-scaling property
- ▲ Training still fully differentiable
- ▲ Testing explicitly 'ArgMax hard'

Future directions :

- ▲ Small-RL and attention games
- ▲ Revisit MNIST by saccades
- ▲ Other uses of 'internal dialogue'

Key References

- "A simple neural network module for relational reasoning" - Santoro et al. (2017)
- "Attention is all you need" - Vaswani et al. (2017)
- "The symbol grounding problem" - Harnad (1990)

Contact

martin@RedDragon.ai
+65 8585 1750
<http://RedDragon.ai>

