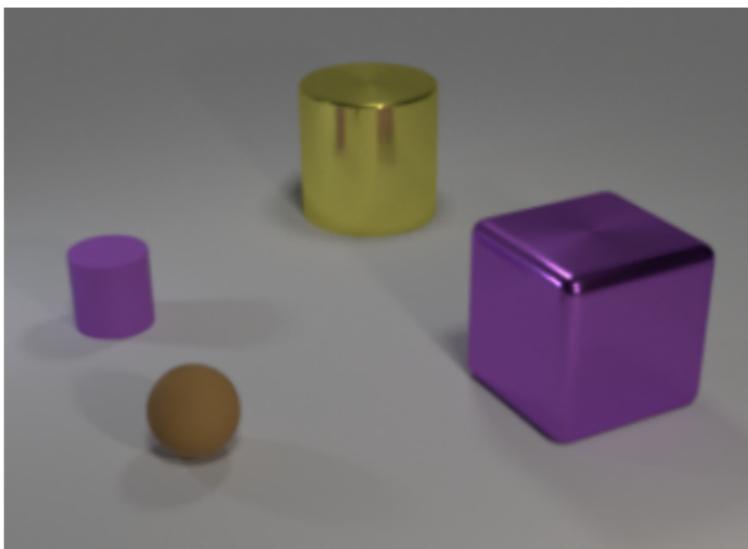


Relational recurrent neural networks

Soboleva Natalia, 151

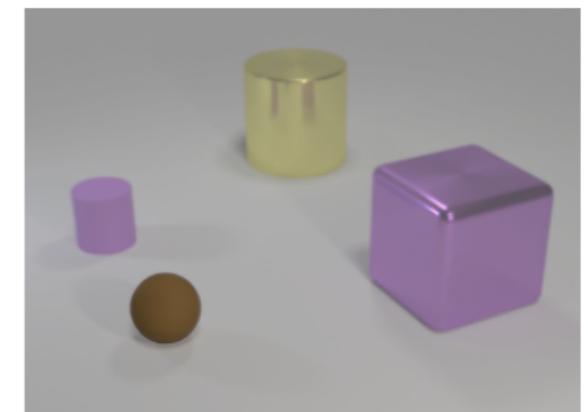
Relational Reasoning

Original Image:



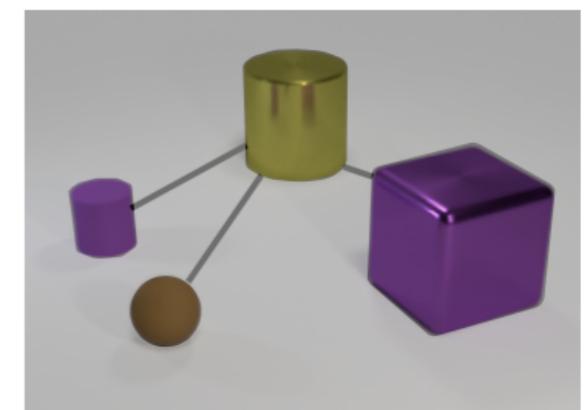
Non-relational question:

What is the size of
the brown sphere?



Relational question:

Are there any rubber
things that have the
same size as the yellow
metallic cylinder?



An illustrative example from the CLEVR dataset of relational reasoning.
Source: <https://papers.nips.cc/paper/7082-a-simple-neural-network-module-for-relational-reasoning.pdf>

Relation Networks

$$O = \{o_1, o_2, \dots, o_n\}, o_i \in \mathbb{R}^m$$

f_ϕ and g_θ are MLPs.

$$\text{RN}(O) = f_\phi \left(\sum_{i,j} g_\theta(o_i, o_j) \right)$$



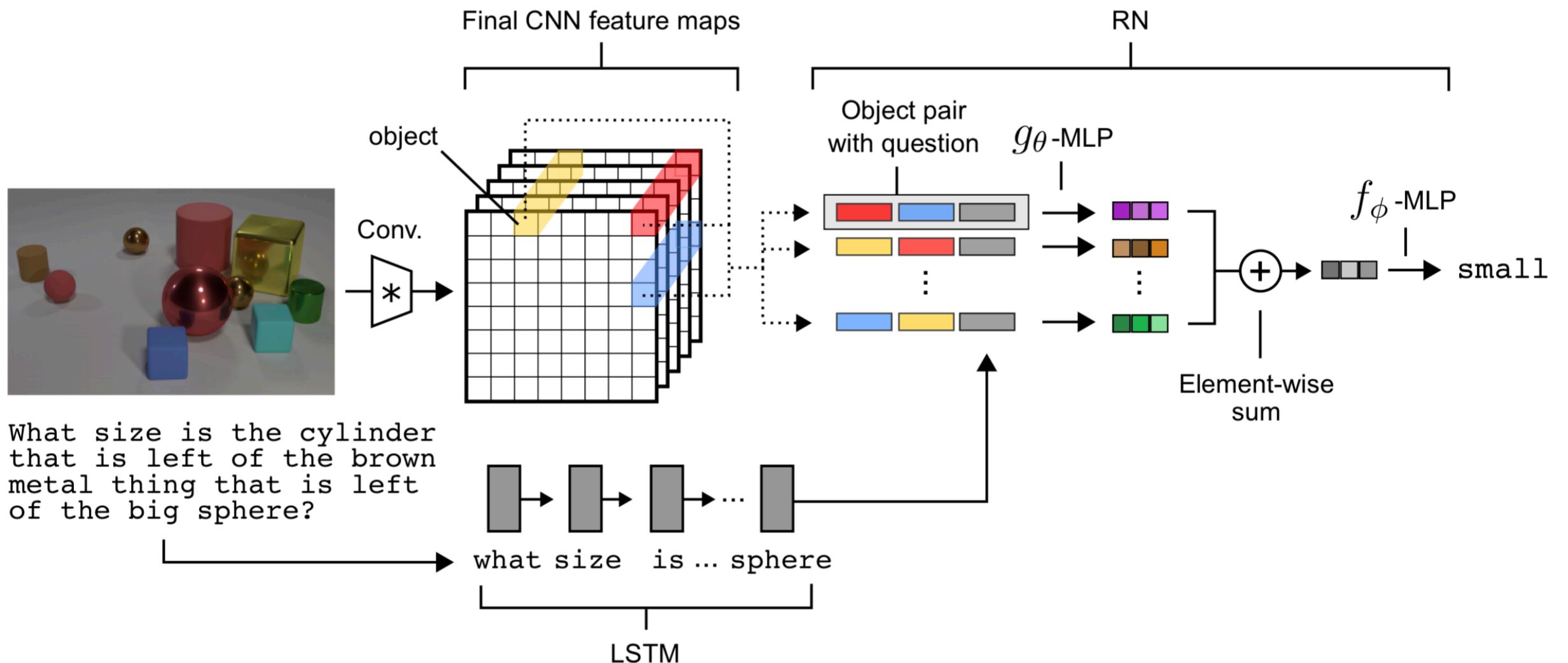
ARTIFICIAL INTELLIGENCE

DeepMind Develops a Neural Network That Can Make Sense of Objects Around It

The better to see you with.

Source: <https://futurism.com>

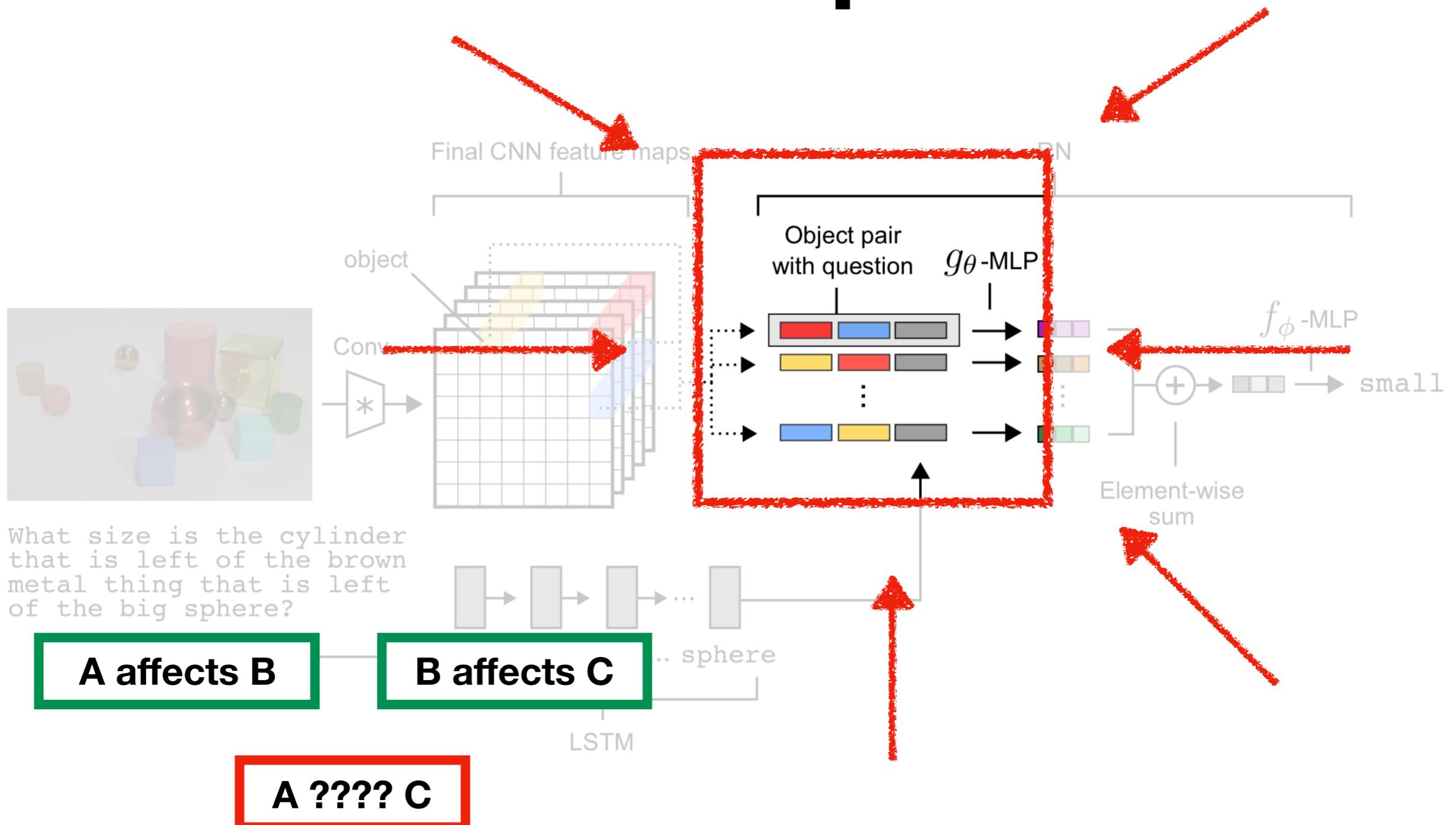
— end to end —



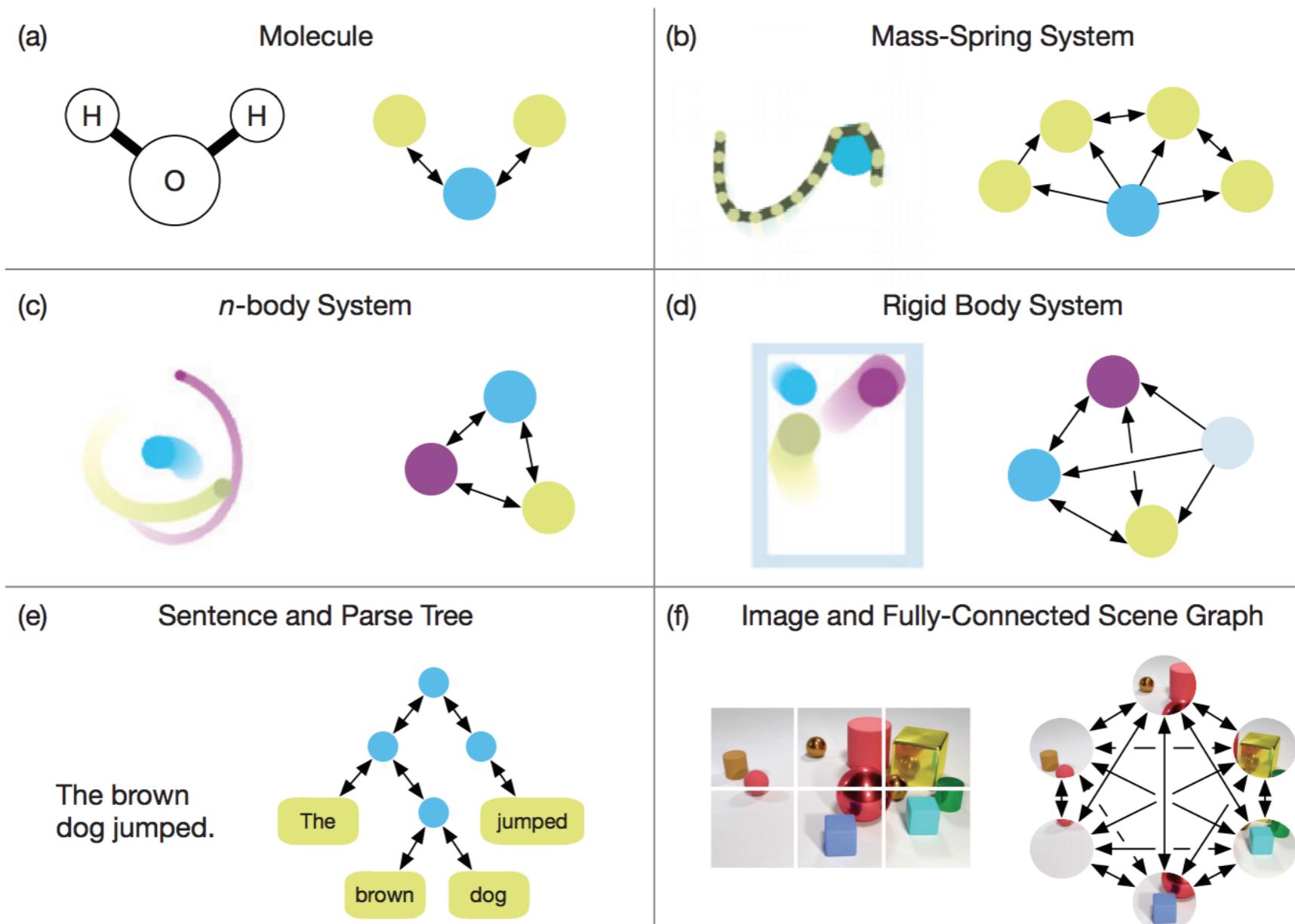
Visual QA architecture.

Source: <https://papers.nips.cc/paper/7082-a-simple-neural-network-module-for-relational-reasoning.pdf>

Where is the problem?

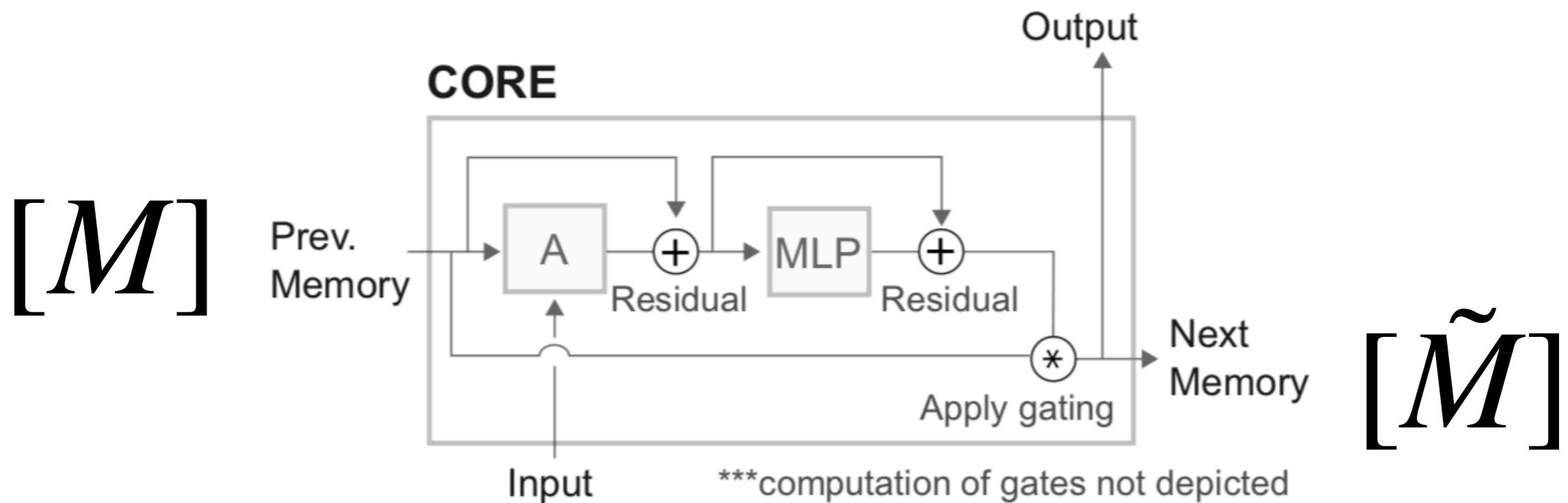


Relational Reasoning



*Universality of Graph Representations.
Source: <https://arxiv.org/pdf/1806.01261.pdf>*

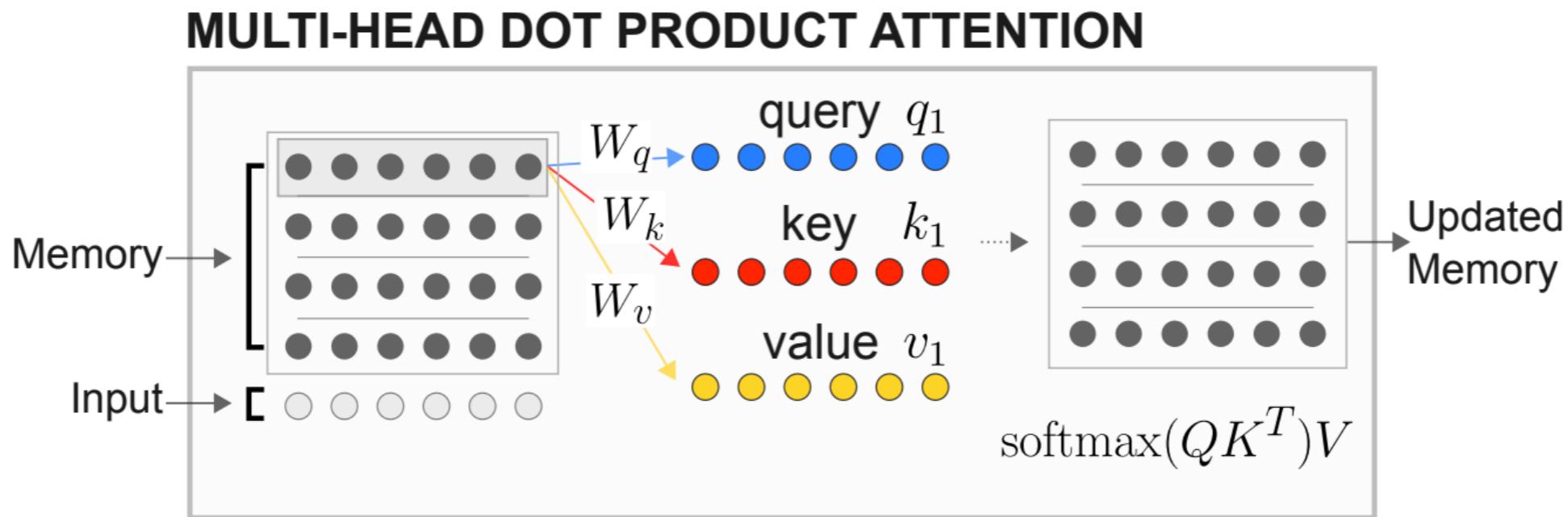
Relational Memory Core



Relational Memory Core.

Source: <http://papers.nips.cc/paper/7960-relational-recurrent-neural-networks.pdf>

Multi-head dot product attention (MHDPA)

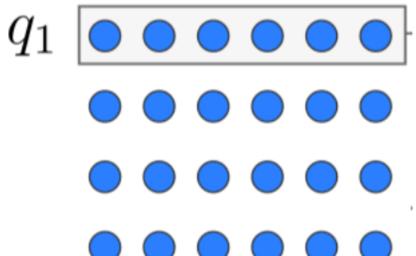


Multi-head dot product attention (MHDPA).

Source: <http://papers.nips.cc/paper/7960-relational-recurrent-neural-networks.pdf>

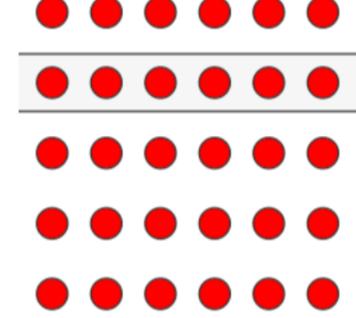
$$Q = MW_q \quad K = MW_k \quad V = MW_v$$

Queries

q_1 

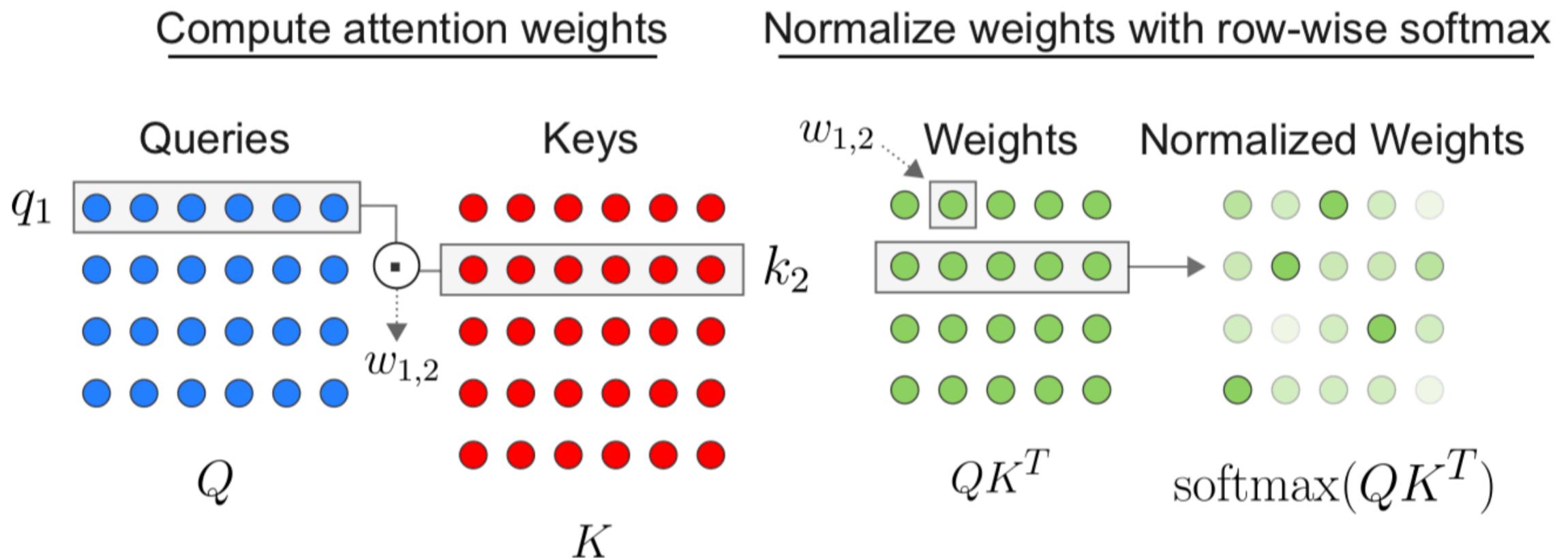
Q

Keys

k_2 

K

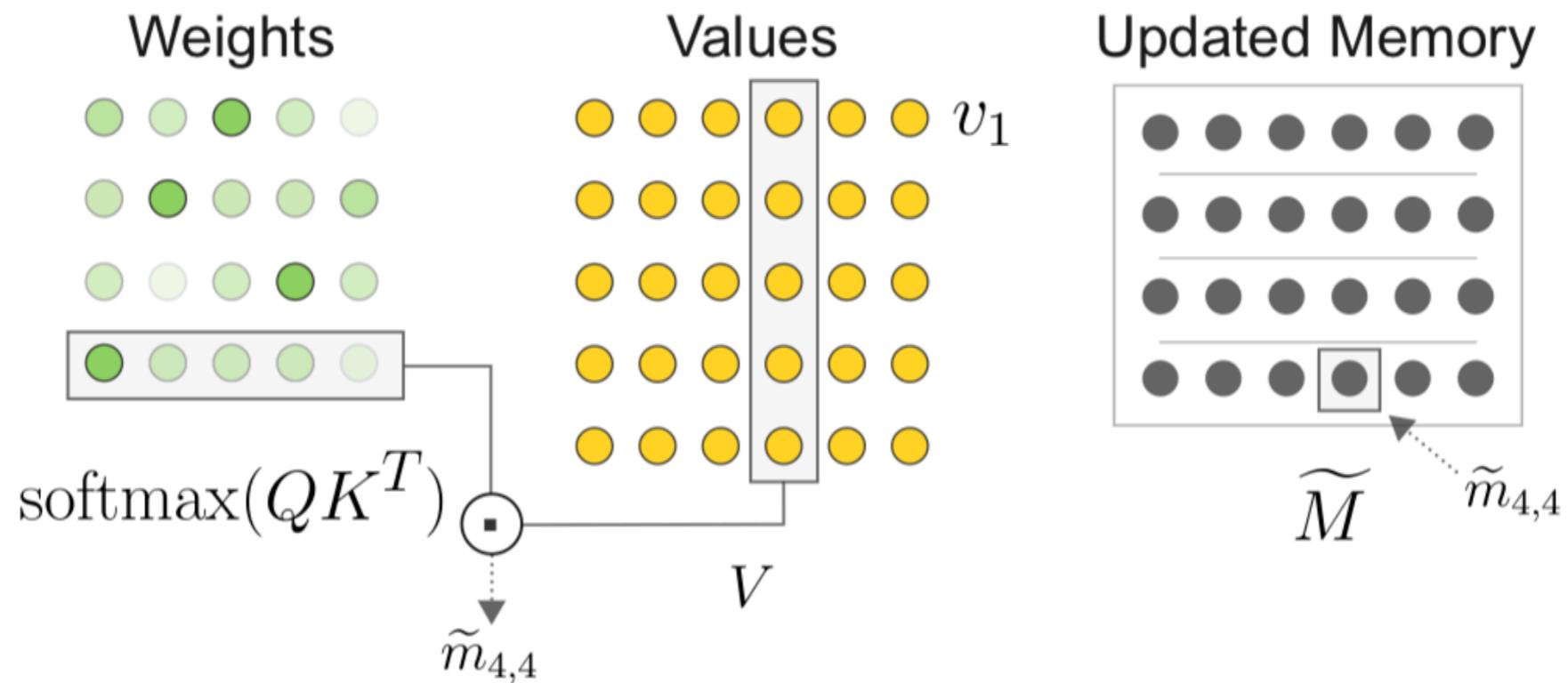
$$A(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V$$



Multi-head dot product attention (MHDPA).

Source: <http://papers.nips.cc/paper/7960-relational-recurrent-neural-networks.pdf>

Compute weighted average of values Return updated memory



Multi-head dot product attention (MHDPA).

Source: <http://papers.nips.cc/paper/7960-relational-recurrent-neural-networks.pdf>

$$A_\theta(M) = \text{softmax} \left(\frac{MW^q(MW^k)^T}{\sqrt{d_k}} \right) MW^v, \text{ where } \theta = (W^q, W^k, W^v)$$

Encoding new memories

(Into the LSTM)

$$A_\theta(M) = \text{softmax} \left(\frac{MW^q(MW^k)^T}{\sqrt{d_k}} \right) MW^v, \text{ where } \theta = (W^q, W^k, W^v)$$



$$\widetilde{M} = \text{softmax} \left(\frac{MW^q([M; x]W^k)^T}{\sqrt{d^k}} \right) [M; x]W^v$$

Examples

```

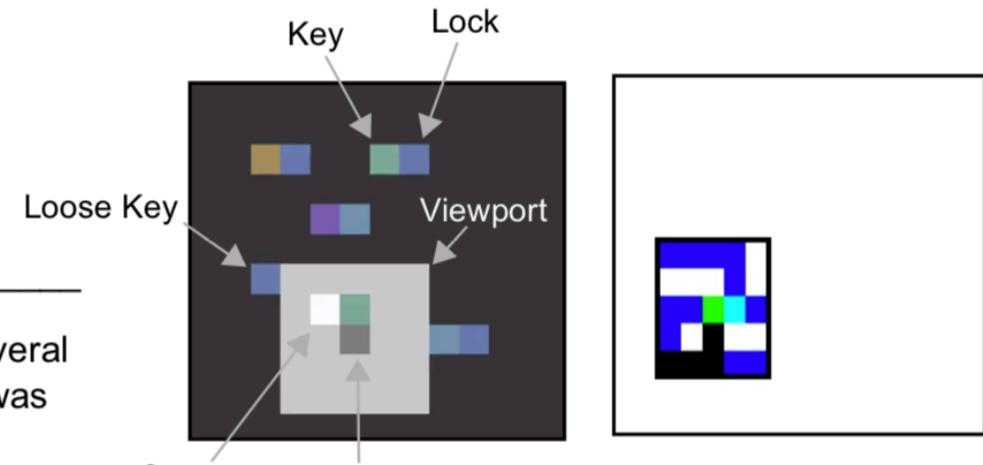
x = 339
for [19]:
    x += 597
        for [94]:
            x += 875
x if 428 < 778 else 652
print(x)

```

Super Mario Land is a 1989 side scrolling platform video _____

It had 24 step programming abilities, which meant it was highly _____

A gold dollar had been proposed several times in the 1830s and 1840s , but was not initially _____



Program Evaluation

Language Modeling

BoxWorld

Mini-Pacman

Supervised Learning

Reinforcement Learning

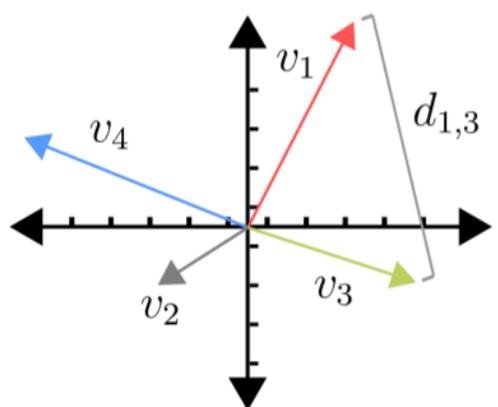
Table 1: Test per character Accuracy on Program Evaluation and Memorization tasks.

Model	Add	Control	Program	Copy	Reverse	Double
LSTM [3, 37]	99.8	97.4	66.1	99.8	99.7	99.7
EntNet [38]	98.4	98.0	73.4	91.8	100.0	62.3
DNC [5]	99.4	83.8	69.5	100.0	100.0	100.0
Relational Memory Core	99.9	99.6	79.0	100.0	100.0	99.8

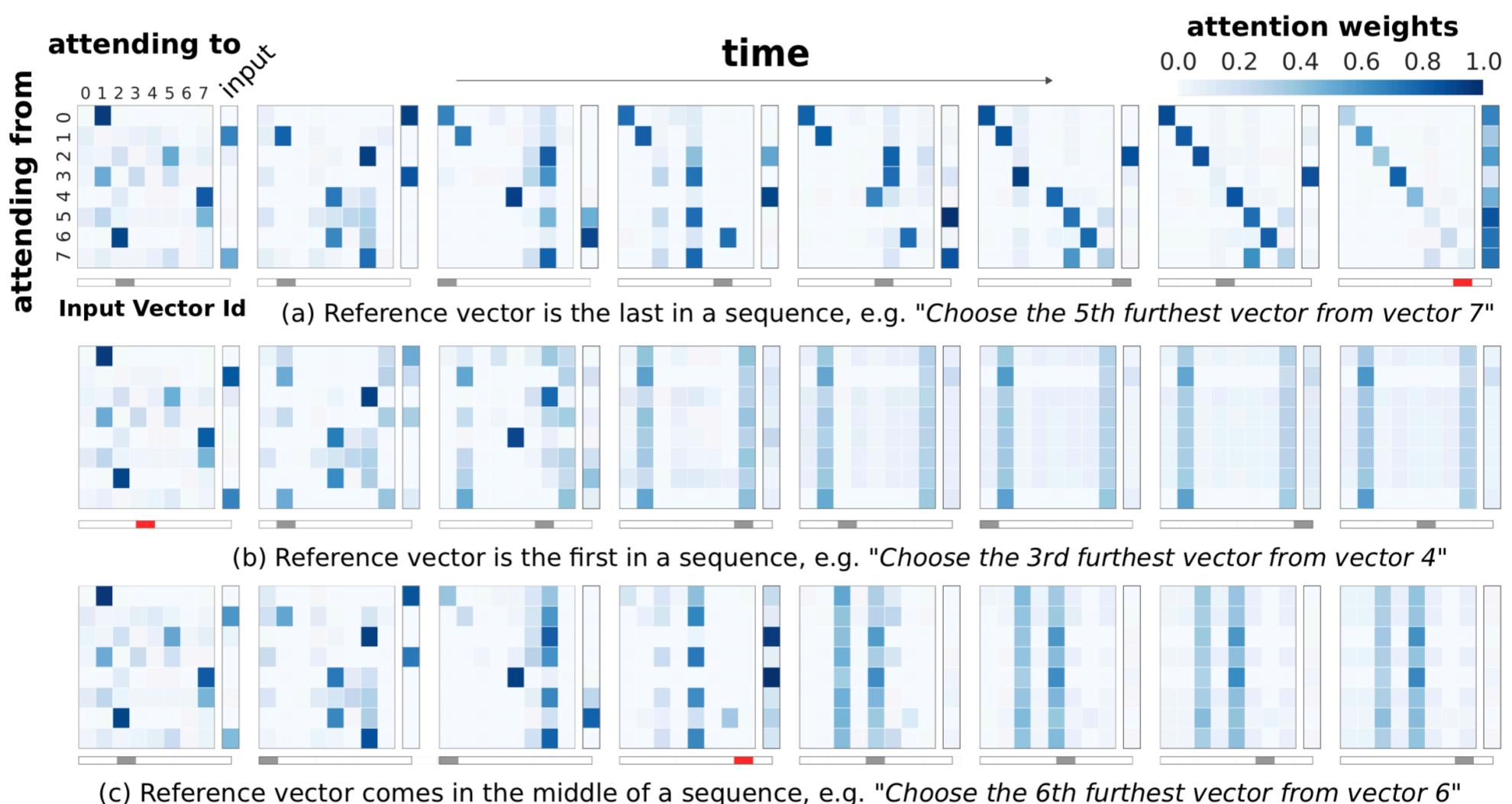
Table 2: Validation and test perplexities on WikiText-103, Project Gutenberg, and GigaWord v5.

	WikiText-103 Valid.	WikiText-103 Test	Gutenberg Valid	Gutenberg Test	GigaWord Test
LSTM [40]	-	48.7	-	-	-
Temporal CNN [41]	-	45.2	-	-	-
Gated CNN [42]	-	37.2	-	-	-
LSTM [32]	34.1	34.3	41.8	45.5	43.7
Quasi-RNN [43]	32	33	-	-	-
Relational Memory Core	30.8	31.6	39.2	42.0	38.3

What is the N^{th} farthest from vector m ?



N^{th} Farthest



Sudoku

5	3			7				
6		1	9	5				
	9	8			6			
8		6			3			
4		8	3		1			
7		2			6			
6			2	8				
	4	1	9			5		
	8			7	9			



5	3	4	6	7	8	9	1	2
6	7	2	1	9	5	3	4	8
1	9	8	3	4	2	5	6	7
8	5	9	7	6	1	4	2	3
4	2	6	8	5	3	7	9	1
7	1	3	9	2	4	8	5	6
9	6	1	5	3	7	2	8	4
2	8	7	4	1	9	6	3	5
3	4	5	2	8	6	1	7	9

Source: <https://rasmusbergpalm.github.io/recurrent-relational-networks/>

Source: <https://rasmusbergpalm.github.io/recurrent-relational-networks/>

	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3			1 2 3
4	4 5 6	4 5 6	4 5 6	4 5 6	4 5 6	4 5 6	6	9	4 5 6
	7 8 9	7 8 9	7 8 9	7 8 9	7 8 9	7 8 9			7 8 9
	1 2 3	1 2 3	3	2	1 2 3		1 2 3	1 2 3	1 2 3
	4 5 6	4 5 6			4 5 6	4	4 5 6	4 5 6	4 5 6
	7 8 9	7 8 9			7 8 9		7 8 9	7 8 9	7 8 9
	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3	1	1 2 3
	4 5 6	4 5 6	4 5 6	4 5 6	4 5 6	4 5 6	4 5 6		4 5 6
	7 8 9	7 8 9	7 8 9	7 8 9	7 8 9	7 8 9	7 8 9		7 8 9
	1 2 3	3		1 2 3	1 2 3	1 2 3	1 2 3	1 2 3	2
	4 5 6			4 5 6	4 5 6	4 5 6	4 5 6	4 5 6	
	7 8 9		9	7 8 9	7 8 9	7 8 9	7 8 9	7 8 9	
	1 2 3	1 2 3	1 2 3		1 2 3	1 2 3		1 2 3	1 2 3
	4 5 6	4 5 6	4 5 6		4 5 6	4 5 6	5	4 5 6	4 5 6
	7 8 9	7 8 9	7 8 9	8	7 8 9	7 8 9		7 8 9	7 8 9
	1 2 3	1		1 2 3	1 2 3	1 2 3		1 2 3	1 2 3
	4 5 6			4 5 6	4 5 6	4 5 6	4 5 6	4 5 6	4 5 6
	7 8 9			7 8 9	7 8 9	7 8 9	7 8 9	7 8 9	7 8 9
	1 2 3	1 2 3	1 2 3	1 2 3		1	1 2 3	1 2 3	1 2 3
	4 5 6	4 5 6	4 5 6	4 5 6			4 5 6	4 5 6	4 5 6
	7 8 9	7 8 9	7 8 9	7 8 9	9		7 8 9	7 8 9	7 8 9
	1 2 3	1 2 3	1 2 3	1 2 3			7 8 9	7 8 9	7 8 9
	4 5 6	4 5 6	4 5 6	4 5 6			7 8 9	7 8 9	7 8 9
	7 8 9	7 8 9	7 8 9	7 8 9			7 8 9	7 8 9	7 8 9
	1 2 3	1 2 3	1 2 3	1 2 3		1	1 2 3	1 2 3	1 2 3
	4 5 6	4 5 6	4 5 6	4 5 6		6	4 5 6	4 5 6	4 5 6
	7 8 9	7 8 9	7 8 9	7 8 9		7 8 9	7 8 9	7 8 9	7 8 9
	1 2 3	1 2 3	1 2 3	1 2 3		1 2 3	1 2 3	1 2 3	1 2 3
	4 5 6	4 5 6	4 5 6	4 5 6			4 5 6	4 5 6	4 5 6
	7 8 9	7 8 9	7 8 9	7 8 9			7 8 9	7 8 9	7 8 9
	1 2 3	1 2 3	1 2 3	1 2 3		1	1 2 3	1 2 3	1 2 3

- <https://papers.nips.cc/paper/7082-a-simple-neural-network-module-for-relational-reasoning.pdf>
- <http://papers.nips.cc/paper/7960-relational-recurrent-neural-networks.pdf>