

# Numerical Reasoning over Financial Data

Aparimeya Taneja, Zihao Lin

Nov. 17th 2021



# The Problem

- Large amounts of Financial statements can not be analysed at scale by humans.
- Takes substantial knowledge to ask meaningful financial questions, challenging for NLP

Page 91 from the annual reports of GRMN (Garmin Ltd.)

The fair value for these options was estimated at the date of grant using a Black-Scholes option pricing model with the following weighted-average assumptions for 2006, 2005 and 2004:

	2006	2005	2004
Weighted average fair value of options granted	\$20.01	\$9.48	\$7.28
Expected volatility	0.3534	0.3224	0.3577
Distribution yield	1.00%	0.98%	1.30%
Expected life of options in years	6.3	6.3	6.3
Risk-free interest rate	5%	4%	4%

... The total fair value of shares vested during 2006, 2005, and 2004 was \$9,413, \$8,249, and \$6,418 respectively. The aggregate intrinsic values of options outstanding and exercisable at December 30, 2006 were \$204.1 million and \$100.2 million, respectively. ( ... abbreviate 10 sentences ... )

**Question:** Considering the weighted average fair value of options , what was the change of shares vested from 2005 to 2006?

**Answer:** - 400

**Calculations:**

$$\left( \frac{9413}{20.01} \right) - \left( \frac{8249}{9.48} \right) = -400$$

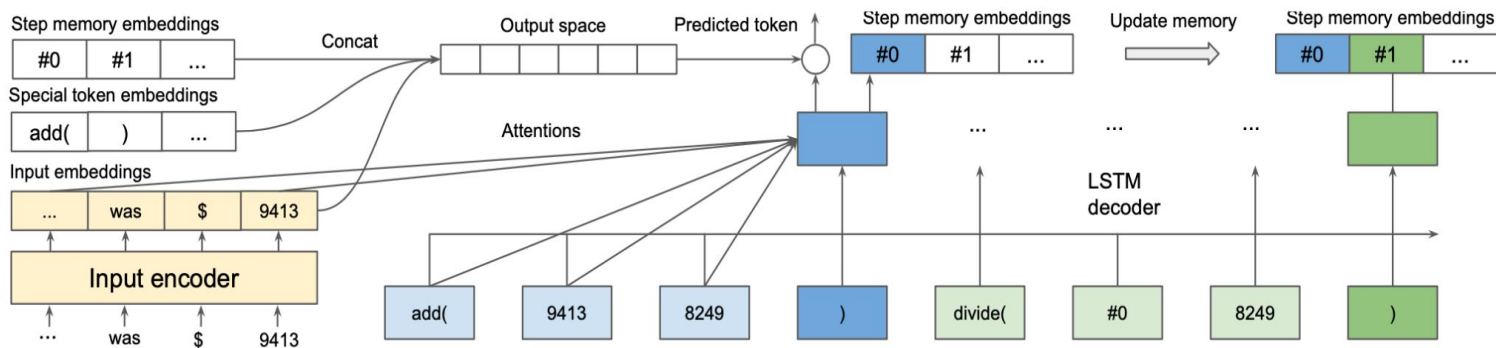
**Program:**

$$\frac{\text{divide} ( 9413, 20.01 )}{\text{divide} ( 8249, 9.48 )} \quad \text{subtract} ( \#0, \#1 )$$

# Current research

## 1) FinQA

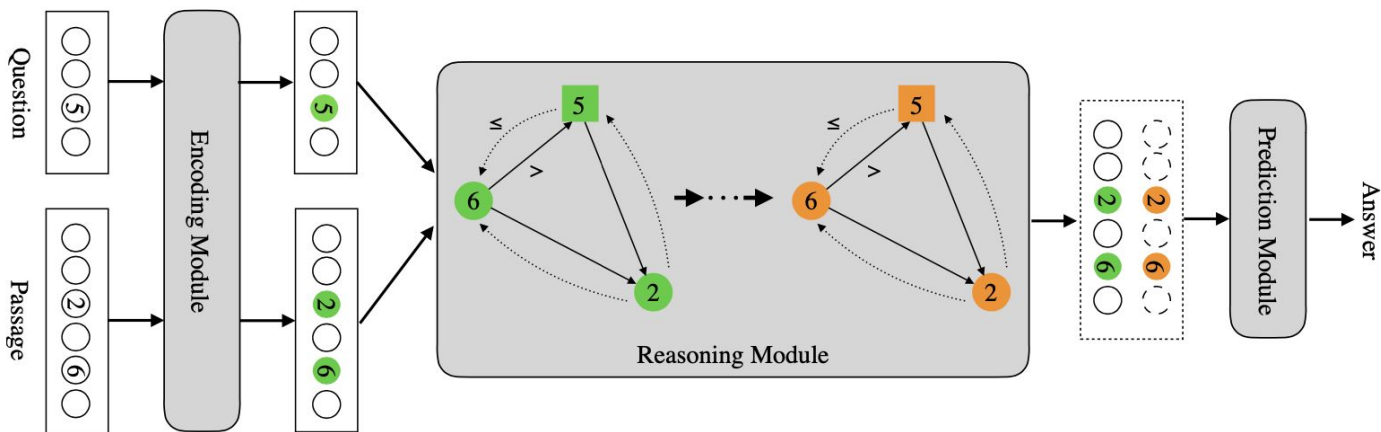
- The baseline of this task.



# Current research

## 2) NumNet

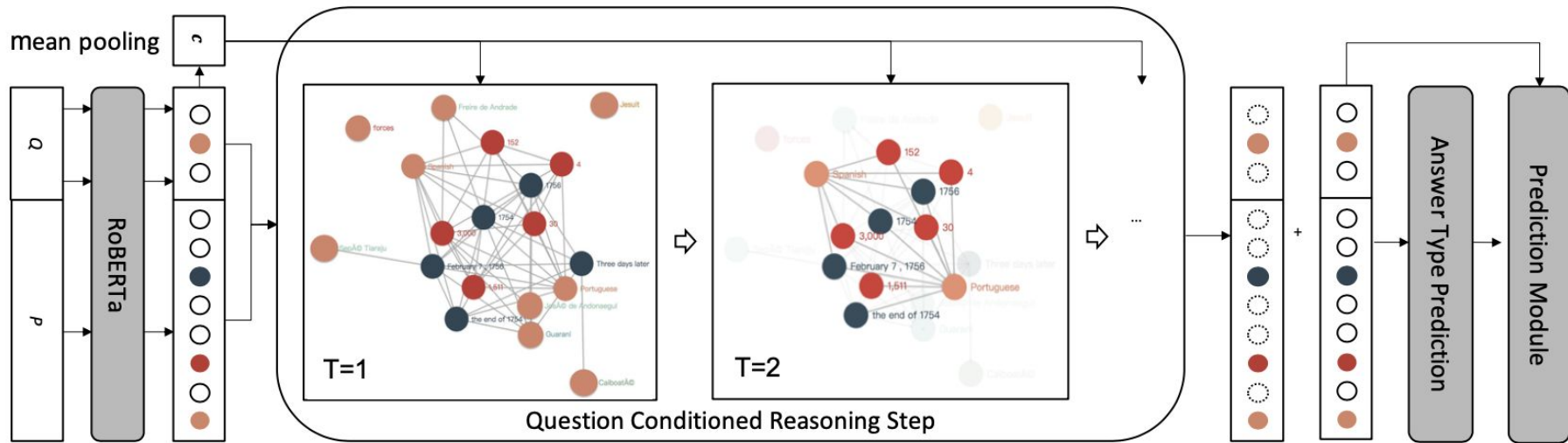
- Utilize a numerically-aware graph neural network to consider the comparing information and performs numerical reasoning over numbers in the question and passage.



# Current research

## 3) QDGAT

- SOTA algorithm for DROP which is a QA dataset which tests comprehensive understanding of paragraphs.





# Results & Challenges: FinQA

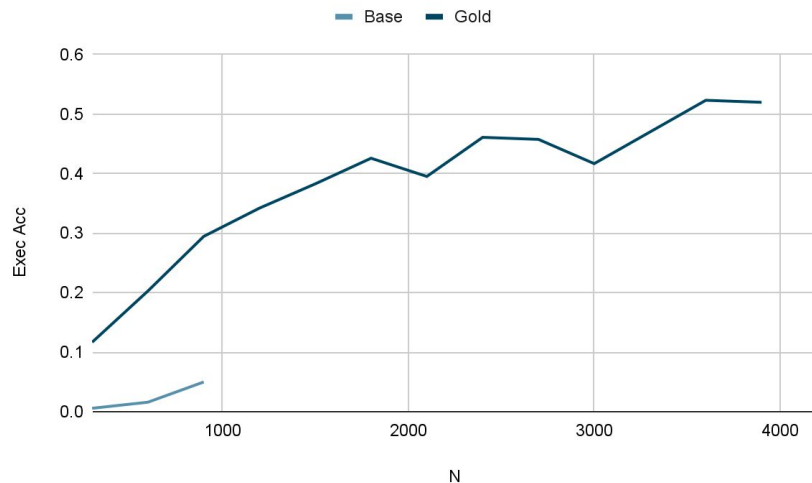
Generator

retrieve	exe acc:	prog acc:	N:
Base	4.98%	4.41%	900
Gold	52.32%	50.28%	4100

Retriever (recall)

Top 3: 0.930203040529391

Top 5: 0.9533482896704795



Challenges: Extremely slow to train & test, too many moving parts



# Results & Challenges: NumNet+

## Challenges:

- Running time is very long. Need more than three and a half hours to run just one epoch on a toy dataset (with only 135 training items).
- Google Cloud Platform, 1 NVIDIA Tesla P4
- NumNet works bad for predicting multisteps numerical reasoning problems.

# Questions