Measuring and Improving BERT Mathematical Abilities by Predicting the Order of Reasoning Supplementary Material

Piotr Piekos
University of Warsaw

Henryk Michalewski
University of Warsaw, Google

Mateusz Malinowski
DeepMind

1 AQuA-RAT examples

Question: *Two ants are moving from their farms towards each other. Ant A is moving at a speed of 9 cm per hour and ant B is moving at a speed of 6 cm per hour. If the farms are 75 cm away from each other, what will be the distance (in cm) that ant A travels until meeting ant B?*
Options: *A)45, B)48, C)51, D)54, E)57*
Rationale:

The two ants move a total of 15 cm per hour.
The time it takes until they meet is 75/15=5 hours.
*In that time, the distance that ant A travels is 5*9=45 cm.*
The answer is A.

Question: *At a certain factory, 10 percent of the staplers produced on Monday were defective and 2 percent of the non-defective staplers were rejected by mistake. If 72 of the non-defective staplers were rejected, what was the number of staplers produced that day?*
Options: *A)4000, B)4200, C)4500, D)4800, E)5000]*
Rationale:

We're told that 10% of staplers in a factory are defective.
X = Total staplers
0.1X = defective staplers
0.9X = normal staplers
Next, we're told that 2% of the normal staplers were rejected by
mistake and that this = 72 staplers.
$0.9X(0.02) = 72$
$0.018X = 72$
$18X = 72000$
$X = 4000$
Final Answer:
A

2 Question difficulty

At this section we present an example from each difficulty group for BERT+NROP and BERT. We have described the grouping procedure in the main paper.

2.1 BERT+NROP

D_5: *How many ways A boy can reach the top of stairs which contain 10 steps, when he can take either one or two steps every time?*
Answers: *A)88, B)89, C)90, D)91, E)92*
Correct Answer: *B*
Model Answer: *D*

D_4: *A square piece of cloth is trimmed by 4 feet on one edge to form a rectangular piece, which is then cut diagonally in half to create two triangles. If the area of each of triangle is 70 square feet, what was the perimeter (in feet) of the original piece of square cloth?*
Options: *A)56, B)58, C)60, D)62, E)64*
Correct Answer: *A*
Model Answer: *B*

D_3: *Train A leaves a station every 16 minutes and Train B leaves every 17 minutes. If both trains just left the station simultaneously, how long until they do so again?*
Options: *A)272 minutes, B)304 minutes, C)190 minutes, D)70 minutes, E)35 minutes*
Correct Answer: *A*
Model Answer: *B*

D_2: *10kg of a mixture contains 30% sand and 70% clay. In order to make the mixture contain equal quantities of clay and sand how much of the mixture is to be removed and replaced with pure sand?*
Options: *A)10/7, B)20/7, C)30/7, D)40/7, E)50/7*
Correct Answer: *B*
Model Answer: *C*

D_1: *If one third of 3/4 of a number is 21. Then, find the number?*
Options: *A)84, B)66, C)28, D)19, E)11*
Correct Answer: *D*
Model Answer: *D*

2.2 BERT

D_5: *The length of the ribbon was originally 30 cm. It was reduced in the ratio 5 : 3. What is its length now?*
Answers: *A)18, B)30, C)6, D)15, E)12*
Correct Answer: *A*
Model Answer: *B*

D_4: *An electric pole, 14 metres high, casts a shadow of 10 metres. Find the height of a tree that casts a shadow of 15 metres under similar conditions.*
Options: *A)21, B)22, C)20, D)23, E)24*

Correct Answer: *A*
Model Answer: *C*

D_3**:** *A rope 20 meters long is cut into two pieces. If the length of one piece of rope is 3 meters shorter than the length of the other, what is the length, in meters, of the longer piece of rope?*
Options: *A)7.5, B)8.9, C)9.9, D)11.5, E)11.7*
Correct Answer: *D*
Model Answer: *B*

D_2**:** *Jerry purchased a 1-year $5,000 bond that paid an annual interest rate of 12% compounded every six months. How much interest had this bond accrued at maturity?*
Options: *A)$5102, B)$618, C)$216, D)$202, E)$200*
Correct Answer: *B*
Model Answer: *A*

D_1**:** *I have a money pouch containing Rs. 700. There are equal number of 25 paise coins, 50 paise coins and one rupee coins. How many of each are there?*
Options: *A)453, B)651, C)400, D)487, E)286*
Correct Answer: *C*
Model Answer: *C*