

**INPUT:**

&lt;DEMONSTRATIONS&gt;

Question: Jim spends 2 hours watching TV and then decides to go to bed and reads for half as long. He does this 3 times a week. How many hours does he spend on TV and reading in 4 weeks?

**OUTPUT:**

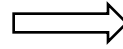
Answer: Let's think step by step.

Step1: Jim spends 2 hours ... , so he reads for  $2/2 = 1$  hour. 1.70

Step2: Jim does this 3 times a week, so he spends a total of  $2+1=3$  hours per week on TV and reading. 1.57

Step3: In 4 weeks, Jim spends a total of  $3*4=12$  hours on ... 1.71

Step4: The answer is  $\boxed{12}$ . 1.76



While not **stop\_criteria()**:

**INPUT:**

&lt;DEMONSTRATIONS&gt;

Question: Jim spends 2 hours watching TV and then decides to go to bed and reads for half as long. He does this 3 times a week. How many hours does he spend on TV and reading in 4 weeks?

Answer: Let's think step by step.

Step1: Jim spends 2 hours ... , so he reads for  $2/2 = 1$  hour.

Step2:

**OUTPUT:**

So, he spends  $2+1=3$  hours each time. 1.74

Step3: Jim does this 3 times a week, so he spends a total of  $3*3=9$  hours per week on TV and reading each week. 1.72

Step4: In 4 weeks, Jim spends a total of  $4*9=36$  hours on ... 1.69

Step5: The answer is  $\boxed{36}$ . 1.89



Iteratively appending  
**correct steps** to INPUT

Step2: Jim does this 3 times a week, so he spends a total of  $2+1=3$  hours ...

transition score

avg. confidence

divergence score

step confidence = transition + avg. conf - divergence