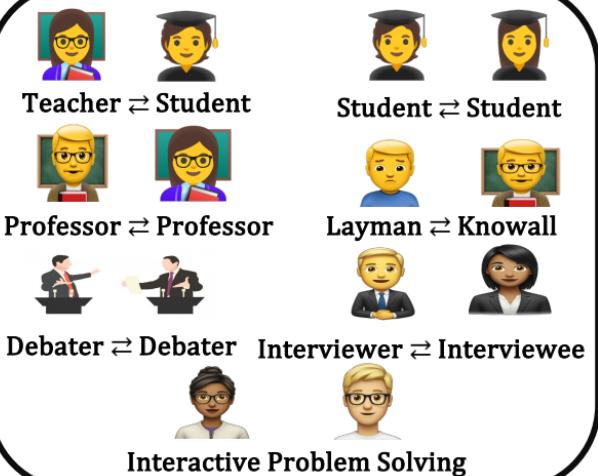
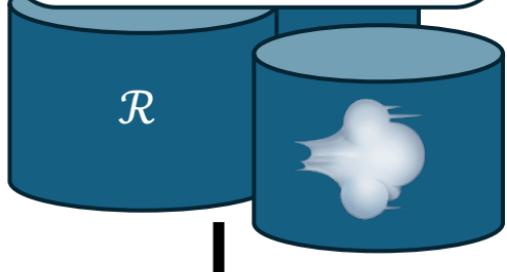


Prompts, $\mathcal{P} \in [1, 7]$

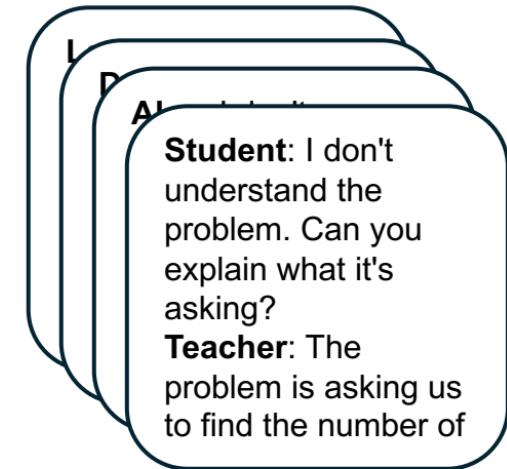


Raw Data, $r_j \in \mathcal{R}$

Browse Questions #
How many numbers are there between 20000 and 30000 in which the digits are 2,3,5,6,7 and each digit can be repeated any number of times.



Conversations, $s_{\{i,j\}} \in S$



Conversation Generator, \mathcal{M}

Generate Synthetic Dialogue Corpora using MIND
 $\mathcal{R} \rightarrow \text{MIND} \rightarrow S'$

Language Model, \mathcal{E}



Continuous Pretraining

Language Model, C



Heuristic Filtering, \mathcal{H}

Evaluate MIND generated data
 $\mathcal{E} \leftarrow \text{pretrain}(C, \mathcal{D});$
 $\mathcal{D} = \{S' \cup R_{pt}\}$