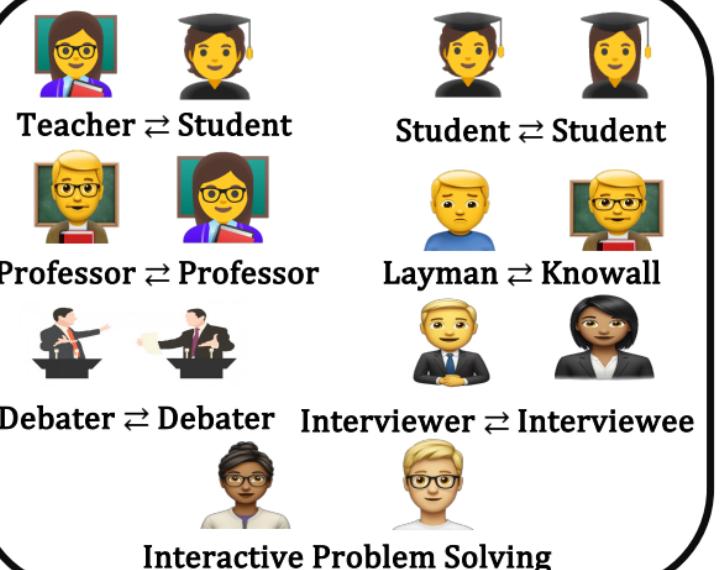


## Prompts, $\mathcal{P} \in [1, 7]$



## 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.

Open Web Math,  $\mathcal{R}$



## Conversation Generator, $\mathcal{M}$

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

Student: I don't understand the problem. Can you explain what it's asking?  
Teacher: The problem is asking us to find the number of

Syn Data,  $S$

Heuristic Filtering,  $\mathcal{H}$

Syn Data,  $S'$

Language Model,  $\mathcal{E}$



Continuous Pretraining

Language Model,  $C$



Syn Data,  $S'$