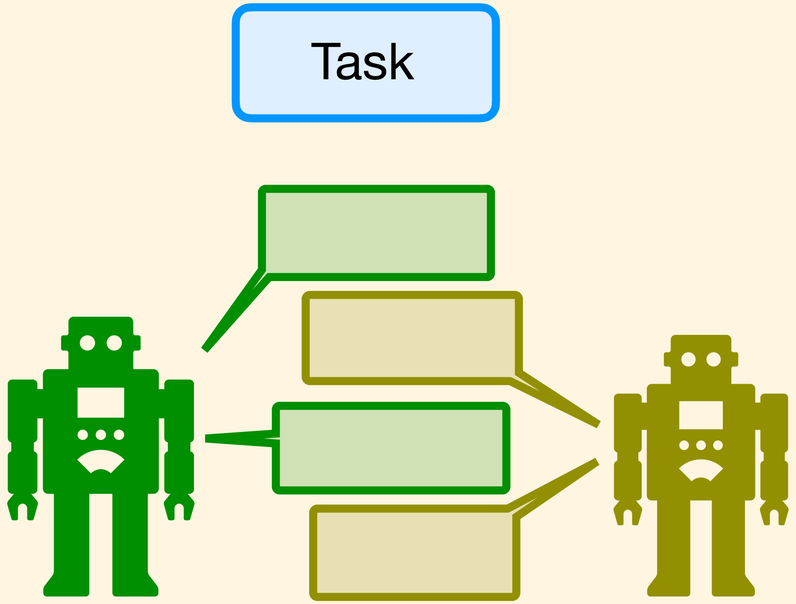
ACC-Collab: An Actor Critic Approach to Multi-Agent LLM Collaboration

Andrew Estornell, Jean-François Ton*, Yuanshun Yao, Yang Liu*

ByteDance

LLM Collaboration



<u>Question:</u> How can we enhance the collaborative abilities of LLMs?

What is $\int \tan^2(x)dx$?

$\tan^2(x) = \sin^2(x) - 1$ so

$$\int \tan^2(x)dx = -\frac{x}{2} - \frac{\sin(2x)}{4} + C$$

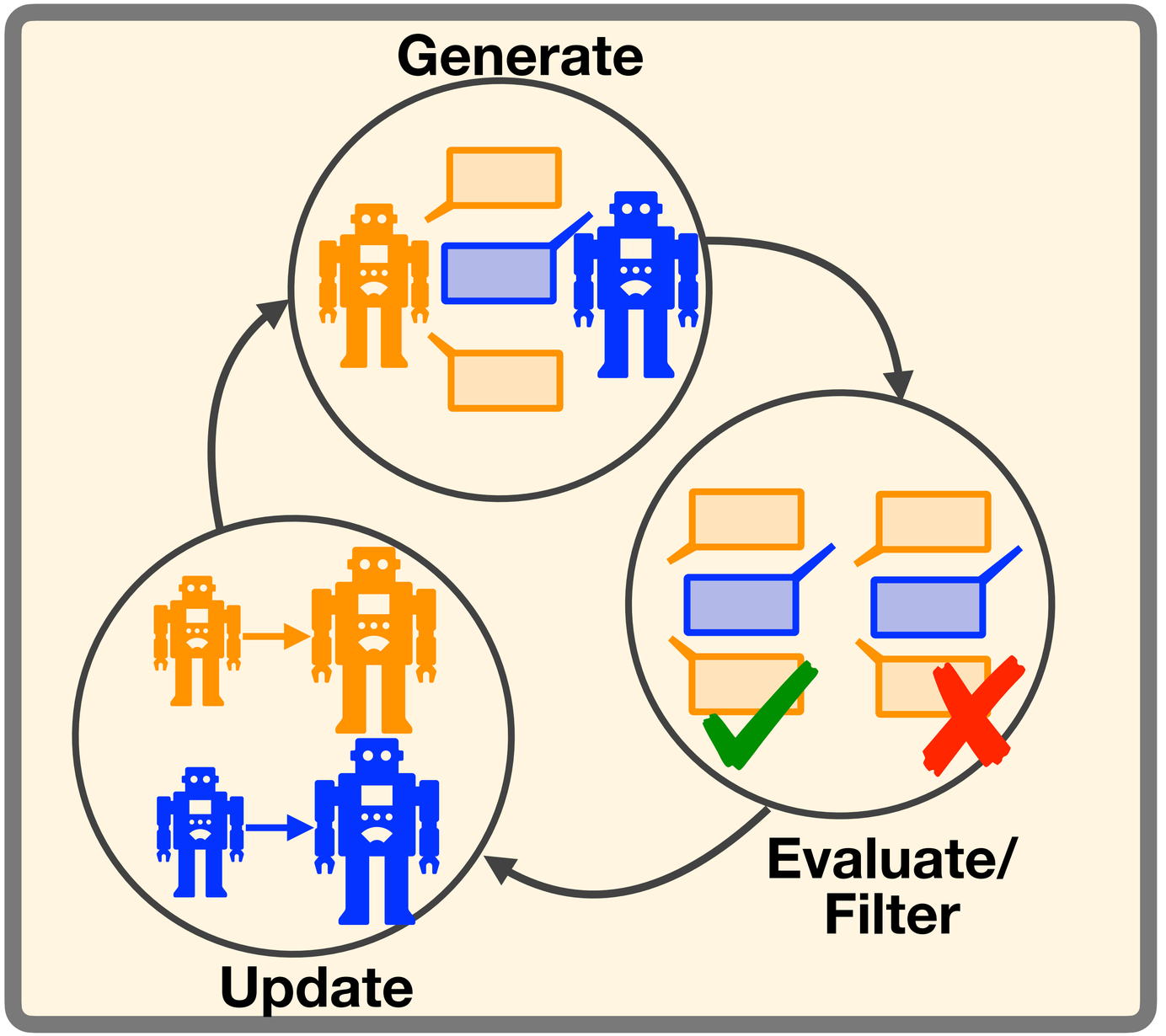
Actor

Actually,
$\tan^2(x) = \sec^2(x) - 1$

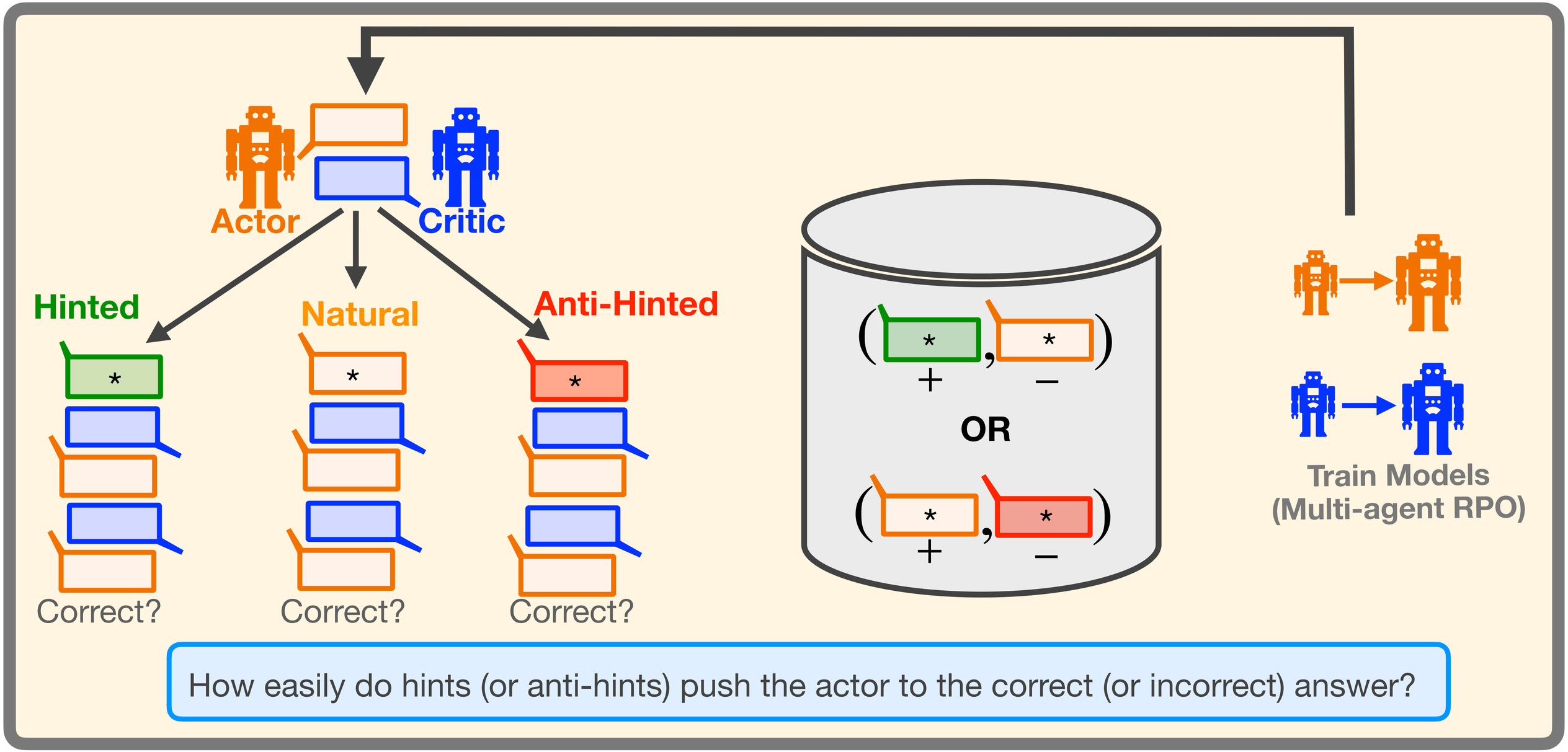
Critic

Oh, then
$$\int \tan^2(x)dx = \tan(x) - x + C$$

How to train a good team?

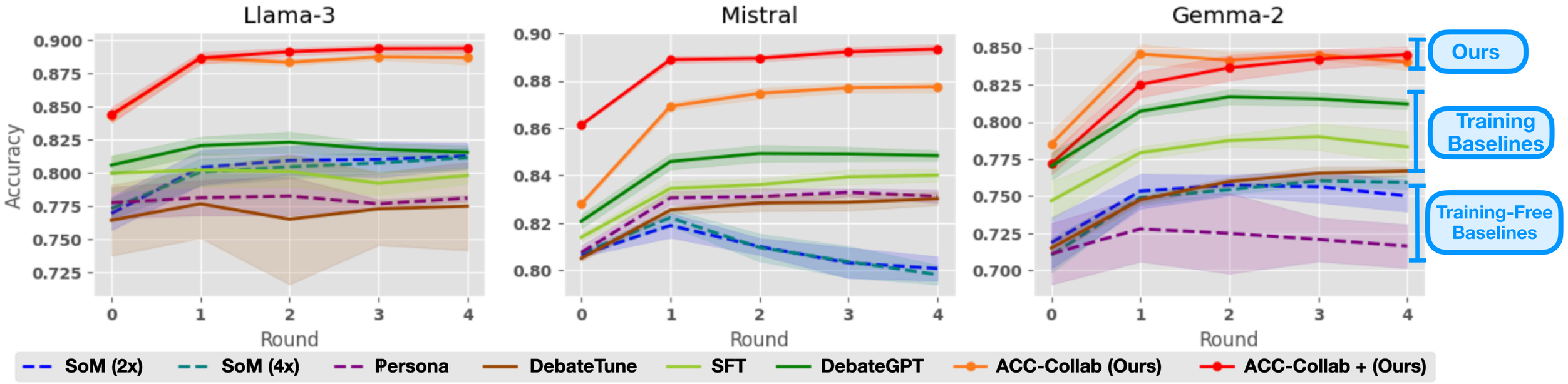


How to train a good team?

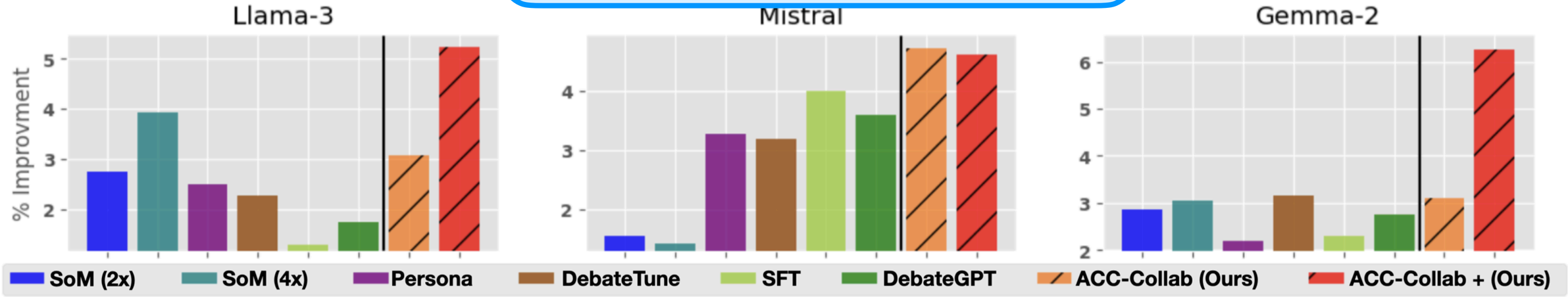


How easily do hints (or anti-hints) push the actor to the correct (or incorrect) answer?

Team Performance on Benchmarks

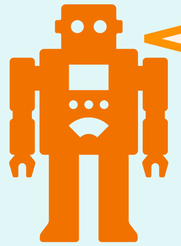


Average % improvement from round 0 to final round



What Does the Critic Actually Learn?

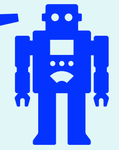
If bee populations continue to decline, how might plants adapt to ensure pollination?



Actor

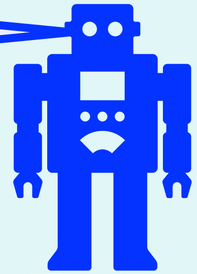
With the decline of honey bee population . . . . .
Final Answer: (B) change their flowers so that wind will fertilize them.

Your response is mostly correct . . .
{continues to agree with the actor}.



Untrained Critic

While your response is a valid explanation there are a few issues,
. . .{Counterpoint to for why B could be wrong}
. . .{Evidence for why C is correct}



Trained Critic

Check out the paper!

ACC-Collab: An Actor Critic Approach to Multi-Agent LLM Collaboration

Andrew Estornell, Jean-François Ton*, Yuanshun Yao, Yang Liu*