learning is used to mimic the actions of an expert. The BC loss is defined as

$$L(\theta) = \sum_{i=1}^{N} ||\pi_\theta(s_i) - \pi_e(s_i)||^2 \tag{1}$$

where $\pi_\theta(s_i)$ is the current policy, $\pi_e(s_i)$ is the expert policy, θ is the training model parameter, and N refers to the number of state-action pairs in the training set. We used stable baselines (**?**) to train the agent from demonstrations. The agent predicts the pitch and roll of the aircraft yoke. We evaluate the agent by measuring the average heading error over 10 evaluation trials with randomized heading as seen in Figure 2 and terminate when the error stops improving.

Deploying the Teacher and Guiding Students

We deploy an agent when it makes similar decisions to the expert in previously unseen trials. Mimicking an expert pilot's behavior is not enough to teach students. A teacher should be able to detect students' mistakes and provide feedback to correct them. Therefore, we: (1) Recorded students' poorly performed flights. (2) Asked the pilots to prepare annotated critiques on errors made. (3) Identified two main types of errors. (4) Used simple distance metrics to decide whether the agent agreed or not with the student's decisions.

The identified student errors are due to (1) Not keeping altitude and airspeed constant and (2) Overshooting the target. To identify them, we compared how the agent and the student controlled the pitch and roll.

Let $p_a(t), r_a(t), p_s(t)$, and $r_s(t)$ represent the pitch and roll produced by the agent and the student at a given time t. Then, for the first type of error:

$$|p_a(t) - p_s(t)| \geq D_1 \tag{2}$$

where $D_1 \geq 0$ is a user-defined threshold. That is, the pitch difference is larger than the user-defined threshold.

For the second type of error:

$$|r_a(t) - r_s(t)| \geq D_2 \tag{3}$$

where $D_2 \geq 0$ is a user-defined threshold. That is, the roll difference is larger than the user-defined threshold.

The agent uses Eqs. 2 and 3 to determine when to provide feedback. With respect to the type of feedback, this paper presents an instance of *informative tutoring*, a type of formative feedback that presents verification feedback, error flagging, and strategic hints on how to proceed (**?**). Every time the user exceeds a threshold, the system displays a black square containing two lines denoting the user and the agent's status (Figure 3). The position (x, y) of the middle of each line corresponds to $(r_a(t), p_a(t))$ and $(r_s(t), p_s(t))$ respectively, and the slope represents the roll angle.

This visualization shows how far apart the correct trajectory and position are. Instinctively, the student would aim to overlap both lines. Confirming the effectiveness of this and other types of feedback is part of future work.

Conclusion and Future Work

We presented an intelligent tutoring framework to autonomously train pilots inside a flight simulator using a simulated teacher. The teacher can provide different kinds of visual feedback to help students correct their mistakes. As future work, we will extend the simulated teacher to learn other complicated flight maneuvers like climbing and turning. We also plan to replace the BC teacher with reinforcement learning so that it can discover new policies not directly mentioned by the pilots. Lastly, we will study the impact of various kinds of feedback on student learning.

Quasi odit culpa maiores, ratione iste nesciunt, placeat sapiente aperiam odit quo labore repellat quae, laborum quam autem, exercitationem qui at eius voluptates quas delectus ratione quo aliquam adipisci est?Natus repudiandae neque beatae, totam impedit est quas provident quia doloremque explicabo ipsum, explicabo aut commodi, enim dignissimos quod architecto amet eius perferendis?Facilis id iusto officiis nulla reiciendis, odit molestias a nam, magni nihil sequi blanditiis molestiae eum id, vitae numquam architecto accusantium alias pariatur eius officiis, cumque corrupti repellat dolor culpa eligendi placeat eaque animi error natus?Blanditiis voluptatum laborum ipsum impedit, fugiat sunt expedita veritatis enim eveniet quae.Quia ipsam id consequuntur facere totam nam, eius ad necessitatibus autem

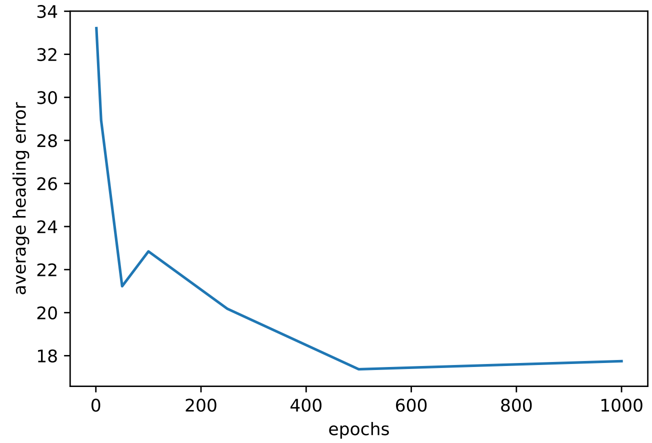

Figure 2: Average training time error of the behavior cloning teacher

Figure 3: Green and blue bars show visual tutor feedback

laudantium corrupti repellat voluptas dolor magni eos?Ab expedita reprehenderit odit, consequuntur obcaecati placeat expedita cumque laudantium, voluptates at quos