ARC Evaluation pass@2 During Post-training

Legend:
- Full fine-tune
- Embeddings only
- Embed+full (quarter)
- Embed+full (half)
- LoRA

X-axis: Training step ($\times 10^3$)
Y-axis: Evaluation pass@2