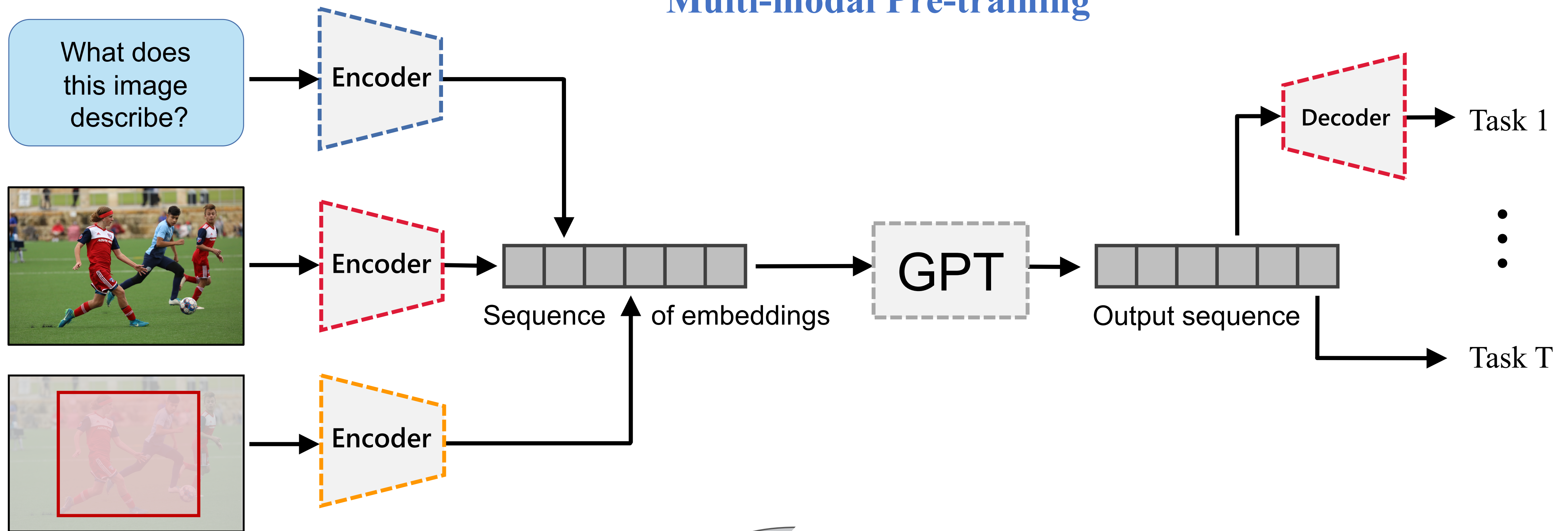
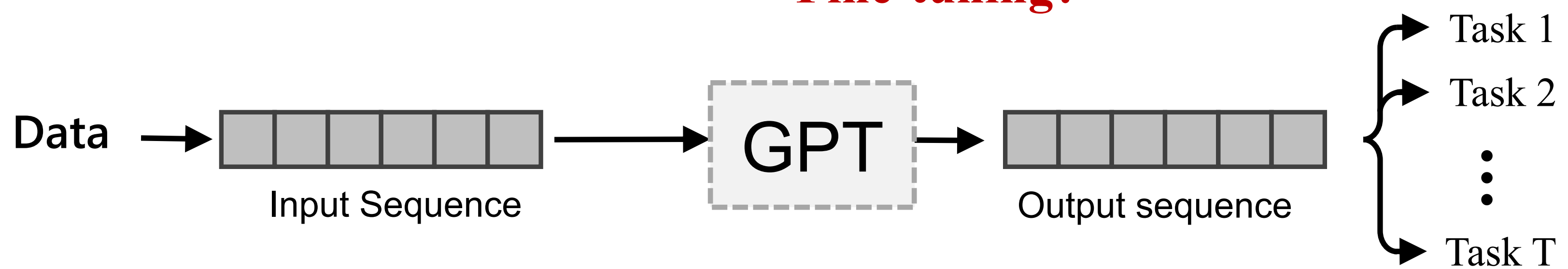


Multi-modal Pre-training



Fine-tuning?



Prediction on Multiple tasks

