



Human

VLaMP



I want to build shelf.

"What should I do next?"

untrimmed
visual history



Your next steps:
1. cut shelf
2. sand shelf
3. paint shelf

Segmentation

Forecasting



Split
History

Observations

$[o_1, \dots, o_t]$

Actions

$[a_1, \dots, a_t]$

Obs.
Encoder

Action
Encoder

Pretrained
Language
Model

Goal Prompt: "Build a Shelf"