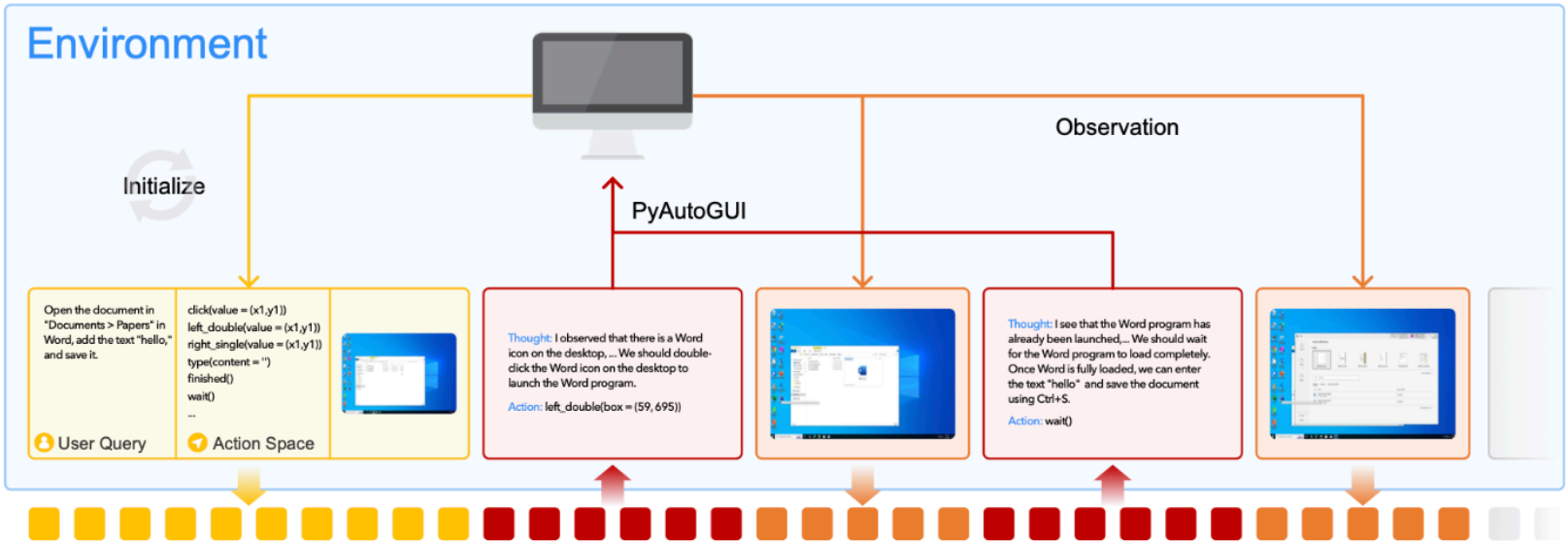
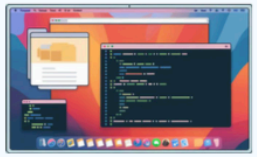


Capability

Perception



- Element Description
- Dense Captioning
- Transition Captioning
- Question Answering
- Set-of-Mark

Action

- Unified Action Space
  - Click
  - Type
  - ...
- \+ Our Annotated Dataset
- \+ Open-Source Data
  - •AITZ ...
  - •AITW ...
- → Multi-Step Trajectory Data

System-2 Reasoning

- Reasoning Enrichment with GUI Tutorials
- Reasoning Stimulation with Thought Augmentation

Learning from Prior Experience

- Online Trace Bootstrapping & Reflection Tuning
- ⟳ Agent DPO