

This repository contains:

1. Code for all models
2. User experiment data

(1) Code for all models

Each Python file begins with `[map_name]`, where `[map_name]` corresponds to one of the four maps we use. The functionality is consistent across maps. Below we use **MapA** as an example to describe the code:

Training TrustPOMDP-related models:

- `[MapA]train_trustee_agent_highB_highI_version1.py`: trains the first trustee agent with high Benevolence and high Integrity.
- `[MapA]train_trustee_agent_highB_highI_version2.py`: trains the second trustee agent with high Benevolence and high Integrity.
- `[MapA]train_trustee_agent_highB_lowI.py`: trains a trustee agent with high Benevolence and low Integrity.
- `[MapA]train_trustee_agent_lowB_highI.py`: trains a trustee agent with low Benevolence and high Integrity.
- `[MapA]train_trustee_agent_lowB_lowI.py`: trains a trustee agent with low Benevolence and low Integrity.

All trained trustee agents are saved under the `final_trained_models` directory.

Once the trustee agents are trained:

- `[MapA]train_ABI_inference_model.py` runs the trustee agents to train the ABI inference model. Outputs are saved under `[MapA]abi_model_outputs_Boltzmaan_10times_adaptivelength_FFFFFFFF FFFFFFFF FINAL_with_uncertainty`.
- `[MapA]train_TrustPOMDP.py` trains the TrustPOMDP agent, with the trained models saved under `final_trained_models`.

Training baseline models:

- `[MapA]train_SP.py`: trains the SP agent population required for the first stage of the FCP model.
- `[MapA]train_FCP.py`: trains the FCP model using the SP agents.
- `[MapA]train_MEP_population.py`: trains the agent population required for the first stage of the MEP model.
- `[MapA]train_MEP.py`: trains the MEP model.

- `[MapA]train_POMDP.py`: trains the basic POMDP model.

Model testing:

- `[MapA]test_TrustPOMDP_with_rule_based_agents.py`: tests the TrustPOMDP model using rule-based agents.
- `[MapA]test_baselines_with_rule_based_agents.py`: tests the baseline models using rule-based agents.

(2) User experiment data

The raw data from the human-subject experiment is stored under the Human Experiment Data directory, containing 102 JSON files.

You can run `extract_overcooked_logs.py` to analyze and visualize the user experiment data.

The trained models exceed the file size limit for supplementary materials. Please train the models yourself based on the provided code, or access them later when we upload them to an external repository.