

Example files of each category of code run during the project are shared in this zip file.

1. ConsensusGame\_DecisionBased folder contains example code for collecting LLM answers to decision-based questions, performing the Consensus Game on these answers, and calculating the accuracy of the LLMs before and after the Consensus Game is performed.
  1. Collecting\_LLMDecisions folder contains three files (CorrectGenerator\_decisionsA.py, InCorrectGenerator\_decisionsA.py, and Disc\_decisionsA.py) that call the code for collecting answers from the correct Generator, incorrect Generator, and Discriminator (03\_14\_26\_CorrectGen\_decisionsA.py, 03\_14\_26\_IncorrectGen\_decisionsA.py, and 03\_14\_26\_Disc\_decisionsA.py)
  2. ConsensusGame folder contains a file (03\_26\_26\_policies\_decisions.py) for calling the code (policies.py) that performs the consensus game on the data gathered by the Collecting\_LLMDecisions code.
  3. getaccuracies\_decisionsA.py contains sample code for calculating the accuracy of initial LLM answer and the updated policies.
2. Dido\_Aeneas folder contains code for generating a sample story based on the opera 'Dido and Aeneas' by Henry Purcell.
  1. Generative\_Algorithm.py runs contains the overall workflow for A) generating possible choices at each decision node, B) collecting LLM decisions from the correct Generator, incorrect Generator, and Discriminator, and C) running the Consensus Game to find the optimal decision based on the LLMs' answers.
3. pixi-py folder contains the code for validating the rest of the code through a series of unit tests.
  1. tests/test\_me.py contains functions that tests whether a certain value is obtained by calling functions from our code. Functions that we tested in our project are contained in src/pixi\_py/\_\_init\_\_.py.