**LLMs and ToM in context outline**

**Simple prompt engineering implementation:**

**Outline:** Build a model that gives advice on employees’ workplace problems, via integrating POMDP reasoning. Initially, this reasoning will be implemented via prompt engineering, where the prior distributions will not be informed from data. There will be 2 sets of states to inform

**States:** Tuple of mental states of the employee, e.g. “angry”, “confused”, etc. and beliefs of the employee about a different character, e.g. “employee thinks his manager is angry”. Beliefs about a different character will be LLM generated.

**Observations:** input or response from employee, or LLM generated observation from input.

**Action:** System response.

**Reward:** Implicitly given, through prompting system to give constructive, sensitive responses

**Prompt engineering implementation with 2nd order ToM reasoning:**

**Outline:** Extend this model to include 2nd order reasoning, by including an additional set of states for the mental state that the employee believes another person is in (e.g. employee believes their boss is angry)