

# 1 Current Formulation

## 1.1 POMDP

- $S = \langle \omega, \mathcal{O}, ds \rangle$
- $A = \{\text{wait}, \text{pick}(\text{object}), \text{ask}(\text{property})\}$
- $T = p(\omega', \mathcal{O}', ds' | \omega, \mathcal{O}, ds, a)$

If  $a$  is a pick action and  $a.\text{object} == \omega$ :

$$p(\omega', \mathcal{O}', ds' | \omega, \mathcal{O}, ds, a) = \begin{cases} 1/|\mathcal{O} \setminus \omega| & \text{if } \omega'! = \omega, ds = \text{None}, \mathcal{O}' = \mathcal{O} \setminus \omega \\ 0 & \text{otherwise} \end{cases}$$

If  $a$  is a pick action and  $a.\text{object}! = \omega$ , we transition to the state where  $s' = \langle \omega', \mathcal{O}', ds' \rangle = \langle \omega, \mathcal{O} \setminus a.\text{object}, ds \rangle$ .

•

$$R(\omega, \mathcal{O}, ds, a) = \begin{cases} 10 & \text{if } a == \text{pick and } a.\text{object} == \omega \\ -80 & \text{if } a == \text{pick and } a.\text{object} != \omega \\ -1 & \text{if } a == \text{wait} \\ -3 & \text{if } a == \text{ask} \end{cases}$$

- $\Omega = \langle l, g \rangle$
- $O = p(l, g | \omega, \mathcal{O}, ds, a) = p(l | \omega, \mathcal{O}, ds, a) p(g | \omega, \mathcal{O}, ds, a)$

$$p(l | \mathcal{O}, \omega, a) = p(c | \omega^{t+1}, a^t) \prod_{w \in l} p(w | \omega, a) \quad (1)$$

$$p(c | a^t) = \begin{cases} 0.8 & \text{if } a^t \text{ is a question} \\ 0.2 & \text{otherwise} \end{cases} \quad (2)$$

$$p(w | \omega, a) = \frac{\text{Number of times } w \text{ is used to describe } \omega}{\text{Total counts for words describing } \omega}$$

etc etc etc

Currently this is being solved with a hacked together cached method, where Belief Sparse Sampling is being run in simulation, and belief states and their corresponding actions are cached and reused for running interactively.

This can also probably be done using POMCP.

## 2 Incremental Actions

## 3 Theory of Mind

### 3.1 Robot Model

The robot acts according to the following model:

- $S = \langle \omega, \mathcal{O}, ds, B_h \rangle$
- $A = \{\text{wait}, \text{pick}(\text{object}), \text{ask}(\text{property})\}$
- $T = p(\omega', \mathcal{O}', ds' | \omega, \mathcal{O}, ds, a)$

If  $a$  is a pick action and  $a.\text{object} == \omega$ :

$$p(\omega', \mathcal{O}', ds' | \omega, \mathcal{O}, ds, a) = \begin{cases} 1/|\mathcal{O} \setminus \omega| & \text{if } \omega' = \omega, ds = \text{None}, \mathcal{O}' = \mathcal{O} \setminus \omega \\ 0 & \text{otherwise} \end{cases}$$

If  $a$  is a pick action and  $a.\text{object} \neq \omega$ , we transition to the state where  $s' = \langle \omega', \mathcal{O}', ds' \rangle = \langle \omega, \mathcal{O} \setminus a.\text{object}, ds \rangle$ .

•

$$R(\omega, \mathcal{O}, ds, a) = \begin{cases} 10 & \text{if } a == \text{pick and } a.\text{object} == \omega \\ -80 & \text{if } a == \text{pick and } a.\text{object} \neq \omega \\ -1 & \text{if } a == \text{wait} \\ -3 & \text{if } a == \text{ask} \end{cases}$$

- $\Omega = \langle l, g \rangle$
- $O = p(l, g | \omega, \mathcal{O}, ds, a)$

Instead of having to guess what this is, we instead can sample from the policy of a POMDP/MDP that describes how the human behaves in response to the human.

### 3.2 Human Model

- $S = \langle \omega, B_h, \mathcal{O}ds \rangle$   
 $\omega$  is the item the human wants,  $B_h$  is the item the human believes the robot will hand them.
- $A = \{\text{wait}, \text{say}(\text{property}), \text{point}(\text{target})\}$