We have 4 steps:

raw input -> processed input -> thinking -> deciding

Audio: making a sound -> interpreting speech -> solving POMDP -> choosing best action

Visual: being in front of robot -> interpreting a gesture -> solving POMDP -> choosing best action

At the thinking part, all processd input comes together. There are essentially two stages, one of refining input, and the other of processing it.

Currently, we have feedback at the choosing best action stage (These are looking around, pointing, asking questions)

We would like to integrate feedback at all 4 stages, not just last

Raw input:

seeing:

if user is not present, robot is looking downward. When a person becomes present, the face title upward/eyes look upward

if user is not present, robot's screen shoulld be off. When a person becomes present, the face should dim on.

hearing:

if user starts to speak and volume is detected, the sides of baxter screen start to glow as if he is listening

if user starts to speak and volume is detected, the eyes begin to grow and shrink

Processed input:

hearing:

whenever baxter hears a word in it's vocabulary said, it's head beeps once.

seeing:

if user makes any valid gesture, baxter moves head to look at place and gaze (monkey brain)

Thinking:

when Baxter's max probability changes, a light starts blinking on his head for a few seconds to represent a change in understanding

when Baxter's max probability changes, baxter open and closes it's grippers.

Example of humans:

raw input:

- seeing/hearing:

change posture/eye gaze in order to best find you. If I drop this continual feedback of dynamic posture, it indicates i no longer acknowledge you

processed input:

seeing:

- if you gesture somewhere, my eye follow where you point, even if i know it doesn't matter

- hearing:

- based on rhythm and logical understanding