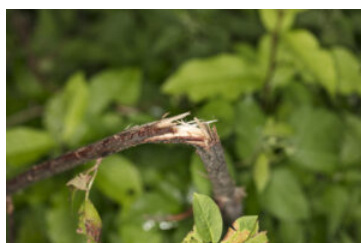


Motivation: pointing triggers richer inferences of observations

Hunting: when someone points to an observation to you, you are more likely to think its caused by your target.



Individual: caused by wind.

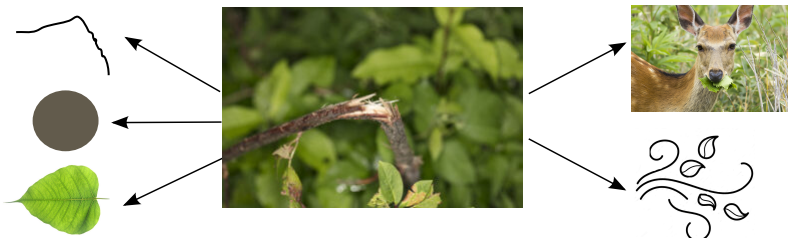
Joint: caused by target.

Goal: model pointing as foundation of human intelligent visual communication

Challenge: pointing is indirect and overloaded

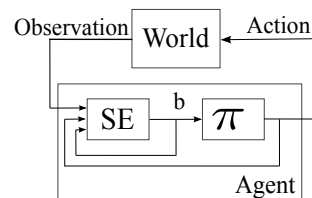
Overloadedness:
What Is she pointing to?

Indirectness:
What is she pointing to suggest?



Our model should be able to interpret from possible meanings.

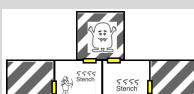
Individual perception model: POMDP



1. Bayesian update of belief (b) based on observation
2. Use belief to derive a policy (π) to maximize expected utility (V)

Task: Wumpus hunt

Wumpus: static, invisible, emits stench nearby
Hunter: can smell stench
Hunter gets a reward if shoots and hits Wumpus, a penalty if misses
Hunter can complete the task using POMDP.



Assumption: Pointing must be Relevant

The **signaler** and the **receiver** should have the mutual assumption: pointing should provide information with high relevancy.

If the hunter thinks the branch is broken by a wind,



The branch being broken by the target is of high relevancy to the situation. ✓



The branch being broken by a wind is of low relevancy to the situation. ✗

The pointing must indicate that the branch is broken by the target.

Smithian coordination of beliefs

Relevancy of Receiver's belief b and Signaler's belief s

$$R(b, s) = V^*(s) - V_{\pi(b)}(s)$$

$V_{\pi(b)}(s)$ is the evaluation of **receiver's** action using **signaler's** belief

$\pi(b)$ is the prediction of **receiver's** action using **receiver's** belief

Use **Receiver's** belief b to predict action, then use **Signaler's** knowledge s to evaluate the action

The name comes from Adam Smith's discussion of empathy: you should feel bad for a mentally deficient person, even he though cannot feel it



Smithian coordination of mind is extended to Paternalistic helping: Provide what I think is good to you, not what you want

Smithian value of information

$$SVI(u) = V_{\pi(b'(u))}(s) - V_{\pi(b)}(s)$$

Signaler evaluates **Receiver's** belief change, then calculate how much a pointing signal can help

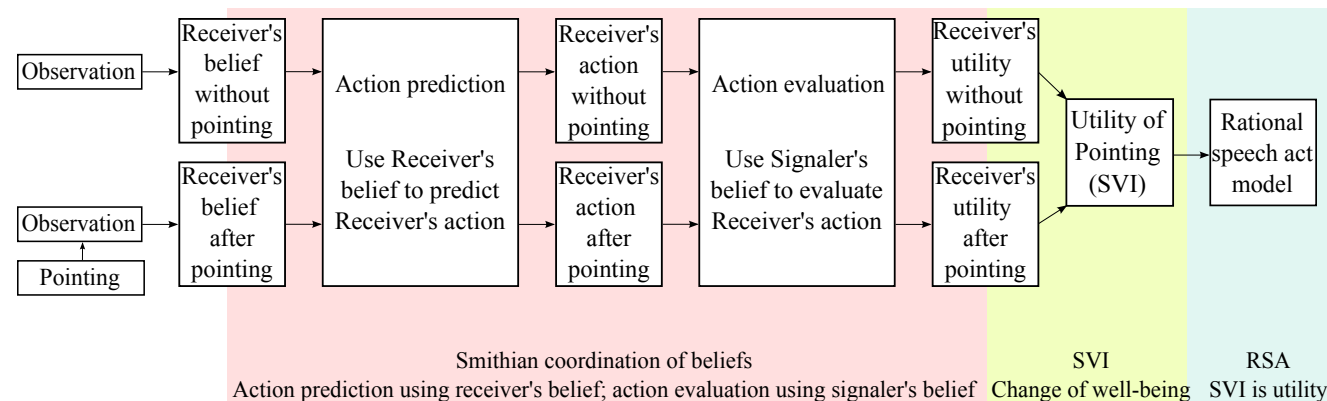
Relevancy as utility

$$P_{Sig}(u|s) \propto \exp\{\alpha SVI(u)\}$$

$$P_{Rec}(s|u) \propto P_{Sig}(u|s)P(s)$$

SVI is the utility of pointing. We can use RSA to model the generation and interpretation of pointing.

Joint perception model: Smithian pointing model



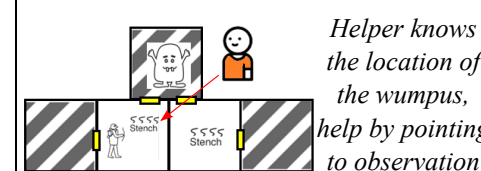
Connection to POMDP

- **Receiver's** belief update without pointing is calculated by POMDP
- **Receiver's** action is predicted by the **signaler** using a POMDP solver on **receiver's** belief

Connection to RSA

- Relevancy calculated using Smithian coordination of beliefs is independent of RSA
- The Smithian value of information can be used as utility function of the pointing signal in RSA

Experiment



Experiment 1: Relevancy

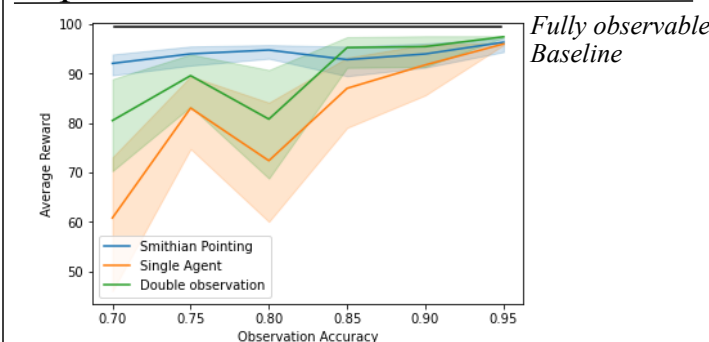
- Individual perception Use POMDP model
- Joint perception Use Smithian pointing model
- Double observation Give hunter another observation Control information received

Experiment 2: Recursion

- Use Smithian pointing model Manipulate level of recursion
- Level-1 signaler, Level-1 receiver
- Level-2 signaler, Level-2 receiver
- Level-2 signaler, Level-1 receiver

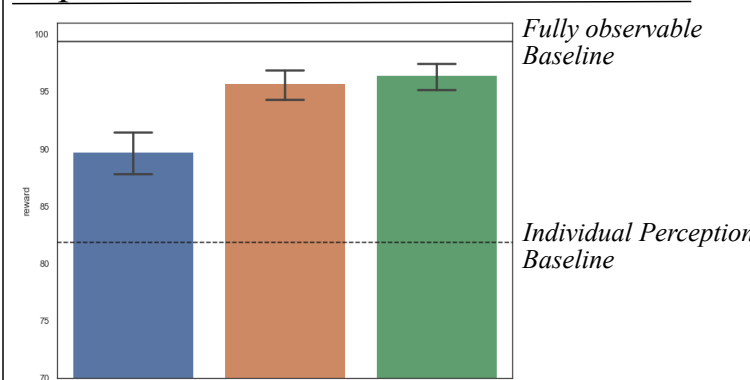
Results

Experiment 1 Relevancy calculation in pointing



Smithian model of pointing improves performance, suggesting the power of pointing comes from inference on relevancy.

Experiment 2 Recursion level



- Even shallow recursion can improve performance.
- Deeper recursion can further improve performance.

Conclusion

Relevancy model of pointing

- We provide a model in pointing using Smithian definition of **relevancy**.
- Relevancy calculation enables **overloaded**, **indirect** and **impromptu** visual communication like pointing, without a pre-defined codebook.

Smithian coordination of beliefs

- The signaler should use the receiver's belief to predict receiver's actions, and her own mind to evaluate actions