Pragmatic Reasoning in Structured Signaling Games

CLASP June 8, 2022

Emil Carlsson, Devdatt Dubhashi
Chalmers, CSE Data Science and AI

To appear at CogSci 2022

Pragmatic Reasoning

- Finding the intended meaning of an utterance based on the context
- Reasoning about the other agent's belief given the context

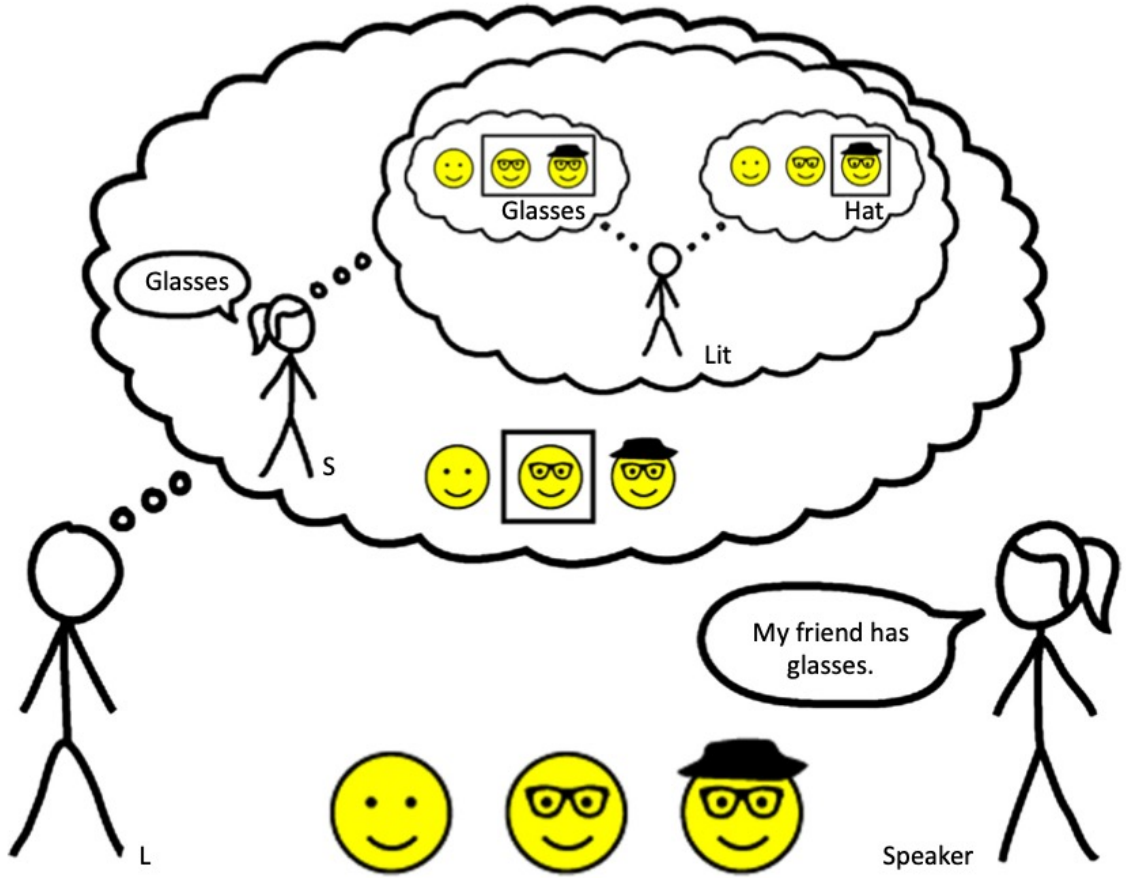
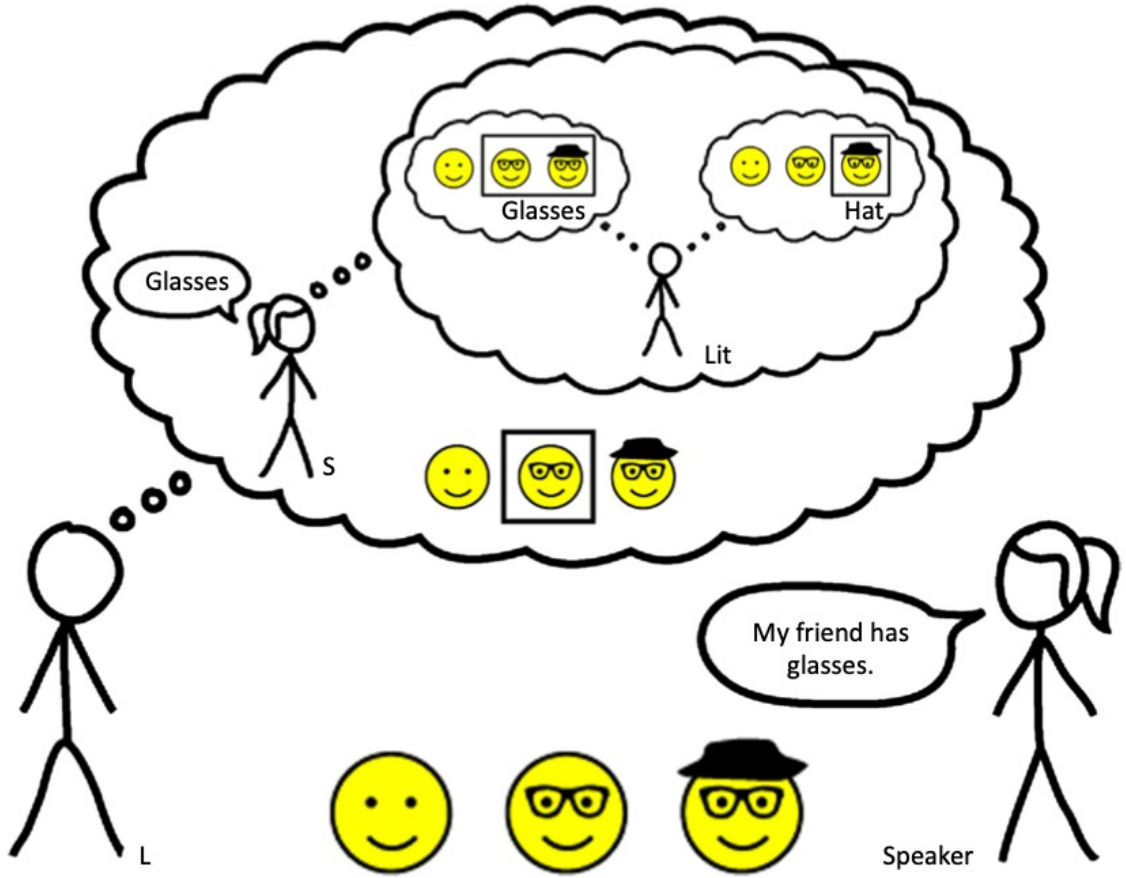


Figure from Goodman and Frank 16

Rational Spech Act (RSA)

- Agent's recursively reason about each other's intentions
  - Using Bayesian inference
  - To derive the contextual meaning of a message



Frank and Goodman 12

Figure from Goodman and Frank 16

Rational Spech Act (RSA)

- Formalization of Grice (1975)
  - *"**The maxim of quantity**, where one tries to be as informative as one possibly can, and gives as much information as is needed, and no more."*
- Game Theoretic Approach
  - Parikh (2001)
  - Franke and Jäger (2006)

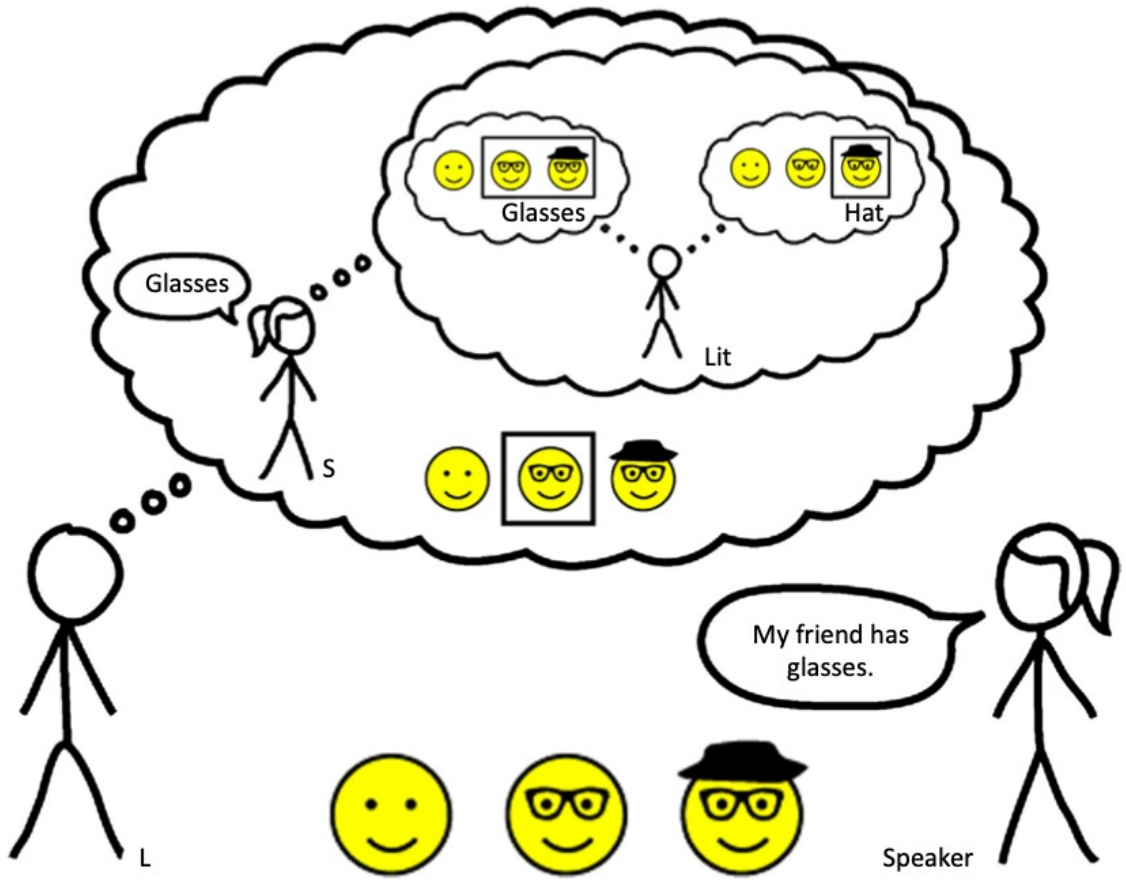
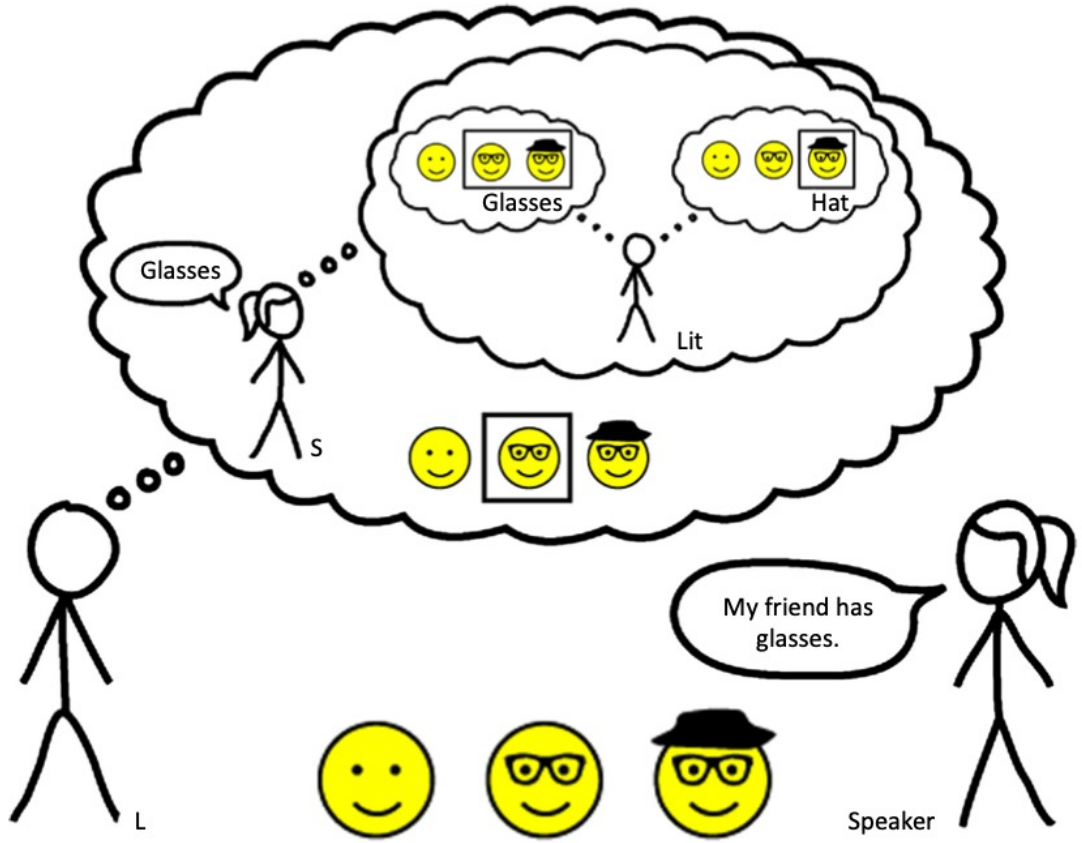


Figure from Goodman and Frank 16

Rational Spech Act (RSA)

Meaning function L(m, w):

- Binary (True-False statement)
- **Graded with values between 0 and 1**
- Context-independent



Rational Spech Act (RSA)

$$L_0(m|w) \propto \mathcal{L}(m, w)$$

$$S_t(w|m, C) \propto e^{\alpha U_t(m,w,C)}$$

$$L_t(m|w, C) \propto S_t(w|m, C)p(m|C)$$

$$U_t(w, m, C) = \log L_{t-1}(m|w, C)$$

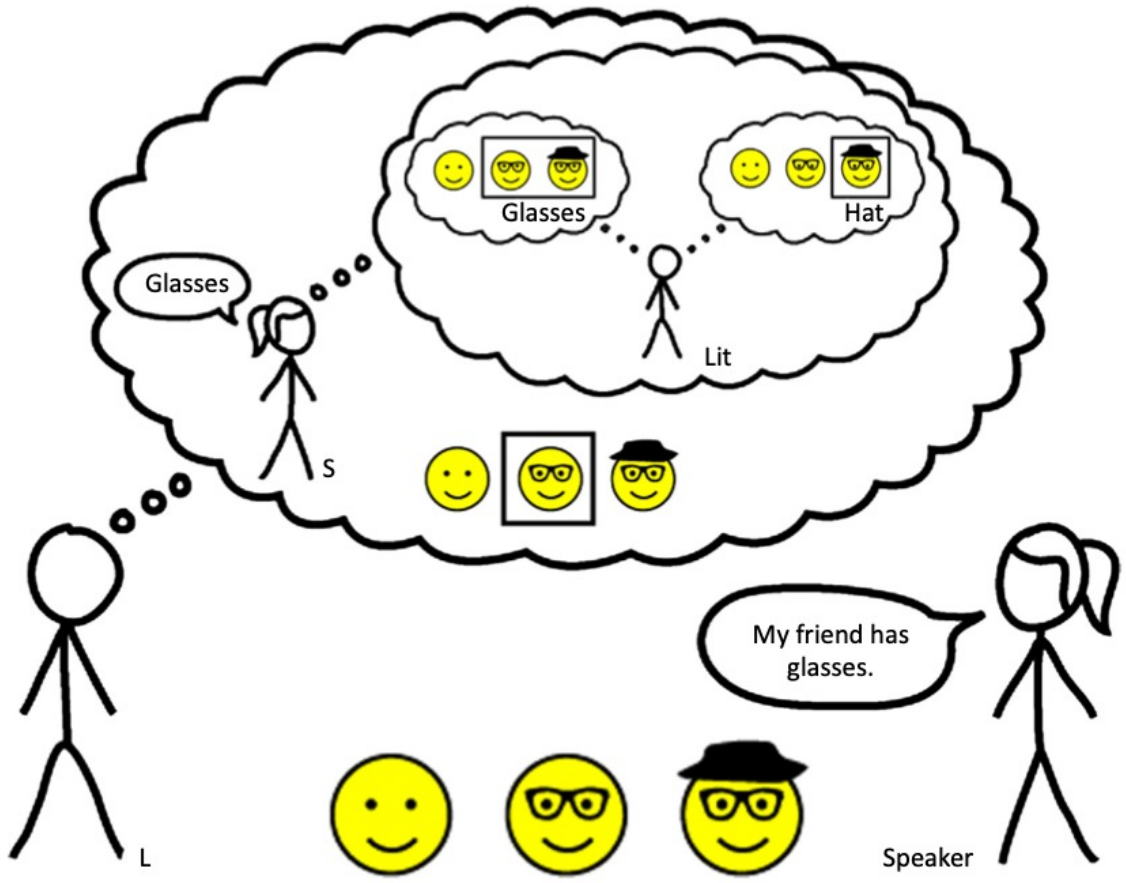


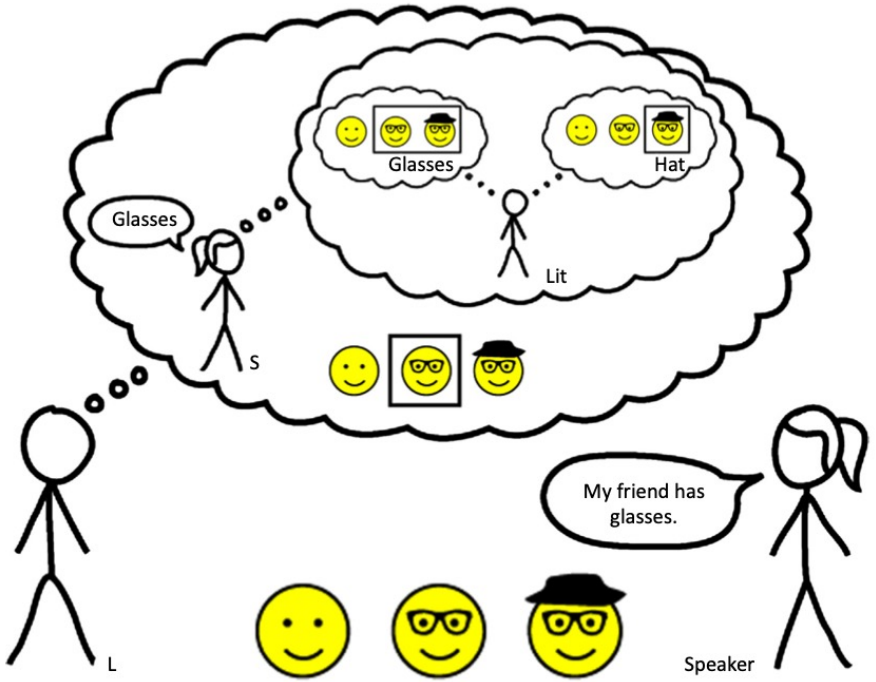
Figure from Goodman and Frank 16

Rational Spech Act (RSA)

$$U_t(w, m, C) = \log L_{t-1}(m|w, C)$$

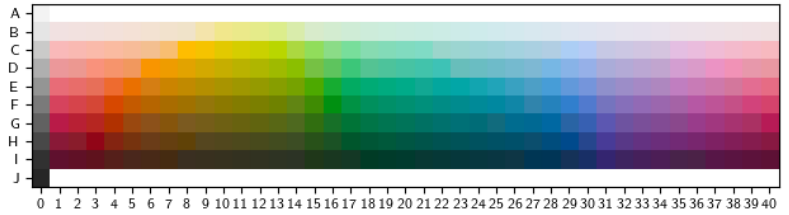
=

Minimizing the epistemic uncertainty of the Listener



World Color Survey

- Human representations from the World Color Survey(WCS)
- Speaker and listener initialized with the meaning function for each WCS language
- Entire Munsell Chart as context



Berlin and Kay 1969

Speaker

a)

Meaning Function: Culina (Peru)

Colors	w_1	w_2	w_3	w_4
	0.0	0.64	0.0	0.36
	0.35	0.05	0.05	0.55
	0.0	0.0	1.0	0.0
	0.714	0.0	0.0	0.286
	0.913	0.0	0.0	0.087
	1.0	0.0	0.0	0.0

c)

$RSA(\infty, 5)$

Colors	w_1	w_2	w_3	w_4
	0.0	1.0	0.0	0.0
	0.0	0.0	0.0	1.0
	0.0	0.0	1.0	0.0
	0.0	0.0	0.0	1.0
	1.0	0.0	0.0	0.0
	1.0	0.0	0.0	0.0

Listener

Structured Signaling Games

Speaker

a)

Meaning Function: Culina (Peru)

Colors	w_1	w_2	w_3	w_4
	0.0	0.64	0.0	0.36
	0.35	0.05	0.05	0.55
	0.0	0.0	1.0	0.0
	0.714	0.0	0.0	0.286
	0.913	0.0	0.0	0.087
	1.0	0.0	0.0	0.0

b)

Similarity Matrix

Color						
	1.0	0.001	0.002	0.062	0.015	0.003
	0.001	1.0	0.0	0.0	0.0	0.0
	0.002	0.0	1.0	0.0	0.0	0.0
	0.062	0.0	0.0	1.0	0.793	0.344
	0.015	0.0	0.0	0.793	1.0	0.624
	0.003	0.0	0.0	0.344	0.624	1.0

Listener

- Similarity measure between meanings $Z_{mm'}$
- Perfect communication not possible
  => Need to minimise the total distortion

Similarity Sensitive Surprisal

$$I^Z(m, w, C) = -\log \sum_{m'} Z_{mm'} L(m'|w, C)$$

- Z(m,m') is the similarity between the two meanings m and m’.
- Listener shouldn't be as surprised if a speaker used the same word for two similar colors compared to if the speaker used the same word for two very different colors.

Leinster, *Entropy and Diversity The Axiomatic Approach, 2021*.

sRSA

$$S_t(w|m,\mathcal{C}) \propto \left(\sum_{m' \in \mathcal{C}} Z_{mm'} L_{t-1}(m'|w) \right)^{\alpha}.$$

sRSA

$$S_t(w|m,\mathcal{C}) \propto \left(\sum_{m' \in \mathcal{C}} Z_{mm'} L_{t-1}(m'|w) \right)^{\alpha}.$$

$$Z_{mm'} := e^{-0.001||x_m - x'_m||^2}$$

RSA vs sRSA

Speaker

a)

Meaning Function: Culina (Peru)

Colors	w_1	w_2	w_3	w_4
	0.0	0.64	0.0	0.36
	0.35	0.05	0.05	0.55
	0.0	0.0	1.0	0.0
	0.714	0.0	0.0	0.286
	0.913	0.0	0.0	0.087
	1.0	0.0	0.0	0.0

b)

Similarity Matrix

Color						
	1.0	0.001	0.002	0.062	0.015	0.003
	0.001	1.0	0.0	0.0	0.0	0.0
	0.002	0.0	1.0	0.0	0.0	0.0
	0.062	0.0	0.0	1.0	0.793	0.344
	0.015	0.0	0.0	0.793	1.0	0.624
	0.003	0.0	0.0	0.344	0.624	1.0

Listener

c)

$RSA(\infty, 5)$

Colors	w_1	w_2	w_3	w_4
	0.0	1.0	0.0	0.0
	0.0	0.0	0.0	1.0
	0.0	0.0	1.0	0.0
	0.0	0.0	0.0	1.0
	1.0	0.0	0.0	0.0
	1.0	0.0	0.0	0.0

d)

$sRSA(\infty, 5)$

Colors	w_1	w_2	w_3	w_4
	0.0	1.0	0.0	0.0
	0.0	0.0	0.0	1.0
	0.0	0.0	1.0	0.0
	1.0	0.0	0.0	0.0
	1.0	0.0	0.0	0.0
	1.0	0.0	0.0	0.0

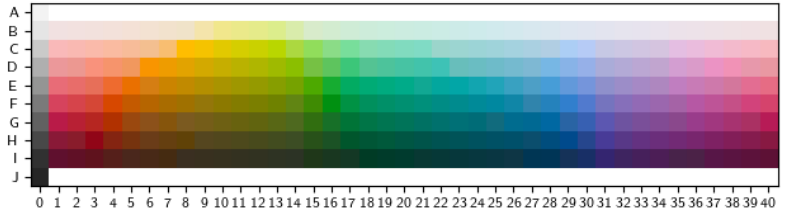
Figure 1: An example of a structured signaling game in the color domain.

Signaling Game

- Entire Munsell Chart as context
  $\Rightarrow$ More colors in the context than words
- Only single word utterances allowed
- No utterance cost
- Meaning functions derived from the WCS naming data

Speaker

Listener



Well-formedness

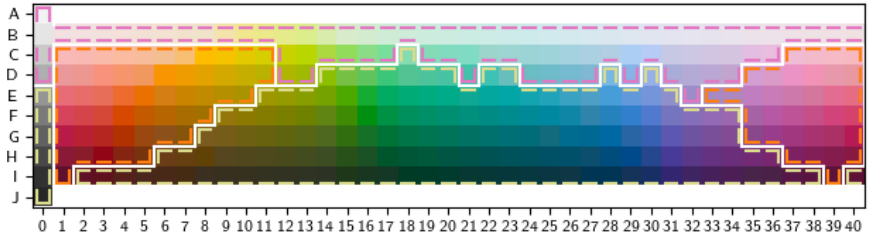
Well-formedness

$$S = \sum_{cat(x)=cat(y)} sim(x, y)$$

$$D = \sum_{cat(x)\neq cat(y)} 1 - sim(x, y)$$

$$W = S + D$$

$$\text{sim}(m, m') = e^{-0.001||x_m - x_{m'}||^2}$$



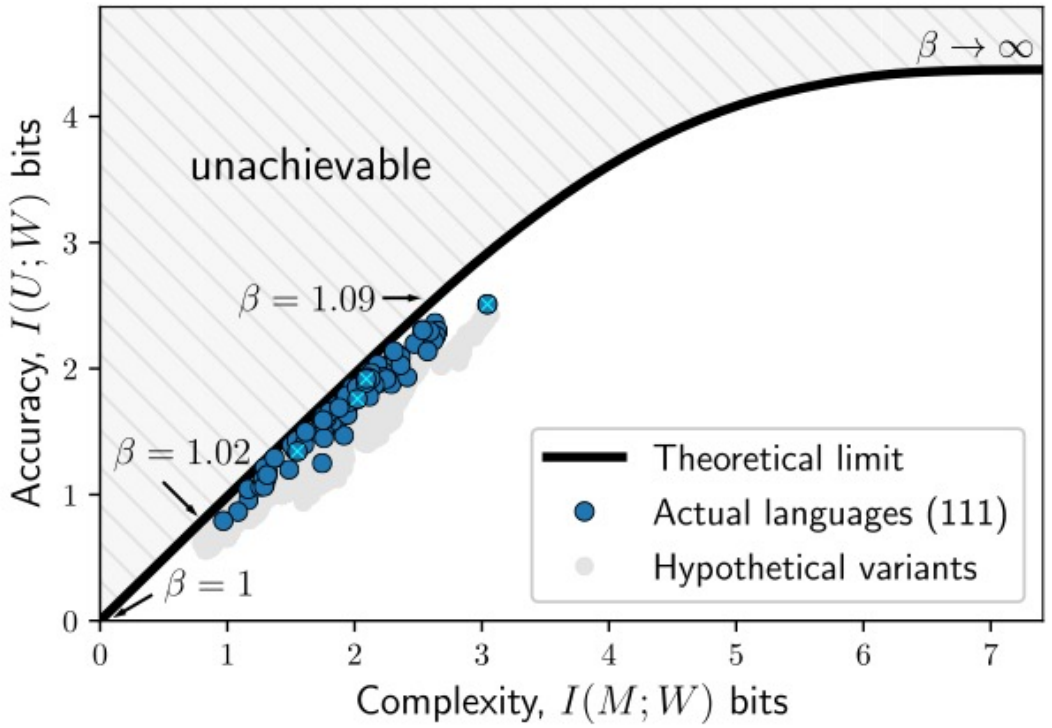
Regier et al. 07

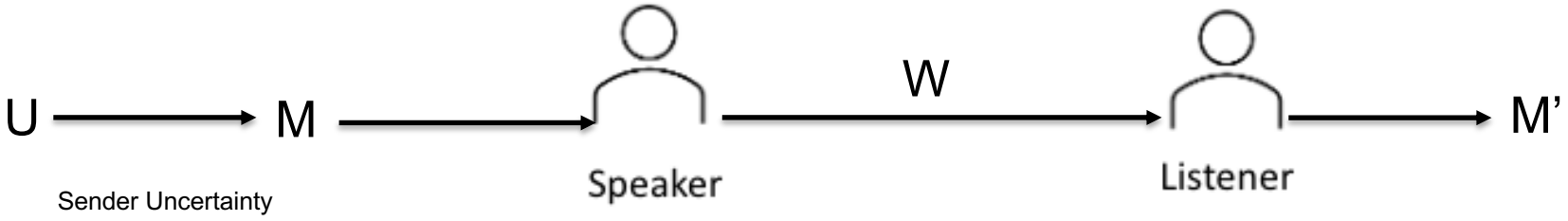
Information-bottleneck

$I(M;W) = H(M) - H(M|W)$

Complexity = I(M;W)

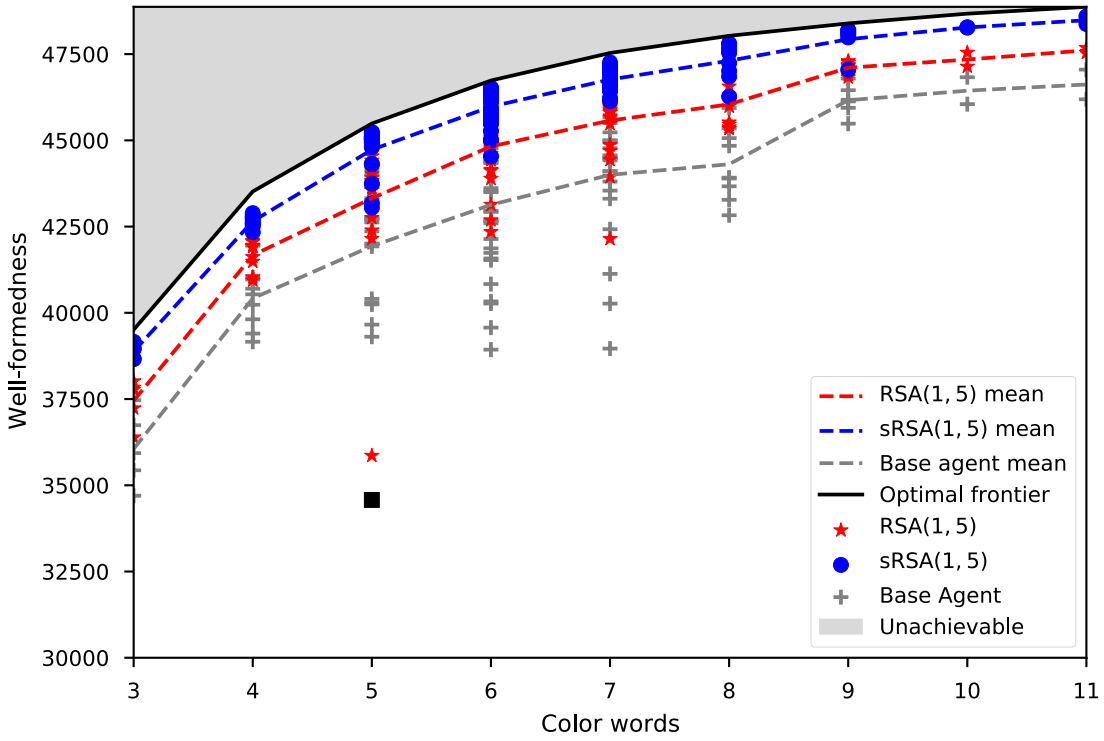
Accuracy = I(U;W)





Zaslavsky et al. 18

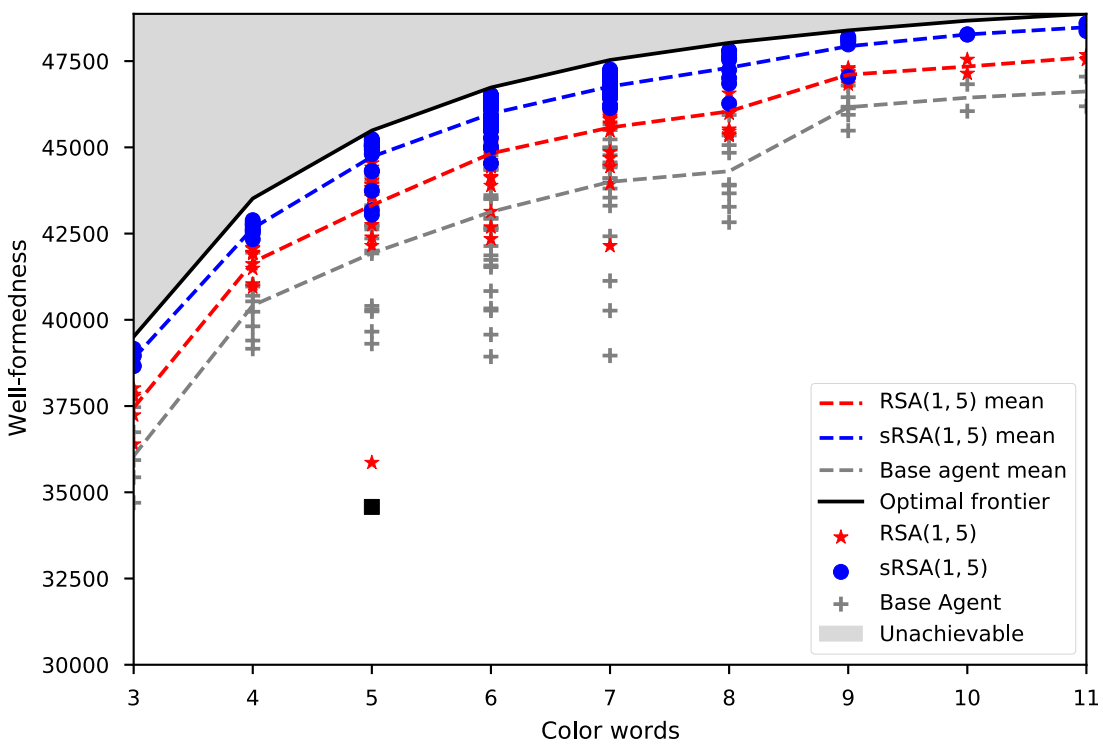
Well-formedness



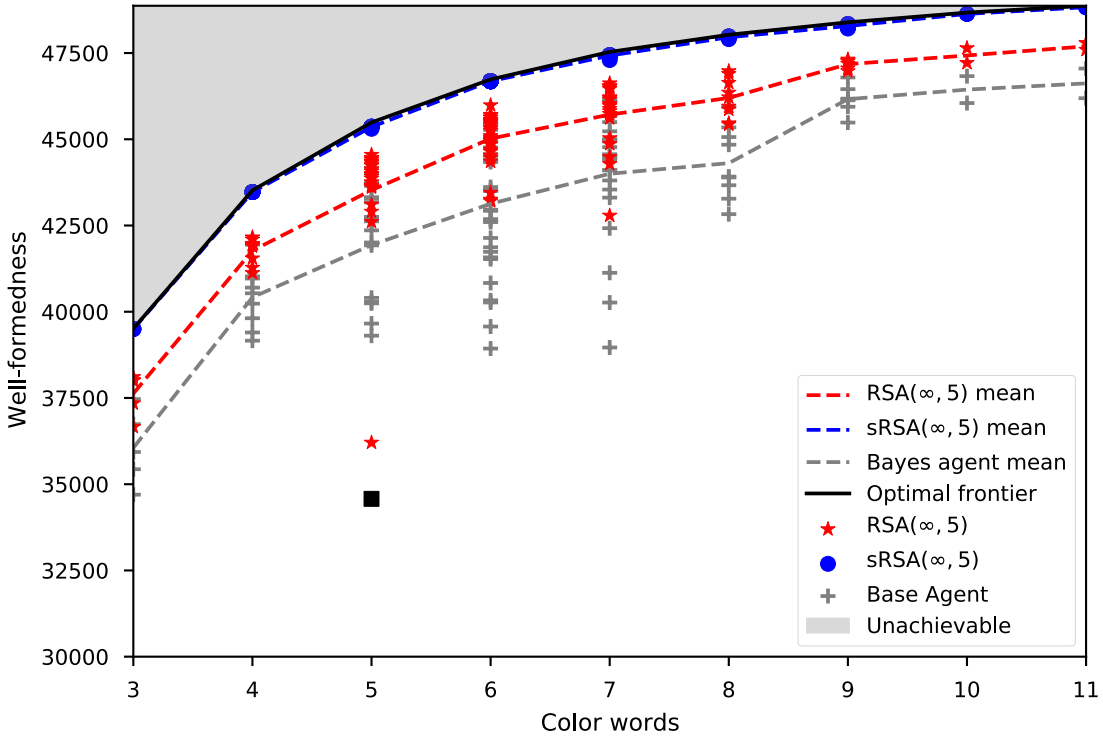
Depth 1

- Alpha = 5
- Very close to the optimal frontier after only one recursion
- Expected that sRSA is more well-formed than RSA

Well-formedness

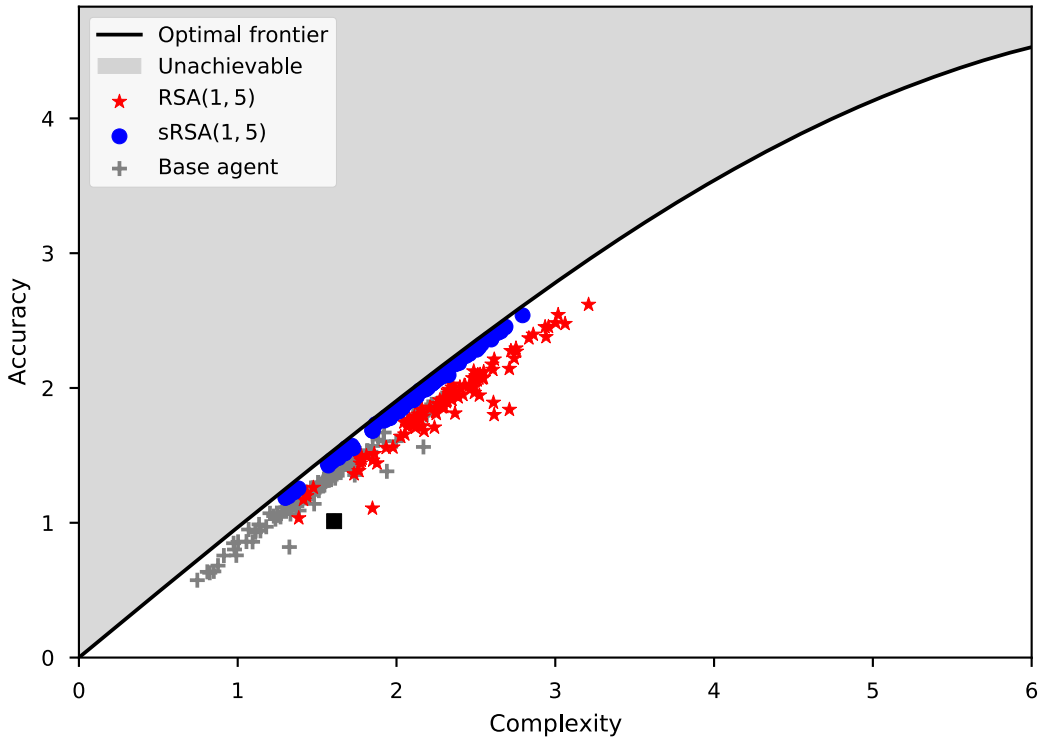


Depth 1

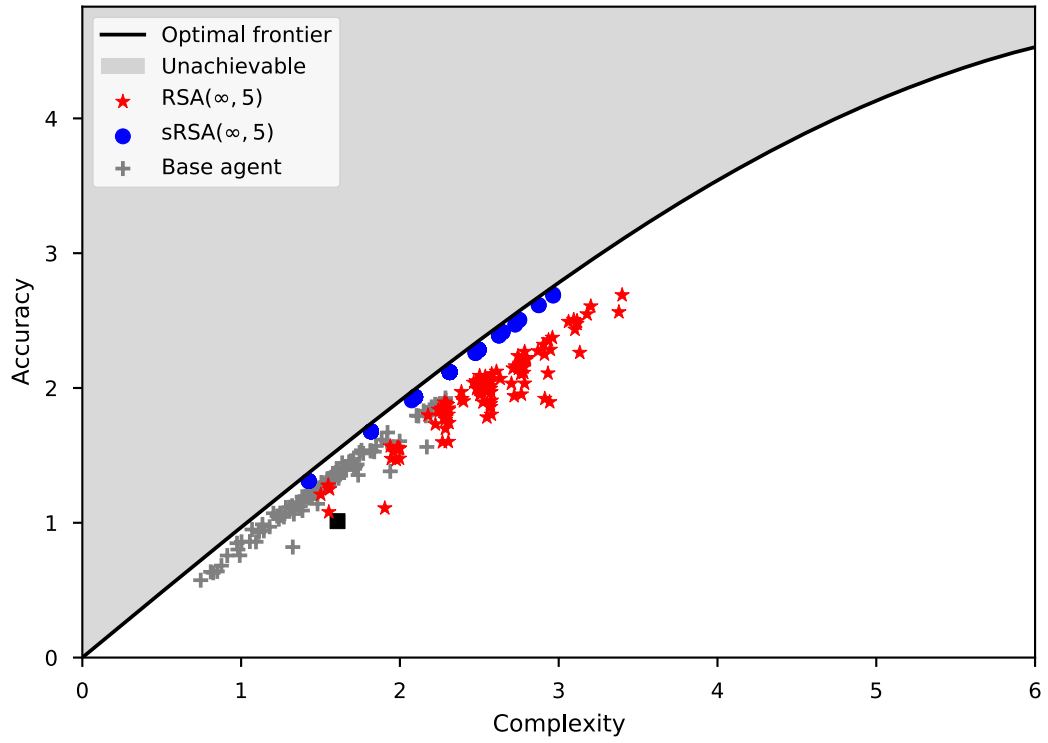


Limit

Information-theoretic tradeoff

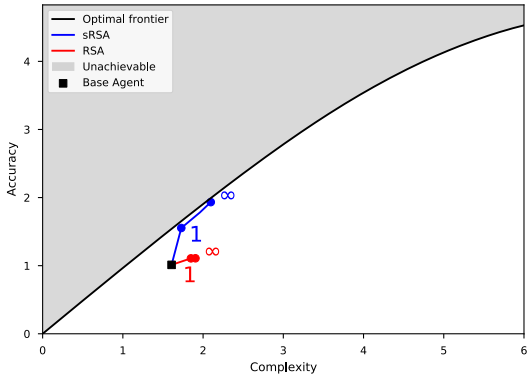


Depth 1

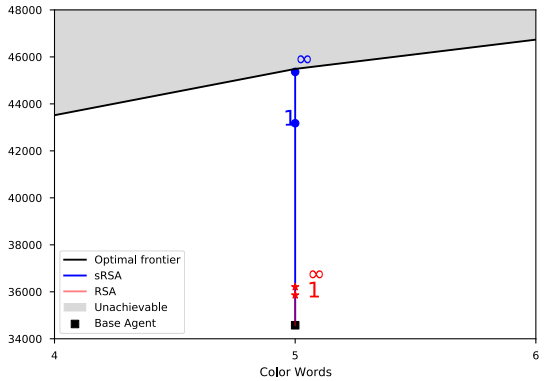


Limit

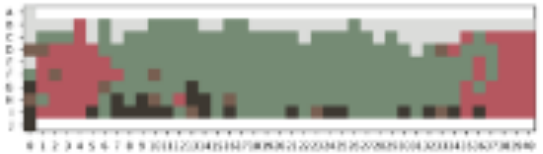
Karajá (Outlier)



IT tradeoff

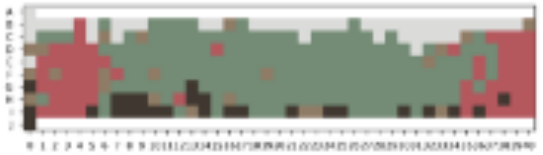


Well-formedness

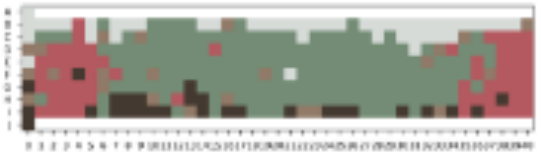


(a) Base agent

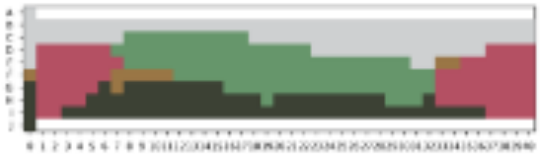
(b) CC agent



(c) RSA(1,5)



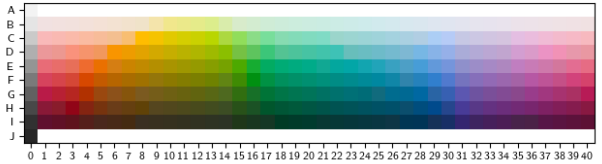
(d) RSA(∞,5)

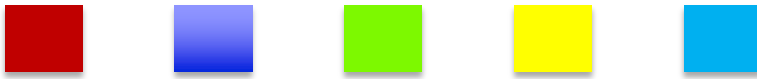


(e) sRSA(1,5)

(f) sRSA(∞,5)

Artificial Agents



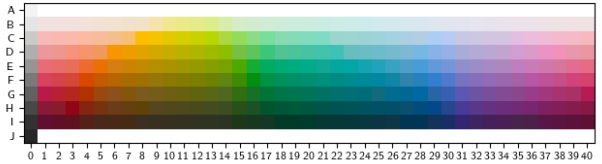


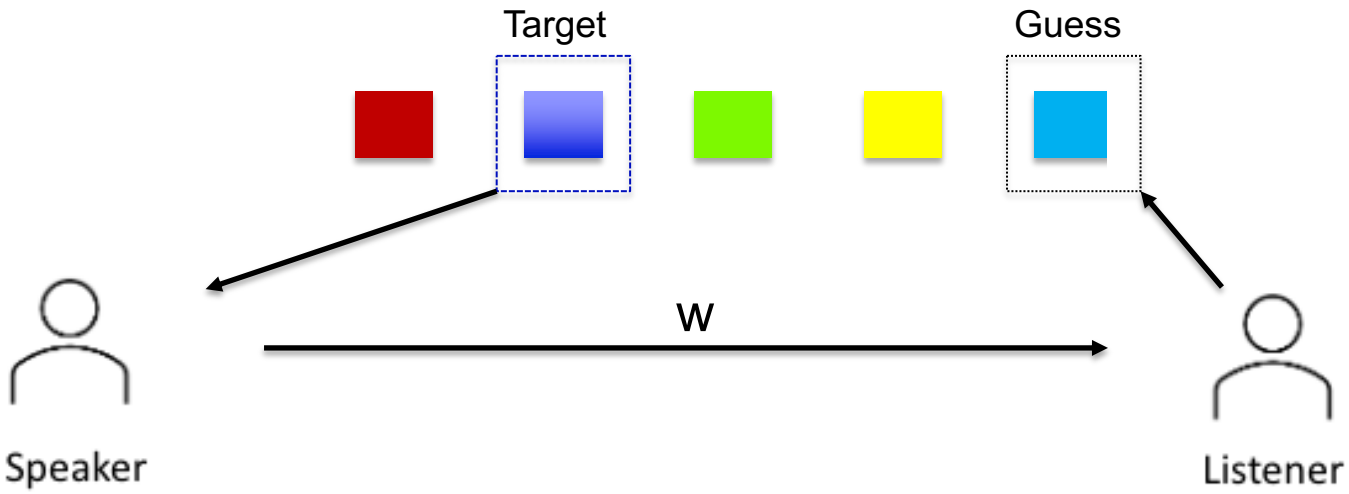
Speaker

Listener

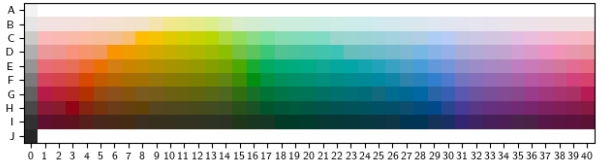
Earlier paper: Kågebäck, Carlsson, Dubhashi and Sayeed 2020, PlosOne

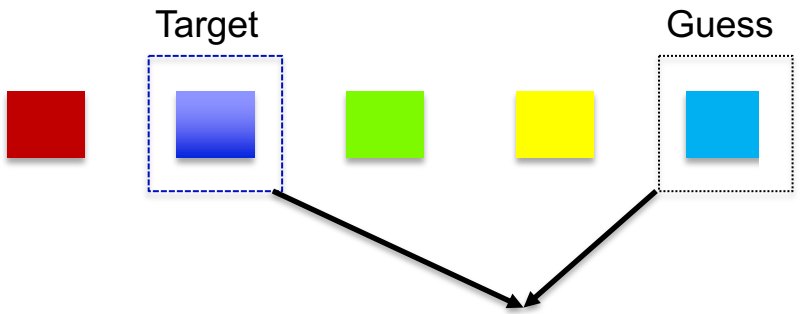
Artificial Agents





Artificial Agents





Artificial Agents

- Meaning Function as Neural Network
- Agents develop colorwords via Reinforcement learning
- Gradient over the recursion using REINFORCE

NN

$$L_0(m|w) \propto \mathcal{L}(m,w)$$

$$S_t(w|m,C) \propto e^{\alpha U_t(m,w,C)}$$

$$L_t(m|w,C) \propto S_t(w|m,C)p(m|C)$$

Artificial Agents

- Meaning Function as Neural Network

$$L_0(m|w) \propto \mathcal{L}(m, w)$$

$$S_t(w|m, C) \propto e^{\alpha U_t(m,w,C)}$$

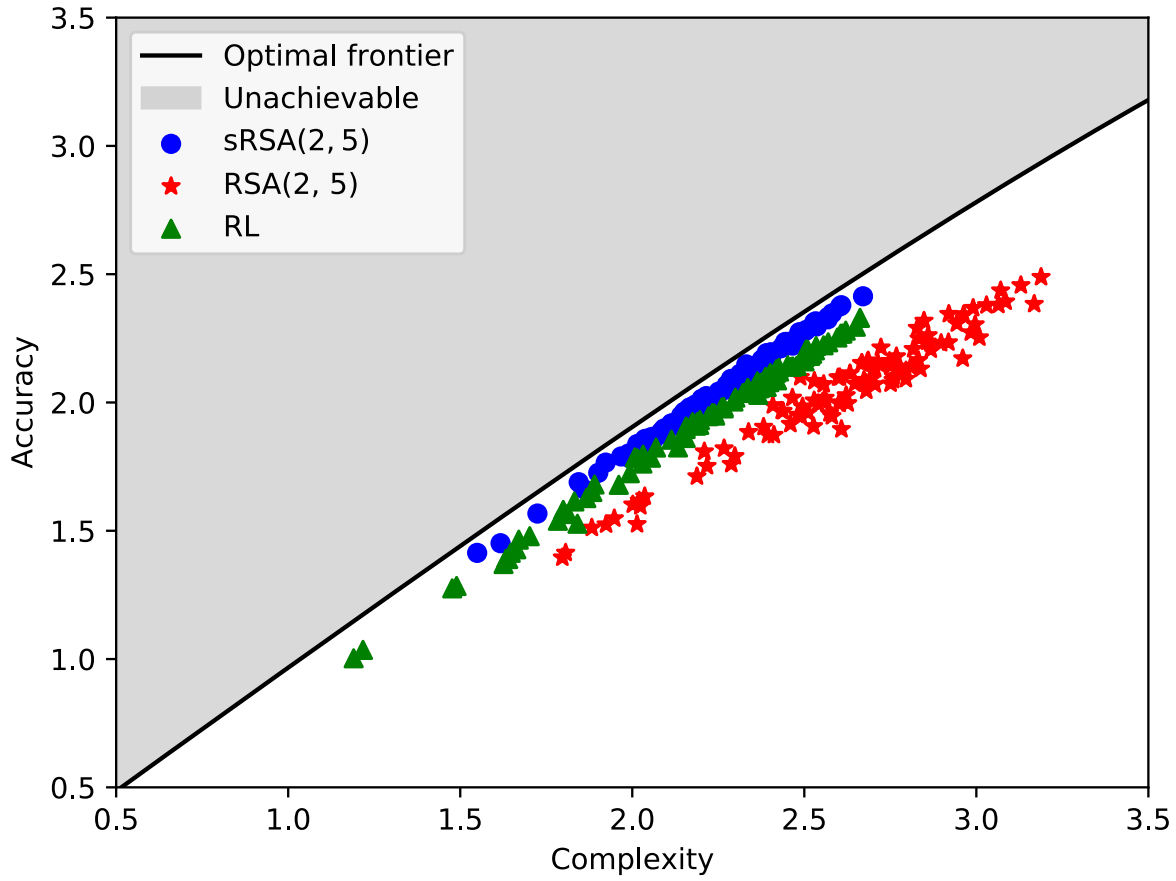
- Agents develop colorwords via Reinforcement learning

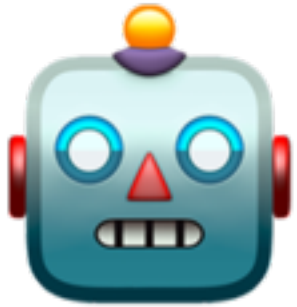
$$L_t(m|w, C) \propto S_t(w|m, C)p(m|C)$$

- Gradient over the recursion using REINFORCE

$$r \log S_{t,\theta}(w|m, C)$$

IT Bottleneck

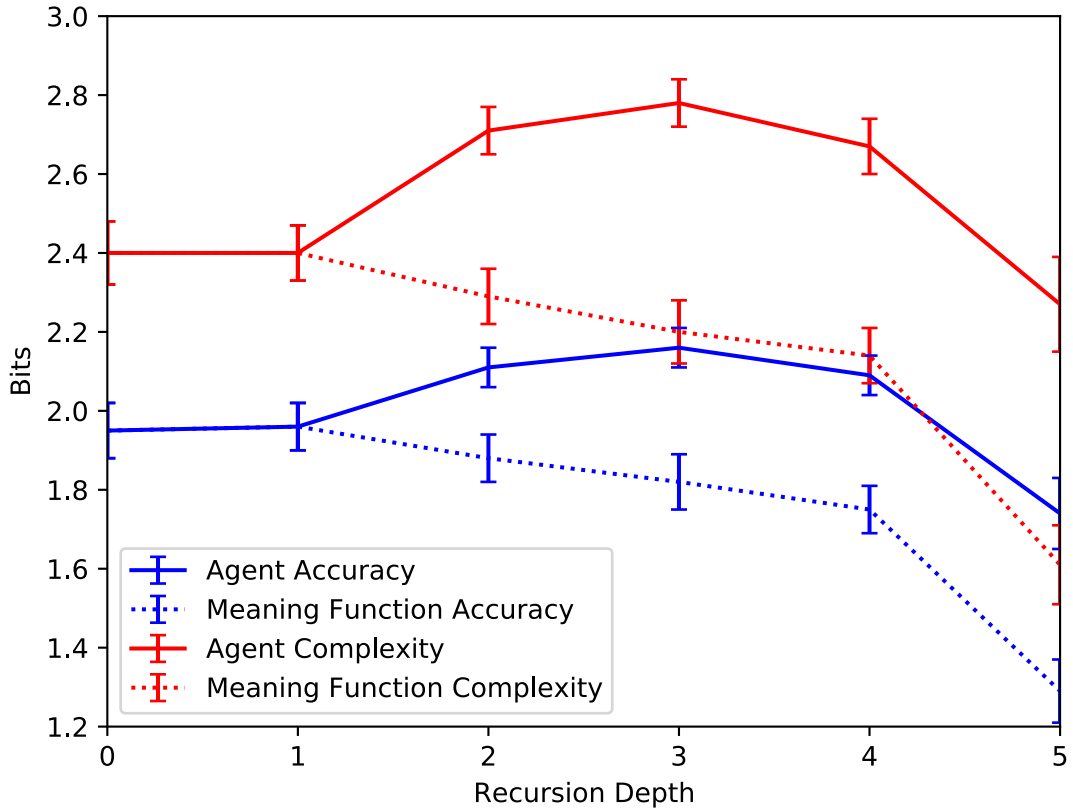




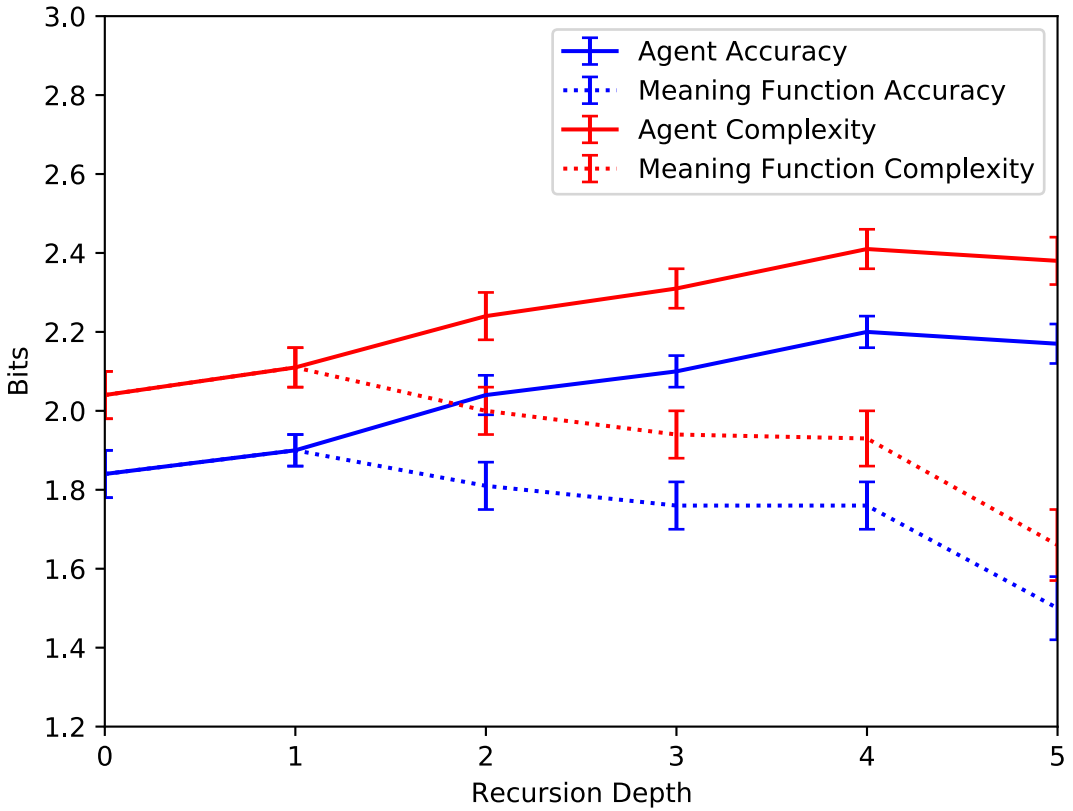
Representation

Reasoning (RSA)

Ambiguity and Recursion Depth



RSA



sRSA

Summary

- Introduced structured version of RSA
- Accounting for the structure greatly improves the efficiency and well-formedness. Demonstrated with both human and artificial representations
- Trade-off between complexity and accuracy of semantic representation and recursion depth. Needs to be explored more.