Supplementary Material

A Rate–Distortion view of human pragmatic reasoning

1 Understanding RSA dynamics as Alternating Maximization

Here we prove the claim that RSA recursion implements and alternating maximization algorithm for optimizing the tradeoff

$$\mathcal{G}_\alpha[S, L] = H_S(U|M) + \alpha \mathbb{E}_S[V_L]\,. \tag{1}$$

Before doing so, we introduce several required definitions and notations. First, we formally define the lexicon as a mapping $l : \mathcal{M} \times \mathcal{U} \to [0, 1]$, such that $l(m, u) > 0$ if u can be applied to m and $l(m, u) = 0$ otherwise. Next, we define the set of speaker and listener distributions that do not violate the lexicon. Denote by $\triangle(\mathcal{U})$ the simplex of all probability distributions over $\mathcal{U}$, and by $\triangle(\mathcal{U})^{\mathcal{M}}$ the set of all conditional probability distributions of U given M. Similarly, denote by $\triangle(\mathcal{M})^{\mathcal{U}}$ the set of all conditional probability distributions of M given U. The set of all possible speakers that do not violate the lexicon is then

$$\mathcal{S}_l = \{S \in \triangle(\mathcal{U})^{\mathcal{M}} : S(u|m) = 0 \;\text{ if } l(m, u) = 0\}\,, \tag{2}$$

and the set of all possible listeners that do not violate the lexicon is

$$\mathcal{L}_l = \{L \in \triangle(\mathcal{M})^{\mathcal{U}} : L(m|u) = 0 \;\text{ if } l(m, u) = 0\}\,. \tag{3}$$

It is easy to verify that $\mathcal{S}$ and $\mathcal{L}$ are convex sets.

Proposition 1 (RSA optimization)**.** *Let $\alpha \geq 0$. The following hold for RSA:*

- *RSA recursion implements an alternating maximization algorithm: for all $t \geq 1$, for a fixed L_{t-1} it holds that*

  $$S_t = \underset{S \in \triangle(\mathcal{U})^{\mathcal{M}}}{\mathrm{argmax}}\; \mathcal{G}_\alpha[S, L_{t-1}]\,, \tag{4}$$

  and for a fixed S_t it holds that

  $$L_t = \underset{L \in \triangle(\mathcal{M})^{\mathcal{U}}}{\mathrm{argmax}}\; \mathcal{G}_\alpha[S_t, L]\,, \tag{5}$$

  where S_t and L_t are RSA's speaker and listener distributions at recursion depth t.

- *If $L_{t-1} \in \mathcal{L}_l$ then $S_t \in \mathcal{S}_l$, and if $S_t \in \mathcal{S}_l$ then $L_t \in \mathcal{L}_l$. That is, RSA iterations do not violate the hard lexicon constraints.*

- *The fixed points of the RSA recursion are stationary points of $\mathcal{G}_\alpha$.*

Proof. First, fix L_{t-1} and notice that the function $g(S) = \mathcal{G}_\alpha[S, L_{t-1}]$ is concave in S. To find a maximizer for g we define the Lagrangian

$$\mathbb{L}[S; \lambda] = g(S) - \sum_m \lambda(m) \sum_u S(u|m)\,,$$

where $\lambda(m)$ are the normalization Lagrange multipliers.[1] Note that if for some m and u it holds that $L_{t-1}(m|u) = 0$ and $S(u|m) > 0$, then $g(S) = -\infty$. Therefore, at the maximum, it necessarily holds that if $L_{t-1}(m|u) = 0$ then also $S(m|u) = 0$. In particular, this implies that if $L_{t-1} \in \mathcal{L}_l$ then $\operatorname{argmax} g(S) \in \mathcal{S}_l$. That is, if L_{t-1} does not violate the lexicon, then maximizing $g(S)$ is guaranteed to give a speaker that also does not violate the lexicon. Taking the derivative of $\mathbb{L}[S; \lambda]$ with respect to $S(u|m)$, for every m and u such that $L_{t-1}(m|u) > 0$, gives

$$\frac{\partial \mathbb{L}}{\partial S(u|m)} = P(m)\left[-\log S(u|m) - 1 + \alpha V_{t-1}(m,u)\right] - \lambda(m)\,.$$

Equating this derivative to zero gives S_t, RSA's speaker equations (equation 2 in the main text), as a necessary condition for optimality. Because $g(S)$ is concave, this is also a sufficient condition.

Next, fix S_t and consider the function $h(L) = \mathcal{G}_\alpha[S_t, L]$. This function is concave in L. To find a maximizer for h over $\mathcal{L}_l$, as before we define the corresponding Lagrangian and take its derivative with respect to $L(m|u)$. This gives

$$\frac{\partial \mathbb{L}}{\partial L(m|u)} = \alpha P(m) S_t(u|m) \frac{1}{L(m|u)} - \lambda(u)\,.$$

Equating this derivative to zero gives L_t, RSA's Bayesian listener (equation 3 in the main text) as a necessary condition for optimality. Because $h(L)$ is concave, this is also a sufficient condition. It is also easy to verify that if $S_t(u|m) = 0$ then $L_t(m|u) = 0$, and therefore if $S_t \in \mathcal{S}_l$ then $L_t \in \mathcal{L}_l$.

Finally, at a fixed point (S^*, L^*) both the derivatives w.r.t. S and w.r.t. L are zero. Since these are also the derivatives of $\mathcal{G}_\alpha$, it holds that (S^*, L^*) is a stationary point of $\mathcal{G}_\alpha$. Note the $\mathcal{G}_\alpha$ is not jointly concave in S and in L. Therefore, (S^*, L^*) is not necessarily a global maximum of $\mathcal{G}_\alpha$. □

2 Derivation of RD-RSA

In this section we derive the RD-RSA update equations from the minimization of

$$\mathcal{F}_\alpha[S] = I_S(M;U) + \alpha \mathbb{E}_S[d(M,U)]\,.$$

Proposition 2 (RD-RSA). *Let $S \in \triangle(\mathcal{U})^{\mathcal{M}}$ and $L \in \triangle(\mathcal{M})^{\mathcal{U}}$. Given $\alpha > 0$, S and L are stationary points of $\mathcal{F}_\alpha$ if and only if they satisfy the following self-consistent conditions:*

$$S(u|m) \;\propto\; S(u)\exp\left(\alpha V_L(m,u)\right) \tag{6}$$

$$S(u) \;=\; \sum_m S(u|m)P(m) \tag{7}$$

$$L(m|u) \;=\; \frac{S(u|m)P(m)}{S(u)} \tag{8}$$

Proof. The main idea of the proof is to take the derivatives of $\mathcal{F}_\alpha$ w.r.t. $S(u|m)$, $S(u)$ and $L(m|u)$, and equate these derivatives to zero. This gives the RD-RSA equations (6)-(8). This derivation is similar to the derivation in the proof of Proposition 1. Therefore, we do not repeat it here. □

3 Asymptotic behavior and the criticality of $\alpha = 1$

In this section we analyze the asymptotic behavior of RSA and RD-RSA dynamics. In cases, we focus mainly on the basic RSA setup discussed in the main text. We also present preliminary analysis of the influence of the cost function.

[1] We omit the non-negativity constraints because these constraints are inactive.

3.1 RSA

Proposition 3. *Let $C(u)$ be a constant function, $P(m)$ be the uniform distribution over $\mathcal{M}$, and assume $K = |\mathcal{M}| = |\mathcal{U}|$. Then the following holds:*

1. $\mathcal{G}^*_\alpha = \max\{(1-\alpha)\log K, 0\}$
2. *For $\alpha \in [0,1]$, $S^*_\alpha(u|m) = \frac{1}{|\mathcal{U}|}$ and $L^*_\alpha(m|u) = P(m)$.*
3. *For $\alpha \geq 1$, S^*_α and L^*_α are deterministic distributions defined by a bijection from $\mathcal{M}$ to $\mathcal{U}$.*

Proof. We prove these claims by first deriving an upper bound on $\mathcal{G}_\alpha$ and then showing that the given S^*_α and L^*_α attain this bound. Assume w.l.o.g. that $C(u) = 0$, and let $S(m|u)$ be the posterior distribution with respect to $S(u|m)$ and $P(m)$. For any S and L is holds that

$$\begin{aligned} \mathcal{G}_\alpha[S,L] &\leq H_S(U|M) + \alpha\mathbb{E}_S\left[\log S(m|u)\right] && (9)\\ &= H_S(U|M) + \alpha\mathbb{E}_S\left[\log\frac{S(u|m)P(m)}{S(u)}\right] && (10)\\ &= (\alpha-1)I_S(M;U) + H_S(U) - \alpha H(M)\,. && (11)\end{aligned}$$

For $\alpha \in [0,1]$ it holds that $\mathcal{G}_\alpha[S,L] \leq \log K - \alpha H(M)$, and this upper bound is attained by $S(u|m) = \frac{1}{K}$ and $L(m|u) = S(m|u) = P(m)$. Specifically, when $P(m)$ is uniform then $\mathcal{G}^*_\alpha = (1-\alpha)\log K$. For $\alpha \geq 1$ it holds that $\mathcal{G}_\alpha[S,L] \leq \alpha\log K - \alpha H(M)$, and this upper bound is 0 when $P(m)$ is uniform. Let $\phi : \mathcal{M} \to \mathcal{U}$ be a bijection, and set $S(u|m) = \delta_{u,\phi(m)}$ and $L(m|u) = S(m|u) = \delta_{u,\phi(m)}$. In this case, $H_S(U|M) = 0$ and $\mathbb{E}_S[V_L] = 0$, based on the convention that $0\log 0 = 0$. Therefore, these distributions attain the upper bound for $\alpha \geq 1$. □

In the more general case, where $C(u)$ is not necessarily constant, the first critical value of α after which the non-informative solution looses optimality is greater or equal to 1. To see this, notice that adding cost to the bound in (11) gives

$$\begin{aligned} \mathcal{G}_\alpha[S,L] &\leq (\alpha-1)I_S(M;U) + H_S(U) - \alpha H(M) - \alpha\mathbb{E}_S[C(U)] && (12)\\ &= (\alpha-1)I_S(M;U) - D[S(u)\|Q_\alpha(u)] + \log Z_\alpha - \alpha H(M) && (13)\\ &\leq \log Z_\alpha - \alpha H(M)\,, && (14)\end{aligned}$$

where

$$Q_\alpha(u) = \frac{e^{-\alpha C(u)}}{Z_\alpha}\,, \quad Z_\alpha = \sum_u e^{-\alpha C(u)}$$

and $D[\cdot\|\cdot]$ is the Kullback-Leibler (KL) divergence between two probability distributions. It is now easy to see that for $\alpha \in [0,1]$ this upper bound is attained by $S^*_\alpha(u|m) = Q_\alpha(u)$ and $L^*_\alpha(m|u) = P(m)$. In this case, $S^*_\alpha(u|m)$ changes continuously at least for $\alpha \leq 1$, however all of these solutions effectively optimize the tradeoff between communicative effort and utterance cost, while ignoring the listener's surprisal. In other words, in this regime, the RSA model predicts that a pragmatic speaker will not try to convey any information to the listener ($I_S(M;U) = 0$), but will rather seek the minimal deviation from random utterance production that reduces the expected utterance cost to a tolerable degree, determined by α.

3.2 RD-RSA

The following proposition characterizes the asymptotic behavior of RD-RSA in the basic setup discussed in the main text.

Proposition 4. *Let $C(u)$ be a constant function, then the following holds for RD-RSA:*

1. *For $\alpha \in [0, 1)$,*
$$\min_{S,L} \mathcal{F}_\alpha[S, L] = \min_S I_S(M; U) = 0$$

2. *For $\alpha > 1$,*
$$\min_{S,L} \mathcal{F}_\alpha[S, L] = \max_S I_S(M; U)$$

3. *For $\alpha = 1$, all stationary points are optimal.*

4 Comparison with human behavior

In the main text we have shown that both RSA and RD-RSA produce listener distributions that are highly correlated with the empirical human listener estimated from the experimental data of Vogel et al. (2014). Here we supplement the evaluation with Figure 1, that shows that the best RSA listener and the best RD-RSA are indeed very similar to the estimated human listener.

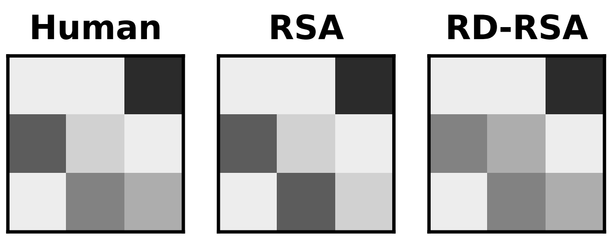

Figure 1: Left: Human listener distribution estimated from Vogel et al. (2014). Middle: RSA's listener distribution for $\alpha = 0.9$ and recursion depth 1. Right: RD-RSA's listener distribution for $\alpha = 1.2$ and recursion depth 5.