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 an 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 a 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 all speaker and listener distributions that do not violate the lexicon. Denote by $\triangle(\mathcal{U})$ the simplex of 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 statements hold for RSA:*

- *RSA recursion implements an alternating maximization algorithm for optimizing $\mathcal{G}_\alpha$. That is, for all $t \geq 1$, for a fixed L_{t-1} it holds that*

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

  and for a fixed S_t it holds that

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

  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 note that the function $g(S) = \mathcal{G}_\alpha[S, L_{t-1}]$ is concave in S. To find a maximizer for $g(S)$ over $\triangle(\mathcal{U})^{\mathcal{M}}$ 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

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

$L_{t-1}(m|u) = 0$ 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 $S(m|u) = 0$ (following the convention that $0 \log 0 = 0$). 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 RSA's speaker equation (S_t, 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 $\triangle(\mathcal{M})^{\mathcal{U}}$, we define as before 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 RSA's Bayesian listener (L_t, 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 with respect to S and to L are zero. Because these are also the derivatives of $\mathcal{G}_\alpha$ over $\triangle(\mathcal{U})^{\mathcal{M}} \times \triangle(\mathcal{M})^{\mathcal{U}}$, it holds that (S^*, L^*) is a stationary point of $\mathcal{G}_\alpha$. Note that $\mathcal{G}_\alpha$ is not jointly concave in S and in L, and therefore, (S^*, L^*) is not necessarily a global maximum. □

2 Derivation of RD-RSA

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

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

This can be seen as a Rate–Distortion optimization problem with distortion $d(m, u) = -V_L(m, u)$.

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:*

$$S(u|m) \propto S(u) \exp\left(\alpha V_L(m, u)\right) \quad (6)$$
$$S(u) = \sum_m S(u|m) P(m) \quad (7)$$
$$L(m|u) = \frac{S(u|m) P(m)}{S(u)} \quad (8)$$

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. □

Note that $\mathcal{F}_\alpha$ is convex in $S(u|m)$, $S(u)$, and $L(m|u)$. Therefore, similar to the proof of Proposition 1, it holds that $\mathcal{F}_\alpha$ can be optimized via an alternating minimization algorithm that iteratively updates equations (6)-(8), as described in the main text. However, because $\mathcal{F}_\alpha$ is not jointly convex in these variables, it will not necessarily converge to a global optimum.

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

In this section we analyze the asymptotic behavior of RSA and RD-RSA dynamics. In both 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.

3.1 RSA

Denote by $\mathcal{G}^*_\alpha$ the maximal value of $\mathcal{G}_\alpha$ given α, and by S^*_α and L^*_α the optimal speaker and listener distributions that attain $\mathcal{G}^*_\alpha$. That is, $\mathcal{G}^*_\alpha = \mathcal{G}_\alpha[S^*_\alpha, L^*_\alpha]$. The following proposition characterizes $\mathcal{G}^*_\alpha$, S^*_α, and L^*_α, as a function of α, in a basic RSA setup.

Proposition 3 (Asymptotic behavior of RSA)**.** *Let $C(u)$ be a constant function, $P(m)$ be the uniform distribution over $\mathcal{M}$, and assume $K = |\mathcal{M}| = |\mathcal{U}|$. In addition, assume a graded lexicon l with no structural zeros. Then the following statements hold:*

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 in the two regimes of α. 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 it holds that:

$$\begin{aligned}
\mathcal{G}_\alpha[S,L] &\leq H_S(U|M) + \alpha \mathbb{E}_S\left[\log S(m|u)\right] && (9)\\
&= H_S(U|M) + \alpha \mathbb{E}_S\left[\log \frac{S(u|m)P(m)}{S(u)}\right] && (10)\\
&= H_S(U|M) + \alpha\left(-H_S(U|M) + H_S(U) - H(M)\right) && (11)\\
&= (\alpha-1)I_S(M;U) + H_S(U) - \alpha H(M)\,. && (12)
\end{aligned}$$

Step by step explanation: (9) follows from the fact that for any two distributions $p(x)$ and $q(x)$ it holds that $\sum_x p(x)\log q(x) \leq \sum_x p(x)\log p(x)$. In particular, this holds for $S(m|u)$ and $L(m|u)$. (10) follows from substituting Bayes' rule, and (11) from the definition of entropy. (12) follows from the identity $I_S(M;U) = H_S(U) - H_S(U|M)$.

For $\alpha \in [0,1]$, this bound is maximal when $I_S(M;U) = 0$ and when $H_S(U) = \log|\mathcal{U}| = \log K$. Therefore, in this regime it holds that $\mathcal{G}_\alpha[S,L] \leq \log K - \alpha H(M)$. It easy to verify that $S(u|m) = \frac{1}{K}$ and $L(m|u) = S(m|u) = P(m)$ attain this upper bound (simply substitute these distributions in the definition of $\mathcal{G}_\alpha$). When $P(m)$ is uniform then $\mathcal{G}^*_\alpha = (1-\alpha)\log K$.

For $\alpha \geq 1$ it holds that $(\alpha-1)I_S(M;U) \leq (\alpha-1)\log K$, and therefore $\mathcal{G}_\alpha[S,L] \leq \alpha\log K - \alpha H(M)$. When $P(m)$ is uniform, this upper bound becomes $\mathcal{G}_\alpha[S,L] \leq 0$. 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 $\mathbb{E}_S[V_L] = 0$, following the convention that $0\log 0 = 0$. Therefore, these distributions attain the upper bound on $\mathcal{G}_\alpha$ for $\alpha \geq 1$. Putting everything together gives $\mathcal{G}^*_\alpha = \max\{(1-\alpha)\log K, 0\}$, which concludes the proof. □

Proposition 3 shows that in this basic RSA setup there is only one critical value $\alpha_c = 1$, which determines the global optimum of $\mathcal{G}_\alpha$ and the asymptotic tendency of the RSA dynamics. When $C(u)$ is not necessarily constant, there could be multiple critical values α_c. We next show that in this case the first critical value α_1, at which the non-informative solution $L(m|u) = P(m)$ looses its optimality, is greater than or equal to 1. To see this, notice that adding a cost to the bound in (12) gives

$$\mathcal{G}_\alpha[S, L] \leq (\alpha - 1)I_S(M; U) + H_S(U) - \alpha H(M) - \alpha \mathbb{E}_S[C(U)]\,. \tag{13}$$

Let $Q_\alpha(u)$ be the maximum-entropy distribution of u with respect to $C(u)$, that is

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

Now, the bound in (13) can be rewritten as

$$\mathcal{G}_\alpha[S, L] \leq (\alpha - 1)I_S(M; U) - D[S(u)\|Q_\alpha(u)] + \log Z_\alpha - \alpha H(M)\,, \tag{14}$$

where $D[\cdot\|\cdot]$ is the Kullback-Leibler (KL) divergence. For $\alpha \in [0, 1]$, due to the non-negativity of the mutual information and the KL-divergence, the right-hand side of (14) can be further bounded by above by omitting the first two terms. This yields

$$\mathcal{G}_\alpha[S, L] \leq \log Z_\alpha - \alpha H(M)\,, \tag{15}$$

and it is easy to verify that this upper bound is attained by $S^*_\alpha(u|m) = Q_\alpha(u)$ and $L^*_\alpha(m|u) = P(m)$ (again, by substituting these distribution in $\mathcal{G}_\alpha$). In this case, $S^*_\alpha(u|m)$ changes continuously for $\alpha \leq \alpha_1$, even though these changes don't convey any information to the listener. 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. This is a further demonstration of RSA's bias toward random utterance production.

3.2 RD-RSA

Next, we characterize the asymptotic behavior of RD-RSA in the basic setup discussed in the main text. Denote by $\mathcal{F}^*_\alpha$ the minimal value of $\mathcal{F}_\alpha$ given α, and by S^*_α and L^*_α the corresponding optimal speaker and listener distributions.

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

1. *For $\alpha \in [0, 1)$,*
$$\mathcal{F}^*_\alpha = \min_S I_S(M; U) = 0$$
2. *For $\alpha > 1$,*
$$\mathcal{F}^*_\alpha = \max_S I_S(M; U)$$
3. *For $\alpha = 1$, all stationary points are optimal.*

Proof. The idea of the proof is similar to the proof of Proposition 3. Here, however, we derive a lower bound (rather than an upper bound) on $\mathcal{F}_\alpha$. To derive this bound we follow similar steps as in (9)-(11), but for $\mathcal{F}_\alpha$ we replace the conditional entropy in the first term by $I_S(M; U)$ and change the sign of the second term in (9). This gives the following lower bound

$$\mathcal{F}_\alpha[S, L] \quad \geq \quad (1 - \alpha)I_S(M; U) - \alpha H(M)\,. \tag{16}$$

For $\alpha < 1$, minimizing this lower bound amounts to minimizing $I_S(M; U)$. The minimum is attained by a non-informative speaker, e.g. $S^*_\alpha(u|m) = \frac{1}{|\mathcal{U}|}$. For $\alpha > 1$, minimizing this lower bound turns into to *maximizing* $I_S(M;U)$. If a there exists a bijection $\phi : \mathcal{M} \to \mathcal{U}$ that does not violate the lexicon, then $S(u|m) = \delta_{u,\phi(m)}$ and $L(m|u) = S(m|u) = \delta_{u,\phi(m)}$ attain this bound.

Finally, for $\alpha = 1$, any fixed point of the RD-RSA equations (S^*, L^*) gives

$$\mathbb{E}_{S^*}\left[\log L^*(m|u)\right] = \mathbb{E}_{S^*}\left[\log S^*(m|u)\right] = -H^*_S(M|U)\,, \tag{17}$$

and therefore

$$\mathcal{F}_\alpha[S^*, L^*] = I_{S^*}(M;U) + H^*_S(M|U) = -\alpha H(M)\,. \tag{18}$$

This means that all fixed points are equally good in this regime. □

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 that evaluation with the figure below. This figure shows that the best RSA listener and the best RD-RSA listener are indeed very similar to each other and to the estimated human listener.

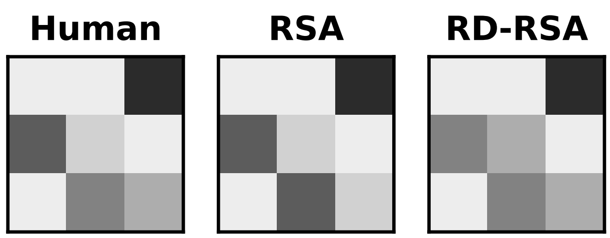

Figure 1: Left: Human listener distribution estimated from Vogel et al. (2014). Middle: RSA's listener distribution for $\alpha = 0.9$ and recursion depth 1. Right: RD-RSA's listener distribution for $\alpha = 1.2$ and recursion depth 5.