Rational Speech Act models are utterance-independent updates of world priors

Jean-Philippe Bernardy **Julian Grove** **Christine Howes**
Centre for Linguistic Theory and Studies in Probability
Department of Philosophy, Linguistics and Theory of Science
University of Gothenburg
`firstname.lastname@gu.se`

Abstract

A popular framework for modelling pragmatic effects is the "rational speech act" (RSA) model introduced by Frank and Goodman (2012). The idea behind RSA is that, to interpret an utterance, a rational (pragmatic) listener reasons about a speaker who chooses their utterance by reasoning about the listener, using a literal semantic model. In the present work, we take the RSA model at face value, but we reformulate it in information-theoretic terms. We find that the pragmatic listener model can be reconceived as an update of the prior over worlds that can be provided independently of the speaker's actual utterance. This update consists in a preference for world states which are the most specific to a given utterance in the set of possible utterances given by the pragmatic context. Our reformulation allows us to deduce general properties of pragmatic reasoning problems. As an example, we show that RSA does not predict certain quantity implicatures in the presence of bell-curve priors.

1 Introduction

The "rational speech act" (RSA) model introduced by Frank and Goodman (2012) recasts a broadly Gricean view of language in Bayesian probabilistic terms. As in the work of Grice (1975), the core ideas underlying RSA are that dialogue participants are rational agents who communicate efficiently by reasoning over each other's beliefs and the shared communicative goal. The core assumption in the RSA model is "…that listeners view speakers as having chosen their words informatively — that is, relative to the information that they would transfer to a naive listener" (Frank and Goodman, 2014, p.84).

The basic RSA model claims that a rational (pragmatic) speaker will take into account how a naive (literal) listener interprets an utterance, assuming it is true. The ideal (pragmatic) listener reasons, in turn, about the pragmatic speaker, thus also taking into account the nested reasoning over the literal listener.

This model is meant to account for human decision making. Much of the support for RSA comes from restrictive communication games in which participants must pick a speaker's intended referent from a set of objects which may match or differ on particular attributes (such as shape or colour) given only a one word utterance. In certain circumstances, a pragmatic listener must take into account both the shared features across objects that are consistent with a given utterance, as well as those features which are not shared, in order to disambiguate among referents for which the utterance is ambiguous. According to Frank and Goodman (2012), the predictions of the RSA model correlate strongly with human behaviour in such one shot referential games. The RSA framework has since been applied to a variety of linguistic puzzles of ambiguity and optionality, including whether plural predications will receive a distributive or collective reading (Scontras and Goodman, 2017), and whether null versus overt pronouns are chosen in constructions which may feature pro-drop (Chen et al., 2018), among others.

In the present work, we take the RSA model at face value, but we reformulate it in explicitly information-theoretic terms by calling on the notion of information gain between the prior and posterior distributions. Our reformulation provides the following insights:

- While a common, algorithmic interpretation of the RSA model suggests that agents reason over each other's reasoning states (listener-speaker-listener), this formulation is not only implausible, but unnecessary, as we show. That is, one can reason in much more direct terms.

- RSA does not, in fact, make correct predictions about the implicatures expected in par-

Proceedings of the 26th Workshop on the Semantics and Pragmatics of Dialogue, August, 22-24, 2022, Dublin.

ticular conversational contexts, according to the Gricean underpinnings of RSA, and given reasonable (bell-curve) priors.

In particular, we show that any given RSA model, whether of a pragmatic listener or a pragmatic speaker, may be presented merely as a *filter* on what we will call a "pragmatic prior"; that is, a prior over worlds or utterances which has been reconceptualised in information-theoretic terms, in order to incorporate notions of specificity and informativeness. Given such a pragmatic prior distribution, any given occasion of interpreting an utterance (or choosing an utterance, given some intended message) requires only that the listener/speaker renormalise this distribution with any incompatible values removed.

There is some precedent for our proposal in work by Scontras et al., which provides an information-theoretic reformulation of the pragmatic speaker model. More precisely, the authors characterise the model in terms of the following formula (which, as we will explain in the next section, incorporates a parameter α setting the model's temperature):

$$\begin{aligned} P_{S_1}(u|w) \;\propto\; & \mathrm{Truth}(u,w) \\ & * \mathrm{Informativeness}(u)^{\alpha} \\ & * \mathrm{Economy}(u)^{\alpha} \end{aligned}$$

Here, the term $\mathrm{Truth}(u, w)$ is a filter on the distribution determined by the other two terms; that is, it is valued as 1 if utterance u is true at world w, and as 0 if it is false. In the present work, we take the next logical step by showing that the pragmatic listener model can be subject to the same kind of reformulation. Consequently, both the pragmatic speaker model and the pragmatic listener model can be reduced to mere filters on their respective pragmatic priors.

2 Background: RSA

RSA, as proposed by Frank and Goodman (2012), assumes a set of possible utterances $\mathcal{U}$ and a set of world states $\mathcal{W}$. World states w come with a prior probability $P(w)$, and utterances u come with a cost $C(u)$. Additionally, we have a relation l on $\mathcal{U}$ and $\mathcal{W}$ such that $l(u, w) = 1$ if utterance u is true at world state w, and $l(u, w) = 0$ otherwise. We say that the tuple $(\mathcal{U}, \mathcal{W}, P, C, l)$ constitutes a *pragmatic interpretation problem*. A solution to such a problem consists in a specification of a pragmatic listener, which is a function from $\mathcal{U}$ to distributions over $\mathcal{W}$. Given an utterance $u \in \mathcal{U}$, it is assumed that the posterior distribution of the pragmatic listener is computed on the assumption that u is literally true. Thus we model the pragmatic listener as taking for granted that its interlocutor is adhering to the Maxim of Quality.

In its most common formulation, RSA models a pragmatic listener as an agent which reasons about a speaker, which, in turn, reasons about a literal listener. To illustrate how this works, let us consider a situation in which there is a box which, at one point, contained 7 cookies, and it is known that John ate at least 5 of them. Thus $\mathcal{U} = \{\text{‘John ate } x \text{ cookies'} \mid x \in [5, 7]\}$. (We let the cost function C be constant across utterances.) The set of possible world states corresponds to those where some number w of cookies has actually been eaten. We choose the literal semantics to allow for more cookies to have actually been eaten than stated:

$$l(\text{‘John ate } x \text{ cookies'}, w) = w \geq x$$

Thus considering only relevant values of w, the literal meaning l can be represented by the following table.

w	5	6	7
‘John ate 5 cookies’	1	1	1
‘John ate 6 cookies’	0	1	1
‘John ate 7 cookies’	0	0	1

2.1 The literal listener model

The literal interpretation of u is given by a Bayesian update of P by l, which thus acts as a filter on P:

$$P_{L_0}(w \mid u) \propto l(u, w) \times P(w) \qquad (1)$$

In our example, we consider the prior P to be uniform, and thus, the family of distributions $P_{L_0}(w \mid u)$ is obtained by normalising each row of the above table:

w	5	6	7
‘John ate 5 cookies’	1/3	1/3	1/3
‘John ate 6 cookies’	0	1/2	1/2
‘John ate 7 cookies’	0	0	1

2.2 The speaker model

According to RSA, the speaker **S** is modelled as an agent which produces a distribution over utterances for each world state w that **S** might wish to convey:

$$P_{S_1}(u \mid w) \propto \exp[\alpha \times (\log(P_{L_0}(w \mid u)) - C(u))]$$

Proceedings of the 26th Workshop on the Semantics and Pragmatics of Dialogue, August, 22-24, 2022, Dublin.

or, equivalently:

$$P_{S_1}(u \mid w) \propto \frac{P_{L_0}(w \mid u)^{\alpha}}{e^{\alpha C(u)}} \qquad (2)$$

Each parameter $C(u)$ represents the cost of uttering u. The role of the parameter α (where it is assumed that $\alpha > 1$), is to exacerbate the differences of literal fit among utterances. For α tending to infinity, **S** chooses the utterance with the highest utility $U(u, w) = \log(P_{L_0}(w \mid u) - C(u))$ with a probability of 1 (i.e., stochastic certainty).

In our example, if we let $\alpha = 4$, then we obtain the family of distributions $P_{S_1}(u \mid w)$ by, first, exponentiating by 4, and, second, normalising each *column*. Dividing by the exponentiated cost has no effect on the resulting distribution because it is a constant that vanishes after normalising.

w	5	6	7
'John ate 5 cookies'	1	0.16	0.01
'John ate 6 cookies'	0	0.84	0.06
'John ate 7 cookies'	0	0	0.93

2.3 The pragmatic listener model

The pragmatic listener model P_{L_1} refers to the above speaker model when updating the distribution over world states:

$$P_{L_1}(w \mid u) \propto P_{S_1}(u \mid w) \times P(w) \qquad (3)$$

Since P is uniform in our example, we obtain the family of distributions $P_{L_1}(w \mid u)$ by once more normalising each row.

w	5	6	7
'John ate 5 cookies'	0.85	0.14	0.01
'John ate 6 cookies'	0	0.93	0.07
'John ate 7 cookies'	0	0	1

This example illustrates some noteworthy points. First, *the core aspect of computing RSA models is the application of normalisation steps*. While the normalisation factors of the nested speaker and listener models are therefore crucial, they are left implicit by the usual formulaic presentation of RSA (Eqs. (1) to (3)). We will see below that making these factors explicit brings insight.

Second, the formulation of RSA in terms of a listener who reasons about a speaker who reasons about a literal listener makes it difficult to build an intuition of what the model predicts. For instance, in our example, RSA predicts, to a large extent, that 'John ate x cookies' implicates 'John ate exactly x cookies'. But does it predict a similar implicature for variations of the same example, for instance using another prior? One might intuitively expect it to do so, but, as we demonstrate below, it does not always make this prediction.

Third, a pragmatic listener is conceived of as reasoning about all possible combinations of world states and utterances simultaneously, which is large for any non-trivial example. Indeed, it is psychologically implausible that such a process is at play in the listener's mind.

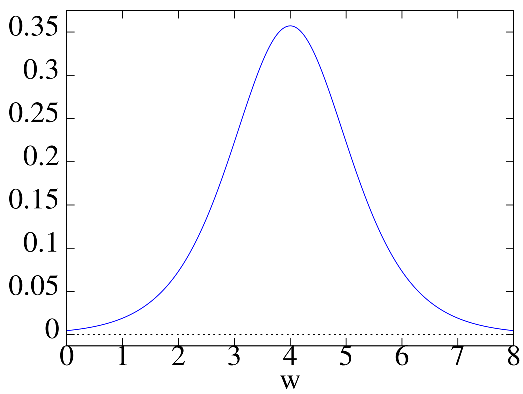
The purpose of the next section is to reformulate RSA in terms that are easier to grasp, while addressing these weaknesses.

3 Information-theoretic reformulation

We carry out our reformulation in terms of information theoretic concepts; in particular, information gain. To illustrate our points, we use a variation of the example from the previous section, in which the alternative utterances differed along some numerical value which provided a lower bound on compatible world states. The main differences in our current example will be the following:

- The relevant numerical variable is now continuous. Thus alternative utterances now have the form 'John ran x kilometres'.

- The prior distribution over world states (i.e., over the number of kilometres John ran) is no longer uniform. We instead use a logistic distribution, defined below. (A logistic distribution is similar to a normal distribution, but it simplifies our calculations.[1] Our qualitative conclusions will hold just as well in the case of a prior which is normally distributed.)



[1]The logistic distribution is leptokurtic. That is, it has fatter tails than the normal distribution, i.e., more outliers.

The above plot represents a prior for the number of kilometres John ran given by a logistic distribution with a mean of 4. Thus to obtain $P_{L_0}(w \mid u)$ for any u of the form 'John ran x kilometres', one should crop this distribution at x (on the left, so only the right part remains) and renormalise.

3.1 Literal information gain

The instrumental concept underlying our information-theoretic reformulation is the Kullback–Leibler (K-L) divergence, which measures the information gained when updating a prior belief taking a distribution P to a posterior belief taking distribution Q. It is defined as follows:[2]

$$D_{\mathrm{KL}}(Q \parallel P) = -\sum_{x \in \mathcal{X}} Q(x) \log \left(\frac{P(x)}{Q(x)} \right)$$

In terms of this definition, we may compute the *literal* information gain provided by an utterance u as the K-L divergence between the prior on worlds P, and the posterior $Q_u(w) = P_{L_0}(w \mid u)$ computed by L_0 (which takes u literally):

$$Q_u(w) \propto l(u, w) \times P(w)$$
$$G_{L_0}(u) = D_{\mathrm{KL}}(Q_u \parallel P)$$

Because $l(u, w)$ takes 0 or 1 values, the following reformulation of G_{L_0} is possible, by Theorem 1 (given in Appendix A):

$$G_{L_0}(u) = -\log \sum_{w \in \mathcal{W}} l(u, w) \times P(w) \quad (4)$$

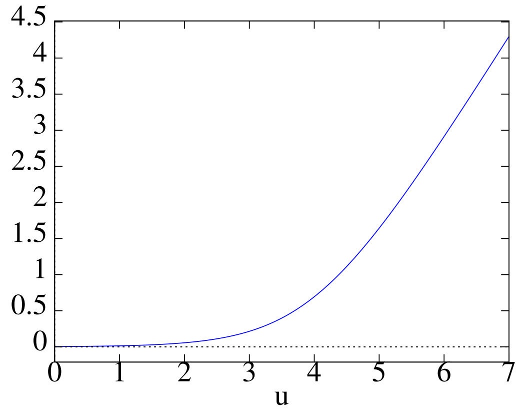
For an alternative, more compact presentation of G_{L_0}, one may first define the following prior over utterances associated with the literal listener:

$$P_{L_0}(u) = \sum_{w \in \mathcal{W}} l(u, w) \times P(w) \quad (5)$$

That is, P_{L_0} is the probability associated with u by L_0, given the prior over world states. G_{L_0} may then, instead, be rendered as follows:

$$G_{L_0}(u) = -\log P_{L_0}(u) \quad (6)$$

In our running example, which uses a logistic prior, we then have the following information gain for the literal listener (G_{L_0}) for utterances of the form 'John ran x kilometres':



The flat regime toward the left of the plot is explained by the fact that utterances of 'John ran x kilometres', where x is lower than the mean of the prior, do not provide much information: they are compatible with most world states. The steadily increasing gain after the mean is explained by the converse: these utterances are *incompatible* with most world states. Furthermore, in this part of the plot, a given increase in the value of x leads to a roughly constant increase in information gain. Thus the information gain here increases in a roughly linear relationship with utterance strength. Such an increase, in turn, rules out a roughly constant *proportion* of the remaining possible world states, given the log scale associated with information gain.

3.2 The reformulated speaker model

With the above notion of information gain in mind, we can make the normalisation factor in P_{L_0} of Eq. (1) explicit, thus turning the proportionality relation into an equality:

$$\begin{aligned} P_{L_0}(w \mid u) &= \frac{l(u, w) \times P(w)}{\sum_{w_1 \in \mathcal{W}} l(u, w_1) \times P(w_1)} \\ &= \frac{l(u, w) \times P(w)}{P_{L_0}(u)} \\ &= l(u, w) \times P(w) \times e^{G_{L_0}(u)} \end{aligned}$$

One can now substitute $P_{L_0}(w \mid u)$ by $l(u, w) \times P(w) \times e^{G_{L_0}(u)}$ in the definition of P_{S_1} (Eq. (2)), and simplify the result. Note that making the normalisation factor explicit was necessary to carry out such a substitution, which is only valid for strictly

[2] The notation '$P(w)$' normally suggests that the described distribution is discrete, in which case weighted averages are computed with a sum. We will use these notations throughout, since they are easier to present than density functions and integrals, a choice we make even though our running example uses a continuous variable.

equal (not just proportional) terms.

$$\begin{aligned} P_{S_1}(u \mid w) &\propto \frac{P_{L_0}(w \mid u)^\alpha}{e^{\alpha \times C(u)}} \\ &\propto \frac{\left(l(u,w) \times P(w) \times e^{G_{L_0}(u)}\right)^\alpha}{e^{\alpha \times C(u)}} \\ &\propto l(u,w)^\alpha \times P(w)^\alpha \times e^{\alpha \times (G_{L_0}(u) - C(u))} \\ &\propto l(u,w) \times e^{\alpha \times (G_{L_0}(u) - C(u))} \end{aligned} \quad (7)$$

The last step of rewriting (7) is justified because (i) the exponent α of $l(u,w)$ has no effect, since $l(u,w) \in \{0,1\}$, and (ii) the term $P(w)$ does not depend on u, and thus does not affect the proportionality relation.

At this point, we may introduce our reformulation of the speaker model as a filter on a *pragmatic prior*. We do so by first defining this prior; in particular, we may view the second factor in (7) as proportional to a speaker's *pragmatic prior over utterances*:[3]

$$P_{S_1}(u) \propto e^{\alpha \times (G_{L_0}(u) - C(u))} \quad (8)$$

Note that this reformulation shows that the speaker *a priori* favours utterances whose information gains are larger than their costs, a preference which is exacerbated by high values of α.[4] We may now formulate $P_{S_1}(u \mid w)$ as a *filter* on the above prior, provided by $l(u,w)$:

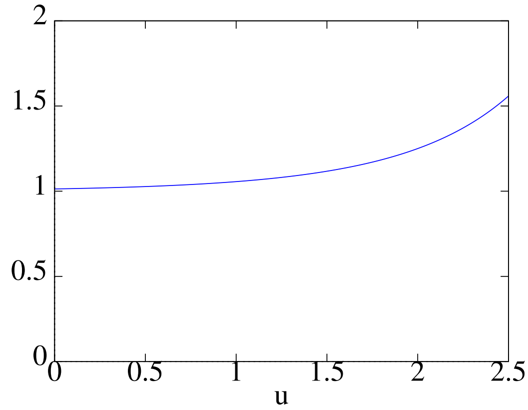
$$P_{S_1}(u \mid w) \propto l(u,w) \times P_{S_1}(u) \quad (9)$$

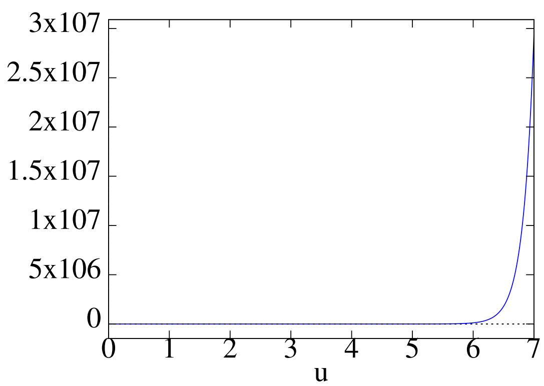
As an algorithmic model of the speaker's reasoning, the above proportionality relation can be seen as implying that a "table" of utterances and their relative degrees of preferredness (according to informativeness and cost) has been constructed *a priori*. Upon choosing a world state w to communicate, the speaker may then filter out those utterances incompatible with w, in order to then select an utterance among those that remain. In other words, the utility of an utterance ($U(u) = G_{L_0}(u) - C(u)$) is *independent of the world state w that the speaker wishes to communicate*.

[3] Given an infinite set of possible alternative utterances (and, indeed, in our running example), $P_{S_1}(u)$ need not define a probability distribution. This is not a problem in practice, as the speaker and listener posteriors, $P_{S_1}(u \mid w)$ and $P_{L_1}(w \mid u)$, will nevertheless be proper probability distributions.

[4] One can additionally choose cost to be proportional to utterance length, following Lassiter and Goodman (2013). Given such a definition of cost, the speaker will prefer utterances with a high (literal) information density.

In our running example, $P_{S_1}(u \mid w)$ is thus obtained by cropping $P_{S_1}(u)$ on the right (and then renormalising). The following plots exemplify $P_{S_1}(u \mid w)$ for two different values of w, prior to normalisation: 2.5, and 7 (where $\alpha = 4$ in each case).





As can be seen, $P_{S_1}(u|w)$ is nearly constant for low values of u, but it shoots up exponentially once u exceeds the mean of the prior over world states. In other words, if choosing an utterance whose strength exceeds the mean of the prior is at all possible, then the speaker will most definitely do so. If only utterances whose strength is below the mean are possible, then the speaker will still be biased towards stronger utterances, but not to the same degree.

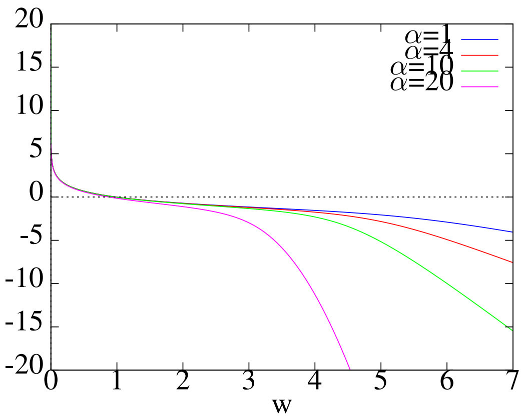
3.3 Normalisation via information gain

The above formulation of P_{S_1} provides only a proportionality relation, which needs to be normalised, in order to obtain a full definition. To do so, we may apply the same information-theoretic treatment to the speaker model as we did to the literal-listener model. First, we make explicit the normalisation factor in P_{S_1}; then, we encode this factor in terms of information gain, relying again on Theorem 1:[5]

$$P_{S_1}(u|w) = \frac{l(u,w) \times P_{S_1}(u)}{\sum_{u_1} l(u_1,w) \times P_{S_1}(u_1)}$$
$$= l(u,w) \times P_{S_1}(u) \times e^{G_{S_1}(w)} \quad (10)$$

Here, $G_{S_1}(w)$ is the information gain on the distribution P_{S_1} provided by w. This gain is high if $l(u,w)$ allows the speaker to discard many utterances u, where $P_{S_1}(u)$ is high. We refer to $G_{S_1}(w)$ as the *specificity* of w; in general, the function G_{S_1} is determined by the pragmatic interpretation problem $(\mathcal{U}, \mathcal{W}, P, C, l)$, together with the model temperature α.

In our running example, we obtain the following contours of specificity for various values of α:[6]



These curves can be analysed as sequences of three different regimes. First, there is asymptotic behaviour around 0: values near 0 are nearly impossible by virtue of excluding most *a priori* possible utterances, and thus they provide an information gain tending to infinity. The transition to the next regime happens very quickly, around 0.2. The middle regime is a small slope with a roughly flat decrease, which lasts up to around the mean of the prior distribution. Above the mean of the prior, the third regime kicks in: there is another roughly flat decrease, but, this time, with a much larger slope. The difference in slope is explained by the following two facts: (i) that the literal information gains associated with utterances increase drastically above the mean of the prior (see the plot of (Eq. (6))), and (ii) that these information gains enter into the calculation of specificity for world states above the mean, as these world states become compatible with more utterances. Moreover, for large values of α, this slope is more pronounced.

The reader may find it odd that there are negative specificities in these plots, given that the distributions that these specificities come from are obtained as *filters* of the pragmatic prior over utterances. Negative values appear because P_{S_1} is not, strictly speaking, a *probability distribution* over utterances (see Footnote 3). Fortunately, negative specificities do not pose a problem in practice. For example, once we get to the pragmatic listener model, they may be seen as having been shifted by a positive constant during normalisation (given that the posterior itself will be multiplied by a constant).

3.4 The reformulated pragmatic listener model

If we now substitute the definition of the speaker model of Eq. (10) into the definition of the pragmatic listener model of Eq. (3), we may obtain the following new definition of the latter:

$$\begin{aligned} P_{L_1}(w \mid u) &\propto P_{S_1}(u \mid w) \times P(w) \\ &\propto l(u,w) \times P_{S_1}(u) \times e^{G_{S_1}(w)} \times P(w) \\ &\propto l(u,w) \times e^{G_{S_1}(w)} \times P(w) \quad (11) \end{aligned}$$

The justification for removing the term $P_{S_1}(u)$ in the fourth line is the fact that u is fixed, and thus the proportionality relation does not depend on it. Now note that we may define the following *pragmatic prior* for the listener model:

$$P_{L_1}(w) = e^{G_{S_1}(w)} \times P(w) \quad (12)$$

Given Eq. (11), the pragmatic listener model may therefore instead be presented as a *filter* on this prior:

$$P_{L_1}(w|u) \propto l(u,w) \times P_{L_1}(w) \quad (13)$$

Equation (13) constitutes the fully reformulated RSA model. Reading it out, we see that L_1 chooses the distribution over world states in a way very

[5]Note that in practice, P_{S_1} in (10) may be substituted by the right-hand side of the proportionality relation in (8), since the relevant normalisation factor is cancelled out.

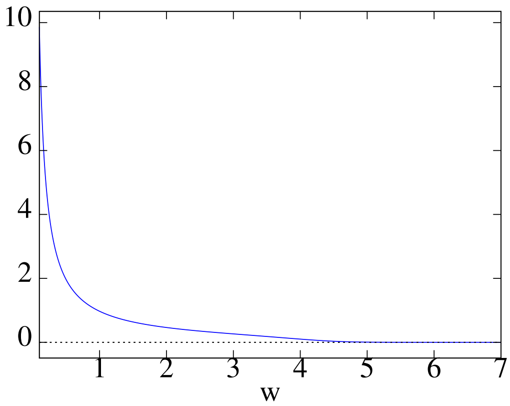
[6]Computed by taking

$$\begin{aligned} & G_{S_1}(w) \\ = & -\log(\sum_{u_1} l(u_1,w) \times P_{S_1}(u_1)) \\ = & -\log(\sum_{u_1} l(u_1,w) \times \frac{e^{\alpha \times (G_{L_0}(u_1) - C(u_1))}}{k}) \quad \text{(by (8))} \\ = & -\log(\sum_{u_1} l(u_1,w) \times e^{\alpha \times (G_{L_0}(u_1) - C(u_1))}) + \log(k) \end{aligned}$$

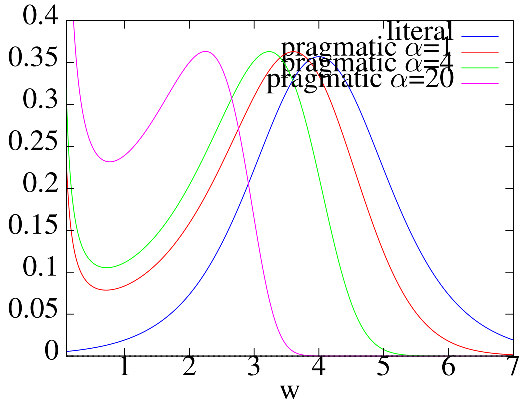
and choosing $\log(k) = 0$ (or $k = 1$). Note that the resulting contours are therefore independent of the normalisation factor k.

similar to L_0. Namely, both merely apply a filter to some prior. In the case of L_0, the relevant prior over world states w is $P(w)$, i.e., the "literal" prior; L_1, instead, uses the pragmatic prior $P_{L_1}(w)$, which multiplies the literal prior by a measure of specificity, defined as the exponentiated information gain associated with the pragmatic speaker. In sum, the RSA model has it that all pragmatic effects are attributable to the relative specificity of world states, relative to the set of possible utterances.

In our running example, the factor $e^{G_{S_1}(w)}$ contributing specificity has the following shape:

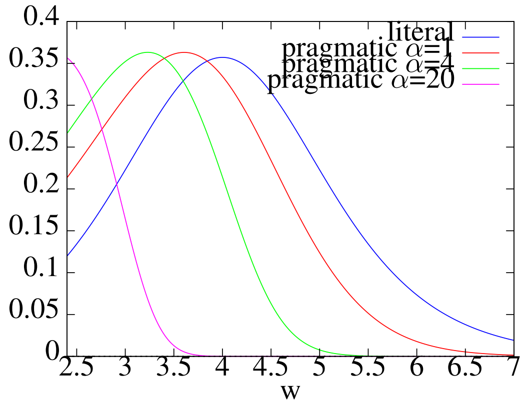


The following plot of P_{L_1} illustrates the effect that this factor has on the prior $P(w)$ over world states w (for various values of α):



The different regimes of specificity can thus be seen to have the following effects on P_{L_1}. Toward the left of the plot, there is an asymptote greatly favouring tiny values of w. Meanwhile, the right of the plot is essentially zeroed out, with a smooth transition, and the peak has also shifted leftward. Thus the bulk of the distribution is shifted to the left, in comparison to the prior P. This shift is larger when α is large; indeed, we would expect that for large enough values of α, the peak will "merge" with the asymptotic behaviour around zero. Unfortunately, our numeric tool cannot handle very large values of α, so we are not able to produce the corresponding plot.

Analogous to the literal listener, one can obtain $P_{L_1}(w \mid u)$ for any u by cropping the pragmatic listener's prior distribution on the left of the plot and renormalising. Consider, for instance, the utterance 'John ran 2.4 kilometres'. In this case, we crop the prior distribution at 2.4:



We indeed observe the effect of a Gricean implicature when $\alpha = 20$, insofar as the mode of the resulting distribution is the value uttered. This effect occurs due to the fact that the pragmatic prior distribution for this value of α happens to have a sharp drop after the uttered value. For lower values of α, however, the observed pragmatic effect is merely to shift the distribution to the left, relative to the literal prior. What results does not, in any obvious way, reflect a Gricean implicature.

4 Discussion

4.1 Algorithmic plausibility

At first glance, the psychological plausibility of RSA as an algorithmic model (in the sense of Marr (1982)) seems highly suspect; for example, it requires the pragmatic listener to consider all compatible combinations of world states and utterances on each occasion of utterance interpretation (though see, e.g., White et al. (2020); Zaslavsky et al. (2021) for recent attempts to address the psychological principles grounding RSA models). In principle, the space of world states includes all those literally compatible with the observed utterance, requiring the listener to deal with a very large space of possibilities, in order to interpret a single utterance. Because our reformulation of RSA as a mere filter on a prior is *functionally equivalent* to RSA as traditionally conceived, it provides a new lens into the issue of algorithmic plausibility. Neither the computation of the literal listener model, nor the computation of the pragmatic speaker model need directly enter into the pragmatic listener's computation of a posterior distribution. Instead, the contributions of the literal listener and pragmatic speaker in the original formulation of RSA are now *repackaged* as part of the prior. As a result, these contributions may be learned and then "memorised";[7] utterance interpretation, meanwhile, becomes a process of merely eliminating alternatives from this re-conceived prior. (Likewise for the pragmatic speaker, whose prior over utterances need not depend on the world state that it wishes to communicate.)

A consequence of this fact is that literal interpretations and pragmatic interpretations (as according to RSA) may be viewed as updates of the same kind: given an utterance and the right prior, both styles of interpretation involve the elimination of world states incompatible with the utterance, followed by a renormalisation step. From this perspective, the RSA framework is not committed to a particular algorithmic implementation of the pragmatic interpretation process beyond what would be required for literal interpretation.

Relatedly, we do not require 3-d (or density) plots with axes representing utterance strength and world state, respectively, in order to illustrate the effect obtained by a pragmatic listener from sequential renormalisation steps. Our theoretical result may thus be framed as the observation that, in such a 3-d plot, and for a semantics of the sort $u \leq w$, any 2-d *slice* acquired by fixing a value for the utterance u is just like the slices associated with weaker utterances, but for a step of cropping and renormalisation.

4.2 Implicature

As mentioned in Section 3.4, the implicature expected based on Grice's Cooperative Principle (in particular, Quantity) is not obtained by the pragmatic listener model in our running example. The expected implicature is an "exactly" interpretation associated with the numeral occurring in the utterance, while what the model obtains is merely a decrease in the mode of the posterior, in comparison to the prior. (This result persists even for large values of α.)

Nevertheless, we can show that the expected implicature occurs when α tends to infinity.[8] This is because the utterance 'John ran $x + \varepsilon$ kilometers' is always, if only slightly, more informative than 'John ran x kilometers'; thus it will always be preferred by the pragmatic speaker for legitimate values of α, if only by a small amount. As a result, only $w = u$ will be admissible by a pragmatic listener, in the limit, where probabilistic choice becomes categorical.

Intriguingly, the theoretical result that the implicatures expected are not always generated may, in fact, reflect some aspects of real human behaviour in certain settings. For instance, Sikos et al. (2021a) found that, even in the non-interactive one-shot games against which RSA models have been most extensively tested, consistency with human performance was driven by cases in which non-Gricean behaviour was, in fact, predicted by the model. In such cases, the RSA model's prior overrode the pragmatic effects associated with specificity. Thus perhaps ironically, RSA's failure to predict Gricean implicatures may sometimes contribute to its empirical successes. As Sikos et al. note, however, RSA does not reflect human behaviour better than a literal semantic model does on such tasks, making it difficult to consider this property a boon.

5 Conclusions and future Work

RSA models can be critiqued on a number of fronts, both theoretical and empirical. In the present work, we have focused on the algorithmic nature of the model, in order to show that it may be reformulated in a manner which appears relatively attractive from a psychological perspective (and which is easier to compute in simulations). We have also shown that doing so brings to the fore the model's unexpected (i.e., non-Gricean) behaviour when it is faced with certain priors: a single plot illustrates the relationship between utterance and posterior in the pragmatic listener model, revealing limits on the conditions under which expected implicatures are actually generated.

On the empirical side, critics of RSA have emphasised the artificial and non-interactive nature of the tasks used to verify the model's predictions, pointing out, for example, that in less constrained contexts, people often produce non-optimal utter-

[7] A reviewer points out that our reformulation of RSA in terms of a pragmatic prior generates questions about how such a prior might be learned in the first place. While we won't provide an account of semantic learning here, we note that the two terms in (12) representing information gain and the prior over worlds suggest that they may be learned independently of one another.

[8] That is, as α tends to infinity, $P_{S_1}(u \mid w) = \delta_{u-w}$, where δ is the Dirac δ function.

ances, e.g., by over-specification (Gatt et al., 2013). Sikos et al. (2021a) have provided evidence suggesting that, even in the restricted reference game domain, people's judgements do not always accord with rational choice, as defined by the model. These authors show that, while speakers behave as the model would predict, listeners do not, and a baseline literal listener model outperforms RSA. It has also been argued that, even where RSA provides a good fit to human data, it does so when other parameters, such as utterance cost, are implausible (Wilcox and Spector, 2019). Moreover, simulation models have called into question whether or not reasoning over an interlocutor's intentions is generally necessary, if, for example, a repair mechanism is available (Van Arkel et al., 2020).[9]

By providing a functionally equivalent reformulation of RSA, we have shown that, for both the pragmatic listener and speaker models, the *merit* of a given world state or utterance can be expressed and evaluated in its own terms, making both models analogous to the literal listener model. Succinctly, RSA models are just *filters* of some chosen prior, in which merit is predetermined. It is thus straightforward to imagine a generalisation of our reformulation in which information gains are not always computed "rationally". Rather, according to such hypothetical alternative models, a speaker might compute a more approximate information gain for each utterance and act accordingly. Similarly, a listener might compute a more approximate notion of specificity with respect to a set of possible utterances, which may then be used to tweak the prior. Such "approximately rational" priors might then be refined over time as more pragmatic problems are encountered.[10] We leave this possibility for future work.

Acknowledgements

The research reported in this paper was supported by grant 2014-39 from the Swedish Research Council, which funds the Centre for Linguistic Theory and Studies in Probability (CLASP) in the Department of Philosophy, Linguistics and Theory of Science at the University of Gothenburg. We thank the anonymous reviewers for their useful comments on an earlier draft of the paper.

[9]Such repair mechanisms are well-studied and highly prevalent in natural dialogue (Schegloff et al., 1977; Dingemanse et al., 2015).

[10]Such a generalised model is supported by findings in human-human reference games where "familiarity with the communicative setting can influence the degree of rationality that listeners realise" (Sikos et al., 2021b, p.1471).

References

Guanyi Chen, CJ van Deemter, and Chenghua Lin. 2018. Modelling pro-drop with the rational speech acts model. In *Proceedings of the 11th International Conference on Natural Language Generation*, pages 57–66. Association for Computational Linguistics (ACL).

Mark Dingemanse, Seán G Roberts, Julija Baranova, Joe Blythe, Paul Drew, Simeon Floyd, Rosa S Gisladottir, Kobin H Kendrick, Stephen C Levinson, Elizabeth Manrique, Giovanni Rossi, and N. J. Enfield. 2015. Universal principles in the repair of communication problems. *PloS one*, 10(9):e0136100.

Michael C. Frank and Noah D. Goodman. 2012. Predicting pragmatic reasoning in language games. *Science*, 336:998 – 998.

Michael C. Frank and Noah D. Goodman. 2014. Inferring word meanings by assuming that speakers are informative. *Cognitive Psychology*, 75:80–96.

Albert Gatt, Roger PG van Gompel, Kees van Deemter, and Emiel Krahmer. 2013. Are we Bayesian referring expression generators? In *Proceedings of the Annual Conference of the Cognitive Science Society*. Cognitive Science Society.

H.P. Grice. 1975. Logic and conversation. *Syntax and Semantics*, 3:41–58.

Daniel Lassiter and Noah D. Goodman. 2013. Context, scale structure, and statistics in the interpretation of positive-form adjectives. *Semantics and Linguistic Theory*, 23(0):587–610. Number: 0.

David Marr. 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. MIT Press, Cambridge.

E.A. Schegloff, Gail Jefferson, and Harvey Sacks. 1977. The preference for self-correction in the organization of repair in conversation. *Language*, 53(2):361–382.

Gregory Scontras and Noah D. Goodman. 2017. Resolving uncertainty in plural predication. *Cognition*, 168:294–311.

Gregory Scontras, Michael Henry Tessler, and Michael Franke. A practical introduction to the rational speech act modeling framework. manuscript.

Les Sikos, Noortje J Venhuizen, Heiner Drenhaus, and Matthew W Crocker. 2021a. Reevaluating pragmatic reasoning in language games. *PloS one*, 16(3):e0248388.

Les Sikos, Noortje J Venhuizen, Heiner Drenhaus, and Matthew W Crocker. 2021b. Speak before you listen: Pragmatic reasoning in multi-trial language games.

In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 43.

Jacqueline Van Arkel, Marieke Woensdregt, Mark Dingemanse, and Mark Blokpoel. 2020. A simple repair mechanism can alleviate computational demands of pragmatic reasoning: Simulations and complexity analysis. In *the 24th (Virtual) Conference on Computational Natural Language Learning (CoNLL 2020)*, pages 177–194. ACL.

Julia White, Jesse Mu, and Noah D. Goodman. 2020. Learning to refer informatively by amortizing pragmatic reasoning. ArXiv:2006.00418 [cs].

Ethan Wilcox and Benjamin Spector. 2019. The role of prior beliefs in the rational speech act model of pragmatics: Exhaustivity as a case study. In *Proceedings of the Annual Conference of the Cognitive Science Society*, pages 3099–3105.

Noga Zaslavsky, Jennifer Hu, and Roger P. Levy. 2021. A Rate–Distortion view of human pragmatic reasoning? In *Proceedings of the Society for Computation in Linguistics 2021*, pages 347–348, Online. Association for Computational Linguistics.

A Proofs

Theorem 1. *If $P(x) \propto f(x) \times Q(x)$, and the codomain of f is $\{0, 1\}$, then*

$$D_{KL}(P \parallel Q) = -\log\left(\sum_x f(x) \times Q(x)\right)$$

Proof. Let $P(x) = \alpha f(x) \times Q(x)$, with α constant.

First, we have

$$\alpha = \frac{1}{\sum_x f(x) \times Q(x)} \quad (14)$$

Indeed, P is a distribution, and we have

$$\begin{aligned} \sum_x P(x) &= 1 \\ \sum_x \alpha f(x) \times Q(x) &= 1 \\ \alpha \sum_x f(x) \times Q(x) &= 1 \end{aligned}$$

Second, we have

$$\alpha f(x) \times \log(\alpha f(x)) = \alpha f(x) \times \log(\alpha) \quad (15)$$

This can be seen by case analysis.

- If $f(x) = 0$, then

$$\begin{aligned} & \alpha f(x) \times \log(\alpha f(x)) \\ = \; & 0 \\ = \; & \alpha f(x) \times \log(\alpha) \end{aligned}$$

  The first equality follows from the fact that, in general, $\lim_{a \to 0}(a \times log(a)) = 0$.

- If $f(x) = 1$, then

$$\begin{aligned} & \alpha f(x) \times \log(\alpha f(x)) \\ = \; & \alpha \log(\alpha) \\ = \; & \alpha f(x) \times \log(\alpha) \end{aligned}$$

Using the above two facts, we can compute:

$$\begin{aligned} D_{\text{KL}}(P \parallel Q) &= \sum_x P(x) \times \log(\frac{P(x)}{Q(x)}) \\ &= \sum_x Q(x) \times \alpha f(x) \times \log(\alpha f(x))) && \text{by def of } P \\ &= \sum_x Q(x) \times \alpha f(x) \times \log(\alpha) && \text{by Eq. (15)} \\ &= \alpha \log(\alpha) \times \sum_x f(x) \times Q(x) \\ &= \alpha \log(\alpha) \times \alpha^{-1} && \text{by Eq. (14)} \\ &= \log(\alpha) \\ &= \log(\frac{1}{\sum_x f(x) \times Q(x)}) && \text{by Eq. (14)} \\ &= -\log(\sum_x f(x) \times Q(x)) \end{aligned}$$

□